



UnB

UNIVERSIDADE DE BRASÍLIA

Faculdade de Medicina

Programa de Pós-Graduação em Ciências Médicas

DISSERTAÇÃO DE MESTRADO

**Integração de Estratégias Computacionais e Experimentais para a
Investigação da Proteína Ligante de Odor (OBP) Salivar de
Rhodnius neglectus, Vetor da Doença de Chagas**

Gabriel dos Santos Silva

Brasília

2023

**Integração de Estratégias Computacionais e Experimentais para a
Investigação da Proteína Ligante de Odor (OBP) Salivar de
Rhodnius neglectus, Vetor da Doença de Chagas**

Dissertação submetida ao Programa de Pós-Graduação em Ciências Médicas da Universidade de Brasília como requisito parcial para a obtenção do título de Mestre em Ciências Médicas, na área de concentração Ciências Aplicadas em Saúde.

Orientador(a): Professora Dra. Carla Nunes de Araújo

Coorientador(a) Professora Dra. Melina Mottin

Brasília – DF
Setembro, 2023

Ficha de identificação da obra [Elemento obrigatório.]

[Insira neste espaço a ficha de identificação da obra.]

Gabriel dos Santos Silva

Integração de Estratégias Computacionais e Experimentais para a Investigação da Proteína Ligante de Odor (OBP) Salivar de *Rhodnius neglectus*, Vetor da Doença de Chagas

O presente trabalho em nível de Mestrado foi avaliado e aprovado, em 06 de outubro de 2023 pela banca examinadora composta pelos seguintes membros:

Presidente: Prof.(a) Dr.(a) Carla Nunes de Araújo
Universidade de Brasília

Membro 1: Prof. Dr. Clever Gomes Cardoso
Universidade Federal de Goiás

Membro 2: Prof.(a) Dr.(a) Paula Beatriz de Medeiros Santiago
Universidade de Brasília

Suplente: Prof.(a) Dr.(a) Aisel Valle Garay
Instituição Universidade de Brasília

Certificamos que esta é a versão original e final do trabalho de conclusão que foi julgado adequado para obtenção do título de Mestre em Ciências Médicas.

Insira neste espaço a
assinatura digital

Coordenação do Programa de Pós-Graduação

Insira neste espaço a
assinatura digital

Prof.(a) Dr.(a) Carla Nunes de Araújo
Orientador(a)

Brasília, 2023.

Dedico esse trabalho em memória ao Rafão, descanse em paz
E ao meu Grande Amor.

AGRADECIMENTOS

À minha orientadora, Profa. Dra. Carla Nunes de Araújo, por todos os anos em que me orientou, sendo essa pessoa maravilhosa que muito me inspira. Obrigado pela confiança, pelo exemplo, pelo carinho e por todos os ensinamentos.

À minha coorientadora, Profa. Dra. Melina Mottin, que me introduziu à bioinformática e sempre esteve por perto para me ajudar com minhas dúvidas.

A Dra. Paula Beatriz Santiago, por toda ajuda, dicas, conselhos e palavras de motivação que foram dadas.

A Profa. Dra. Flavia Silva Nader Motta e Prof. Dr. Gabriel, por todas as dicas sobre metodologias, experimentos e conselhos.

A Profa. Dra. Izabela Marques Dourado Bastos Charneau, por ter me apresentado à Dra. Carla e ao LIPH.

Ao Prof. Dr. Jaime Martins de Santana, pelos conselhos e acolhimento.

Ao Professor Dr. Eric Faudry, por ter me dado inspiração, dicas e verdadeiras aulas durante o curto período em que estive no Laboratório de Interação Parasito Hospedeiro.

Ao Prof. Dr. Aisel Valle Garay, por toda a orientação, os ensinamentos e pela paciência; o senhor me inspira muito.

Ao Prof. Napoleão Fonseca Valadares, por ter despertado em mim um grande interesse em ciência e por mostrar outra forma de enxergar o trabalho em laboratório.

À Me. Caroline Barreto Vieira, que representou o melhor primeiro contato possível dentro de um laboratório, sempre demonstrando paciência e carinho ao explicar cada metodologia.

Ao Dr. Lucas Oliveira, por ter me ensinado e dado ideias valiosas sobre bioquímica e identificação de proteínas.

Aos preferidos da professora, Me. Kaio Luis da Silva Bentes e Dra. Yanna Praça Reis, por toda ajuda e todos os conselhos dados durante minhas iniciações científicas e mestrado.

Aos Professores do curso de Ciências Biológicas.

Aos técnicos Nuno e Jaques Miranda, do laboratório de bioquímica, pela ajuda com o servidor e com os espectrômetros de massas.

Aos amigos de bancada, Me. Alessandro, Me. Alexandra, Me. Allan, Andrey, Me. Arthur, Beatriz, Dra. Clênia, Dani, Dr. Gilberto, Me. Giulia, Guilherme, Me.

Gabriela, Juliana, Natalia, Dra. Natália, Renan, Sophia, Veronica, Me. Tarcila, Me. Tati, por todos os momentos de parceria no LIPH.

Aos meus tios e prima, Sandro, Lara e Luiza, que são verdadeiros anjos na minha vida; amo vocês!

À minha velhinha e ao meu irmão, Gislaíne e Caleb, as pessoas que mais contribuíram para eu me tornar quem sou. Agradeço por tudo que fizeram por mim e por tudo que representam; amo vocês!

Aos meus amigos mais ou menos: Roberta, Miguel, Carol, Pirex, Gabi, Gustavo, Danilo, Chris, Julliana Eloi, Suzane, Adriel, Matheus, Bruno, Lara, Paulinho, Nathan, Giovane, Diego, Juliano, André, Xavier, Chris², Nina, Hagui, Anderson, Paty, Matta, Eric.

Aos meus amigos de verdade, Fontenele e Renato.

À minha segunda família, Wesley, Lucivânia, Sophia, Eric e Karen; vocês são muito importantes para mim. Obrigado por me aguentarem todo esse tempo e por me aceitarem na família.

À minha companheira, Karen, quero expressar minha gratidão por estar ao meu lado em todos os momentos e por me apoiar em todas as decisões; meu amor por você é imenso!

Às agências de fomento, FAPDF (00193-00000825/2021-19), CAPES, CNPQ e à PPGCM, desejo agradecer pelo apoio fundamental para a realização do meu mestrado.

Aos membros da banca, quero reconhecer o tempo dedicado à leitura do meu trabalho e às valiosas contribuições que ofereceram.

*“Tá pensando o que? Rapadura é doce, mas
não é mole, não!”.*

– Yusuke Urameshi.

RESUMO

O *Rhodnius neglectus*, um inseto da família Reduviidae (Hemiptera), é um potencial vetor do *Trypanosoma cruzi*, o protozoário responsável pela doença de Chagas. Na saliva do *R. neglectus*, encontram-se moléculas importantes que neutralizam os mecanismos de coagulação do hospedeiro, auxiliando na alimentação sanguínea e, indiretamente, na transmissão do protozoário. Por isso, a saliva desse inseto hematófago tem sido alvo relevante de estudos devido à presença de moléculas com alto potencial biotecnológico. Este trabalho teve como objetivo estudar uma Proteína de Ligação a Odorantes da saliva do *R. neglectus* (RnOBP) visando contribuir na compreensão de suas características estruturais e funcionais, utilizando abordagens computacionais e experimentais. Foram realizadas modelagens moleculares da RnOBP e ensaios computacionais de dinâmica molecular e clusterização para obter informações sobre a flexibilidade da proteína. Em seguida, foram realizados ensaios de *docking* molecular para verificar a afinidade da RnOBP com alguns ligantes relacionados à coagulação de vertebrados. Por fim, observou-se uma possível forma de oligomerização da RnOBP. Para a obtenção da RnOBP recombinante, a sequência nucleotídica da proteína foi obtida a partir do transcriptoma do *R. neglectus*, clonada no vetor de expressão pET100/D-TOPO e produzida por bactérias *Escherichia coli* Rosetta(DE3)pLysS. A RnOBP recombinante foi purificada por cromatografia de afinidade, renovelada e utilizada em ensaios de caracterização biofísica. O modelo gerado pelo AlphaFold mostrou que a RnOBP é formada por hélices α e possui uma cavidade interna significativa, potencialmente um sítio de ligação. Em comparação com outras OBPs de insetos, independentemente da composição ou classificação da OBP, a região do sítio de ligação é bem conservada. A RnOBP recombinante purificada a partir de corpos de inclusão revelou uma composição de estrutura secundária diferente da esperada. Os ensaios funcionais iniciais mostraram que a rRnOBP pode atuar na inibição da agregação plaquetária, mas não a coagulação. Este estudo mostrou propriedades estruturais da RnOBP e pode contribuir para a compreensão de suas funções e potencial biotecnológico.

Palavras-chave: Proteína de Ligação a Odorantes; *Rhodnius neglectus*; Dinâmica Molecular; Biologia Estrutural.

ABSTRACT

Rhodnius neglectus, an insect from the Reduviidae family (Hemiptera), is a potential vector of *Trypanosoma cruzi*, the causative protozoan of Chagas disease. Triatomine saliva comprises a rich mixture of molecules synthesized by the salivary glands to counter host responses, aid the uptake of blood and, indirectly, *T. cruzi* transmission. Therefore, the saliva of this hematophagous insect has been a relevant subject of studies due to the presence of these molecules with high biotechnological potential. The goal of this work was to study an Odorant Binding Protein from *R. neglectus* saliva (RnOBP) using computational and experimental approaches aiming, to contribute to the understanding of its structural and functional characteristics. RnOBP molecular modeling, computational molecular dynamics and clustering tests were performed to obtain information about the protein's flexibility. Next, molecular docking assays were carried out to verify the affinity of RnOBP with agonists of vertebrate coagulation pathways. Finally, putative RnOBP oligomers were observed. To obtain the recombinant RnOBP, its nucleotide sequence was obtained from the *R. neglectus* transcriptome, cloned into the pET100/D-TOPO expression vector and produced by *Escherichia coli* Rosetta(DE3)pLysS bacteria. The recombinant RnOBP was purified by affinity chromatography, refolded, and used in biophysical characterization assays. This protein is modeled to be entirely formed by α -helices and have a significant internal cavity, potentially a ligand-binding site. Compared with insect OBPs, regardless of the composition or classification of the OBP, the binding site region is well conserved. The recombinant RnOBP purified from inclusion bodies revealed a secondary structure composition different than expected. Initial functional indicated rRnOBP could act to inhibit platelet aggregation, but not coagulation. This study showed RnOBP structural properties and may contribute to the understanding of its functions and biotechnological potential.

Keywords: Odorant-binding protein; *Rhodnius neglectus*; Molecular dynamics; Structural biology

LISTA DE FIGURAS

Figura 1 Ciclo de vida do <i>Rhodnius neglectus</i>	20
Figura 2. Estruturas tridimensionais de OBPs.	24
Figura 3. Exemplos de espectros de DC.....	32
Figura 4 Fluxograma geral utilizado.	36
Figura 5. Mapa do vetor recombinante pET100/D-TOPO..	42
Figura 6 Modelo da RnOBP gerado pelo programa Alphafold2.	55
Figura 7 . Modelos e propriedades da superfície eletrostática da RnOBP ...	56
Figura 8. Sobreposição de estruturas de OBPs com o modelo da RnOBP.	57
Figura 9 Modelo de dimerização Alpha Multimer.....	57
Figura 10. Raio de giro da triplicata da dinamica molecular.....	59
Figura 11. RMSD.....	59
Figura 12. Flexibilidade e RMSF.	60
Figura 13 Superposição dos frames representativos da dinâmca molecular	61
Figura 14 Ligantes selecionados para ensaio de <i>ensemble docking</i>	62
Figura 15 Contatos do <i>ensemble docking</i> da RnOBP com ligantes	65
Figura 16. Expressão da RnOBP em linhagens diferentes.....	66
Figura 17. Expressão da rRnOBP em diferentes condições.	67
Figura 18 Ensaio de obtenção das proteínas induzidas.....	68
Figura 19. Sobreposição dos cromatogramas	68
Figura 20 (SDS-PAGE) Purificação FI.	69
Figura 21 SDS-PAGE Purificação da FS.	70
Figura 22 : Velocidade de sedimentação em ultracentrifuga analítica.	71
Figura 23. Purificação da rRnOBP por cromatografia de exclusão	72
Figura 24 Exclusão molecular da rRnOBP renovelada.	73
Figura 25. Espectrometria de massas (MALDI-TOF)	75
Figura 26. Cromatogramas exclusão molecular rRnOBP desnaturante	76
Figura 27. Dicroísmo circular no espectro UV proximo da luz visível.	77
Figura 28. Dicroísmo circular no espectro UV longe da luz visível.	78
Figura 29. Perda de estruturas secundárias.	79
Figura 30. Avaliação da inibição da agregação plaquetária pela rRnOBP... ..	80
Figura 31 Avaliação de inibição da coagulação.....	81
Figura 32 Modelagem da interação trimérica e tetramérica da rRnOBP	91

LISTA DE TABELAS

Tabela 1. Estruturas de proteínas OBPs da família D7	27
Tabela 2. Propriedades calculadas da RnOBP	53
Tabela 3. Pontuações de qualidade dos principais modelos da RnOBP	54
Tabela 4. Afinidades de ligação e RMSD dos ligantes	63
Tabela 5. Pontuações de afinidade do <i>Ensemble docking</i> da RnOBP	64
Tabela 6 Resíduos que estabilizaram os ligantes	65
Tabela 7. Espalhamento de Luz Dinâmico (DLS)	74
Tabela 8. Estruturas secundárias experimental e teórica	78
Tabela 9. Composição de estruturas secundárias	79
Tabela 10. Clusterização dinâmica molecular 1	90
Tabela 11. Clusterização dinâmica molecular 2	90
Tabela 12. Clusterização dinâmica molecular 3	90

LISTA DE ABREVIATURAS, SIGLAS E SIMBOLOS

ϵ Coeficiente de extinção

ϵ_1 Coeficiente de extinção com as cisteínas oxidadas

ϵ_2 Coeficiente de extinção com as cisteínas reduzidas

Å angstroms

A Amperes

A280 Absorbância a 280 nm

A350 Absorbância a 350 nm

ABP Proteína Ligante de antena (*Antennae-Binding Protein*)

Acorr Absorbância corrigida

AUC Ultra centrifugação analítica (*Analytical Ultracentrifugation*)

BSA Albumina de soro bovino (*Bovine Serum Albumin*)

DC Doença de Chagas

CD Dicroísmo circular (*Circular Dichroism*)

JHBP Proteína ligante ao hormônio juvenil (*Juvenile Hormone- Binding Protein*)

IMAC (*Immobilized Metal Affinity Chromatography*)

Kav relação do volume da eluição sobre o volume de retenção da coluna.

kDa Kilo Dalton

OBP Proteína-ligante ao odorante (*Odorant-Binding Protein*)

PBS Tampão de fosfato salino (*Phosphate-Buffered Saline*)

PDB (*Protein Data Bank*)

rRnOBP OBP recombinante de *Rhodnius neglectus* (*recombinant Rhodnius neglectus Odorant-Binding Protein*)

RnOBP OBP nativa de *Rhodnius neglectus* (*Native Rhodnius neglectus Odorant-Binding Protein*)

RMSD Desvio médio quadrático (*Root Mean Squared Deviation*)

RMSF Flutuação média quadrática (*Root Mean Squared Fluctuation*)

SEC Cromatografia de exclusão de tamanho (*Size Exclusion Chromatography*)

TBS Tampão tris salino (*Tris-Buffered Saline*)

TBST TBS com tween (*Tris-Buffered Saline Tween 0.1%*)

TXA2 (Tromboxano A2)

V Volts

SUMÁRIO

1	INTRODUÇÃO	19
1.1	DOENÇA DE CHAGAS	19
1.2	TRIATOMÍNEOS, OS VETORES DO <i>TRYPANOSOMA CRUZI</i> , O AGENTE ETIOLÓGICO DA DOENÇA DE CHAGAS	20
1.3	HEMATOFAGIA E MOLÉCULAS SALIVARES DE VETORES HEMATÓFAGOS.....	21
1.4	PROTEÍNAS DE LIGAÇÃO A ODORANTES (<i>ODORANT-BINDING PROTEINS</i> – OBP)	23
1.5	ABORDAGENS COMPUTACIONAIS E SUAS APLICAÇÕES	27
1.5.1	Predição de estruturas tridimensionais de proteínas	27
1.5.2	Dinâmica molecular	28
1.5.3	<i>Docking</i> molecular	29
1.6	ABORDAGENS BIOFÍSICAS E SUAS APLICAÇÕES	30
1.6.1	Ultracentrifugação analítica	30
1.6.2	Espalhamento de Luz Dinâmico (DLS)	31
1.6.3	Dicroísmo Circular (CD)	32
2	JUSTIFICATIVA	34
3	OBJETIVOS	35
3.1	OBJETIVO GERAL.....	35
3.2	OBJETIVOS ESPECIFICOS	35
4	MATERIAIS E MÉTODOS	36
4.1	METODOLOGIA COMPUTACIONAL.....	36
4.1.1	Análise <i>in-silico</i> da sequência de aminoácidos da RnOBP	36
4.1.2	Modelagem molecular da estrutura 3D de RnOBP	37
4.1.3	Predição de sítios de ligação	37
4.1.4	Dinâmica molecular do modelo RnOBP e clusterização das conformações	38
4.1.5	Preparação da proteína e dos ligantes	39
4.1.6	<i>Docking</i> molecular	39
4.2	METODOLOGIA EXPERIMENTAL	41
4.2.1	Expressão e purificação da RnOBP	41
4.2.1.1	<i>Clonagem da sequência da RnGOBP no vetor de expressão</i>	41

4.2.1.2	<i>Produção de células competentes</i>	42
4.2.1.3	<i>Transformação de linhagens bacterianas</i>	43
4.2.1.4	<i>Expressão da RnOBP recombinante</i>	43
4.2.1.5	<i>Lise das bactérias</i>	44
4.2.1.6	<i>Purificação da RnOBP recombinante presente na fração solúvel por cromatografia de afinidade</i>	44
4.2.1.7	<i>Purificação da RnOBP recombinante presente na fração insolúvel por cromatografia de afinidade</i>	45
4.2.1.8	<i>Purificação da RnOBP recombinante por cromatografia de exclusão molecular</i>	45
4.2.2	Análise da RnOBP recombinante em SDS-PAGE	46
4.2.3	Confirmação da expressão da RnOBP por western blotting	46
4.2.4	Renovimento e quantificação	47
4.2.5	Testes biofísicos	48
4.2.5.1	<i>Ultracentrifugação analítica</i>	48
4.2.5.2	<i>Espalhamento de luz dinâmico</i>	49
4.2.5.3	<i>Dicroísmo Circular</i>	49
4.2.6	MALDI-TOF (<i>Matrix-Assisted Laser Desorption Ionization–Time Of Flight</i>) 50	
4.2.7	Testes de inibição da agregação plaquetária e de coagulação	50
4.2.7.1	<i>Manutenção e coleta de glândulas salivares de <i>Rhodnius neglectus</i></i>	50
4.2.7.2	<i>Teste de coagulação</i>	51
4.2.7.3	<i>Teste de agregação plaquetária</i>	51
5	RESULTADOS	53
5.1	ENSAIOS COMPUTACIONAIS	53
5.1.1	Análise da sequência	53
5.1.2	Predição da estrutura tridimensional da RnOBP	54
5.1.2.1	<i>Sítios de ligação da RnOBP</i>	55
5.1.3	Ensaio de oligomerização da rRnOBP	57
5.1.4	Dinâmica molecular da RnOBP	58
5.1.5	Clusterização das conformações da RnOBP	60
5.1.6	Seleção de compostos promissores	62
5.1.7	Docking molecular da RnOBP	63
5.1.7.1	<i>Redocking dos ligantes selecionados</i>	63

5.1.7.2	<i>Ensemble docking das conformações da RnOBP com os possíveis ligantes</i>	63
5.2	ENSAIOS EXPERIMENTAIS	66
5.2.1	Expressão e purificação da rRnOBP	66
5.2.2	Ensaio de caracterização biofísica	71
5.2.2.1	<i>Ultracentrifugação Analítica</i>	71
5.2.2.2	<i>Cromatografia de exclusão molecular</i>	72
5.2.2.3	<i>Espalhamento de Luz Dinâmico</i>	73
5.2.2.4	<i>Análise de espectrometria de massas por MALDI-TOF da rRnOBP</i>	74
5.2.2.5	<i>Cromatografia de exclusão molecular na presença de uréia</i>	75
5.2.2.6	<i>Caracterização estrutural da rRnOBP</i>	76
5.2.3	Atividade inibitória da agregação plaquetária e da coagulação da rRnOBP	80
6	DISCUSSÃO	82
7	CONCLUSÕES	88
8	PERSPECTIVAS	89
	APÊNDICE	90
9	REFERÊNCIAS BIBLIOGRÁFICAS	94

1 INTRODUÇÃO

1.1 DOENÇA DE CHAGAS

A doença de Chagas (DC), também conhecida como Tripanossomíase Americana, recebeu esse nome em homenagem ao médico brasileiro Carlos Chagas, que descobriu o agente etiológico e caracterizou a doença no ano de 1909 (4). A DC é causada pelo protozoário *Trypanosoma cruzi*, e é transmitida aos humanos principalmente por insetos vetores hematófagos pertencentes à subfamília Triatominae (Hemiptera: Reduviidae), popularmente conhecidos como barbeiros (5). A transmissão vetorial ocorre quando o triatomíneo se alimenta de sangue e elimina as formas infectivas do parasito em suas fezes/urina próximo ao local da picada. A pessoa ou animal picado pode coçar a área, levando as fezes do inseto para uma ferida aberta, olhos, boca, nariz ou outras mucosas do corpo, permitindo a entrada do parasito no organismo do hospedeiro (5).

Além da transmissão por triatomíneos hematófagos, que são os únicos vetores naturais do *T. cruzi*, o parasito também pode ser transmitido aos humanos por transfusões de sangue contaminado, transplantes de órgãos infectados, acidentes de laboratório envolvendo o parasito, ingestão de alimentos ou bebidas contaminados, e de mãe para filho durante a gravidez. A prevenção envolve o controle dos insetos vetores, triatomíneos, melhorias na habitação e triagem de doadores de sangue e órgãos, tendo em vista que ainda não há vacinas tratamentos ou antiparasitários eficazes disponíveis para curar a doença de Chagas na fase crônica (6)

Em resposta a essa preocupação, programas governamentais foram estabelecidos nas décadas de 70 e 90 para combater a doença de Chagas. O programa *Southern Cone Initiative* (Iniciativa do Cone Sul) e o programa de controle da doença de Chagas foram criados com o objetivo de eliminar a transmissão da doença por meio de transfusões de sangue e controlar o principal vetor da doença no Brasil, o *Triatoma infestans* (7).

A Iniciativa do Cone Sul foi uma colaboração entre os países Brasil, Chile, Bolívia, Argentina, Paraguai e Uruguai, que resultou em melhorias significativas no controle de bancos de sangue(7).

Graças ao sucesso desses programas governamentais nas décadas de 80 e 90, houve uma significativa redução na transmissão vetorial nos países envolvidos, resultando em uma diminuição de 94% na incidência da doença (8). No entanto,

estima-se que aproximadamente 7 milhões de pessoas em todo o mundo ainda sejam afetadas por essa doença, com mais de 75 milhões em risco de infecção. A doença de Chagas permanece classificada como uma das principais doenças tropicais negligenciadas, especialmente na América Latina, onde é endêmica, de acordo com a Organização Mundial da Saúde (9).

1.2 TRIATOMÍNEOS, OS VETORES DO *TRYPANOSOMA CRUZI*, O AGENTE ETIOLÓGICO DA DOENÇA DE CHAGAS

Triatominae é uma subfamília de insetos da família Reduviidae, que agrupa os insetos hematófagos. Em seu ciclo de vida esses insetos passam por cinco estágios ninfais até se tornarem imagos (Figura 1). Tanto os machos quanto as fêmeas necessitam ingerir sangue de animais vertebrados para completarem seu ciclo de vida e se reproduzirem. Atualmente são descritas 156 espécies de triatomíneos distribuídas em 18 gêneros e agrupadas em cinco tribos (10), sendo todas elas potenciais vetores de *Trypanosoma cruzi* e alguns outros tripanossomatídeos, como o *Trypanosoma rangeli*. A descrição das espécies é feita principalmente baseada na taxonomia tradicional com base na morfologia do inseto, entretanto também é baseada em diferenças fisiológicas e genéticas (11).

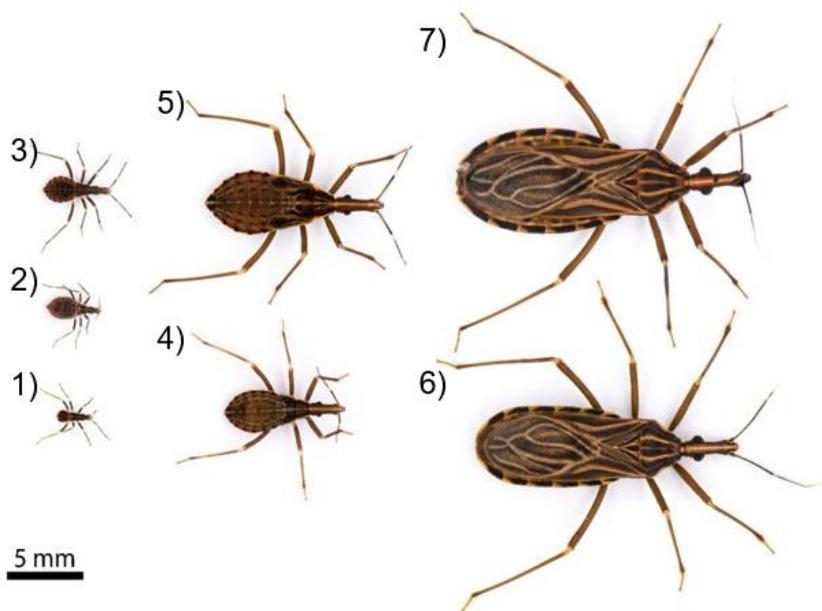


Figura 1 Ciclo de vida do *Rhodnius neglectus*. Os triatomíneos são hemimetábolos que possuem 6 fases de vida. Ovo, 5 estágios ninfais (1-5) e a fase adulta, sendo 6) o macho e 7) a fêmea. Todas as fases, com exceção do ovo, são hematófagos. Fotos de Praça. Retirado e modificado de (1).

A maioria dos insetos triatomíneos são espécies silvestres, que se alimentam de hospedeiros vertebrados específicos, como mamíferos e aves silvestres. Entretanto, algumas espécies ocupam habitats peridomiciliares como galinheiros e currais, e outras se adaptaram completamente aos ambientes domésticos. Essas espécies que ocupam esses ambientes peridomiciliares e domiciliares podem transmitir o *T. cruzi* zoonótico para os humanos, e entre os humanos (12).

O gênero *Rhodnius* compreende 20 espécies e é considerado um dos gêneros taxonomicamente complexos da subfamília Triatominae. A espécie *Rhodnius neglectus* (13), é bem distribuída pelo país, mas é encontrado principalmente em regiões de cerrado e caatinga. Apesar de ser uma espécie silvestre, o *R. neglectus* é um vetor secundário que desempenha um papel importante na manutenção silvestre do *T. cruzi* na América do Sul e pode transmitir a doença a humanos tendo em vista que estudos recentes têm relatado que esta espécie também é encontrada em áreas urbanas, domicílios e peridomicílios (8, 14).

As glândulas salivares do gênero *Rhodnius* apresentam uma coloração avermelhada que é característica, isso se deve a alta concentração de ferro encontrado no grupo heme carregado por proteínas nitroforinas. Juntamente com essa proteína, muitas outras com funções relevantes dentro da hematofagia são expressas e armazenadas dentro destas glândulas (15).

1.3 HEMATOFAGIA E MOLÉCULAS SALIVARES DE VETORES HEMATÓFAGOS

Outra característica de suma importância na biologia dos triatomíneos é o processo de hematofagia. Esses insetos apresentam várias modificações morfológicas e fisiológicas que os capacitam a se tornarem especialistas em alimentação sanguínea. Entre as adaptações morfológicas, destacam-se a presença de uma probóscide, um par de mandíbulas serrilhadas e um par de maxilas flexíveis. Essas estruturas desempenham o papel de lesar o tecido epitelial do hospedeiro até atingir um vaso sanguíneo (16). Nesse ponto, o triatomíneo libera saliva e injeta moléculas produzidas em suas glândulas salivares (17).

Essas moléculas altamente adaptadas ingressam na lesão causada pelo aparelho bucal do inseto através da salivação do triatomíneo. O composto salivar exerce ação na contra regulação da hemostasia e imunomodulação do hospedeiro. A

hemostasia dos hospedeiros vertebrados engloba os processos fisiológicos que evitam o extravasamento sanguíneo (18). Diversas proteínas foram identificadas com funções específicas no contexto da hematofagia em triatomíneos: Dipetalodipinas (vasodilatadoras), Nitroforinas (vasodilatação), Triplatinas (antiagregação plaquetária), Nitroforinas (anti-histamínicos), lipocalinas (imunomoduladoras), Serpinas e outros inibidores de proteases (anticoagulantes), entre outras funções (18, 19)

Proteínas salivares de insetos vetores hematófagos podem realizar funções que acabam por influenciar indiretamente em outros fatores relacionados com a interação patógeno-hospedeiro, como o aumento da taxa de infecção por alguns parasitos. Em um estudo foi observado que uma hialuronidase salivar de *Lutzomyia longipalpis* aumenta a virulência de *Leishmania major* através de sua atividade de quebra de matriz extracelular (20). Outro caso em que a proteína de vetores que acaba influenciando positivamente na infectividade de parasitos é o da Apirase de *Anopheles gambiae*, que ativa a conversão de plasminogênio em plasmina que é uma proteína humana necessária para a transmissão do *Plasmodium* (21).

Estudos demonstram que o uso de proteínas salivares de vetores para a formulação de vacinas também pode influenciar positivamente na redução da transmissão de doenças. No estudo de Martin-Martin e colaboradores (20) a imunização contra uma hialuronidase e uma endonuclease de *Lutzomyia longipalpis* mostrou uma diminuição da infiltração de neutrófilos que servem como veículo para o estabelecimento da infecção de *Leishmania major*. Em *Phlebotomus papatasi* as proteínas Yellow-related protein e uma apirase foram elencadas como potenciais vacinas, devido a suas capacidades de evocar respostas do tipo Th1 em humanos previamente expostos pela saliva do flebotomíneo (22). Nestes últimos dois casos a vacina funcionaria como uma forma de diminuir o fenótipo gerado pela leishmaniose, que é uma doença que gera feridas e cicatrizes que acompanham o hospedeiro pelo resto da vida e podem ser amenizados por uma resposta mais direcionada (20, 22).

Desse modo, algumas das funções desempenhadas por essas moléculas salivares de animais hematófagos também estão sendo objeto de estudo devido ao seu potencial farmacológico. Um exemplo ocorre na sanguessuga *Hirudo medicinalis*, na qual uma proteína com alta afinidade de ligação e inibição da trombina tem sido explorada para o desenvolvimento de fármacos anti-trombóticos. (23). Outra proteína salivar que é utilizada como farmaco e que já é comercializada por sua função vaso-

dilatadora é o maxadilan, um peptídeo de 7kDa do mosquito-palha (*Lutzomyia longipalpis*). (24).

Um grupo de proteínas que se mostra promissor é o das Odorant-binding proteins salivares de insetos hematófagos, que aparecem com diversas funções dentro da contraposição da hemostasia e da modulação da resposta imune (25)

1.4 PROTEÍNAS DE LIGAÇÃO A ODORANTES (*ODORANT-BINDING PROTEINS* – OBP)

As Proteínas de Ligação a Odorantes (*Odorant-Binding Proteins* - OBPs) constituem uma superfamília de proteínas com diversas funções em diferentes tecidos. Sua descoberta ocorreu no sistema de quimiorrecepção olfatória de animais. Essas proteínas apresentam um bolso central hidrofóbico que lhes permite encapsular e proteger ligantes hidrofóbicos do ambiente aquoso circundante, o que facilita o transporte hidrofílico de odorantes para receptores de membrana neuronal (26). Este mecanismo ocorre em vertebrados e insetos, porém as estruturas das proteínas são diferentes, em insetos ela é totalmente constituída por α -hélices e em vertebrados ela forma uma estrutura de lipocalina (Figura 2).

Em 1981, registrou-se o primeiro relato da presença de uma OBP em insetos, com a descoberta de uma proteína linfática sensilar na mariposa *Antheraea polyphemus*, capaz de se ligar a um feromônio sexual, sendo classificada como uma *Pheromone-Binding Protein* (PBP) (27, 28). Investigando mais a fundo, estudos subsequentes revelaram a natureza generalista dessa classe de proteínas. Elas não se limitam a um único tecido, sendo identificadas em diversos locais, como a hemolinfa, onde se ligam a hormônios juvenis (do inglês, *Juvenile Hormone-Binding Protein* - JHBP) (29); sensilas gustativas, facilitando a transferência de moléculas para receptores gustativos (30, 31); e órgãos como ovários e antenas (do inglês, *Antennae-Binding Protein* - ABP) (32-34) e proteínas que se ligam a odorantes gerais (do inglês, *General Odorant-Binding Proteins* - GOBP) (28).

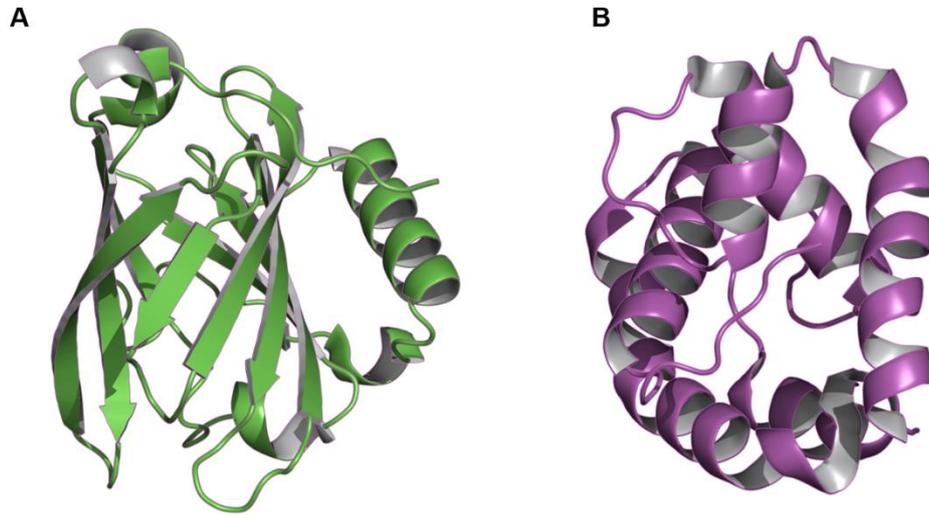


Figura 2. Estruturas tridimensionais de OBPs. A) OBP de *Rattus norvegicus* (PDB ID 3FIQ). B) OBP LUSH de *Drosophila melanogaster* (PDB ID 2QDI). Estruturas cristalográficas retiradas do banco de dados PDB e as imagens foram editadas no Pymol.

A regulação da ligação e liberação de ligantes pelas OBPs pode ser influenciada pela alteração do pH, um fenômeno observado em OBPs que se conectam a feromônios em insetos lepidópteros, coleópteros e dípteros (35). Vale destacar também a quantidade de genes variável que codificam essas proteínas entre diferentes espécies. Por exemplo, *Aedes aegypti* apresenta 66 genes putativos de OBPs, enquanto em *Drosophila melanogaster* possui 51 genes potenciais (36, 37).

Quanto à classificação das OBPs, elas podem ser categorizadas com base no número de resíduos de cisteína. As OBPs clássicas contêm seis cisteínas, formando três pontes dissulfeto cruciais para a estabilidade de sua estrutura tridimensional (38). Além dessas, foram identificados outros tipos de OBPs, incluindo as não clássicas, que têm contagens variadas de cisteínas. As OBPs C-plus possuem oito cisteínas conservadas, as C-minus possuem quatro, as OBPs dímeras contam com doze cisteínas conservadas e, adicionalmente, as OBPs atípicas apresentam de nove a dez cisteínas conservadas (38). As OBPs clássicas de insetos possuem seis α -hélices interconectadas por 3 pontes dissulfeto, formando uma estrutura altamente estável que possui um sítio hidrofóbico que pode abrigar ligantes (39). As sequências aminoacídicas de OBPs são altamente variadas, porém isso não interfere na estrutura tridimensional destas proteínas, que é altamente conservada(33).

Inicialmente, as OBPs foram agrupadas em cinco subfamílias, estabelecidas com base nas sequências de aminoácidos, nos tecidos onde ocorre a sua expressão e em suas funções supostas: uma subfamília de PBPs (Fromone-binding proteins), duas de GOBP1 e GOBP2 (general odorant-binding proteins), e duas de ABP1 e ABP2 ou ABPX (33)

Dentro do grupo das OBPs salivares, destaca-se o conjunto de proteínas D7. Essas proteínas são expressas em insetos hematófagos e têm sido amplamente estudadas em flebotomíneos e mosquitos. Elas são classificadas em duas categorias: D7r (proteínas relacionadas à D7) e D7L (proteínas longas D7). As primeiras são mais curtas, com um domínio OBP, enquanto as segundas são mais longas, com dois domínios OBP (25).

As proteínas D7 salivares (tanto as D7r quanto as D7L), integrantes da vasta super família das Odorant-Binding Proteins (OBPs), estão sendo minuciosamente exploradas por pesquisadores. Uma gama diversificada de proteínas D7 presentes em dípteros foi caracterizada a fundo, revelando funções multifacetadas. A D7r2 de *Anopheles gambiae*, por exemplo, tem a habilidade de capturar mensageiros moleculares que influenciam a vasoconstrição, como serotonina e epinefrina. A D7r3 da mesma espécie atua de maneira sinérgica com a D7r2, ligando-se também à noradrenalina, além de se conectar com a serotonina e epinefrina. Adicionalmente, a D7r4, além das afinidades já mencionadas, também manifesta afinidade pela triptamina, todos sendo ligantes de aminas biogênicas, frequentemente encontrados em interações com proteínas D7, especialmente nas variantes curtas (40).

Em mosquitos da espécie *Anopheles stephensi*, a D7r1 demonstrou a capacidade de inibir a cascata de coagulação, atuando na prevenção da conversão do fator XII em fator XIIa, além de sequestrar o cininogênio de alto peso molecular (41). Nas proteínas D7 longas, que contêm dois domínios característicos das superfamílias OBPs, observou-se em *Anopheles gambiae* que a proteína D7L3 se conecta com a serotonina, histamina e noradrenalina através do seu domínio C-terminal. No entanto, o domínio N-terminal não manifesta afinidade por eicosanoides, moléculas que normalmente se ligam ao domínio N-terminal das OBPs anofelinas (42). Além disso, em *Anophles stephensi*, a D7L2 é capaz de sequestrar o U46619 (43) um análogo de tromboxano A₂, assim como outros eicosanoides da via metabólica do ácido araquidônico, como LTC₄, LTD₄ e LTE₄, que desempenham papéis na constrição da musculatura lisa.

No contexto dos triatomíneos, apenas duas espécies, *Rhodnius prolixus* (GenBank: GCA_000181055.3) e *Triatoma infestans* (GenBank: GCA_011037195.1), têm seus genomas totalmente sequenciados. Apesar disso, os estudos voltados para a caracterização do sialoma, que engloba o transcriptoma e o proteoma das glândulas salivares, têm se estendido a outras espécies de triatomíneos. Por exemplo,

investigações desse gênero foram conduzidas em *Rhodnius neglectus* (44), *Triatoma dimidiata* (45), *Triatoma sordida* ((46), *Panstrongylus lignarius* (47) e mesmo em espécies que tem seus genomas totalmente sequenciados, como *Rhodnius prolixus* (48) e *Triatoma infestans* (49).

Dentre todos os sialomas (estudos transcriptômicos e proteômicos de glândulas salivares) examinados em triatomíneos, as OBPs putativas têm sido objetos de investigação, embora sua expressão geralmente seja discreta. No entanto, existe uma notável lacuna na pesquisa relacionada a essas sequências. É importante ressaltar que essa lacuna também persiste em hemípteros hematófagos, onde não há registro de nenhuma estrutura de OBP salivar em bancos de dados, como o PDB. No âmbito do estudo das proteínas D7, algumas estruturas com diversas funções na modulação da hemostasia foram resolvidas e depositadas no PDB, como ilustrado na Tabela 1. Para investigar proteínas com estruturas não resolvidas, estão disponíveis técnicas de modelagem e predição que demonstram uma boa acurácia e podem sugerir funções para estudos experimentais.

Tabela 1. Estruturas de proteínas OBPs da família D7 resolvidas e depositadas no PDB.

Espécie	Proteína	Ligantes	Referência	PDB ID
<i>Culex quinquefasciatus</i>	D7r		(42)	7TVY
<i>Anopheles darlingi</i>	D7L	U46619*, serotonina	(42)	7TX8
<i>Aedes aegypti</i>	D7r	N/A	(42)	7TVC
<i>Anopheles darlingi</i>	D7L		(42)	7UIN
<i>Culex quinquefasciatus</i>	D7L	ADP	(50)	6V4C
<i>Phlebotomus duboscqi</i>	D7L	N/A	(50, 51)	6MTF
<i>Anopheles gambiae</i>	D7r	Triptamina, Serotonina, Histamina, Norepinefrina	(52)	2PQL 2QEB 2QEH 2QEO 2QEV
<i>Anopheles stephensi</i>	D7L	U46619*, Leucotrieno C4	(43)	3NHT 3NGV 3NHI
<i>Aedes aegypti</i>	D7L	Brometo, Cloreto (Cl-), Leucotrieno E4	(53)	3DXL 3DYE 3DY9 3DZT

*Análogo de tromboxano A2

1.5 ABORDAGENS COMPUTACIONAIS E SUAS APLICAÇÕES

1.5.1 Predição de estruturas tridimensionais de proteínas

A modelagem de estruturas tridimensionais de proteínas é uma área da biologia que se concentra na determinação de estruturas usando métodos como a difração de raios-X, espectroscopia de ressonância magnética e crio-microscopia eletrônica. Embora esses métodos sejam eficazes, eles podem ser dispendiosos (54). Como alternativa, a bioinformática estrutural oferece a capacidade de prever as estruturas tridimensionais de proteínas de forma mais econômica, utilizando cálculos computacionais para deduzir a estrutura nativa de uma proteína com base em sua sequência de aminoácidos ou em estruturas de proteínas semelhantes (55). Isso é de suma importância, pois a conformação tridimensional de uma proteína está intimamente relacionada às suas funções e interações com outras moléculas.(56).

Várias abordagens são empregadas na modelagem de proteínas. A modelagem por homologia, também conhecida como modelagem comparativa, é a técnica mais comumente utilizada. Nesse método, a estrutura de uma proteína é modelada com base na estrutura previamente conhecida de uma proteína homóloga, ou seja, que possui uma estrutura semelhante (57).

Além disso, a modelagem *Ab initio*, também conhecida como modelagem livre [*Free modeling (FM)*] que emprega cálculos direcionados para deduzir a estrutura tridimensional com base exclusivamente na sequência de aminoácidos da proteína em questão, calculando contatos resíduo-resíduo, orientações e distâncias entre os átomos que compõe a proteína como um todo (58).

Diversos programas e ferramentas estão disponíveis para a predição de estruturas proteicas. Entre eles, destacam-se o SwissModel (59), que se baseia na modelagem por homologia; o I-Tasser (60), que combina uma variedade de técnicas, incluindo homologia, threading e simulações de Monte Carlo; o TrRosetta (61), que recorre ao aprendizado profundo e redes neurais convolucionais para gerar estruturas com base na sequência primária das proteínas; e o AlphaFold 2 (62), desenvolvido pela DeepMind (Google), uma empresa de inteligência artificial. O AlphaFold utiliza técnicas avançadas de aprendizado profundo, principalmente redes neurais, para prever estruturas tridimensionais de proteínas com notável precisão. É importante expor que há um empenho da ciência em alavancar a área da predição de estruturas tridimensionais de proteínas, e um dos indícios disso são os resultados do CASP (*Critical Assessment of Protein Structure Prediction*) que é uma competição bianual mundial que acontece desde 1994 onde os grupos avaliam o estado da arte da capacidade preditiva das principais metodologias desenvolvidas (63), e o grupo desenvolvedor do AlphaFold foi o vencedor nas edições de 2018 e 2020. Em 2020 a precisão na predição da posição dos átomos atingiu um valor acima de 80% para as estruturas consideradas mais difíceis (64, 65)

A modelagem de proteínas desempenha um papel de destaque na pesquisa biomédica, permitindo a previsão do comportamento molecular das proteínas, a compreensão de seu envolvimento em doenças e auxiliando no desenvolvimento de novos fármacos. Contudo, é essencial ressaltar que esses modelos são predições teóricas e, sempre que possível, devem ser validados por meio de experimentos laboratoriais.

1.5.2 Dinâmica molecular

A simulação de dinâmica molecular (MD) é uma técnica computacional usada para estudar o comportamento de moléculas ao longo do tempo (66). Isso envolve a modelagem das trajetórias das partículas (geralmente átomos) em um sistema molecular, levando em consideração forças como ligações químicas e interações intra

e intermoleculares, e para isso foi utilizado campos de força newtonianos que levarão em conta as propriedades mecânicas dos átomos do sistema. As simulações de MD consistem em representações em alta resolução, com intervalos de tempo na ordem de femtossegundos (10^{-15} segundos). Nesse processo, em cada intervalo de tempo, as forças decorrentes do campo de força, que atua de acordo com a mecânica clássica e ajusta os cálculos da mecânica quântica, são aplicadas a cada átomo, provocando alterações em suas posições. Após esse passo, as forças são recalculadas com base nas novas posições e o ciclo se repete (67). Em última análise, a trajetória final é o resultado da interação do sistema com o campo de força aplicado, gerando um filme que retrata a configuração do sistema em nível atômico em cada ponto do tempo simulado (68). As simulações de MD são aplicadas em várias áreas, desde química e biologia até o design de fármacos. Essas simulações oferecem insights detalhados sobre a dinâmica molecular em escala atômica, fornecendo pistas sobre os possíveis mecanismos de ação de proteínas e as interações entre proteínas e ligantes..

1.5.3 *Docking* molecular

O *docking* molecular é uma técnica computacional bastante utilizada nas áreas de química e biologia computacional (69). Seu propósito é prever como uma molécula, chamada ligante, se ajusta a uma proteína específica, conhecida como receptor, formando um complexo chamado complexo ligante-proteína.

Nesta abordagem, parte-se do princípio da estrutura rígida da proteína, que deve ser previamente ajustada, junto com o ligante, em seus estados ideais de protonação. Isso é crucial para garantir uma interação precisa. O processo de *docking* molecular começa com o algoritmo de acoplamento, o qual explora conformações do ligante que se adaptem ao sítio da proteína onde o *docking* ocorrerá. Essa fase pode ser desafiadora, especialmente para moléculas pequenas e flexíveis, que podem adotar várias conformações possíveis, dependendo de seu grau de liberdade (70).

Na etapa final do *docking*, ocorre a pontuação das interações identificadas na fase anterior. Isso envolve a consideração dos ângulos das interações formadas, cálculos para avaliar interações eletrostáticas, interações de van der Waals e grandezas termodinâmicas, como entropia e entalpia. Esses dados são então traduzidos em escores de afinidade que refletem a força da interação entre a proteína e o ligante (71).

É importante notar que, embora o *docking* molecular ofereça várias vantagens, ele também possui limitações significativas. Estas incluem a necessidade de uma estrutura de proteína de alta qualidade (em termos de resolução e pré-tratamento), a consideração da proteína como rígida, o que não permite a modelagem do fenômeno de ajuste induzido (*induced fit*), onde o sítio de ligação se adapta ao ligante, e a complexidade na elaboração de funções de pontuação confiáveis (69)

Redocking é a capacidade de replicar a geometria de ligação co-cristalizada e a orientação do ligante associado, considerando um estado rígido da proteína. Essa prática é comumente empregada como uma estratégia de validação para avaliar a eficácia de funções de pontuação e algoritmos de acoplamento

Por outro lado, o *crossdocking* envolve o uso de várias estruturas de ligantes obtidas de diferentes arquivos PDB, os quais são submetidos à mesma proteína-alvo. Essa abordagem visa testar a capacidade de um único modelo de proteína rígida em interagir com múltiplas estruturas de ligantes (72).

Finalmente, o *ensemble docking* é uma técnica avançada que considera a flexibilidade das proteínas. Em vez de usar uma única estrutura de proteína, ele utiliza várias conformações diferentes da proteína para prever como ela interage com ligantes. Isso proporciona previsões mais precisas de interações proteína-ligante em sistemas biológicos complexos (73).

1.6 ABORDAGENS BIOFÍSICAS E SUAS APLICAÇÕES

1.6.1 Ultracentrifugação analítica

A ultracentrifugação analítica é uma técnica poderosa e versátil utilizada para estudar propriedades físicas e interações de macromoléculas, como proteínas e ácidos nucleicos, em solução. Ela permite a análise da velocidade de sedimentação das partículas em um campo centrífugo, fornecendo informações valiosas sobre tamanho, forma, massa molecular, e interações das moléculas.

Dentre os dois tipos principais de ultracentrifugação, a ultracentrifugação analítica concentra-se principalmente em dois tipos de experimentos, sendo um deles a Sedimentação em Velocidade (UC-SV).

O princípio fundamental da UC-SV é que a taxa de sedimentação das partículas é influenciada por suas propriedades, como tamanho e forma. Moléculas maiores ou mais densas sedimentam mais rapidamente do que as menores. A partir

dos dados de sedimentação, é possível calcular os coeficientes de sedimentação, que são relacionados com as propriedades hidrodinâmicas das moléculas.

Os coeficientes de sedimentação podem ser usados para calcular a massa molecular das moléculas em estudo, mas essa conversão depende de várias considerações, incluindo a forma da molécula, as propriedades do solvente e a temperatura. Essa técnica é particularmente útil para estudar interações macromoleculares, determinar a pureza de amostras e avaliar o estado oligomérico de proteínas e complexos (74).

No entanto, é importante notar que a interpretação precisa dos resultados da ultracentrifugação analítica requer conhecimento de fatores como a correção de dispersão, a viscosidade do solvente, densidade do solvente, volume específico parcial e coeficiente de fricção que podem afetar a sedimentação (75).

Em resumo, a ultracentrifugação analítica, com foco nos testes de velocidade de sedimentação, é uma técnica fundamental para a caracterização de macromoléculas em solução. Ela permite determinar coeficientes de sedimentação que podem ser usados para estimar propriedades moleculares, como massa molecular, forma e interações, sendo especialmente valiosa para estudos bioquímicos e biológicos.

1.6.2 Espalhamento de Luz Dinâmico (DLS)

A técnica de Espalhamento de Luz Dinâmico (DLS, do inglês *Dynamic Light Scattering*) é uma técnica analítica poderosa usada para investigar o tamanho de partículas ou moléculas em solução (76). Ela é amplamente empregada em diversas áreas, como química, biologia, farmacologia e ciência dos materiais, para determinar o tamanho médio das partículas em suspensão.

O princípio básico do DLS é baseado na análise da mudança de intensidade da luz dispersa por partículas em movimento Browniano (aleatório) em uma solução. À medida que essas partículas se movem, elas causam flutuações na intensidade da luz dispersa. Essas flutuações são relacionadas ao tamanho das partículas e à taxa de difusão, que, por sua vez, está relacionada ao seu volume (77).

A partir dos padrões de flutuações na luz dispersa, o DLS calcula a função de correlação temporal (ou função de autocorrelação), que descreve como as flutuações na intensidade da luz mudam ao longo do tempo. Através dessa função, é possível extrair informações sobre o tamanho das partículas em suspensão (77).

1.6.3 Dicroísmo Circular (CD)

A técnica de Dicroísmo Circular (CD, do inglês *Circular Dichroism*) é uma ferramenta valiosa para analisar estrutura de proteínas e outras moléculas quirais, como ácidos nucleicos. O CD mede a diferença na absorção de luz circularmente polarizada em um espectropolarímetro por moléculas que interagem com a luz de diferentes maneiras dependendo da sua estrutura tridimensional (3).

O princípio básico do CD baseia-se na interação entre a estrutura quiral das moléculas e a polarização da luz. Quando uma molécula quiral, como uma proteína, interage com luz polarizada circularmente (com polarizações opostas para esquerda e direita), ocorre uma diferença na absorção dessas duas polarizações (3).

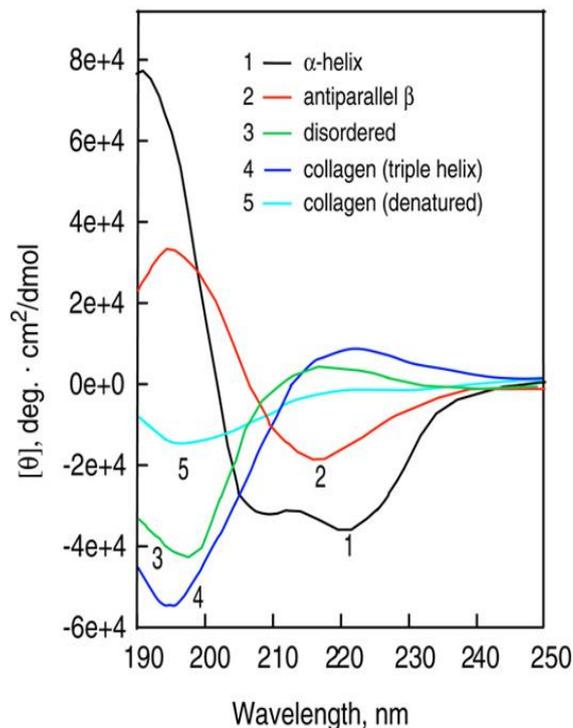


Figura 3. Exemplos de espectros de DC. Imagem ilustra espectros referentes a determinadas estruturas secundárias. Modificado de (3)

O espectro de CD resultante fornece informações sobre a estrutura secundária e terciária da molécula, ajudando a determinar a quantidade de estruturas de hélice, folha e estruturas desordenadas (loops) em uma proteína de acordo com o espectro gerado (Figura 3).

Além disso, os espectros que utilizam comprimentos de onda mais próximos do espectro visível também são empregados para extrair informações acerca das assimetrias em torno dos anéis aromáticos presentes nos resíduos, como triptofano, fenilalanina e tirosina. Essas assimetrias fornecem indícios sobre a posição dos anéis, que, quando expostos ao solvente, perdem essa característica. Portanto, o experimento de CD-NEAR revela um sinal quando a proteína está estruturada. Por outro lado, quando não há estruturação ou quando os resíduos aromáticos hidrofóbicos estão expostos ao solvente, o sinal é próximo de zero (74).

2 JUSTIFICATIVA

O *R. neglectus* é um vetor secundário da DC, com o hábito de buscar abrigo em ambientes domiciliares e peridomiciliares. O estudo das proteínas salivares desses triatomíneos contribui para uma compreensão mais abrangente das interações entre eles, os parasitas e os hospedeiros vertebrados. Essas interações têm implicações práticas na prevenção e controle da DC, além de serem de interesse acadêmico.

As características estruturais e funcionais das OBPs são fundamentais nesse contexto e podem ajudar no desenvolvimento de métodos de controle desses insetos. Notavelmente, a RnOBP, uma proteína de ligação a odorantes de *R. neglectus*, carece de informações na literatura sobre sua conformação e atividade. Dado que as OBPs têm a capacidade de interagir com uma ampla gama de compostos hidrofóbicos e apresentam características como tamanho reduzido e estabilidade, elas representam alvos promissores para pesquisas aplicadas em áreas como pesticidas e farmacologia, o que justifica uma investigação mais aprofundada na RnOBP.

3 OBJETIVOS

3.1 OBJETIVO GERAL

Estudar os possíveis ligantes da *Odorant-Binding Protein* de *R. neglectus* (RnOBP), um dos vetores da doença de Chagas, para assim contribuir na compreensão de suas características estruturais e funcionais, utilizando abordagens computacionais e experimentais.

3.2 OBJETIVOS ESPECIFICOS

1. Predizer a estrutura tridimensional da RnOBP, utilizando ferramentas computacionais de modelagem molecular;
2. Realizar a análise da estrutura obtida utilizando ferramentas de bioinformática com ênfase nos sítios relevantes presentes nos membros da família OBP;
3. Realizar simulações de dinâmica molecular da estrutura modelada em meio fisiológico, para estudar o comportamento dinâmico desta;
4. Verificar a interação, através de docking molecular, de possíveis ligantes relacionados à hemostasia com a RnOBP;
5. Expressar e purificar a RnOBP recombinante (rRnOBP) e iniciar os testes de caracterização biofísica.

4 MATERIAIS E MÉTODOS

A Figura 4 esquematiza as etapas metodológicas utilizadas para a realização desse estudo.

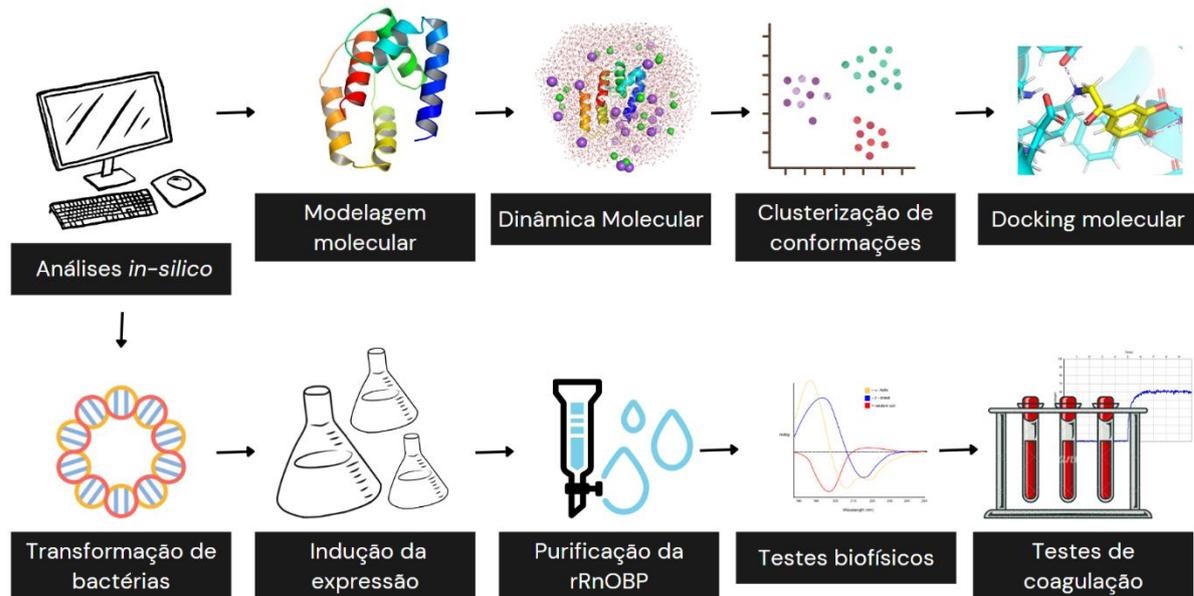


Figura 4 Fluxograma geral utilizado. A OBP salivar de *Rhodnius neglectus* foi estudada utilizando abordagens computacionais (A) e experimentais (B). A sequência da RnOBP putativa (uniprotID: A0A0N7Z921) foi clonada em um vetor de expressão bacteriano para a realização dos métodos experimentais.

4.1 METODOLOGIA COMPUTACIONAL

4.1.1 Análise *in-silico* da sequência de aminoácidos da RnOBP

A caracterização da sequência de aminoácidos da RnOBP (ID GenBank: GDKW01001763.1) foi inicialmente realizada por meio de uma análise *in silico*. Os seguintes programas foram utilizados para essa finalidade: **ProtParam** (78): Este programa foi empregado para calcular parâmetros físicos e químicos da proteína, tais como peso molecular, ponto isoelétrico teórico, composição de aminoácidos e índice de instabilidade. **SignalP 5.0** (79): Foi utilizado o SignalP 5.0 para prever a presença de peptídeos sinal e determinar os locais de clivagem em proteínas provenientes de diferentes domínios. **NetPhos 3.1** (80): Este programa foi aplicado para identificar sítios de fosforilação de serina, treonina ou tirosina em proteínas de natureza eucariótica. **NetNGlyc 1.0** (81) e **NetOGlyc 4.0** (82): Ambos os programas foram

utilizados para prever sítios de N-glicosilação e O-glicosilação, respectivamente, nas sequências das proteínas analisadas.

4.1.2 Modelagem molecular da estrutura 3D de RnOBP

Para a modelagem da estrutura 3D de RnOBP, a sequência primária desta foi obtida do banco de dados Uniprot (ID Uniprot: A0A0N7Z921). A fim de melhorar a análise, a porção N-terminal foi eliminada, a qual corresponde ao peptídeo sinal, seguindo as indicações do servidor SignalP 5.0 (79). Em seguida, foi utilizado o programa **AlphaFold** (62) no ColabFold (2) hospedado na plataforma Google Colab, para a modelagem molecular. O modelo foi gerado utilizando 20 reciclagens para o modelo de monômero e 3 para os modelos de multímeros, o alinhamento múltiplo de sequências baseado no modo *mmseqs2_uniref_env* e modo de pareamento *unpaired_paired* no modo “ganancioso”. O programa utilizado para realizar a modelagem do monômero foi o **AlphaFold2** e o para multímeros foi o **Alphafold2-multimer**. O servidor GalaxyWEB (83) foi utilizado para o refinamento e minimização de energia dos modelos, o modo utilizado foi o Refine2. Para avaliar a qualidade dos modelos, o servidor MolProbity (84) foi utilizado para calcular as pontuações *clashscore*, *Molprobity score* e *Ramachandram favored* dos modelos. Os estados de protonação e carga líquida da proteína em pH 7.4 foram calculados usando o servidor PDB2PQR (85) em campo de força PARSE.

4.1.3 Predição de sítios de ligação

Para a análise preliminar dos sítios de ligação, empregou-se o servidor de predição de cavidades potenciais de ligação, conhecido como CASTp (86). Que utiliza do mapeamento de anotações de sítios de ligação de proteínas depositadas no banco de dados swissprot e comparado ao arquivo PDB depositado (87). Os dados de predição foram comparados com utilizando estruturas cristalográficas depositadas no PDB (3q8i de *Anophles gambiae*, 3r72 de *Apis mellifera*, 3vb1 [PBP] *Anopheles gambiae*, 4f7f [PBP] de *Anophles gambiae*, 5dic (OBP gustativa) de *Phormia regina*, 6jpm [PBP] de *Chrysopa pallens*, 1ooh [PBP] de *Drosophila melanogaster* e 3b6x de *Drosophila melanogaster*) de OBPs de diferentes tecidos e espécies

4.1.4 Dinâmica molecular do modelo RnOBP e clusterização das conformações

A partir da obtenção do modelo 3D da RnOBP, o servidor CHARMM-GUI (88) foi utilizado para gerar arquivos de entrada para realizar simulações de dinâmica molecular. As simulações foram realizadas utilizando o programa NAMD (89), o campo de força utilizado para computar a simulação foi o CHARMM36 (90), utilizando números constantes de átomos, pressão e temperatura (*ensemble* NPT). As simulações de dinâmica molecular foram realizadas em triplicata, em uma caixa octaédrica com bordas de 6 Å, e a carga do sistema foi neutralizada adicionando NaCl 150 mM, o modelo TIP3 de moléculas de água foi utilizado. A etapa de produção da simulação foi executada em 50.000.000 passos de 2 femtossegundos (totalizando 100 ns), com uma etapa de equilibração de energias anterior de 125.000 passos (2 ns). Para esclarecer a localização das pontes dissulfeto, foi utilizado o motivo biológico na estrutura conservada das OBPs. Foi realizado uma sobreposição de estruturas de OBPs do *Protein Data Bank* (PDB) a conservação dos alinhamentos de cisteínas em suas estruturas tridimensionais.

Para visualizar os modelos, dinâmicas e agrupamentos, os softwares Pymol, Chimera (91) foram utilizados. O programa VMD (92) foi usado para visualizar e analisar as dinâmicas moleculares.

Por fim, a clusterização das conformações da RnOBP obtidas através da MD foi realizada usando o programa TtClust (93), utilizando o método de Ward (94) para formação dos clusters. Este método agrupa conformações semelhantes de acordo com a minimização do RMSD entre as estruturas, agrupando estruturas que possuem RMSDs próximos de 0 e separando em grupos diferentes estruturas com RMSDs elevados.

O RMSD é caracterizado pela formula:

$$\begin{aligned} \text{RMSD}(\mathbf{v}, \mathbf{w}) &= \sqrt{\frac{1}{n} \sum_{i=1}^n \|v_i - w_i\|^2} \\ &= \sqrt{\frac{1}{n} \sum_{i=1}^n ((v_{ix} - w_{ix})^2 + (v_{iy} - w_{iy})^2 + (v_{iz} - w_{iz})^2)} \end{aligned}$$

Onde V e W são representações de estruturas sobrepostas;

N é o número de pontos calculados (geralmente os carbonos α de cada aminoácido)

V_{ix} , V_{iy} e V_{iz} são as representações do ponto i da estrutura V em suas coordenadas x, y e z

Wix, Wiy e WiZ são as representações do ponto i da estrutura W em suas coordenadas x, y e z.

4.1.5 Preparação da proteína e dos ligantes

A preparação dos ligantes foi realizada utilizando o software Avogadro (95) que utiliza o Open Babel (96) para realizar otimizações geométricas (campo de força MMFF94, 500 passos, gradiente conjugado). Inicialmente os ligantes foram retirados das suas estruturas cristalográficas depositadas no PDB (IDs 2PQL, 2QEB, 2QEH, 2QEO, 3NHI, 3NHT, 6V4C), utilizando o software PyMOL e depois foram salvos em arquivos separados, que foram abertos utilizando o software Avogadro e submetidos a otimização geométrica com a adição de hidrogênios para o pH 7.4. As conformações da RnOBP foram otimizadas utilizando o servidor PDB2PQR (85), utilizando o campo de forças CHARMM (97) e o programa PROPKA (98) para definir estados de protonação dos aminoácidos e conformações pH-dependente da proteína, levando em conta o pH de 7.4.

4.1.6 Docking molecular

O *docking* molecular dos possíveis ligantes de RnOBP foi realizado usando o servidor DockThor (99). Para assegurar uma análise abrangente, todas as conformações representativas da OBP, derivadas das simulações de dinâmica molecular, foram sobrepostas, e os ligantes foram posicionados no bolso principal da proteína observado na literatura e previsto computacionalmente. A grade de ancoragem para o *docking* foi padronizada com uma dimensão de 22 angstroms. Todas as etapas de ancoramento foram realizadas no modo suave (soft mode), com os ligantes sendo colocados no local primário de ligação. Precedendo o processo de ancoramento, as estruturas proteicas foram processadas pela adição de hidrogênios em pH 7.4 e otimização da geometria por meio do software Avogadro (95). Vale ressaltar que as simulações de *ensemble docking* foram conduzidas sem considerar quaisquer cofatores presentes na estrutura original do PDB, concentrando-se exclusivamente nas interações ligante-proteína.

A pontuação de afinidade serviu como a métrica principal para avaliar as melhores conformações e posições. Um script em Python foi utilizado para identificar os aminoácidos dentro de uma distância de 4 angstroms do ligante no bolso de ligação

(script disponível em: <https://github.com/Morgadogabriel/Bioinfo>). Para analisar os aminoácidos específicos envolvidos nas interações do ligante, o servidor web *Protein-Ligand Interaction Profiler* (PLIP) (100) foi utilizado e os resultados foram comparado para cada interação com os ligantes.

No contexto do experimento de *redocking*, a primeira etapa consistiu na extração do ligante da estrutura proteica, realizada através da edição do arquivo de formato ".pdb". Posteriormente, foi realizada a adição de hidrogênios e à otimização da geometria desses arquivos, levando em consideração um pH de 7, utilizando a ferramenta Avogadro.(95). Após esta etapa, foram realizados os ensaios de *redocking* utilizando o servidor DockThor (99). Nestes ensaios, avaliamos a posição de ligação do ligante e, se a pontuação obtida for menor que 2 angstroms, consideramos isso como um acerto (hit). As pontuações dos hits foram utilizadas como base para julgar as afinidades do *ensemble docking*.

4.2 METODOLOGIA EXPERIMENTAL

4.2.1 Expressão e purificação da RnOBP

4.2.1.1 Clonagem da sequência da RnGOBP no vetor de expressão

A sequência da RnGOBP (ID GenBank: GDKW01001763.1, ID UniProt A0A0N7Z921) foi códon-otimizada para expressão em bactérias e clonada no vetor de expressão pET100/D-TOPO® (Invitrogen, Thermo Fisher Scientific; RnGOBP_pET100/D-TOPO) (Figura 5). Os nucleotídeos que codificam para o peptídeo sinal (60 primeiros pares de bases) foram excluídos da sequência a ser clonada no vetor. O vetor pET100/D-TOPO® possui um promotor para a enzima RNA T7 polimerase, apresenta um local de expressão que abriga um gene beta-lactamase (*bla*), proporcionando resistência a ampicilina às colônias transformadas. Além disso, na tradução da sequência gênica, na porção *upstream* do N-terminal da sequência da RnOBP sem o peptídeo sinal, é adicionada uma sequência que codifica para 36 aminoácidos (MRGSHHHHHHGMASMTGGQQMGRDLYDDDDKDHPFT). Essa região adiciona: 1) uma cauda de 6 histidinas à região N-terminal da proteína, que serve para reconhecimento em *western blotting* e também para purificação por cromatografia de afinidade; 2) uma cauda T7 (MASMTGGQQMGR) que é uma sequência presente no capsídeo do bacteriófago T7 que também pode ser reconhecida por anticorpos comerciais monoclonais anti-T7 *tag*; e 3) uma sequência que codifica para o epítopo X-press (DLYDDDDK), permitindo a detecção por *western blotting* e que contém um sítio com 4 aminoácidos de asparagina seguidos por uma lisina (DDDDK) para a clivagem pela enzima enterocinase.

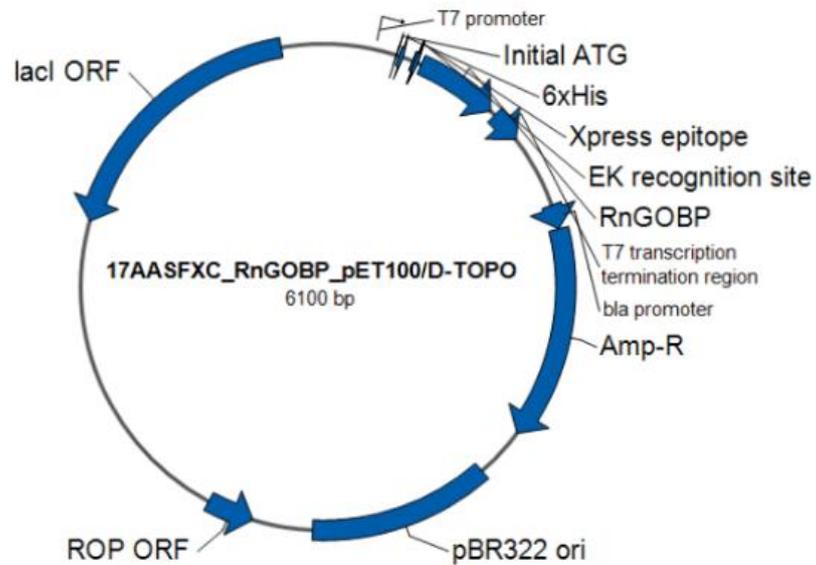


Figura 5. Mapa do vetor recombinante pET100/D-TOPO. O vetor contém a sequência de DNA da RnOBP fusionada com a cauda de histidinas (6xhis), o epítipo Xpress (DLYDDDDK - para reconhecimento imunológico) e o sítio de clivagem específico para enterocinase (*EK recognition site*). O plasmídeo contém o gene para produção de β -lactamase (AMP-R) que confere resistência a ampicilina às linhagens transformadas.

4.2.1.2 Produção de células competentes

Células competentes de *E. coli* das linhagens Rosetta(DE3)pLysS, Lemo21(DE3), Origami(DE3) foram preparadas seguindo o protocolo (101), com adaptações. Resumidamente, uma alíquota de bactérias da linhagem escolhida foi retirada do congelador a $-80\text{ }^{\circ}\text{C}$ e descongelada a 4°C . Posteriormente, $50\text{ }\mu\text{L}$ dessa alíquota foram transferidos para um tubo de ensaio contendo 5 mL de meio LB, o qual foi submetido à agitação a 250 rpm em um agitador padrão InforsTM Multitron (Infors HT) com temperatura ajustada para $37\text{ }^{\circ}\text{C}$ durante a noite.

Em seguida, 1 mL das bactérias foi transferido para um frasco Erlenmeyer contendo 100 mL de meio LB suplementado com o antibiótico relacionado a resistência da bactéria quando necessário (por exemplo, $34\text{ }\mu\text{g/mL}$ de cloranfenicol para a seleção da linhagem Rosetta (DE3) pLysS e Lemo21 (DE3), para a linhagem Origami (DE3) foi utilizado $15\text{ }\mu\text{g/mL}$ de kanamicina e $12.5\text{ }\mu\text{g/mL}$ de tetraciclina. O frasco Erlenmeyer com as bactérias foi agitado a 250 rpm e mantido a 37°C até que a densidade óptica (OD) a 600 nm atingisse o valor de $0,3$.

Posteriormente, o conteúdo dos frascos Erlenmeyer foi dividido em dois tubos Falcon de 50 mL e centrifugado a 4000 xg em uma centrífuga de bancada modelo Z 326 (Hermle LaborTechnik) por 10 min. O sobrenadante foi então descartado, e 50 mL de uma solução 0,1 M de MgCl₂ foram adicionados a cada tubo Falcon contendo o sedimento de células. As células foram ressuspensas e os tubos Falcon foram incubados em gelo por 30 min. Após esse período, os tubos Falcon foram novamente centrifugados a 4000 xg, e o sobrenadante foi descartado.

Cinquenta mililitros de CaCl₂ 0,1M foram adicionados aos tubos Falcon, e o sedimento de células foi novamente ressuspensado. Em seguida, os tubos Falcon foram incubados no gelo por mais 30 min. Após a incubação, os tubos Falcon foram retirados do gelo, centrifugados novamente e o sobrenadante foi descartado.

Finalmente, foi adicionado 1 mL de uma solução 0,1M de CaCl₂ com 15% de glicerol a cada um dos tubos Falcon, e as células foram alíquotadas em microtubos (Axygen; 100 µL cada tubo). Essas amostras foram armazenadas no freezer a -80 °C para uso futuro.

4.2.1.3 *Transformação de linhagens bacterianas*

Para a transformação das diferentes linhagens de *E. coli*, 2 µL de plasmídeo RnGOBP_pET100/D-TOPO (10 ng/µL) foram incubados com 100 µL de células competentes em um micro tubo de 1,5 mL no gelo por 30 min. Subsequentemente, as células foram submetidas ao choque térmico a 42 °C por 1,5 min e incubadas no gelo por 2 min. Após esse período, 1 mL de meio Luria Bertani (LB) líquido foi adicionado ao microtubo e as bactérias cresceram a 37 °C por 1 h sob agitação de 250 rpm. Ao término dessa incubação, a cultura foi plaqueada em meio LB ágar contendo 100 µg/mL de ampicilina (dependendo da linhagem, outro antibiótico necessário para a seleção da linhagem específica foi usado na concentração estabelecida pelo fabricante). As placas foram incubadas à 37 °C por 16 h para o desenvolvimento das colônias.

4.2.1.4 *Expressão da RnOBP recombinante*

Uma colônia transformada de *E. coli* Rosetta (DE3) pLysS_His-RnOBP foi cultivada em um frasco Erlenmeyer de vidro contendo 30 mL de meio líquido LB, suplementado com 100 µg/mL de ampicilina e 34 µg/mL de cloranfenicol. O frasco

contendo as bactérias transformadas foi colocado em um agitador Infors™ Multitron Standard Shaker (Infors HT) a 250 rpm, a uma temperatura de 37 °C, por um período de 16 h. Posteriormente, a cultura foi adicionada a 500 mL de meio LB suplementado com 100 µg/mL de ampicilina e 34 µg/mL de cloranfenicol em um frasco Erlenmeyer com capacidade de 1 L. O frasco Erlenmeyer foi então mantido em agitação a 200 rpm, a 37 °C, até que a densidade óptica em 600 nm (OD_{600nm}) atingisse aproximadamente 0,4. Nesse ponto, isopropyl-β-D-thiogalactopyranosídeo (IPTG) foi adicionado para uma concentração final de 1 mM e a cultura foi incubada a 37 °C por 3 h a 200 rpm. Ao término desse intervalo, o meio contendo as bactérias foi transferido para tubos tipo Falcon de 50 mL e as células foram sedimentadas em uma centrífuga Z 326 (Hermle LaborTechnik) a 10.000 xg por 5 min e o sobrenadante descartado. Após essa etapa, o sedimento de bactérias foi imediatamente congelado a -20 °C.

4.2.1.5 *Lise das bactérias*

O sedimento de bactérias foi ressuspensão em 15 mL do tampão Tris 50mM pH 8 contendo 500mM de NaCl 5mM de Imidazol e um coquetel de inibidores de proteases (cOmplete™, *EDTA-free protease inhibitor cocktail*, Roche). As células foram lisadas por 3 ciclos de sonicação em um banho de gelo por 30 segundos, com a amplitude ajustada para 40 % em um sonicador VCX500 (SONICS & MATERIALS). O lisado de bactérias foi centrifugado a 16.000 xg a 4°C por 20 min na centrífuga Z 326 (Hermle LaborTechnik) e a fração solúvel (sobrenadante) foi separada da fração insolúvel (sedimento).

4.2.1.6 *Purificação da RnOBP recombinante presente na fração solúvel por cromatografia de afinidade*

A purificação da rRnGOBP recombinante presente na fração solúvel foi realizada por cromatografia de afinidade ao níquel. A fração solúvel foi injetada em uma coluna HisTrap FF (GE Healthcare Life Sciences) previamente equilibrada com o tampão de equilíbrio (Tris 50mM pH 8, NaCl 500 mM, 5mM imidazol, pH 8,0), com o auxílio de uma bomba peristáltica. A coluna foi, então, conectada ao sistema AKTA UPLC system (GE Healthcare Life Sciences) para a execução da cromatografia, com os dados cromatográficos sendo registrados através do software Unicorn™ 5.2. A purificação foi realizada a um fluxo constante de 1 mL/min, dividido em quatro etapas,

cada uma com um volume de 10 mL de tampão. A rRnOBP foi eluída com o tampão de equilíbrio contendo concentrações diferentes de imidazol: 25, 50, 200 e 500 mM. Resumidamente, o programa utilizado no software Unicorn consistia em uma lavagem com 5% (25 mM imidazol) do tampão "B" (50 mM Tris, 500 mM NaCl, 500 mM imidazol, pH 8,0), seguida por uma segunda lavagem com 10% (50 mM imidazol), uma terceira com 40% (200 mM imidazol) e uma quarta com 100% (500 mM imidazol). As eluições foram coletadas a partir da terceira etapa, utilizando o fracionador do sistema AKTA (Frac – 950). O tampão "A" utilizado foi composto por 50 mM Tris e 500 mM NaCl, pH 8,0. Por outro lado, o tampão "B" consistiu em 50 mM Tris, 500 mM NaCl e 500 mM imidazol, pH 8,0.

4.2.1.7 *Purificação da RnOBP recombinante presente na fração insolúvel por cromatografia de afinidade*

A fração insolúvel foi solubilizada com 15 mL de tampão de lise (Tris 50 mM, NaCl 500 mM, de Imidazol 5 mM) contendo 8 M de ureia. Após a adição do tampão, o sedimento foi submetido a agitação e sonicação em um banho de gelo por 2 min, com a amplitude ajustada para 40 % por um sonicador VCX500 (SONICS & MATERIALS). Posteriormente, a amostra foi centrifugada a 18.000 xg, à temperatura ambiente, por um período de 30 min. O sobrenadante foi injetado em uma coluna HisTrap FF (GE Healthcare Life Sciences) previamente equilibrada com o tampão de equilíbrio (Tris 50mM pH 8, NaCl 500 mM, 5mM imidazol, pH 8,0, ureia 8 M), com o auxílio de uma bomba peristáltica. O processo empregou o mesmo programa utilizado na purificação da fração solúvel, com a exceção de que os tampões continham ureia com uma concentração de 8 M.

4.2.1.8 *Purificação da RnOBP recombinante por cromatografia de exclusão molecular*

Para determinar a massa molecular da rRnOBP em solução, uma cromatografia de exclusão por tamanho foi conduzida em uma coluna Superdex 200 10/300 GL (GE Healthcare). Para calibrar a coluna, foram empregados padrões de proteínas da GE Healthcare (Aprotinina, 6.500 Da; Ribonuclease A, 13.700 Da; Ovalbumina, 43.000 Da; e Aldolase, 158.000 Da), e uma função linear foi estabelecida para a predição das massas das eluições da rRnOBP. As proteínas foram eluídas da coluna utilizando uma

solução de 50 mM de NaH₂PO₄ a pH 7,5 e a monitorização foi realizada a 280 nm, com uma taxa de fluxo de 0,5 mL min⁻¹. O volume do vazio (V₀) foi determinado por meio do uso de Blue Dextran. O valor de K_{av} foi calculado para cada eluição, empregando a equação $(V_e - V_0) / (V_t - V_0)$, onde V_e representa o volume de eluição da proteína e V_t é o volume total da coluna. Posteriormente, o gráfico de K_{av} em função do logaritmo dos pesos moleculares dos padrões e da RnOBP foi plotado para estimar as massas moleculares dos oligômeros.

4.2.2 Análise da RnOBP recombinante em SDS-PAGE

Em todas as etapas, as amostras foram submetidas a eletroforese em gel de poliacrilamida SDS-PAGE 15% (102). Para a realização da eletroforese, as amostras foram desnaturadas através de sua diluição em tampão Laemmli (Tris 50 mM, pH 6,8, SDS 2%, glicerol 10%, 2-mercaptoetanol 1% e azul de bromofenol 0,1%), seguida pela fervura das amostras a 100 °C por 5 min. Posteriormente, foram aplicados 12 µL das amostras nos géis de poliacrilamida.

A eletroforese foi realizada em sistema Mini-PROTEAN® 3 Cell (Bio-Rad), seguindo as instruções do fabricante, em duplicata. Um dos géis era corado com Coomassie Brilliant Blue R-250 (solução contendo 10% de ácido acético, 45% de metanol e 0,2% de Coomassie Brilliant Blue R-250), o outro gel era destinado para o processo de imuno identificação por Western Blot.

4.2.3 Confirmação da expressão da RnOBP por *western blotting*

Para confirmar a identidade da proteína, um dos géis foi transferido para uma membrana de nitrocelulose (Amersham™ Protran® Supported 0,45 µm, GE Healthcare). Esse procedimento envolveu a imersão da membrana, do gel e de dois papéis esponjosos em tampão de transferência (25 mM de Tris, 192 mM de glicina, 20% de metanol) por 5 min. A transferência a seco foi executada no sistema Trans-blot® Turbo™ Transfer System (BIO RAD), utilizando uma corrente de 1,5 A e uma tensão de 25 V durante 30 min.

Após a transferência, a membrana foi corada com Ponceau (Ponceau S 0.2%, ácido acético 1%) por 5 min, seguido por lavagens com água destilada para verificar a presença das bandas de proteínas transferidas. Em seguida, a membrana foi bloqueada com *Tris Buffered Saline Tween* (TBST) contendo leite desnatado (20 mM

de Tris, 150 mM de NaCl, 0.1% de Tween, pH 7,4, com 5% de leite) por 1 h. Após o bloqueio, a membrana foi lavada com TBST para remover o excesso de leite.

O anticorpo primário anti-His (Sigma Aldrich) foi diluído 1:2000 em TBST contendo 1% de leite desnatado e adicionado à membrana, que foi incubada sob agitação leve por 1 h. Após essa etapa, a membrana foi lavada cinco vezes por 5 min cada, utilizando TBST. O anticorpo secundário GAM (Goat Anti-Mouse IgG horseradish peroxidase conjugated) foi diluído 1:20.000 em TBST com 1% de leite, e a membrana foi novamente incubada sob agitação leve por 1 h para permitir a interação entre os anticorpos.

Após mais cinco lavagens de 5 min cada com TBST, a membrana foi exposta a um substrato quimioluminescente para peroxidase (ECL™ Prime Western Blotting System, GE Healthcare) e colocada na câmara do ImageQuant LAS 4000 (GE Healthcare) para detecção da luminescência pelo sensor do equipamento. Dessa forma, a proteína recombinante foi observada por meio da interação do anticorpo específico com a cauda de 6 histidinas.

4.2.4 Renovelamento e quantificação

O processo de renovelamento foi realizado por meio de diálise em tampão (Tris 20mM, NaCl 150 mM pH 7,4) utilizando um Becker de 4 L. A amostra purificada foi depositada em um saco de diálise (INLAB), submerso no tampão de diálise por 24 h a 4 °C, sob a agitação de uma bailarina magnética. Após isso, as amostras recuperadas após a dialise e os tampões foram coletados, armazenados a 4 °C ou utilizados para o ensaio de quantificação de proteínas pelo método de Lowry. O coeficiente de extinção molar (ϵ) da rRnOBP foi determinado pela absorbância a 280 nm (A_{280}) em função da concentração de proteína, conforme previamente descrito (74). Utilizou-se um Espectrofotômetro UV/Visível Jasco V-530 (Jasco Corporation, Tóquio, Japão) para conduzir as medições de absorbância corrigida, levando em conta o espalhamento de luz a 350 nm (A_{350}) (Eq. 1), bem como a absorbância intrínseca do tampão empregado na solubilização da proteína rRnOBP.

$$A_{corr280} = A_{280} - A_{350} \quad (1)$$

A concentração de proteína foi quantificada através do método de Lowry, empregando uma solução padrão de albumina sérica bovina (0,1–1 mg/ml; Sigma), e

a medição da absorvância a 620 nm ocorreu após 30 minutos em um leitor de SpectraMax M3 (Molecular Devices). A reação, com volume de 200 μL , incluiu proteína rRnOBP em concentrações de 0,18–0,48 mg/mL, p-nitrofenil (ρNP)-butirato em isopropanol de 1 mM, tampão fosfato de sódio 50 mM (pH 7,5) com adição de 0,3% (v/v) de Triton X-100 (PBTx), e ocorreu a 25 °C.

4.2.5 Testes biofísicos

4.2.5.1 Ultracentrifugação analítica

Os experimentos de velocidade de sedimentação foram conduzidos com a utilização de uma ultracentrífuga analítica Beckman XL-I, equipada com um rotor AN-60 TI da Beckman Coulter. A proteína rRnOBP reconstituída, obtida por meio de cromatografia de afinidade, foi testada em várias concentrações (0,68 mg/ml, 0,49 mg/ml, 0,46 mg/ml, 0,37 mg/ml, 0,29 mg/ml e 0,18 mg/ml), em uma solução contendo 25 mM de Tris pH 8,0 e 150 mM de NaCl. Essas concentrações correspondiam a absorvâncias variando de 0,2 a 0,04 na onda de 280 nm. Cada amostra, totalizando 420 μL , foi introduzida em células de trajeto de 1,2 cm e submetida a uma centrifugação a 40.000 rpm. Escaneamentos foram realizados a cada 7 min durante a noite, a 280 nm e 25°C.

Para a determinação dos parâmetros essenciais para a análise, foi realizado o cálculo do volume específico parcial da cadeia polipeptídica (0,723 mL/g), da densidade do solvente ($r = 1,00640$ g/ml) e da viscosidade do solvente ($h = 1,0320$ poise [$\text{g}\cdot\text{cm}^{-1}\cdot\text{s}^{-1}$]) a 25°C. Esses valores foram obtidos por meio do *software* gratuito Sednterp, disponível em: <http://www.jphilo.mailway.com> (acesso em 11/09/2023). Os perfis de sedimentação foram avaliados por meio da análise de distribuição de tamanho utilizando o Sedfit, um *software* gratuito disponível em <https://sedfitsedphat.github.io> (acesso em 11/09/2023). Esse método emprega soluções de elementos finitos da equação de Lamm para representar uma distribuição contínua de tamanhos, com base em um grande número de espécies discretas e independentes. Foram analisados 231 conjuntos de dados, variando a razão de fricção de 0,0 a 20 S. Foi utilizado um nível de confiança de 0,95 no procedimento de regularização. Os gráficos foram plotados por meio do software GUSSI.

4.2.5.2 *Espalhamento de luz dinâmico*

As medidas de Espalhamento de Luz Dinâmico [*Dynamic light scattering* (DLS)] foram realizadas utilizando um Zetasizer Nano-ZS da série Nano-ZS (Malvern Instruments) para determinar o tamanho de partículas/moléculas a 20°C. Um laser He-Ne com comprimento de onda $\lambda = 633$ nm foi empregado nas medições, utilizando um ângulo de 173° e detecção de retrodispersão por meio de uma cubeta de plástico. Os experimentos foram conduzidos com a proteína em um tampão Tris-HCl 20 mM (pH 7,6) contendo 150 mM de NaCl. Antes das medições, as amostras purificadas em uma coluna Superdex 200 10/300 e os picos de diferentes massas moleculares foram separados para a análise. O método de DLS foi empregado para avaliar as populações de proteínas em solução, examinando a polidispersividade, a massa molecular e os diâmetros hidrodinâmicos.

4.2.5.3 *Dicroísmo Circular*

A estabilidade térmica das proteínas foi investigada por meio da técnica de espectroscopia de dicroísmo circular (CD) utilizando o espectropolarímetro Jasco J-815 (JASCO Inc.) equipado com um sistema de aquecimento Peltier. Amostras foram preparadas em um tampão de Tris 5 mM, pH 7,2, nas concentrações de 0,1012 mg/mL para a região do UV distante e 1,2 mg/mL para a região do UV próximo ao visível. Curvas de desdobraimento térmico e espectros de CD no UV distante (195 a 260 nm) foram registrados em uma faixa de temperatura de 25 °C a 95 °C, com uma taxa de aquecimento de aproximadamente 0,2 °C/min, utilizando uma cubeta de quartzo de 1 mm. As curvas de desdobraimento foram registradas a 220 nm para a rRnOBP. Os espectros e curvas de desdobraimento foram suavizados usando o método de Transformada Rápida de Fourier (FFT) com o auxílio do *software* ORIGINPRO 2022 (OriginLab). As elipticidades médias foram convertidas em elipticidade molar antes da análise dos dados, de acordo com a seguinte equação:

$$[\theta] = (\theta * MRW)/(10 * L * C)$$

Onde MRW representa o peso médio de resíduos que constituem a cadeia primária da rRnOBP, θ é a elipticidade medida em mdeg, L é o comprimento ótico da cubeta em cm e C é a concentração da proteína em mg/mL.

Para comparar a análise de UV distante da rRnOBP, realizou-se uma avaliação da composição de estruturas secundárias do modelo usando o algoritmo DSSP (103), por meio do servidor disponível em (<http://bioinformatica.isa.cnr.it/SUSAN/DSSP-web/>). A análise dos padrões de elipticidade molar por nm no UV distante da amostra foi conduzida utilizando o servidor BeStSel (104), tanto para o experimento de desnaturação quanto para a rRnOBP em seu estado nativo.

4.2.6 MALDI-TOF (*Matrix-Assisted Laser Desorption Ionization–Time Of Flight*)

As amostras obtidas da cromatografia de exclusão molecular foram analisadas quanto à sua pureza e oligomerização por espectrometria de massas no equipamento Autoflex II MALDI-TOF/TOF (BrukerDaltonics) sob controle do programa computacional FlexAnalysis v. 2.4. As amostras dos diferentes picos de rRnOBP foram liofilizadas e dissolvidas em água ultrapura (10-20 μ L). Após isto, em um parafilme, 1,0 μ L da amostra analisada foi misturado com 3 μ L de uma solução de matriz saturada de ácido α -cyano-4-hidroxicinâmico contendo 5 mg de ácido α -cyano-4-hidroxicinâmico, 200 μ L de água ultrapura, 250 μ L de acetonitrila e 50 μ L de ácido trifluoroacético). Ao ser realizada a mistura, 1 μ L desta foi depositado sobre uma placa de MTP AnchorChip 600/384 e em temperatura ambiente, e esperou-se a secagem completa das amostras para que as análises por meio da espectrometria de massa fossem realizadas.

Os espectros foram adquiridos a partir de 200 disparos de laser tipo SmartBeam à frequência de 100 Hz em modo linear. Para as análises foi realizada a calibração externa com uma mistura padrão de proteínas em uma faixa m/z 4000 a 100000, sendo que a faixa molecular do calibrante foi escolhida devido à massa da rRnOBP, aproximadamente, 17.000 Da.

4.2.7 Testes de inibição da agregação plaquetária e de coagulação

4.2.7.1 *Manutenção e coleta de glândulas salivares de Rhodnius neglectus*

A colônia de *R. neglectus* é mantida no biotério da UnB, onde os insetos são alojados em frascos de polietileno. A sala mantém a umidade relativa do ar em torno de 70% e segue um ciclo luminoso de 12 h de luz e 12 h de escuridão. Os animais são alimentados a cada quinze dias, durante um período de 30 min, com *Gallus Gallus domesticus*.

O processo de dissecação das glândulas salivares de *R. neglectus* foi conduzido com o auxílio de uma lupa, que ampliou o campo de visão para permitir o deslocamento e a separação da região do pescoço do inseto em relação ao tórax, que foi realizado com o auxílio de pinças. Dessa forma, o par de glândulas salivares foi coletado, lavado (para eliminação de tecido conjuntivo) e posteriormente homogeneizado dentro de microtubos contendo 20 µL PBS. 13 pares de glândulas foram coletados e mantidos em gelo durante a coleta. Após a coleta, os microtubos contendo as glândulas salivares foram centrifugados a 16.000 xg, 4 °C, por 10 min e o sobrenante foi transferido para um novo microtubo. Utilizou-se o conteúdo salivar equivalente a duas glândulas nos experimentos subsequentes, como controles da inibição dos testes.

4.2.7.2 *Teste de coagulação*

Foram pré-aquecidos 100 µL de CaCl² no coagulômetro CLOTimer Laser sensor, à temperatura de 37 °C. Em seguida, 100 µL da amostra (rRnOBP, PBS como controle negativo ou o conteúdo salivar de 2 glândulas salivares de *R. neglectus* como controle positivo da inibição) foram misturados com 100 µL do reagente TTPA CLOT (Teste de Tromboplastina Parcial Ativada) em um tubo separado, que também foi pré-aquecido a 37 °C por 3 min. Após esse período, 100 µL de CaCl² pré-aquecido foram adicionados ao tubo contendo a amostra e o reagente, e o temporizador foi iniciado para registrar o tempo de coagulação da amostra.

4.2.7.3 *Teste de agregação plaquetária*

O procedimento para a coleta de sangue humano usado para a realização desses testes foi aprovado pela Comissão de Ética em Pesquisa Humana (CEP) da Universidade de Brasília (**CAAE**: 48142021.1.0000.8093).

O sangue foi obtido de um doador voluntário saudável do sexo masculino, após um jejum mínimo de 6 h. O doador confirmou que não utilizou qualquer medicamento que pudesse afetar a agregação plaquetária por pelo menos 10 dias. A coleta do sangue foi realizada da veia cubital por meio de punção venosa com agulha de segurança 20 G, em tubos contendo citrato de sódio a 3,8%, invertidos 3 vezes para garantir uma mistura eficaz do anticoagulante. O plasma rico em plaquetas (PRP) e o plasma pobre em plaquetas (PPP) foram obtidos por meio de centrifugação do sangue a 100 x g por 10 min e 1500 x g por 15 min, respectivamente, a uma temperatura de 4°C. O PRP foi deixado em repouso por 30 min à temperatura ambiente antes do início dos testes. Os testes foram conduzidos dentro de um prazo máximo de 4 h após a coleta. O método se baseou na formação de agregados plaquetários após a exposição a um agonista (ADP 5 µM).

No agregômetro, 300 µL de PRP e PPP foram adicionados separadamente a cubetas do agregômetro e incubados por 10 min a 37°C. O PRP foi calibrado em relação ao PPP. Em seguida, rRnOBP (30 µL), ou o extrato de 2 glândulas salivares de *R. neglectus* (controle da inibição; 30 µL), ou PBS (30 µL) foram adicionados ao PRP e incubados por 5 min a 37°C. Após esse período, o agonista ADP (5 µM) foi adicionado às cubetas e a curva de agregação plaquetária foi registrada por 5 min, a 600 nm. O controle positivo da agregação plaquetária foi obtido incubando o PRP com ADP na ausência de saliva (PBS). Todas as reações foram conduzidas no agregômetro EasyAgreg (Qualitem). Esses ensaios iniciais foram realizados em duplicata.

5 RESULTADOS

5.1 ENSAIOS COMPUTACIONAIS

5.1.1 Análise da sequência

A sequência aminoacídica da RnOBP (GenBank: GDKW01001763.1) contém 131 resíduos de aminoácidos, que dão origem a uma proteína de massa molecular igual a 14.913 Da. Os 20 primeiros resíduos na porção N-terminal codificam para um peptídeo sinal, o que indica que a RnOBP é possivelmente secretada pela via clássica. A clivagem do peptídeo sinal dá origem a uma proteína composta por 111 resíduos de aminoácidos e 12.582 Da, com ponto isoelétrico (PI) teórico de 4,97. As sequências exibem 14 resíduos propícios para fosforilação, dois resíduos possíveis para N-glicosilação e nenhum para O-glicosilação. A rRnOBP, embora não apresente o peptídeo sinal, foi expressa como uma proteína de 16,7 kDa, consistindo em 147 resíduos de aminoácidos em decorrência da adição na porção N-terminal *tag* do epítipo Xpress, da cauda T7 e da cauda de histidinas, para reconhecimento imunológico (Tabela 2). Os dados da proteína recombinante foram utilizados para a pesquisa experimental enquanto os da RnOBP sem o peptídeo sinal foram utilizados para os ensaios *in-silico*.

Tabela 2. Propriedades calculadas da RnOBP contendo o peptídeo sinal (A0A0N7Z921), sem o peptídeo sinal e recombinante (rRnOBP).

	RnOBP (A0A0N7Z921)	RnOBP sem peptídeo sinal	rRnOBP
Programas			
ProtParam			
No. de resíduos	131	111	147
Massa molecular	14,9 kDa	12,6 kDa	16,7 kDa
pl	5,11	4,97	5,45
SignalP	1-20 aa	-	-
NetPhos			
Serina		8*	
Treonina		4*	
Tirosina		2*	
NetNGlyc		N55 e N80*	
NetOGlyc			
Serina		-*	
Treonina		-*	

* - dados compartilhados por todas as sequências

5.1.2 Predição da estrutura tridimensional da RnOBP

A sequência primária sem o peptídeo sinal da RnOBP passou por modelagem molecular para predição da estrutura 3D em diversos servidores, incluindo o SwissModel, TrRosetta, iTasser e AlphaFold 2. Cada um desses servidores emprega abordagens distintas de modelagem. Para avaliar a qualidade dos modelos resultantes (originais e minimizados), todos foram submetidos ao servidor MolProbity, que fornece pontuações indicativas de qualidade para os modelos (Tabela 3). A análise das pontuações dos modelos revelou que dois servidores se destacaram em relação aos demais. O TrRosetta, após duas minimizações de energia, e o AlphaFold 2, após uma minimização, obtiveram as maiores pontuações (Tabela 3). A escolha do modelo a ser utilizado levou em consideração o critério de desempate baseado no Molprobity Score, uma pontuação mais abrangente que combina diversos indicadores em um único valor. Notavelmente, o modelo gerado pelo AlphaFold 2 obteve as melhores pontuações globais, o que o tornou a escolha para os estudos estruturais subsequentes.

Tabela 3. Pontuações de qualidade dos principais modelos da RnOBP sem o peptídeo sinal, obtidas no servidor Molprobity e Verify

Servidor (minimização)	MolProbity score(-)	ClashScore(-)	Ramma Favored (>98%)	Verify(>80%)
TrRosetta (original)	0.83 (100th)	0.57 (99th)	106 (97.25%)	67.57%
TrRosetta (1)	1.25 (99th)	2.28 (99th)	105 (96.33%)	77.48%
TrRosetta (2)	0.69 (100th)	0.57 (99th)	109 (100%)	100%
iTasser (original)	2.49 (47%)	2.86 (98th)	100 (91%)	67.57%
iTasser (1)	0.96 (100th)	1.14 (99th)	106 (97.25%)	67.57%
iTasser (2)	0.69 (100th)	0.57(100th)	109 (100%)	78.38%
SwissModel (original)	1.62 (92nd)	4.87 (94th)	98 (95.15%)	49.52%
SwissModel (1)	0.70 (100th)	0.61 (99th)	101 (98.06%)	74.29%
Alphafold2 (1)	0.50 (100th)	0 (100th)	108 (99.08%)	82.75%

O modelo gerado mostrou que a estrutura predita para RnOBP é totalmente formada por α -hélices, contendo 6 dessas estruturas separadas por regiões de loops, a proteína é globular e possui pontes dissulfeto que conectam as α -hélices 1 a 3, 3 a 6 e 6 a 5 (Figura 6).

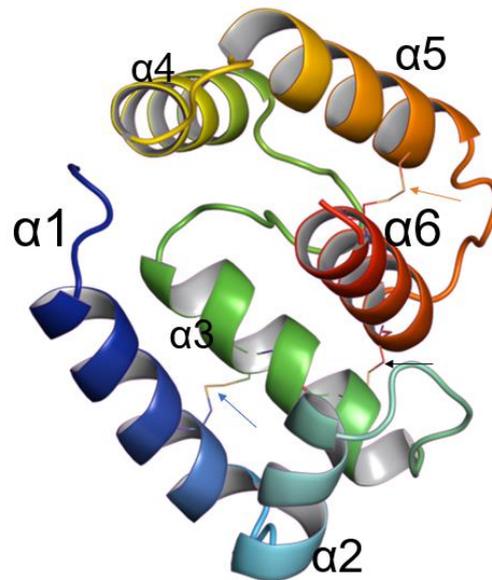


Figura 6 Modelo da RnOBP gerado pelo programa AlphaFold2. A RnOBP possui 6 α -Hélices enumeradas no N-terminal ao C-terminal e interligadas por setas, seta azul mostra a ponte dissulfeto entre as alfas hélices 1 e 3, a seta preta indica a ponte dissulfeto entre as alfas hélices 3 e 6 e em amarelo a ponte dissulfeto entre as alfas hélices 6 e 5.

5.1.2.1 *Sítios de ligação da RnOBP.*

Para a análise preliminar dos sítios de ligação, empregou-se o servidor de predição de cavidades potenciais de ligação, conhecido como CASTp (86). Este servidor identificou uma região de fenda que se revelou altamente conservada nas estruturas gerais de OBPs de insetos. Para avaliar a relevância biológica da região sugerida pelo programa como potencial sítio de ligação para a RnOBP sem o peptídeo sinal, foram realizadas sobreposições de diversos tipos de OBPs disponíveis no servidor PDB, utilizando o programa Pymol. Essas sobreposições confirmaram que os ligantes se acumulam na região prevista (Figura 8).

O sítio de ligação previsto para a RnOBP sem o peptídeo sinal é composto pelos resíduos: Q1, N4, E5, W8, E12, W67, E71, Y105, F108, E109, R111, abrangendo uma área de contato de 80,761 Å² e um volume de 29,265 Å³ (Figura 7A). A região que compreende o pocket apresenta uma fenda interna contendo aminoácidos hidrofóbicos e aromáticos, como W8, Y105 e F108. A entrada do pocket é predominantemente formada por aminoácidos ácidos, como Q1, E5 e E71, e inclui aminoácidos alcalinos, como R111 (Figura 7B e C).

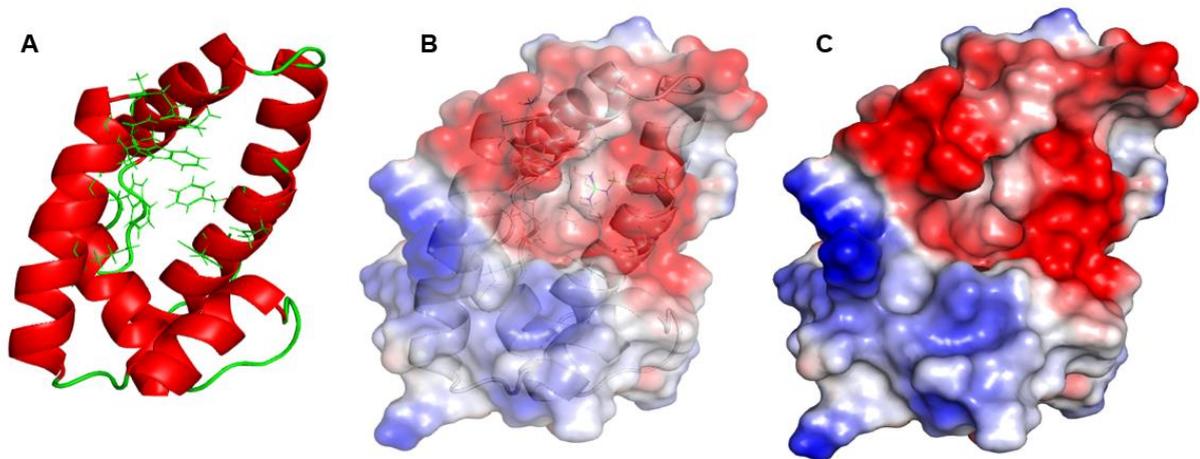


Figura 7 . Modelos e propriedades da superfície eletrostática da RnOBP Ativa. A) modelo representando em cartoon com os aminoácidos pertencentes ao bolso de ligação sendo representados em sticks e coloridos de verde. B) Sobreposição do modelo em cartoon com o modelo em surface, com as cargas negativas representadas pela cor vermelho e as positivas representadas pela cor azul. C) modelo de superfície eletrostática.

As sobreposições das estruturas (3q8i de *Anopheles gambiae*, 3r72 de *Apis mellifera*, 3vb1 *Anopheles gambiae*, 4f7f de *Anopheles gambiae*, 5dic de *Phormia regina*, 6jpm de *Chrysopa pallens*, 1ooh de *Drosophila melanogaster* e 3b6x) mostraram que independentemente da composição ou da classificação da OBP, a região do sítio de ligação é bem conservada (Figura 8).



Figura 8. Sobreposição de estruturas de OBPs com o modelo da RnOBP. As estruturas retiradas do PDB (3q8i, 3r72, 3vb1, 4f7f, 5dic e 6jpm) de *Anopheles gambiae*, *Drosophila melanogaster*, *Phormia regina* *Chrysopa pallens* dos grupos PBP, OBP gustativa e GOBP foram sobrepostas e seus ligantes se ligam na mesma região

5.1.3 Ensaio de oligomerização da rRnOBP

A fim de investigar estruturas oligoméricas que a rRnOBP (OBP de *Rhodnius neglectus* recombinante) poderia assumir, foi utilizada a plataforma ColabFold, que oferece recursos computacionais especializados nesse tipo de análise. (2) utilizando o programa AlphaMultimer v03(105). De maneira surpreendente, identificou-se uma conformação baseada em "domain swapping" interligada por pontes dissulfeto (Figura 9). Nessa configuração, a porção N-terminal da sequência da rRnOBP se separa de forma singular e interage com a porção N-terminal de outra molécula da mesma proteína. O que chama atenção é que esse processo ocorre sem levar em consideração a cauda N-terminal onde estão presentes as *tags* da proteína (T7, cauda de histidinas e epítipo Xpress). Em vez disso, a interação envolve a porção dos primeiros 35 resíduos, partindo do Q47, que representa o primeiro aminoácido nativo da RnOBP.

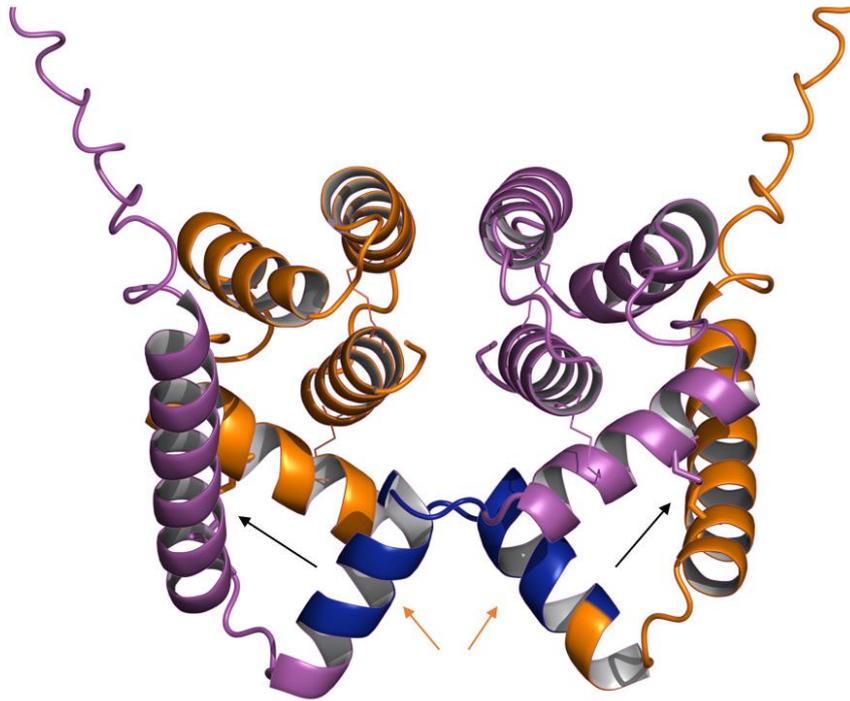


Figura 10 Modelo de dimerização Alpha Multimer. Em azul, apontado por setas laranjas está a porção da proteína que se desdobra para formar o Domain Swap. As setas pretas apontam as pontes dissulfeto formadas pelas cisteínas cruzadas.

De forma interessante, as cisteínas C52 e C83, que normalmente interagem na estrutura prevista para o monômero, também foram observadas interagindo nas conformações homoméricas. No entanto, o fenômeno do "*domain swapping*" parece estabilizar essas estruturas oligoméricas por meio de uma ligação covalente, conferindo um novo nível de estabilidade à configuração.

5.1.4 Dinâmica molecular da RnOBP

A partir do melhor modelo 3D da RnOBP obtido pelo Alphafold e minimizado pelo GalaxyRefine, foram conduzidas três simulações de dinâmica molecular com duração de 100 ns. Esta simulação tinha como objetivo examinar o comportamento flexível da proteína em um ambiente aquoso contendo NaCl 150 mM.

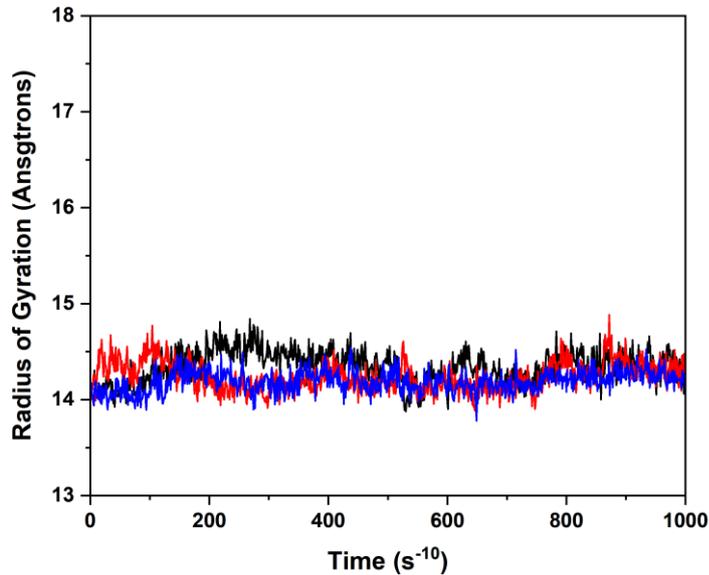


Figura 11. Raio de giro da triplicata da dinâmica molecular realizada com o modelo da RnOBP
Com a finalidade de avaliar a estabilidade e a conformação da RnOBP sem o peptídeo sinal, foram calculados os desvios quadráticos médios (RMSD) para as três dinâmicas. Observou-se que essas dinâmicas apresentaram comportamentos semelhantes, resultando em gráficos de que se estabilizaram em torno de uma diferença de cerca de 2.5 Å em relação ao frame 1, que corresponde ao modelo de referência previamente estabelecido (Figura 11A e B).

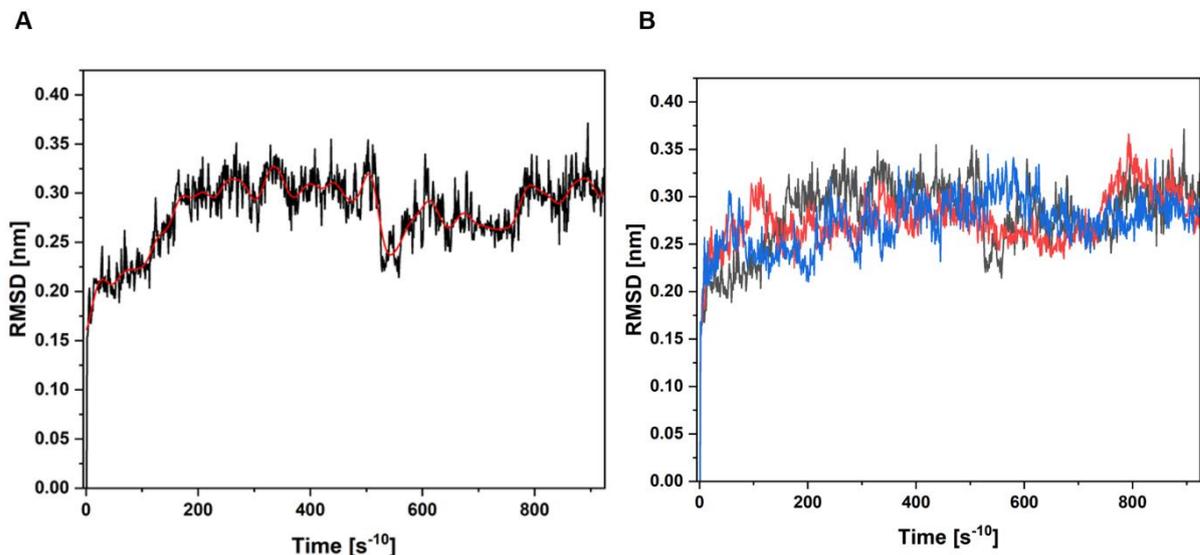


Figura 12. RMSD. A) Gráfico de RMSD pelo tempo da primeira dinâmica molecular realizada mostrando a variação do RMSD em relação ao frame 1, que é o modelo estático do alphafold. A linha vermelha é uma suavização por FFT (fast Fourier transform) da linha preta que representa a variação do RMSD. **B)** Mesmo gráfico representando a triplicata sobreposta das 3 dinâmicas realizadas.

O Raio de Giro (RG), que nos informa se a proteína gira ou realiza movimentos periódicos de contração e expansão, também foi calculado. A flutuação do RG da proteína se manteve estável, por volta de 14 Å, demonstrando que não havia

desenovelamento da proteína durante a trajetória realizada na dinâmica molecular (Figura 10).

O comportamento dinâmico da proteína revelou algumas regiões com maior flexibilidade. No gráfico de flutuação média quadrática (*Root Mean Square Fluctuation* [RMSF]) pode se observar que a região entre o resíduo R25 e o I35 (Figura 12A) apresenta a maior mobilidade da cadeia principal da proteína, sendo esta região um loop que está disposto após a segunda alfa hélice na estrutura da RnOBP sem o peptídeo sinal (Figura 12B).

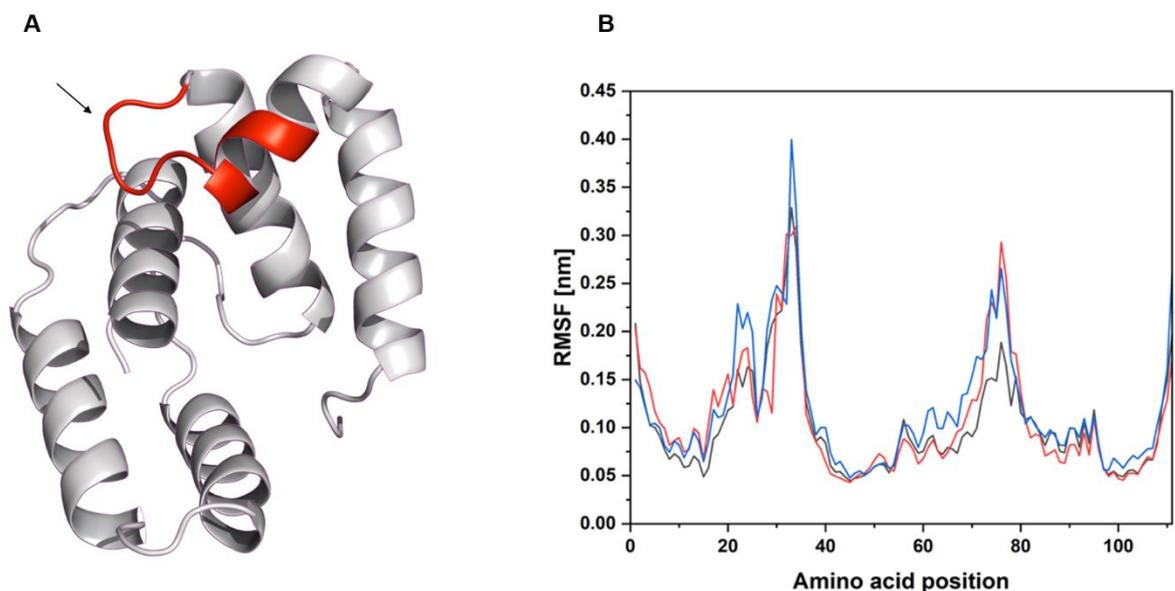


Figura 13. Flexibilidade e RMSF. A) Modelo 3D da rNOBP com os aminoácidos da posição 25 a 35 sendo representados pela cor vermelha. B) gráfico de RMSF pela posição dos aminoácidos mostrando a variação da posição média destes na triplicata da dinâmica molecular, dinâmica 1 sendo representada pelo azul, a 2 pelo vermelho e a 3 pelo preto.

5.1.5 Clusterização das conformações da RnOBP

Após calcular a flexibilidade da RnOBP sem o peptídeo sinal por meio de Dinâmica Molecular (DM), foi realizada a clusterização das várias conformações adotadas pela proteína ao longo da trajetória. Esse processo de agrupar as conformações observadas é essencial para identificar as posições mais frequentemente assumidas e permitir a representação da flexibilidade da proteína de forma aprimorada para os experimentos subsequentes de *docking* molecular, que normalmente tratam a proteína como uma estrutura rígida. As conformações semelhantes revelaram informações valiosas sobre a frequência com que ocorrem durante a trajetória, ao mesmo tempo em que forneceram *insights* sobre o

comportamento do sítio de ligação da proteína, visando melhorar a precisão nas etapas subsequentes de *docking* molecular.

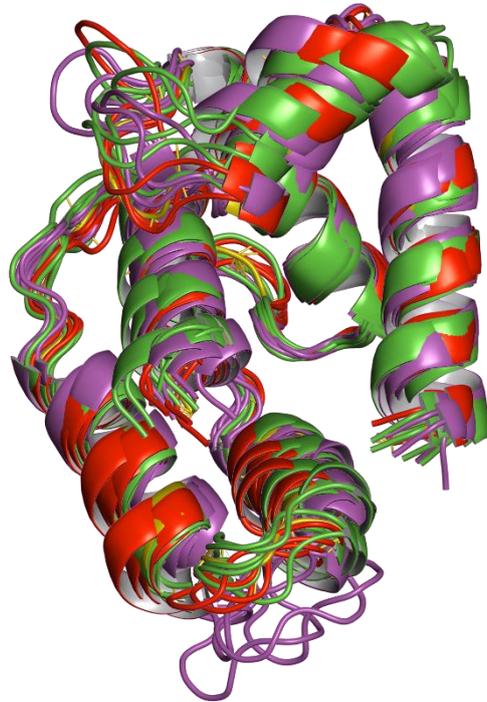


Figura 14 Superposição dos frames representativos da dinâmica molecular da RnOBP. Frames em verde retirados da dinâmica 1, frames em magenta retirados da dinâmica 2 e frames em vermelho retirados da dinâmica 3.

Para realizar a clusterização, foi utilizado o software TTclust (93) e aplicado o método Ward, conforme detalhado na seção de metodologia. A análise de clusterização foi aplicada aos arquivos das dinâmicas moleculares realizadas no

modelo gerado. Todas as conformações obtidas foram comparadas com o modelo estático gerado pelo AlphaFold (Figura 13).

Foram identificadas as conformações representativas que capturaram os movimentos dinâmicos da RnOBP e calculados os valores de RMSD entre cada uma delas, os quais estão listados no (Apêndice A – Tabelas 1, 2 e 3). Esses resultados proporcionam informações sobre a flexibilidade estrutural da proteína, o que altera a capacidade de ligação aos ligantes.

5.1.6 Seleção de compostos promissores

Foram selecionados os seguintes compostos como sendo promissores ligantes para proteínas D7: o análogo de tromboxano (U46619) e um leucotrieno cisteínico (43), bem como aminas biogênicas como triptamina, histamina, norepinefrina e serotonina (52) e ADP (50) (Figura 14). Esses compostos foram selecionados devido a suas capacidades de influenciar na hemostasia de animais vertebrados e também por representarem os 2 grupos principais de moléculas que interagem com proteínas D7, aminas biogênicas e eicosanoides.

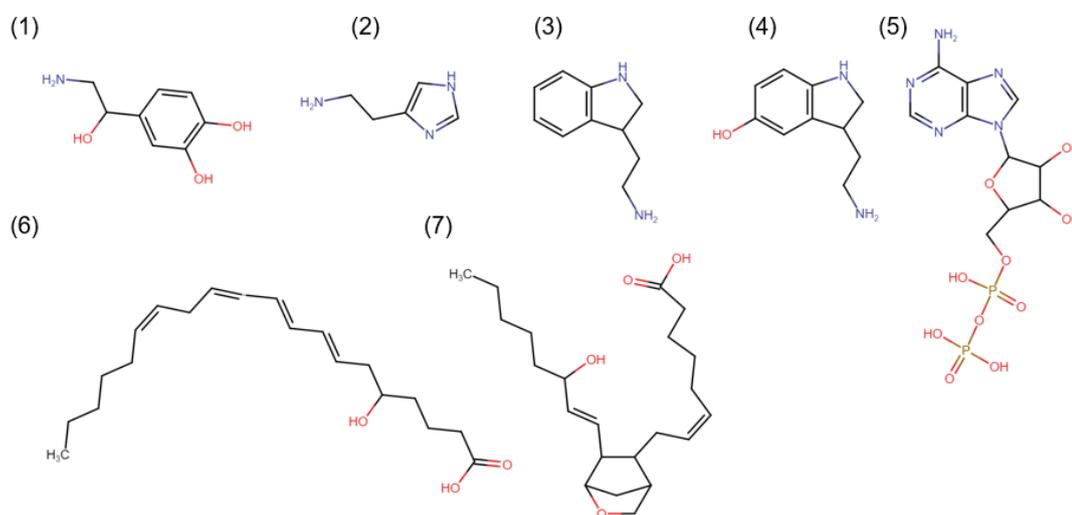


Figura 15 Ligantes selecionados para ensaio de *ensemble docking*. Em (1) está representado a norepinefrina, em (2) histamina, em (3) triptamina, em (4) serotonina, em (5) ADP, em (6) Leucotrieno C4, em (7) U46619, um análogo do tromboxano 2A.

5.1.7 Docking molecular da RnOBP

5.1.7.1 Redocking dos ligantes selecionados.

Para fins de validação, procedeu-se ao *redocking* de cada ligante em suas estruturas originais, utilizando como referência as estruturas 2PQL para triptamina, 2QEB para histamina, 2QEH para serotonina, 2QEO para norepinefrina, 3NHI para leucotrieno C4, 3NHT para U46619 e 6V4C para ADP. O cálculo do valor de RMSD foi realizado e avaliado. A correspondência das posições de ligação encontradas no *redocking* em relação às interações originais nos complexos proteína-ligante depositados no PDB foi alta, como indicado pelo baixo RMSD (Tabela 4). As afinidades resultantes do *redocking* foram usadas para fins de comparação com as energias obtidas no *ensemble docking* (Tabela 4).

Os resultados do *redocking* revelaram afinidades de -8.36, -7.21 e -8.42 Kcal/mol para as aminas biogênicas triptamina, histamina e serotonina, e -10.84 e -10.16 Kcal/mol para os eicosanoides leucotrieno C4 e U46618 (análogo de TXA₂), respectivamente. As poses encontradas no *redocking* mostraram-se muito próximas das observadas nas estruturas cristalinas depositadas no PDB (2PQL, 2QEB, 2QEH, 3NHI, 3NHT). No entanto, para avaliar as pontuações de 2QEO e 6V4C, que correspondem às estruturas dos ligantes norepinefrina e ADP, respectivamente, não foram alcançados valores de RMSD satisfatórios (Tabela 4).

Tabela 4. Afinidades de ligação e RMSD dos ligantes do procedimento de *redocking*.

<i>Redocking</i> PDB ID	RMSD (Å)	Score de <i>docking</i> (Kcal·mol ⁻¹)
2pql	0.70	-8.36
2qeb	1.20	-7.21
2qeh	0.65	-8.42
2qeo	-	-
3nhi	0.97	-10.84
3nht	0.48	-10.16
6v4c	-	-

5.1.7.2 Ensemble docking das conformações da RnOBP com os possíveis ligantes

Para avaliar a afinidade da RnOBP com os ligantes selecionados, foram conduzidos sete procedimentos de *docking*, um para cada ligante, considerando

estruturas representativas da flexibilidade da proteína a partir de três dinâmicas moleculares distintas (Tabela 5). As pontuações resultantes desses procedimentos de *ensemble docking* foram superiores às obtidas no *redocking*, conforme registrado na Tabela 5. Essa diferença na afinidade entre a RnOBP e a estrutura original dos ligantes foi de 0,64 Kcal/mol para a triptamina, 0,42 Kcal/mol para a histamina, 0,75 Kcal/mol para a serotonina, 2,72 Kcal/mol para o leucotrieno C4 e 1,99 Kcal/mol para U46619. Quanto maior a diferença de afinidade, menor é a interação entre a proteína e o ligante.

Tabela 5. Pontuações de afinidade do *Ensemble docking* da RnOBP.

Ligante	Afinidade de ligação (Kcal·mol ⁻¹)		
	MD1 (C6)	MD2 (C5)	MD3 (C4)
Triptamina	-7.72	-7.07	-7.36
Histamina	-6.58	-6.65	-6.79
Serotonina	-7.67	-7.16	-7.21
Norepinefrina	-7.13	-6.57	-7.16
Leucotrieno C4	-8.12	-7.41	-8.06
U46619	-8.17	-8.14	-7.57
ADP	-6.71	-6.96	-7.06

Os resíduos fundamentais para a estabilização das aminas biogênicas, conforme revelado no ensaio de *ensemble docking*, compreendem E109 e E71, os quais estão presentes em todas as interações, formando pontes de hidrogênio, com a exceção de E71 na triptamina. Além disso, no grupo das aminas biogênicas, F108, Y105 e W8 desempenham papéis de interações hidrofóbicas significativas.

Tabela 6 Resíduos que estabilizaram os ligantes nas interações com a RnOBP

	Ligação de hidrogênio	Pontes salinas	Interações hidrofóbicas	π -stacking
Triptamina	E109	-	Y105	W8
Histamina	E71, E109	-	-	-
Serotonina	E71, E5, E109	-	F108	-
Norepinefrina	E71, E109	-	-	-
Leucotrieno	-	K10, K13	G9, F108	-
U46619	R26	K13	-	-
ADP	E12, K13	K13, ARG26	-	-

O resíduo F108 também contribui para a estabilização da cadeia hidrofóbica do leucotrieno C4, em conjunto com o resíduo G9. As interações polares estabilizadoras identificadas no grupo dos eicosanoides estão relacionadas a pontes salinas formadas por resíduos básicos, como L13 e L10, que interagem com os grupos carboxilas do leucotrieno C4 e U416619, como detalhado na Tabela 6 e demonstrado na Figura 16.

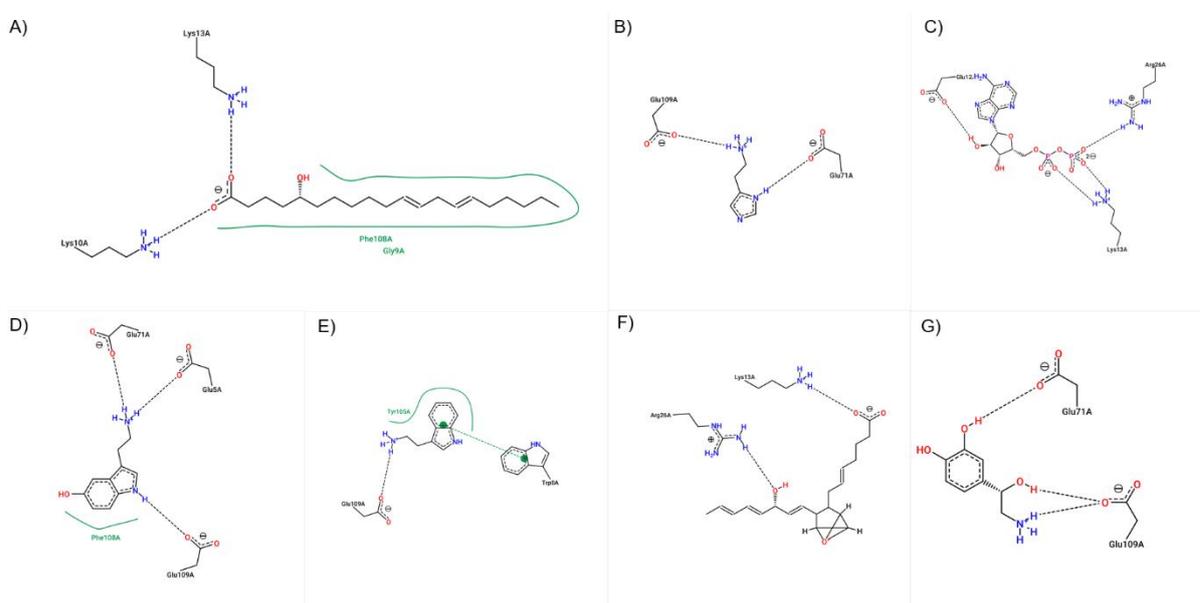


Figura 16 Contatos do *ensemble docking* da RnOBP com ligantes priorizados. O *ensemble docking* foi realizado com os ligantes co-cristalizados nas estruturas A) 3nhi (leucotrieno C4), B) 2qeb (histamina), 6v4c (ADP), D) 2qeh (serotonina), E) 2pql (triptamina), F) 3nht (u46619[análogo de tromboxano A2]), G) 2qeo (norepinefrina). Os contatos com os resíduos estão explicitados pelo código de três letras seguido pela posição do aminoácido.

5.2 ENSAIOS EXPERIMENTAIS

5.2.1 Expressão e purificação da rRnOBP

A tentativa de expressar a proteína recombinante OBP de *R. neglectus* (rRnOBP) foi realizada nas linhagens bacterianas Lemo21(DE3), Rosetta(DE3)pLysS e Origami(DE3). Os extratos bacterianos foram obtidos após a indução da expressão da rRnOBP e separados em frações solúvel e insolúvel, e posteriormente analisados em gel SDS-PAGE corado por azul de Coomassie. Uma banda de proteína mais intensa acima de 15 kDa (massa molecular esperada 16,7 kDa) foi observada nas frações insolúveis de clones das três linhagens transformadas com o plasmídeo contendo a sequência da RnOBP, mas não no clone transformado com um plasmídeo controle (C + AG5). A banda de proteína foi mais intensa no extrato insolúvel da linhagem Rosetta (DE3) pLysS (Figura 16). A banda de proteína de aproximadamente 16,7 kDa não estava presente nas frações antes da indução (AI), indicando que as colônias foram corretamente transformadas e que a banda deveria corresponder à rRnOBP (Figura 16).

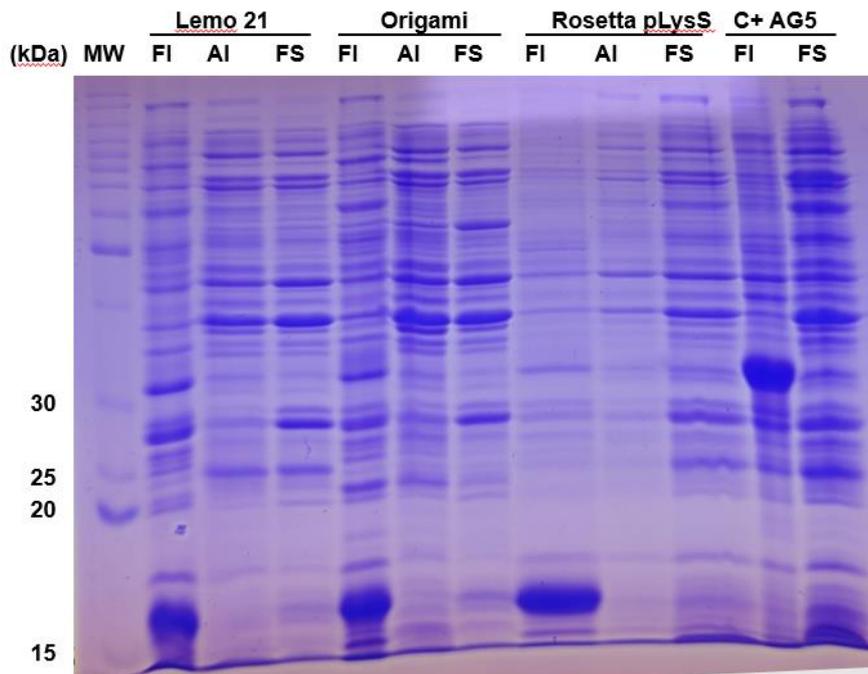


Figura 17. Expressão da RnOBP em linhagens diferentes de *Escherichia coli*. Clones de *E. coli* transformados com o plasmídeo RnGOBP_pET100/D-TOPO foram induzidos com 1mM de IPTG, por 3h a 37°C. Após esse período, as bactérias foram sedimentadas, lisadas e suas frações analisadas por eletroforese em gel SDS-PAGE 15% corado por azul de Coomassie. AI- amostra retirada antes da adição de IPTG, FI- amostra da fração insolúvel, FS amostra da fração solúvel. Lemo 21- linhagem LEMO21 DE3), Origami- linhagem Origami™ (DE3), Rosetta pLysS- linhagem Rosetta (DE3)pLysS, C+ AG5 é o controle com a linhagem Rosetta(DE3)pLysS transformada com o plasmídeo contendo o gene da proteína AG5 (31 kDa).MW – marcador de massa molecular. A seta azul indica a banda super expressa possível para a rRnOBP.

Na tentativa de otimizar a expressão da rRnOBP na fração solúvel, a temperatura, a concentração do indutor e o período de indução foram variados conforme pode ser observado na Figura 17. Os resultados mostraram que a condição de expressão mais eficiente utilizando a linhagem Rosetta (DE3) pLysS foi cultivo a 37 °C, concentração de 1 mM de IPTG com período de indução igual a 3h. Curiosamente, essa condição resultou em uma banda forte com o dobro da massa molecular, visualizada no *western blotting* (Figura 17).

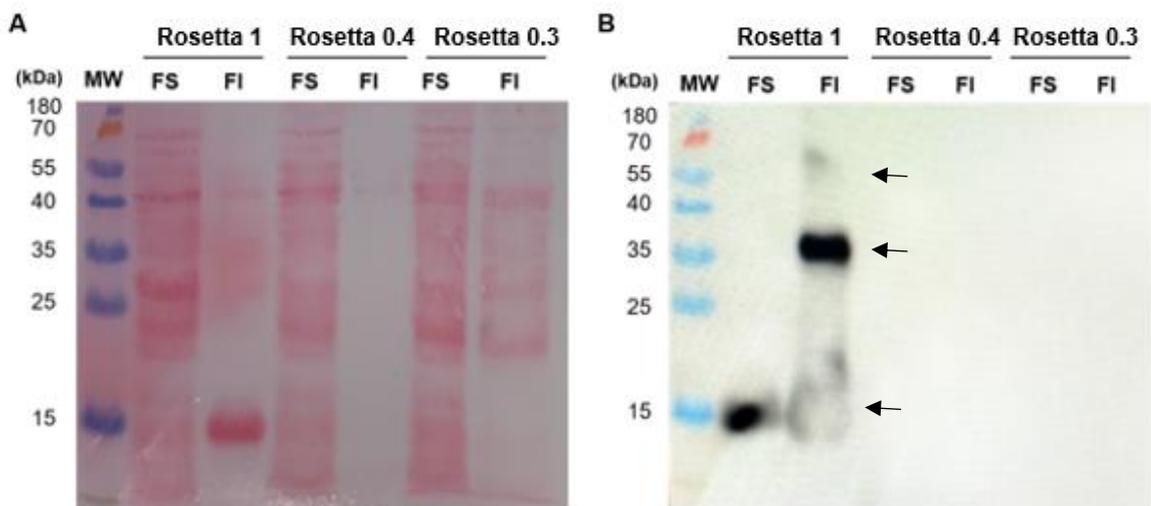


Figura 18. Expressão da rRnOBP na linhagem de *E. coli* [Rosetta (DE3) pLysS] em diferentes condições. A) Membrana de nitrocelulose transferida com os extratos totais das induções corada com ponceau. B) Western blotting da membrana em “A” usando anticorpo anti-His (1:5000). MW – Marcador de massa molecular, FS – fração solúvel, FI – Fração insolúvel, Rosetta 1 – amostras induzidas com 1mM de IPTG por 3 h a 37 °C, Rosetta 0.4 – amostras induzidas com 0.4 mM de IPTG por 15 h a 18 °C, Rosetta 0.3 – Amostras induzidas com 0.3 mM de IPTG por 20 h a 18 °C. Setas indicam as bandas marcadas com massas múltiplas.

Após a otimização das condições de indução, foram conduzidos testes de otimização de protocolo para a lise das bactérias. As bactérias foram lisadas imediatamente após o término do período de indução ou congeladas e lisadas 16 h depois. O número de pulsos de sonicação necessários também foi determinado. Essas avaliações permitiram observar que não houve diferença no padrão de bandas de proteínas nas condições testadas no gel de eletroforese (Figura 18).

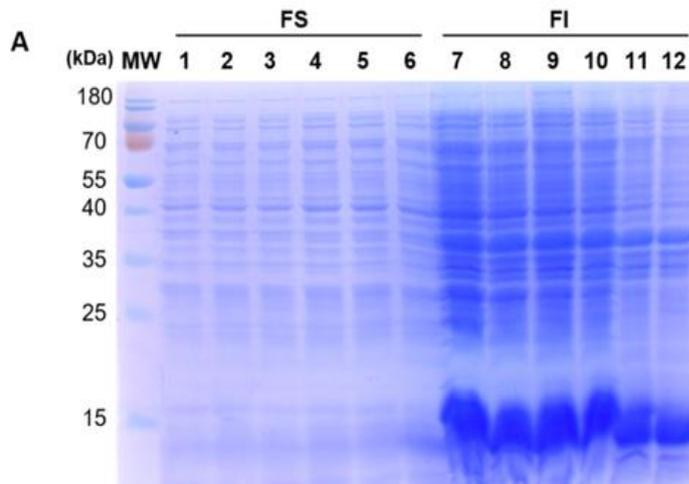


Figura 19 Ensaio de obtenção das proteínas induzidas em *E. coli* da linhagem Rosetta(DE3)pLysS imediatamente após o cultivo ou 16h após congelamento e descongelamento de sedimento de células. SDS-PAGE 15% corado por azul de Coomassie. Mw – marcador de massa molecular, FS – fração solúvel, FI – fração insolúvel, (1 a 3, 7 a 9) frações coletadas imediatamente após a indução, (4 a 6, 10 a 12) frações congeladas e lisadas após 16h, (1,4,7,10) 1 pulso de 30 segundos de sonicação, (2,5,8,11) dois pulsos de sonicação,

Após a elaboração do protocolo de expressão e sonicação, o extrato total bacteriano foi separado nas Frações Insolúvel (FI) e solúvel (FS), que foram submetidas separadamente a cromatografia de afinidade a íons metálicos (IMAC) usando uma coluna niquelada. Após lavagem exaustiva do conteúdo da amostra com concentrações de 25 e 50 mM de imidazol, foi observado um pico de proteínas proeminente quando a concentração de imidazol atingiu 200 mM (24 mL de volume), tanto na FI quanto na FS, como mostrado na Figura 19.

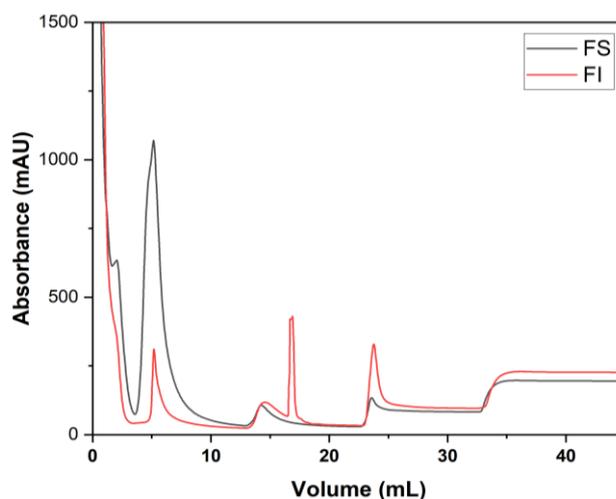


Figura 20. Sobreposição dos cromatogramas da fração solúvel (preto) e da fração insolúvel (vermelho). Cromatografia de afinidade realizada em um AKTA UPLC com uma coluna HisTrap FF em 4 etapas. 0 a 10 mL (lavagem com 25 mM imidazol), 10 a 20 mL (lavagem 50mM de imidazol) 20 a 30 mL (eluição com 200 mM imidazol), 30 a 40 mL (eluição com 500 mM imidazol).

Frações resultantes dessas cromatografias foram então submetidas a análises por eletroforese em gel de SDS-PAGE, o que levou à identificação de um pico eluído em aproximadamente 24 mL. Este pico exibia bandas correspondentes à massa do monômero, bem como bandas com o dobro e triplo dessa massa, conforme evidenciado no *western blotting* (Figura 20). Esses resultados sugerem a possível presença de formas oligoméricas da proteína, com um peso molecular estimado de cerca de 35 e 50 kDa (Figura 20). Além disso, a análise indicou que as bandas apresentavam um grau significativo de pureza.

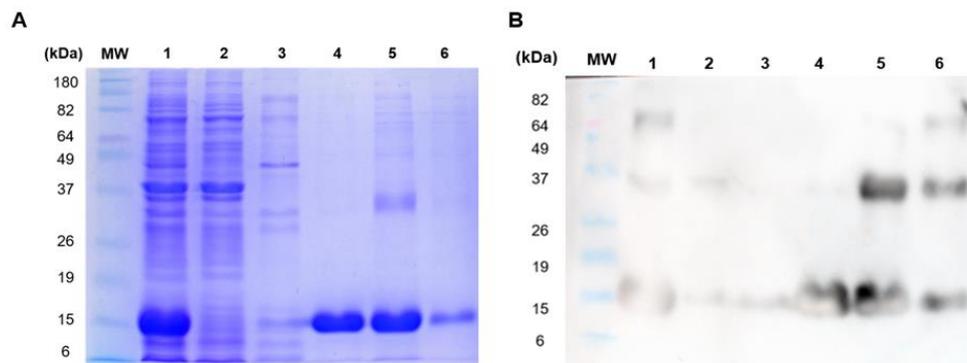


Figura 21 (SDS-PAGE) Purificação FI. Mw – Marcador, 1- extrato total da fração insolúvel, 2 – Lavagem da fração insolúvel (25mM imidazol), sendo representada pelo pico entre o volume 0 e o 10 do cromatograma da figura 13, 3 – Lavagem da fração insolúvel (50mM imidazol) pico entre o volume 10 e 20 do cromatograma da figura 13, - 4 e 5 eluição da RnOBP sendo representada pelo pico de 20 a 30 mL da figura 13, 6 - pico do volume 30 a 40 do cromatograma. A) SDS-PAGE 15%. B) Western Blotting do gel espelho transferido para a membrana de nitrocelulose e tratada com anti-His (1.5000).

No que diz respeito ao rendimento aparente, foi observado que a fração solúvel em todas as linhagens de expressão testadas continha uma quantidade limitada de proteína recombinante e não apresentava pureza. Esse cenário apresentou um desafio para a viabilidade da purificação e subsequente caracterização da rRnOBP usando essa fração (Figura 21).

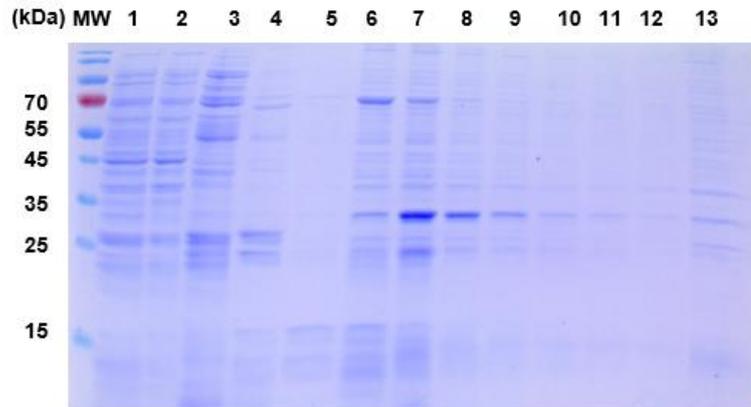


Figura 22 SDS-PAGE Purificação da FS. MW – Marcador, 1 – extrato total da fração solúvel, 2 amostras do não ligado a coluna, 3 - Lavagem com 25 mM de imidazol, 4 - lavagem com 50 mM de imidazol, (5 a 12) - eluições com 200 mM de imidazol e 13 – eluição com 500 mM de imidazol. SDS-PAGE 15% com as amostras explicitadas acima.

Dado que são necessárias quantidades substanciais de proteína para conduzir os experimentos de caracterização, optou-se por purificar a rRnOBP da fração insolúvel bacteriana e realizar o seu renovelamento. As eluições coletadas após a cromatografia de afinidade da FI foram reunidas e dialisadas para remoção da ureia e imidazol, substâncias que podem afetar tanto o enovelamento da proteína quanto interferir nos experimentos de caracterização. A diálise mesmo em tampões com pouca força iônica (Tris 5 mM pH 7.4) não influenciou na solubilidade da proteína.

As quantificações conduzidas com diferentes concentrações de rRnOBP, juntamente com a curva padrão de BSA, permitiram a obtenção do coeficiente de extinção experimental da rRnOBP (ϵ rRnOBP = 18436,86 L / (mol.cm) ou 1,1445 mL/(mgcm)). É relevante notar que esse valor ficou bastante próximo do coeficiente de extinção teórico (ϵ_1 RnOBP = 18825 L / (molcm) com cisteínas oxidadas ou ϵ_2 RnOBP = 18450 L / (mol.cm) com cisteínas reduzidas).

No que diz respeito ao rendimento das purificações, as amostras purificadas e dialisadas de rRnOBP demonstraram uma concentração de 0,69 mg/mL.

5.2.2 Ensaios de caracterização biofísica

5.2.2.1 Ultracentrifugação Analítica

Para obter informações sobre os parâmetros estruturais da rRnOBP em solução, experimentos de velocidade de sedimentação foram realizados por ultracentrifugação analítica. A amostra continha várias espécies de moléculas com coeficiente de sedimentação de 2 a 15 S (Figura 22). Com base nas características dos picos observados, a conversão para a massa molecular foi calculada utilizando a equação de Svedberg, resultando em valores que variaram de 17,763 kDa a 332,470 kDa. O primeiro valor é aproximadamente igual à massa molecular (16,7 kDa) do monômero calculada a partir da sequência de aminoácidos da rRnOBP.

Os dímeros, por sua vez, foram identificados em todas as amostras examinadas, independentemente da concentração. Eles apresentaram valores de sedimentação variando de 3,234 S a 3,969 S (correspondendo a massas entre 33,818 kDa e 43,570 kDa) e exibiram as populações mais significativas, em média. Essas populações variaram de 11,97% na amostra menos concentrada até 32% na amostra com 0,46 mg/mL (Figura 22B). Este resultado indica que a rRnOBP existe predominantemente como um dímero em solução e é propensa a formar oligômeros, como por exemplo tetrâmeros (aproximadamente 5,4 S), octâmeros (aproximadamente 8,5 S) e outras espécies oligoméricas (Figura 22B).

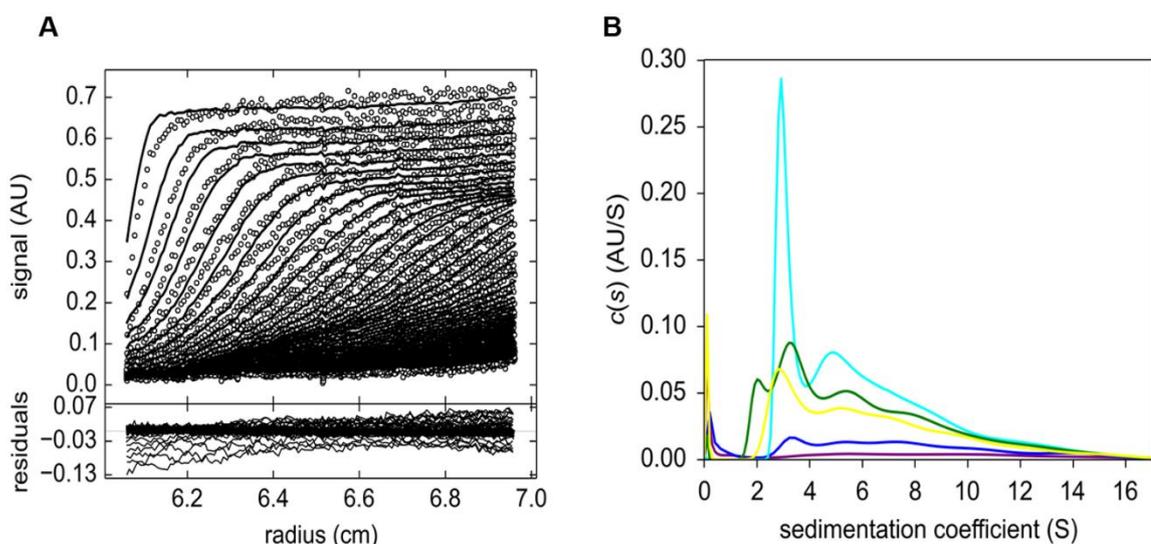


Figura 23 : Velocidade de sedimentação em ultracentrifuga analítica. A) Sobreposição de 231 leituras da sedimentação da rRnOBP realizadas no período de 16 horas, a 40000 rpm, 25 °C, em tampão 50mM tris, 150mM NaCl, ph 8 utilizando absorção a 280nm. Curvas modeladas pela análise $c(s)$ no sedfit. B) Gráfico de residuals correspondente a leitura de "A". C) Sobreposição dos padrões de sedimentação da rRnOBP em diferentes concentrações: 0.68 mg/ml (ciano), 0.49 mg/ml (verde), 0.46 mg/ml (amarelo), 0.37 mg/ml (azul), 0.18 mg/ml (magenta).

5.2.2.2 Cromatografia de exclusão molecular

Devido à observação de que a rRnOBP é propensa a formar polímeros, as frações contendo a rRnOBP obtidas da cromatografia de afinidade foram reunidas e submetidas a cromatografia de exclusão molecular utilizando uma coluna Superdex 200 10/300 GL. Após a cromatografia, os picos detectados a 280 nm foram examinados, e a massa molecular foi estimada empregando a equação K_{av} . O cromatograma gerado revelou quatro picos principais, cujas massas foram determinadas com base em um padrão de referência de uma mistura de quatro proteínas que abrangiam a faixa de 6,5 kDa a 145 kDa (Figura 23), conforme descrito em materiais e métodos. As massas moleculares dos polímeros da rRnOBP foram calculadas como sendo de 26 kDa, 39 kDa, 51 kDa e 260 kDa (Figura 23A).

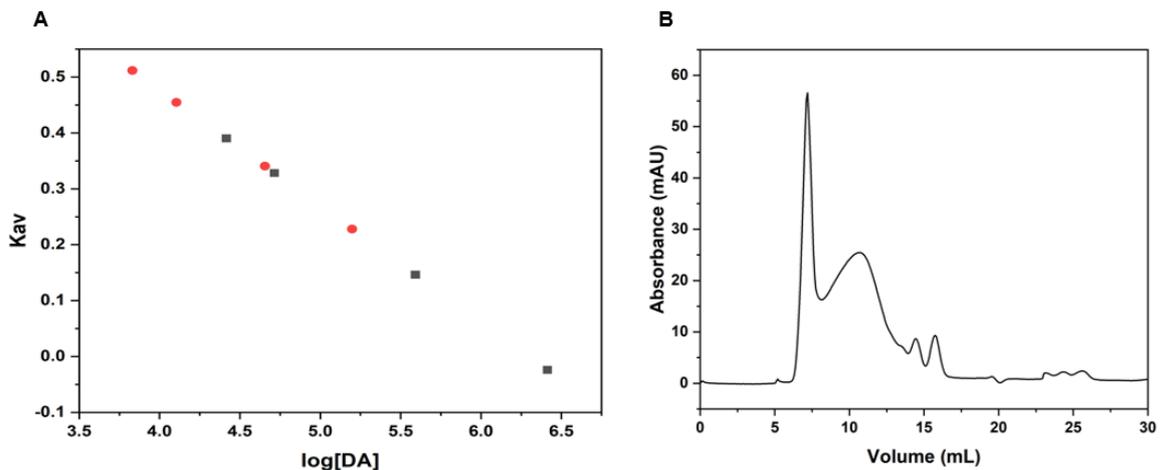


Figura 24. Purificação da rRnOBP por cromatografia de exclusão molecular. A) Quadrados pretos representam as quatro proteínas padrão usadas na curva de calibração da coluna Superdex 200 10/300, os pontos vermelhos representam as eluições dos picos da amostra purificada de RnOBP coletados a partir do cromatograma representado em B. B) Cromatograma da rRnOBP purificada e reenovelada.

5.2.2.3 Espalhamento de Luz Dinâmico

Para realizar obter informações sobre o raio hidrodinâmico e a qualidade da amostra, uma nova cromatografia de exclusão molecular foi realizada e as eluições foram coletadas em micro tubos para a realização do ensaio de DLS, como o ilustrado na Figura 24. Posteriormente, realizou-se um experimento de dispersão de luz para determinar o raio hidrodinâmico de cada espécie recuperada durante o processo cromatográfico.

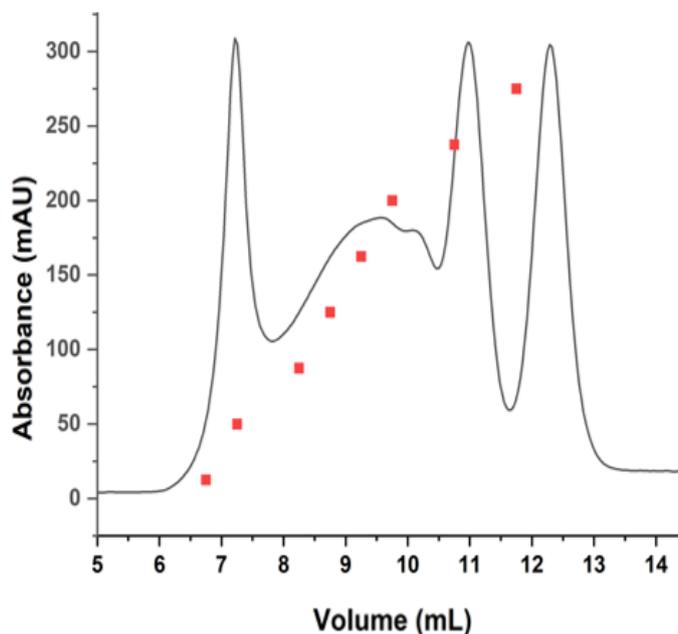


Figura 25 Exclusão molecular da rRnOBP renovelada. O cromatograma mostra o padrão de eluição da amostra em uma coluna de exclusão molecular Superdex 200 10/300. Os pontos em vermelho representam as eluições coletadas e nomeadas como B11, B10, B8, B7, B6, B5, B1 e C1, da esquerda para a direita.

A hipótese de pureza das alíquotas da rRnOBP foi corroborada por meio de análises de DLS, como indicado pela baixa polidispersividade nas amostras coletadas a partir da terceira eluição (Tabela 7). Essa baixa polidispersividade sugere a presença de uma única forma molecular, caracterizando a amostra como monodispersa.

Os resultados obtidos com o DLS revelaram moléculas de considerável tamanho nas frações coletadas nos primeiros picos da cromatografia de exclusão (Tabela 7), chegando a atingir 1,8 MDa, além de demonstrarem polidispersão. Entretanto, nas frações coletadas nos picos subsequentes foi observada uma redução nos diâmetros e massas, aproximando-se dos valores típicos de monômeros da rRnOBP.

Tabela 7. Espalhamento de Luz Dinâmico (DLS) das frações contendo a rRnOBP coletadas da cromatografia de exclusão molecular

Fração	d (nm)	%Pd	KDa	%int	%mass	Poli/monodisperso
B11	29,62±6,475	21,9	1850±756,7	24,2	100	polidisperso
B10	37,37±20,23	50,9	3170±1,870	82,9	99,9	polidisperso
B8	4,455±0,1535	3,5	21,9±3,5	2,3	24,8	monodisperso
B8	6,156±0,2359	3,9	46,7±2,5	9,5	32,4	monodisperso
B8	10,72±1,162	12,2	171±7,4	7,7	23,8	monodisperso
B7	8,909±0,309	3,6	111±13,5	3,2	30,4	monodisperso
B7	11,22±1,248	11	190,6±11	8,8	68,2	monodisperso
B6	8,507±0,6048	7,1	99,6±7,2	58,8	99,4	monodisperso
B5	6,752±0,1763	2,6	58±1,1	46,3	99,9	monodisperso
B3	1,470±0,1205	8,1	1,6±0,3	66	95,3	monodisperso
B3	8,123±1,526	18,5	89,4±15,4	46,4	4,7	monodisperso
B1	3,078±0,1760	5,7	9,2±0,4	24,2	100	monodisperso

5.2.2.4 *Análise de espectrometria de massas por MALDI-TOF da rRnOBP*

Após a análise de dispersão de luz das frações coletadas pela exclusão molecular, as mesmas amostras foram submetidas a espectrometria de massas por MALDI-TOF, e foi revelada a presença em todas as amostras de picos correspondendo aos valores de massa molecular próximos aos esperados, como ilustrado na Figura 25, reforçando a suposição da pureza da proteína. Foram observadas massas correspondentes ao dobro, triplo e quádruplo do tamanho esperado para o monômero da rRnOBP, indicando a ocorrência de oligomerização entre as proteínas rRnOBP (Figura 25).

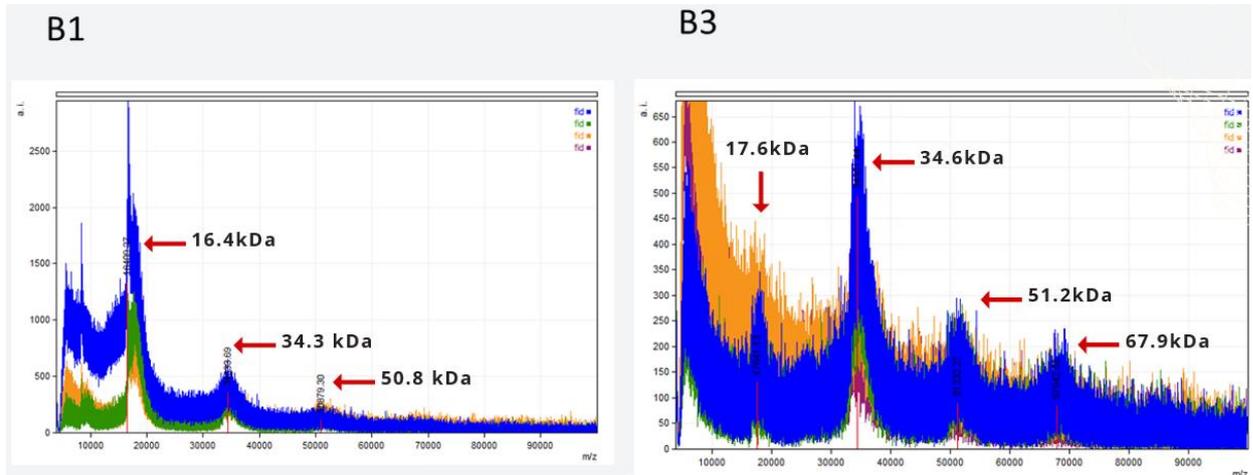


Figura 26. Espectrometria de massas (MALDI-TOF) das amostras coletadas da exclusão molecular realizada na figura 21. B1 – Amostra da fração B1 analisada por DLS na Tabela 7, B3 – Amostra da fração B3 analisada por DLS na Tabela 7.

5.2.2.5 *Cromatografia de exclusão molecular na presença de uréia*

Para validar o resultado anterior, um ensaio de cromatografia de exclusão molecular sob condições desnaturantes foi realizado (Figura 26B). O objetivo desta cromatografia foi confirmar a formação de ligações covalentes entre diferentes cadeias peptídicas. A hipótese era que, se as interações entre as cadeias fossem de natureza covalente, o padrão de picos cromatográficos na presença do agente caotrópico não se diferenciaria do padrão observado na ausência do mesmo (Figura 9). Esse cenário fortalece a viabilidade do modelo predito pelo AlphaFold Multimer, fornecendo mais suporte à conformação estrutural sugerida.

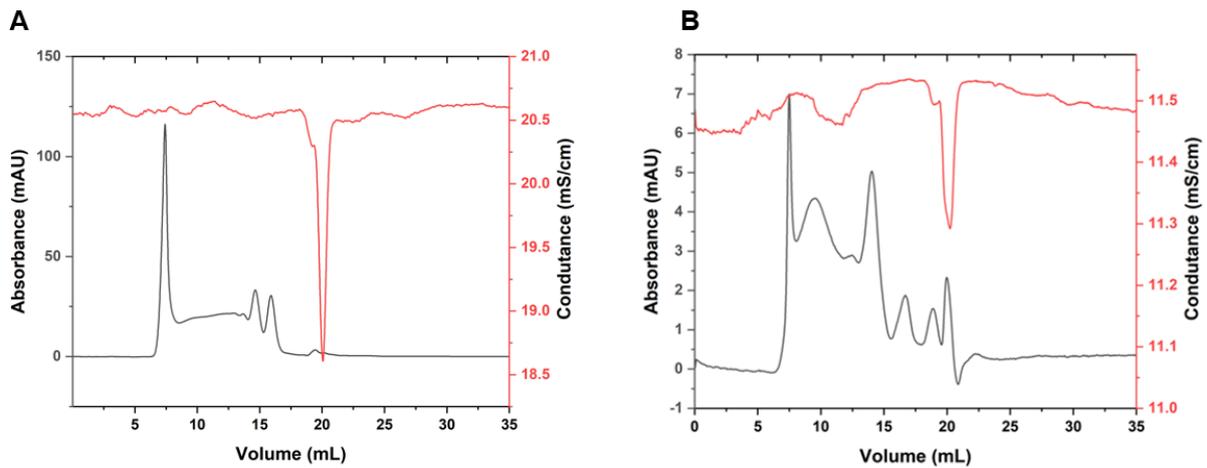


Figura 27. Cromatogramas exclusão molecular rRnOBP em condição desnaturante e nativa. A) Exclusão molecular em condições nativas realizadas com a amostra de RnOBP purificada e concentrada, linha preta representa a absorbância da amostra em 280nm e a linha vermelha representa a condutividade da amostra pelo volume passado pela coluna. B) exclusão molecular realizada em condição desnaturante, com Ureia 8M na coluna de exclusão molecular Superdex 200 10/300.

5.2.2.6 Caracterização estrutural da rRnOBP

Com base no modelo gerado pelo Alphafold Multimer, que abrange dímeros (Figura 9), trímeros, tetrâmeros (Apêndice B), e considerando os resultados previamente obtidos através da cromatografia sob condições desnaturantes (Figura 26B), foram concentrados os esforços na confirmação experimental da validade da estrutura prevista. Para verificar a consonância entre a estrutura experimental e a teórica, a amostra que passou por renovação e extensa diálise em tampão 5mM Tris pH 7.2 foi submetida a análises de dicróismo circular (CD) nos espectros de ultravioleta próximo e distante, bem como ao experimento de estabilidade térmica da rRnOBP.

O experimento de dicróismo circular aproveita as variações na absorção da luz circularmente polarizada em direções opostas para sondar as características estruturais das proteínas. Para avaliar a aproximação entre a estrutura experimental e a teórica, foi utilizado o espectro na região distante do ultravioleta (190 – 260nm), a qual se mostra especialmente apta para investigar as transições eletrônicas envolvendo as amidas presentes nas ligações peptídicas da cadeia principal das proteínas.

Essas abordagens analíticas permitiram aprofundar nossa compreensão acerca da estrutura da rRnOBP, ao mesmo tempo que solidificaram a relação entre as previsões teóricas e os resultados experimentais.(106). Um experimento focalizado no espectro de ultravioleta (UV) próximo à luz visível (250 - 360 nm) foi então conduzido para fornecer insights acerca da estrutura da proteína (Figura 28). Nesse experimento, foi possível notar que nos comprimentos de onda associados à absorção de triptofano (290~295 nm), tirosina (280~285 nm) e fenilalanina (266~272 nm), emergiu um sinal característico de estruturação. Esse achado apontou para a ocorrência de um enterramento e, por conseguinte, para a provável conformação tridimensional da proteína. Esses resultados validaram a formação de uma estrutura tridimensional, alinhando-se com as expectativas para uma proteína globular.

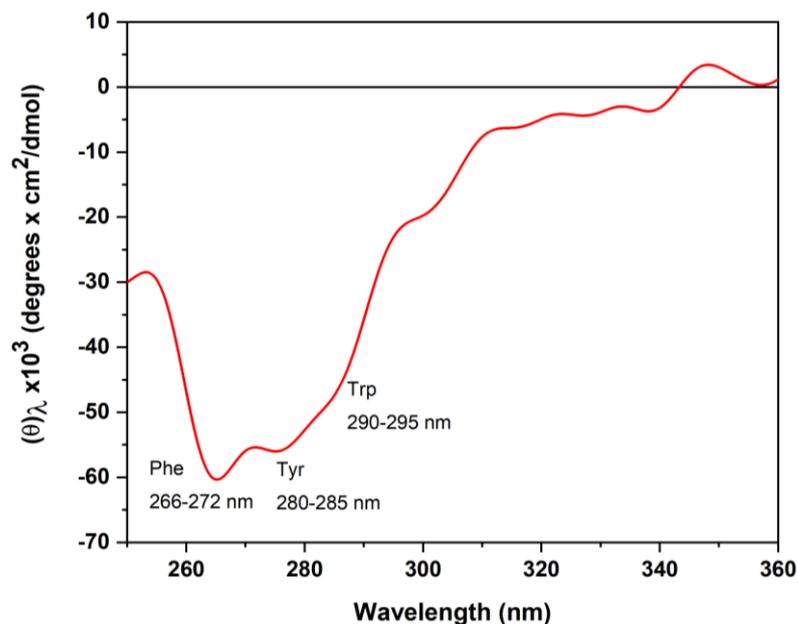


Figura 28. Dicroísmo circular no espectro UV próximo da luz visível. O eixo “Y” dá a elipticidade molar no comprimento de onda descrito no eixo “X”. 266-272, 280-285 e 290-295 mostram os comprimentos de ondas nos quais os aminoácidos da fenilalanina, tirosina e triptofano (respectivamente) absorvem em luz circularmente polarizada.

No ensaio utilizando os espectros UV longe da luz visível (190 – 260nm) (Figura 28) foi observada uma estrutura mista para rRnOBP, composta por α -hélices (17,5%), folhas β antiparalelas (38%), embora o esperado pelo modelo tridimensional gerado fosse apenas α -hélices (58.5%) (Tabela 8). A composição precisa do modelo calculada pelo programa de DSSP (103) e a composição da amostra pelo servidor BeStSel (104), os dados obtidos estão presentes na Tabela 8.

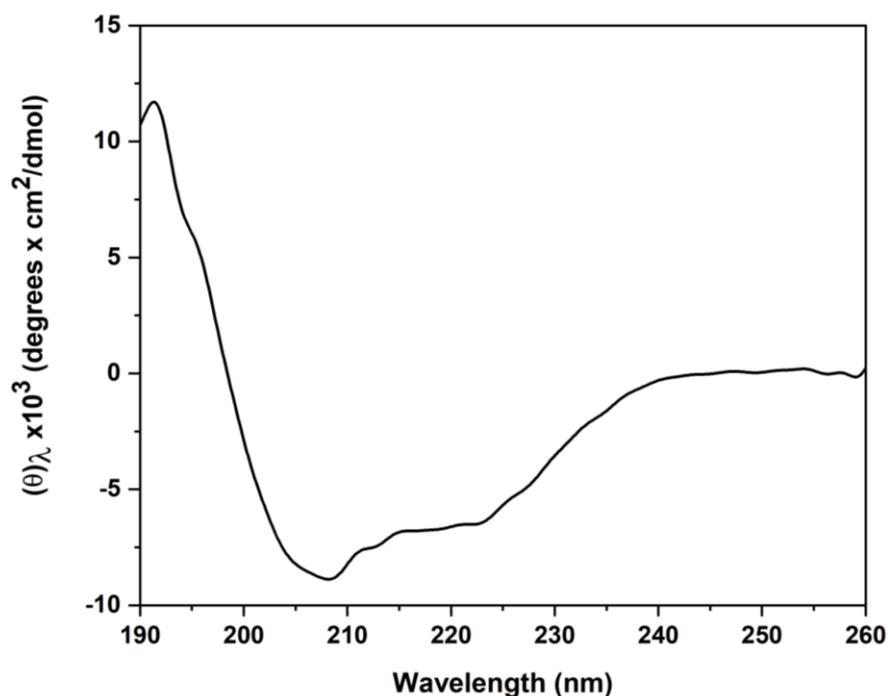


Figura 29. Dicroísmo circular no espectro UV longe da luz visível. O eixo “Y” dá a elipticidade molar no comprimento de onda descrito no eixo “X”.

Tabela 8. Estruturas secundárias experimental e teórica

Estrutura secundária	Modelo AlphaFold 2 (%)	rRnOBP (%)
Alfa hélice	58.5	17.5
3-hélice	0.0	0.0
Pi-hélice	0.0	0.0
Folha beta anti paralela	2.8	38.0
Resíduo isolado em uma beta-bridge	0.0	0.0
Cotovelo beta	7.7	16.3
Bend	3.5	0.0
Loop	27.5	28.1

Devido à composição diferente entre as estruturas secundárias do modelo e da amostra, um experimento de desnaturação térmica foi desenvolvido para verificar a estabilidade das estruturas secundárias. Para isso leituras UV distante foram feitas em diferentes temperaturas (25 °C a 95 °C) para verificar a perda de estruturas mediante ao aumento da temperatura. A Figura 29 mostra várias curvas realizadas em diferentes temperaturas e mostra que algumas estruturas são perdidas neste processo, após a deconvolução da função. A perda de estruturas é ilustrada mais detalhadamente na Tabela 9.

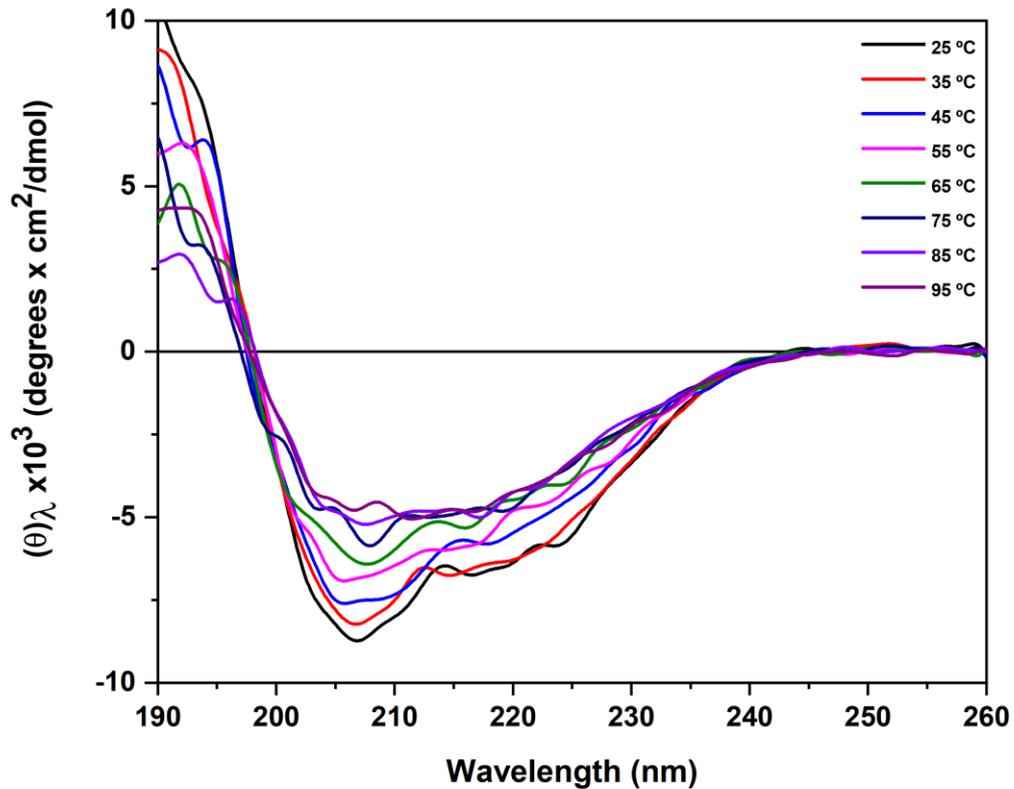


Figura 30. Perda de estruturas secundárias. Espectros de dicroísmo circular na faixa dos 190 a 260 nm da amostra em diferentes temperaturas, de 25 a 95 graus celcius. Curvas geradas a partir da leitura da amostra em concentração de aproximadamente 0.1 mg/ml em um espectropolaritro Jasco V-530, a temperatura foi ajustada para as leituras serem realizadas de 25 a 95 °C, utilizando um aparelho peltier.

Tabela 9. Composição de estruturas secundárias no experimento de estabilidade térmica.

T (°C)	Hélice 1	Hélice 2	Anti 1	Anti 2	Anti 3	Paralela	Cotovelos β	Outros
25	8,8%	6,4%	1,9%	11,5%	18,7%	0%	13,2%	39,5%
35	8,2%	4,3%	2,2%	12%	19,9%	0%	11,8%	41,5%
45	8,3%	7,8%	2,4%	13,1%	12,3%	0%	14,6%	41,4%
55	5%	5,8%	1,3%	12,2%	17,3%	0,9%	14,1%	43,4%
65	5,8%	6,2%	2,9%	11,5%	16,1%	0%	13,5%	43,9%
75	5,4%	3,2%	3,2%	16,5%	18,9%	0%	12,7%	40,1%
85	4,1%	5%	4,3%	13,1%	18,1%	0%	13,1%	42,2%
95	2,8%	4,5%	2%	14,9%	18%	2,1%	13,6%	42,1%

5.2.3 Atividade inibitória da agregação plaquetária e da coagulação da rRnOBP

A atividade inibitória da rRnOBP foi realizada pelo teste de agregação plaquetária induzida por ADP. A rRnOBP, em uma concentração de 8 μM , inibiu a agregação plaquetária induzida por ADP em 50% comparado com o controle na ausência da proteína recombinante (Figura 30). Esses resultados iniciais sugerem que a rRnOBP é capaz de inibir parcialmente a agregação plaquetária induzida por ADP.

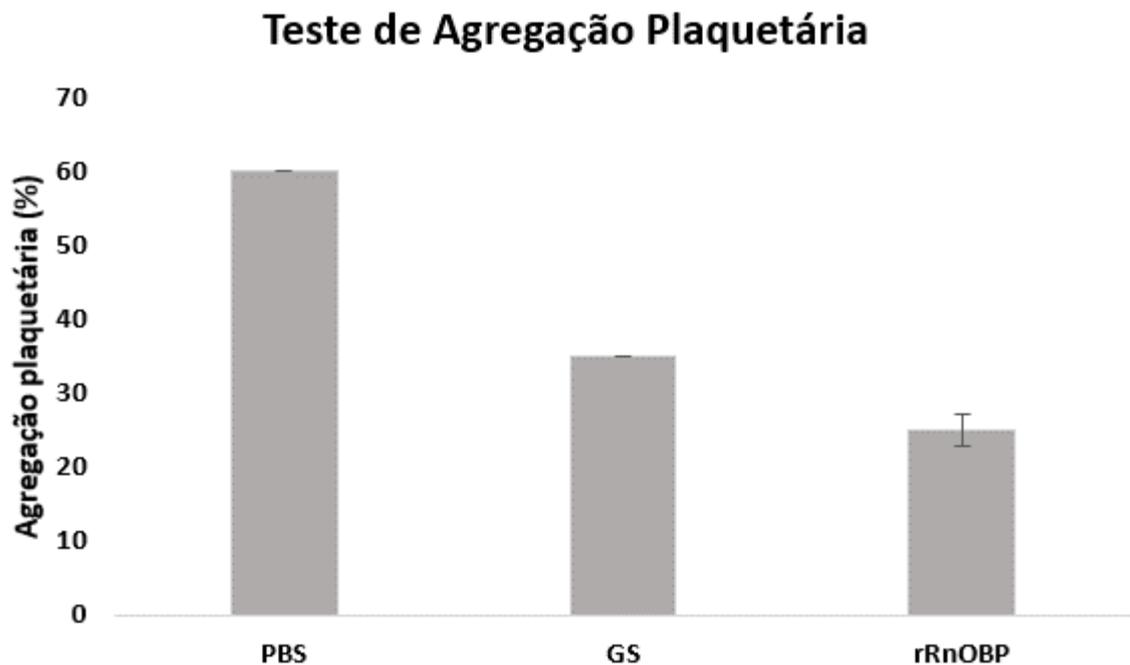


Figura 31. Avaliação da inibição da agregação plaquetária pela rRnOBP. PBS (A), extrato de 2 glândulas salivares de *R. neglectus* (B) e rRnOBP 8 μM (C e D) foram adicionados ao plasma rico em plaquetas (PRP) e incubados por 5 min a 37°C. Após esse período, ADP (concentração final 5 μM) foi adicionado e a curva de agregação plaquetária foi registrada por 5 min, a 600 nm. As reações foram realizadas usando o agregômetro EasyAgreg (Qualitem). Ensaios realizados em duplicata. A porcentagem de agregação está explicitada abaixo de cada gráfico.

Quanto à atividade inibitória da coagulação, a rRnOBP na concentração de 8 μM não aumentou o tempo de tromboplastina parcial ativada (TTPa). O TTPa do controle negativo (PBS) e na presença de rRnOBP foi 35,5 e 36,4 segundos, respectivamente, enquanto o TTPa na presença do extrato de glândulas salivares de *R. neglectus* foi 159,6 segundos (Figura 31).

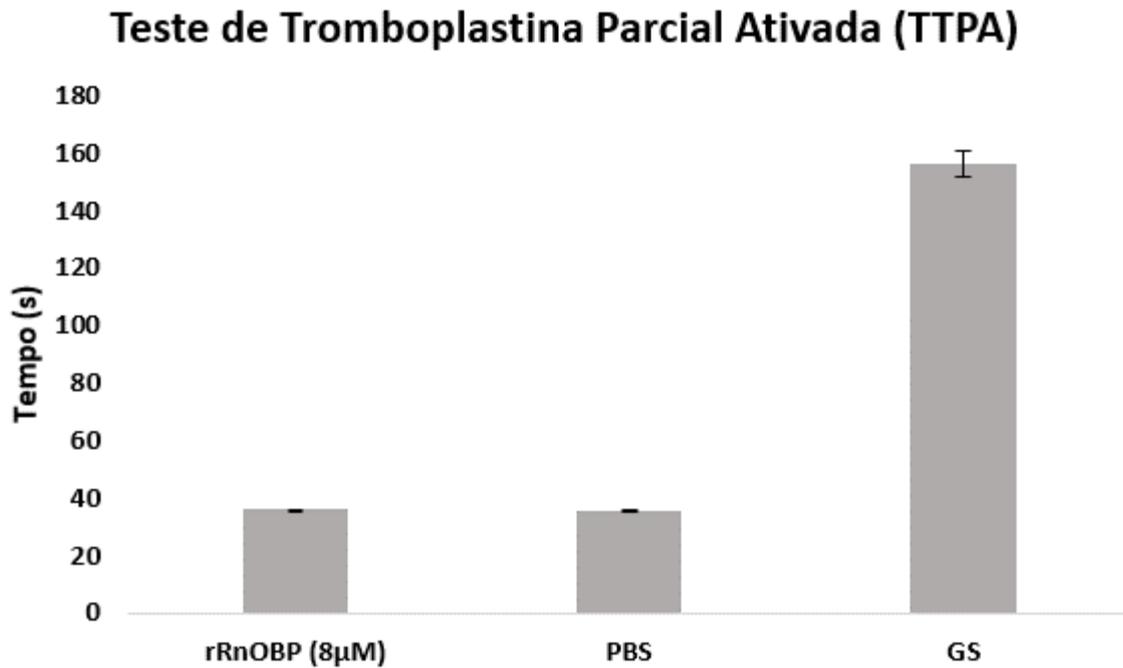


Figura 32 Avaliação de inibição da coagulação no teste de tromboplastina ativada. 100 µL de CaCl₂ foram pré-aquecidos no coagulômetro CLOTimer Laser sensor, à temperatura de 37 °C. Em seguida, 100 µL da amostra (PBS como controle negativo; GS conteúdo salivar de 2 glândulas de *R. neglectus* como controle positivo; rRnOBP proteína recombinante na concentração de 8 micromolar). O teste foi realizado em duplicata, com o desvio padrão ilustrado pelas barras de erro.

6 DISCUSSÃO

As proteínas salivares de insetos hematófagos apresentam grande potencial farmacológico principalmente pelas suas funções anti-hemostáticas, que os possibilitam se alimentar de sangue (107). Neste estudo, as propriedades da RnOBP, uma proteína salivar de triatomíneo, foram detalhadamente exploradas por meio da utilização de servidores e programas computacionais, além de ensaios experimentais. Para atingir esse objetivo, foram investigados potenciais modificações pós-traducionais que poderiam ocorrer, bem como foram examinadas propriedades intrínsecas da sequência primária da proteína, obtida de bancos de dados biológicos. A modelagem 3D da proteína foi realizada para obter *insights* sobre a possível conformação tridimensional da mesma. Além disso, as características de ligação e as propriedades dinâmicas da proteína foram analisadas por meio de programas de simulação. Realizou-se expressão, purificação e, por fim, caracterização da amostra de rRnOBP utilizada no estudo.

A utilização de abordagens computacionais para investigar moléculas salivares de animais hematófagos é vantajosa, pois oferece uma série de benefícios significativos no âmbito da pesquisa biológica. Por exemplo, eles permitem a análise de um grande número de moléculas e interações de maneira rápida e econômica, eliminando a necessidade de realizar experimentos em laboratório para cada molécula individualmente, assim como em estratégias de *design* de fármacos (108). Isso acelera o processo de identificação de moléculas promissoras para investigações mais detalhadas, pois permite a pré-seleção de moléculas relevantes com base em suas propriedades estruturais e características funcionais preditas (109). Dessa forma é possível direcionar os esforços de pesquisa experimental para moléculas com maior probabilidade de atividade funcional, economizando recursos e tempo. Isso é especialmente relevante ao lidar com a complexidade das moléculas salivares de animais hematófagos, onde as interações podem ser variadas e numerosas (18, 110, 111).

Além disso, a modelagem computacional oferece a capacidade de analisar diversos parâmetros, como estrutura tridimensional, propriedades físico-químicas, interações moleculares e possíveis sítios de ligação. Isso proporciona *insights* detalhados que podem orientar a pesquisa experimental subsequente (58). Ainda, a

abordagem *in silico* é útil para lidar com considerações éticas e custos significativos associados à pesquisa com animais hematófagos. Isso permite explorar alternativas sem a necessidade de experimentos invasivos em animais.

No trabalho de (112) foram realizados ensaios de *docking* para identificar potenciais inibidores de uma JHBP presente em mosquitos *Aedes spp.* Essa abordagem busca impedir o desenvolvimento completo dos insetos vetores, uma estratégia promissora para o controle dessas pragas.

Além disso, em (113), foram testados compostos derivados de óleos essenciais que atendem aos requisitos de repelência em uma OBP que influencia o comportamento do *Aedes aegypti*. Essa abordagem utiliza o *docking* molecular para avaliar a interação entre os compostos e a OBP.

No campo da busca por repelentes contra insetos hematófagos transmissores de doenças, as pesquisas sobre OBPs têm produzido resultados promissores, potencialmente gerando moléculas atraentes para o desenvolvimento de produtos comerciais (114).

No contexto específico deste estudo, modelo 3D gerado da RnOBP revelou algumas características cruciais para a estrutura do sítio de ligação da RnOBP. A região adjacente à fenda que acomoda os ligantes exibe uma predominância de cargas negativas (Figura 7), representadas pelos aminoácidos ácidos Q1, E5 e E71. Esses aminoácidos desempenham um papel estabilizador, interagindo com grupos polares de alguns ligantes testados no experimento de *docking*. Essa observação sugere a importância estrutural dessa região para a função efetiva da RnOBP. Além disso, a presença de R111, com características básicas, pode conferir uma variação nas cargas do sítio de ligação da RnOBP.

No ensaio de *ensemble docking*, utilizando as diversas conformações observadas durante a dinâmica molecular, os resultados apresentaram pontuações aproximadamente 0.6 kcal/mol maiores do que aquelas obtidas no *redocking* das estruturas cristalográficas das aminas biogênicas. No entanto, as pontuações para os eicosanoides foram 2 kcal/mol superiores, enquanto chegaram a -8 kcal/mol tanto para os leucotrienos quanto para o análogo de tromboxano U46619. As proteínas D7 curtas, que possuem apenas um domínio OBP e tem preferência por se ligarem a aminas biogênicas (25). Isso pode ser uma ocorrência relevante para a RnOBP, uma vez que está relacionado ao volume do sítio de interação com o ligante da RnOBP. A melhoria nas interações calculadas também pode ser uma perspectiva viável por meio

de futuros ensaios de dinâmica molecular da proteína com ligantes. Esses ensaios permitirão ajustar as angulações entre as cadeias laterais e os ligantes, possivelmente corrigindo a ausência do fenômeno de ajuste induzido, que não é observado em *dockings* de estrutura rígida.

Os resíduos importantes para a estabilização das aminas biogênicas incluem E109, presente em todas as interações e formando pontes de hidrogênio, F108 e Y105, que estabilizam aminas biogênicas por meio de interações hidrofóbicas (Tabela 6). A análise do docking molecular da RnOBP com a triptamina revelou uma interação hidrofílica envolvendo o nitrogênio na extremidade da porção alifática do ligante com o resíduo E109. Esta interação difere da estrutura 2PQL e a estabilização dessa cadeia ocorre por meio da interação com uma molécula de água estrutural e o resíduo D111. Quando comparamos a interação da RnOBP com a histamina com a estrutura 2QEB, notamos que, enquanto o resíduo Y24 interage com a região cíclica da histamina em 2QEB, no docking da RnOBP, o resíduo E71 estabiliza essa região, e a região acíclica é estabilizada pelo resíduo E109, mantendo a mesma natureza das ligações encontradas na estrutura original, que é estabilizada pelo resíduo E114. É interessante notar que ambos os contatos hidrofóbicos são mantidos no ligante, mas na RnOBP, o resíduo W67 substitui o Y24 na estabilização da histamina. Na interação com a serotonina, notamos a presença de duas moléculas de água estruturais importantes na estrutura 2QEH, o que não foi utilizado no estudo de docking com os modelos da RnOBP. No entanto, a natureza das interações de resíduos ácidos estabilizantes permanece a mesma, assim como a interação pi-stacking. No entanto, o número de contatos é menor na RnOBP (52).

Nas interações com os eicosanoides Leucotrieno C4 e o análogo de Tromboxano A2 (U46619), observamos muitas interações semelhantes às encontradas em suas respectivas estruturas cristalográficas (3NHI e 3NHT). No entanto, esses ligantes ficaram expostos ao solvente em regiões que não se encaixaram no sítio de ligação da RnOBP. Isso requer investigações adicionais para determinar a viabilidade dessas interações (43).

Os resultados do *ensemble docking* indicam um potencial para a RnOBP atuar como inibidora da hemostasia, interagindo com os ligantes mencionados anteriormente. No entanto, é crucial conduzir experimentos adicionais para validar essa interação em um contexto experimental.

Em proteínas D7 curtas de anofelinos o que se é observado é uma preferência por ligantes pequenos como aminas biogênicas, mas, interessantemente foi observado que a RnOBP obteve pontuações boas nos ensaios de *docking* com todos os tipos de moléculas testadas, e para verificar este resultado estudos futuros devem ser realizados para expressar e testar a afinidade real da proteína com tais ligantes.

Quanto às modificações pós-traducionais, a sequência de aminoácidos da RnOBP apresenta dois sítios putativos de *N*-glicosilação na (N55, N80), dos quais a N80 possuiu uma pontuação alta (9/9) segundo o servidor NetNGlyc (81). As modificações pós-traducionais realizadas nas proteínas secretadas durante o seu processo de maturação no retículo endoplasmático de células eucarióticas não são realizadas nas proteínas recombinantes expressas em células procarióticas como *E. coli*. Estas modificações possuem caráter importante na função e no envelhecimento de proteínas (115).

Ainda com relação às modificações pós-traducionais, a formação das pontes dissulfeto entre as cisteínas conservadas das OBPs (33) são importantes, pois estabilizam as OBPs de insetos (33). O processo de envelhecimento de proteínas que possuem pontes dissulfeto pode ser difícil, levando em consideração que o processo de controle sobre as ligações das cisteínas não existe e ocorre de uma maneira não natural em células procarióticas em um sistema induzível (116). A maior concentração das proteínas heterólogas dentro das células bacterianas favorece interações de cisteínas de cadeias diferentes o que pode gerar artefatos, como a ligação entre cadeias diferentes, gerando agregados que geralmente se acumulam nos corpos de inclusão (116). Para evitar este tipo de problema, existem metodologias de envelhecimento que utilizam reagentes redutores para quebrar as pontes dissulfeto e diluição da amostra para permitir que o processo de envelhecimento ocorra de acordo com o Dóγμα de Anfinsen, com as suas interações intramoleculares dirigindo o processo de envelhecimento para um mínimo de energia de Gibbs que leva à conformação nativa (117, 118) uma metodologia de reenvelhecimento que obteve uma boa recuperação da estrutura nativa de proteínas D7 pode ser observada nos trabalhos de Ines Martin-Martin e colaboradores(50, 119)

O processo de envelhecimento realizado neste estudo resultou em conformações oligoméricas que, diferentemente do que ocorre em PBPs que formam complexos como a OBP48 de *Anopheles gambiae* que se engaja de com interações polares e apolares (120) a rRnOBP formou homômeros com pontes dissulfeto (Figura

9). O fenômeno de *domain-swapping* (121) mostrado no modelo homomérico (Figura 9), apesar de fazer sentido com os experimentos de exclusão molecular em condição desnaturante (Figura 26), não foi observado em nenhum outro estudo relacionado com OBPs. Isso reforça a hipótese de que provavelmente o que ocorreu com a amostra de rRnOBP foi um artefato, o que é reforçado pelo padrão retirado dos experimentos de dicroísmo circular (Figura 28) que mostram estruturas secundárias que não são observadas nos modelos do monômero, dímeros, trímeros e tetrâmeros (totalmente constituídos por α -hélices) (Tabela 8). Porém é importante citar que embora a amostra tenha apresentado uma grande quantidade de estruturas secundárias não esperadas, a proteína estava solúvel e estruturada (Figura 28), e apresentou função anti agregadora plaquetária (Figura 30).

Este artefato pode ser explicado pela metodologia de reenovelamento utilizada após o processo de purificação da rRnOBP. No estudo de Alvarenga e colaboradores, uma proteína D7 longa de *Anopheles stephensi* foi produzida de forma semelhante, utilizando a linhagem de bactérias BL21(DE3) pLysS (43). Contudo, diferentemente deste estudo, no processo de reenovelamento da proteína, foi incorporada uma etapa com agentes redutores para desfazer as pontes dissulfeto, seguida por uma etapa de diluição para promover o enovelamento das cadeias em forma de monômeros. Após esse procedimento de reenovelamento, a proteína foi concentrada e alcançou um estado de alta qualidade, caracterizado pela sua pureza e conformação nativa (43). A rRnOBP, por não ter sido submetida a esse procedimento com agentes redutores, pode ter mantido conformações artificiais devido a metodologia que forçou um enovelamento em ambiente altamente concentrado (116). A massa molecular do monômero da rRnOBP foi 16,706 kDa. No entanto, é importante destacar que essa medida pode sofrer alterações decorrentes de modificações na estrutura tridimensional da proteína, que podem afetar a resistência à sedimentação (75).

Os transcritos de OBP em triatomíneos são raros. Em *R. neglectus*, as OBPs representam 0,06% das reads obtidas com 11 CDS diferentes (44). No transcriptoma de *Triatoma sordida*, as sequências de OBPs, PBPs e JHBPs possuem um FPKM (*Fragments Per Kilobase Million*) menor do que 0,1% (46). A baixa abundância de transcritos nas glândulas salivares de triatomíneos representa que o número de proteínas não deve ser tão alto quanto o observado em OBPs salivares de outras espécies, como em *Aedes aegypti*, que em glândulas salivares de fêmeas em período

de alimentação em sangue as proteínas D7 estão entre os transcritos mais abundantes, avaliados por PCR quantitativa (122).

A baixa expressão das OBPs na saliva de triatomíneos leva a algumas questões sobre seu papel na hematofagia: *(i)* a OBP consegue atuar com efetividade em baixa concentração; *(ii)* a OBP atua de maneira sinérgica com outras proteínas salivares. A hipótese formulada sugere que a forte afinidade da RnOBP com as aminas biogênicas pode indicar uma função significativa para essa proteína na inibição da comunicação celular induzida por este tipo molecular. Essa descoberta abre novas perspectivas para futuros estudos de caracterização experimental da RnOBP. Para aprofundar a compreensão das características estruturais, é essencial reavaliar certos procedimentos na etapa de renovação da proteína, bem como no processo de purificação das conformações oligoméricas. Esse ajuste se torna ainda mais relevante caso os resultados da oligomerização obtidos *in silico* sejam confirmados pelas conformações previstas.

7 CONCLUSÕES

Neste trabalho, a estrutura tridimensional da RnOBP foi modelada e refinada com sucesso, algo que não havia sido descrito anteriormente na literatura. O sítio de ligação a ligantes predito computacionalmente foi semelhante ao sítio de proteínas homólogas.

As simulações de dinâmica molecular geraram conformações estáveis e representativas da RnOBP. Essas conformações representativas foram utilizadas nos cálculos de ensemble *docking* contra ligantes envolvidos na hemostasia. Destacaram-se as aminas biogênicas, que apresentaram pontuações de docking satisfatórias e próximas às do redocking.

A rRnOBP expressa e purificada a partir de corpos de inclusão revelou uma composição de estruturas secundárias diferente da esperada. Os testes funcionais iniciais para esta recombinante indicaram que ela pode atuar na inibição da agregação plaquetária, mas não na inibição da coagulação.

8 PERSPECTIVAS

1. Conduzir dinâmicas de interação entre a proteína e os ligantes avaliados no estudo de *docking molecular*;
2. Aprimorar o protocolo de renaturação da rRnOBP, incluindo agentes redutores;
3. Realizar a produção da proteína rRnOBP em sistema de expressão eucarioto;
4. Realizar experimentos de cristalização e repetir as análises de caracterização funcional;
5. Produzir anticorpos específicos contra a rRnOBP.

APÊNDICE

Apêndice A – Tabela de clusters para cada dinâmica.

Tabela 10. Clusterização dinâmica molecular 1. Mostrando o RMSD em angstroms entre as estruturas representativas de cada cluster realizado na primeira dinâmica molecular

DM1	C1	C2	C3	C4	C5	C6
C1	0	2,1	3,18	2,64	3,24	2,42
C2	2,1	0	2,85	2,49	2,84	2,56
C3	3,18	2,85	0	2,31	2,37	2,76
C4	2,64	2,49	2,31	0	2,21	1,93
C5	3,24	2,84	2,37	2,21	0	2,66
C6	2,42	2,56	2,76	1,93	2,66	0

Tabela 11. Clusterização dinâmica molecular 2 Mostrando o RMSD em angstroms entre as estruturas representativas de cada cluster realizado na segunda dinâmica molecular

DM2	C1	C2	C3	C4	C5
C1	0	2,1	3,18	2,94	3,24
C2	2,1	0	2,85	2,78	2,84
C3	3,18	2,85	0		2,37
C4	2,94	2,78	2,24	0	2,25
C5	3,24	2,84	2,37	2,25	0

Tabela 12. Clusterização dinâmica molecular 3 Mostrando o RMSD em angstroms entre as estruturas representativas de cada cluster realizado na terceira dinâmica molecular

DM3	C1	C2	C3	C4
C1	0	2,42	1,82	1,55
C2	2,42	0	1,99	1,89
C3	1,82	1,99	0	1,67
C4	1,55	1,89	1,67	0

Apendice B – Predição de interações de sequencias de rRnOBP.

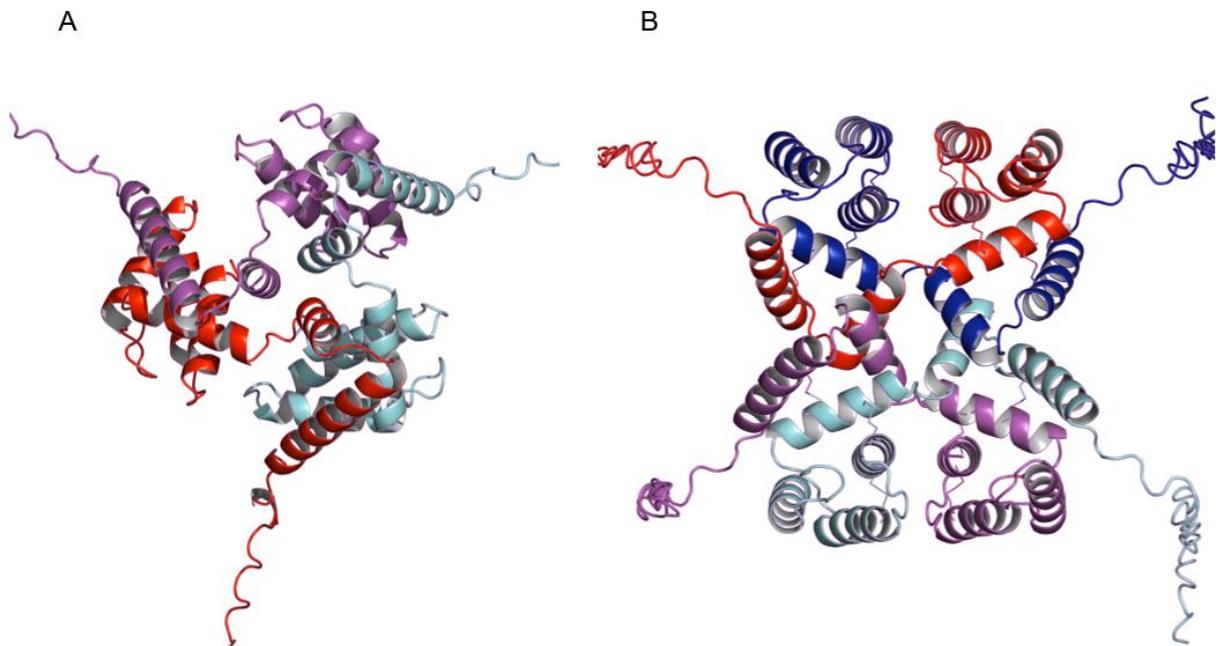


Figura 33 Modelagem da interação trimérica e tetramérica da rRnOBP no servidor AlphaMultimer. Foi utilizado o servidor disponibilizado na página do colabfold (2) para realizar a predição de oligomerização de sequencias de rRnOBP. A) modelo de trímero da rRnOBP. B) modelo de tetrâmero da rRnOBP.

Apêndice C - Código utilizado para identificar aminoácidos pertencentes ao bolso de ligação de eicosanoides e amins biogênicas:

```

lig1 = ['GLU109', 'TYR105', 'TRP8', 'PHE108', 'TRP67', 'GLN1', 'GLU5', 'GLU71', 'GLU12', 'ARG26', 'ASN4', 'ARG111', 'GLY9', 'ILE36', 'LEU30', 'SER32', 'THR31', 'TYR102']
lig2 = ['GLU109', 'ILE29', 'LEU30', 'TYR105', 'ARG111', 'GLN1', 'GLU5', 'GLU71', 'TRP67', 'TRP8', 'ASN4', 'PHE108', 'ILE36', 'TYR102', 'GLU2']
lig3 = ['GLN1', 'GLU109', 'GLU5', 'GLU71', 'TRP67', 'ASN4', 'PHE108', 'TRP8', 'TYR105', 'ARG26', 'GLU12', 'GLY9', 'ARG111', 'LEU30']
lig4 = ['ARG26', 'GLU109', 'GLU12', 'GLU5', 'GLY9', 'LEU30', 'LYS13', 'TRP67', 'TRP8', 'TYR105', 'GLU6', 'GLU71', 'LYS10', 'PHE108', 'ARG111', 'ASN4', 'GLN1', 'ILE29', 'TYR43', 'VAL44', 'GLN16', 'SER23']
lig5 = ['ARG26', 'GLU12', 'GLU5', 'GLY9', 'LEU30', 'LYS13', 'PHE108', 'TRP67', 'TRP8', 'TYR105', 'GLN16', 'GLU6', 'LYS10', 'GLU109', 'ARG111', 'ASN4', 'GLU71']
lig6 = ['ARG26', 'GLU12', 'GLY9', 'LYS13', 'TRP8', 'GLN16', 'SER23', 'GLU2', 'GLU5', 'GLU6', 'LYS10']

aminasB = []

eicosanoides = []

for i in lig1:
    aminasB.append(i)
for i2 in lig2:
    if not i2 in aminasB:
        aminasB.append(i2)
for i3 in lig3:
    if not i3 in aminasB:
        aminasB.append(i3)
print(f'aminoscidos que se ligam a amins biogenicas {aminasB}')

for i4 in lig4:
    eicosanoides.append(i4)
for i5 in lig5:
    if not i5 in eicosanoides:
        eicosanoides.append(i5)

print(f'aminoscidos que se ligam a eicosanoides {eicosanoides}')

fixedAAinAB = []
fixedAAinEI = []

for i7 in lig3:
    if i7 in lig2 and lig1:
        fixedAAinAB.append(i7)

for i8 in lig5:
    if i8 in lig4:
        fixedAAinEI.append(i8)

print(f'aminoscidos presentes em todas interações com amins B {fixedAAinAB}')
print(f'aminoscidos presentes em todas interações com eicosanoides {fixedAAinEI}')

```

Apendice D - Código utilizado para verificar os resíduos a uma distância de 4 angstroms dos ligantes:

```

import math

def lerArquivo(nomeArquivo):
    with open(nomeArquivo, 'r') as arquivo:
        linhas = arquivo.readlines()
    return linhas

def imprimeArquivo(linhas):
    for l in linhas:
        l = l.rstrip('\n')
    return l

def obtemCadeia(cadeia, linhas):
    atoms = {}
    for l in linhas:
        l = l.rstrip('\n')
        if l[0:4] == 'ATOM' and l[21] == cadeia:
            #COLUMNS      DATA  TYPE   FIELD      DEFINITION
            #-----
            # 7 - 11      Integer   serial    Atom serial number.
            serial = int(l[6:11])
            #18 - 20      Residue name  resName   Residue name.
            resName = l[17:20]
            #23 - 26      Integer   resSeq    Residue sequence number.
            resSeq = int(l[22:26])
            #31 - 38      Real(8.3)   x         Orthogonal coordinates for X in Angstroms.
            x = float(l[30:38])
            #39 - 46      Real(8.3)   y         Orthogonal coordinates for Y in Angstroms.
            y = float(l[38:46])
            #47 - 54      Real(8.3)   z         Orthogonal coordinates for Z in Angstroms.
            z = float(l[46:54])

            atoms[serial] = (x, y, z, resName, resSeq)
    return atoms

def obtemMolecula(id, linhas):
    atoms = {}
    for l in linhas:
        l = l.rstrip('\n')
        if l[0:6] == 'HETATM' and l[17:20] == id:
            serial = int(l[6:11])
            name = l[12:16].strip()
            x = float(l[30:38])
            y = float(l[38:46])
            z = float(l[46:54])

            atoms[serial] = (x, y, z, name)
    return atoms

def dist(x1, y1, z1, x2, y2, z2):
    d = math.sqrt((x1-x2)**2+(y1-y2)**2+(z1-z2)**2)
    return d

def obtemSitioligacao(molecula, cadeia, limiar = 4):
    x=0
    y=1
    z=2
    resName = 3
    resSeq = 4
    name = 3
    sitio = []
    for m in molecula:
        for c in cadeia:
            d = dist(molecula[m][x], molecula[m][y], molecula[m][z], cadeia[c][x], cadeia[c][y], cadeia[c][z])
            if d <= limiar:
                #print(d, cadeia[c][resName], cadeia[c][resSeq], molecula[m][name])
                if not cadeia[c][resName] + str(cadeia[c][resSeq]) in sitio:
                    sitio.append(cadeia[c][resName]+str(cadeia[c][resSeq]))
    return sorted(sitio)

def gravarBolsa(arquivo, estrutura, aminoacidos):
    with open(arquivo, 'a') as file:
        file.write(f'{estrutura}-{aminoacidos}\n')

for i in range(5):
    linhas = lerArquivo(f'DockingsClusterizacao\DockingsDM2\C{i+1}D2-2qeb.pdb')
    imprimeArquivo(linhas)
    obtemCadeia("A", linhas)
    cadeia = obtemCadeia('A', linhas)
    molecula = obtemMolecula("LIG", linhas)
    print(obtemSitioligacao(molecula, cadeia, 4))

gravarBolsa('DockingsClusterizacao\D2_2qeb.txt', f'C{i+1}D2_2qeb', obtemSitioligacao(molecula, cadeia, 4))

```

9 REFERÊNCIAS BIBLIOGRÁFICAS

1. Praça YR. Sialoma do *Triatoma sordida*, triatomíneo peridomiciliar vetor da Doença de Chagas: Universidade de Brasília; 2022.
2. Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. ColabFold: making protein folding accessible to all. *Nature Methods*. 2022;19(6):679-82.
3. Greenfield NJ. Using circular dichroism spectra to estimate protein secondary structure. *Nat Protoc*. 2006;1(6):2876-90.
4. Chagas C. Nova tripanozomíase humana: estudos sobre a morfologia e o ciclo evolutivo do *Schizotrypanum cruzi* n. gen., n. sp., agente etiológico de nova entidade morbida do homem. *Memórias do Instituto Oswaldo Cruz*. 1909;1:159-218.
5. Rassi A, Marin-Neto JA. Chagas disease. *Lancet*. 2010;375(9723):1388-402.
6. Nunes MC, Dones W, Morillo CA, Encina JJ, Ribeiro AL, Cardiology CoCDotISo. Chagas disease: an overview of clinical and epidemiological aspects. *J Am Coll Cardiol*. 2013;62(9):767-76.
7. Bello Corassa R, Aceijas C, Alves PAB, Garelick H. Evolution of Chagas' disease in Brazil. Epidemiological perspective and challenges for the future: a critical review. *Perspect Public Health*. 2017;137(5):289-95.
8. Gurgel-Gonçalves R, Galvão C, Costa J, Peterson AT. Geographic distribution of chagas disease vectors in Brazil based on ecological niche modeling. *J Trop Med*. 2012;2012:705326.
9. WHO. Chagas disease (American trypanosomiasis). 2020.
10. Rivas N, Antonio-Campos A, Noguez-García J, Alejandro-Aguilar R. First record of *Triatoma longipennis*, Usinger, 1939 (Hemiptera: Reduviidae: Triatominae) in Tecozautla, Hidalgo. *Rev Soc Bras Med Trop*. 2023;56:e00782023.
11. Falcone R, Ribeiro AR, Oliveira J, Mendonça VJ, Graminha M, Rosa JAD. Differentiation of *Rhodnius neglectus* and *Rhodnius prolixus* (Hemiptera: Reduviidae: Triatominae) by multiple parameters. *Rev Soc Bras Med Trop*. 2020;53:e20190503.
12. Cruz-López L, Malo EA, Rojas JC, Morgan ED. Chemical ecology of triatomine bugs: vectors of Chagas disease. *Medical and Veterinary Entomology*. 2001;15(4):351-7.
13. Lent HW, Pedro W. Revision of the Triatominae (Hemiptera, Reduviidae), and their significance as vectors of Chagas' disease. In: Wygodzinsky PW, editor. *American Museum of Natural History* 1979. p. 125-520.
14. Silva RAD, Maldonado LA, Pessoa GCD, Diotaiuti L. Monitoring *Rhodnius neglectus* (Lent, 1954) populations' susceptibility to insecticide used in controlling actions in urban areas northwest of São Paulo state. *Rev Soc Bras Med Trop*. 2022;55:e0553.
15. de Oliveira J, Ravazi A, Souza EDS, Moreira FFF, Galvão C, da Rosa JA, et al. Study of the Salivary Glands in Triatominae (Hemiptera, Reduviidae, Triatominae): Their Color and Application to the Chagas Disease Vector Evolution. *Am J Trop Med Hyg*. 2017;97(3):771-3.
16. Silva MBA, Barbosa HS, Galvão C, Jurberg J, Carcavallo RU. Comparative study of the stridulatory sulcus, buccula and rostrum of the nymphs of *Triatoma guazu* Lent & Wygodzinsky, 1979 and *Triatoma jurbergi* Carcavallo, Galvão & Lent, 1998 by scanning electron microscopy (Hemiptera, Reduviidae). *Memórias do Instituto Oswaldo Cruz*. 2003;98.

17. LAVOPIERRE MM, DICKERSON G, GORDON RM. Studies on the methods of feeding of blood-sucking arthropods. I. The manner in which triatomine bugs obtain their blood-meal, as observed in the tissues of the living rodent, with some remarks on the effects of the bite on human volunteers. *Ann Trop Med Parasitol*. 1959;53:235-50.
18. Santiago PB, de Araújo CN, Charneau S, Praça YR, Bastos IMD, Ribeiro JMC, et al. The Pharmacopea within Triatomine Salivary Glands. *Trends Parasitol*. 2020;36(3):250-65.
19. Andersen JF, Gudderra NP, Francischetti IM, Ribeiro JM. The role of salivary lipocalins in blood feeding by *Rhodnius prolixus*. *Arch Insect Biochem Physiol*. 2005;58(2):97-105.
20. Martin-Martin I, Chagas AC, Guimaraes-Costa AB, Amo L, Oliveira F, Moore IN, et al. Immunity to LuloHya and Lundep, the salivary spreading factors from *Lutzomyia longipalpis*, protects against *Leishmania major* infection. *PLoS Pathog*. 2018;14(5):e1007006.
21. Pala ZR, Alves E Silva TL, Minai M, Crews B, Patino-Martinez E, Carmona-Rivera C, et al. salivary apyrase regulates blood meal hemostasis and drives malaria parasite transmission. *bioRxiv*. 2023.
22. Tlili A, Marzouki S, Chabaane E, Abdeladhim M, Kammoun-Rebai W, Sakkouhi R, et al. Phlebotomus papatasi Yellow-Related and Apyrase Salivary Proteins Are Candidates for Vaccination against Human Cutaneous Leishmaniasis. *J Invest Dermatol*. 2018;138(3):598-606.
23. Nowak G. Pharmacology of recombinant hirudin. *Semin Thromb Hemost*. 2002;28(5):415-24.
24. Lerner EA, Ribeiro JM, Nelson RJ, Lerner MR. Isolation of maxadilan, a potent vasodilatory peptide from the salivary glands of the sand fly *Lutzomyia longipalpis*. *J Biol Chem*. 1991;266(17):11234-6.
25. Alvarenga PH, Andersen JF. An Overview of D7 Protein Structure and Physiological Roles in Blood-Feeding Nematocera. *Biology (Basel)*. 2022;12(1).
26. Tegoni M, Pelosi P, Vincent F, Spinelli S, Campanacci V, Grolli S, et al. Mammalian odorant binding proteins. *Biochim Biophys Acta*. 2000;1482(1-2):229-40.
27. Vogt RG, Riddiford LM. Pheromone binding and inactivation by moth antennae. *Nature*. 1981;293(5828):161-3.
28. Han WK, Yang YL, Si YX, Wei ZQ, Liu SR, Liu XL, et al. Involvement of GOBP2 in the perception of a sex pheromone component in both larval and adult *Spodoptera litura* revealed using CRISPR/Cas9 mutagenesis. *Insect Biochem Mol Biol*. 2022;141:103719.
29. Kim IH, Pham V, Jablonka W, Goodman WG, Ribeiro JMC, Andersen JF. A mosquito hemolymph odorant-binding protein family member specifically binds juvenile hormone. *J Biol Chem*. 2017;292(37):15329-39.
30. Jeong YT, Shim J, Oh SR, Yoon HI, Kim CH, Moon SJ, et al. An odorant-binding protein required for suppression of sweet taste by bitter chemicals. *Neuron*. 2013;79(4):725-37.
31. Ozaki M, Morisaki K, Idei W, Ozaki K, Tokunaga F. A putative lipophilic stimulant carrier protein commonly found in the taste and olfactory systems. A unique member of the pheromone-binding protein superfamily. *Eur J Biochem*. 1995;230(1):298-308.
32. Costa-da-Silva AL, Kojin BB, Marinotti O, James AA, Capurro ML. Expression and accumulation of the two-domain odorant-binding protein AaegOBP45 in the ovaries of blood-fed *Aedes aegypti*. *Parasit Vectors*. 2013;6:364.
33. Rihani K, Ferveur JF, Briand L. The 40-Year Mystery of Insect Odorant-Binding Proteins. *Biomolecules*. 2021;11(4).

34. Traverso L, Latorre Estivalis JM, da Rocha Fernandes G, Fronza G, Lobbia P, Mougabure Cueto G, et al. Transcriptomic modulation in response to an intoxication with deltamethrin in a population of *Triatoma infestans* with low resistance to pyrethroids. *PLoS Negl Trop Dis*. 2022;16(6):e0010060.
35. Zafar Z, Fatima S, Bhatti MF, Shah FA, Saud Z, Butt TM. Odorant Binding Proteins (OBPs) and Odorant Receptors (ORs) of *Anopheles stephensi*: Identification and comparative insights. *PLoS One*. 2022;17(3):e0265896.
36. Zhou JJ, He XL, Pickett JA, Field LM. Identification of odorant-binding proteins of the yellow fever mosquito *Aedes aegypti*: genome annotation and comparative analyses. *Insect Mol Biol*. 2008;17(2):147-63.
37. Hekmat-Safe DS, Safe CR, McKinney AJ, Tanouye MA. Genome-wide analysis of the odorant-binding protein gene family in *Drosophila melanogaster*. *Genome Res*. 2002;12(9):1357-69.
38. Wang R, Li F, Zhang W, Zhang X, Qu C, Tetreau G, et al. Identification and expression profile analysis of odorant binding protein and chemosensory protein genes in *Bemisia tabaci* MED by head transcriptome. *PLoS One*. 2017;12(2):e0171739.
39. Gonzalez D, Rihani K, Neiers F, Poirier N, Fraichard S, Gotthard G, et al. The *Drosophila* odorant-binding protein 28a is involved in the detection of the floral odour β -ionone. *Cell Mol Life Sci*. 2020;77(13):2565-77.
40. Calvo E, Mans BJ, Andersen JF, Ribeiro JM. Function and evolution of a mosquito salivary protein family. *J Biol Chem*. 2006;281(4):1935-42.
41. Isawa H, Yuda M, Orito Y, Chinzei Y. A mosquito salivary protein inhibits activation of the plasma contact system by binding to factor XII and high molecular weight kininogen. *J Biol Chem*. 2002;277(31):27651-8.
42. Alvarenga PH, Dias DR, Xu X, Francischetti IMB, Gittis AG, Arp G, et al. Functional aspects of evolution in a cluster of salivary protein genes from mosquitoes. *Insect Biochem Mol Biol*. 2022;146:103785.
43. Alvarenga PH, Francischetti IM, Calvo E, Sá-Nunes A, Ribeiro JM, Andersen JF. The function and three-dimensional structure of a thromboxane A₂/cysteinyl leukotriene-binding protein from the saliva of a mosquito vector of the malaria parasite. *PLoS Biol*. 2010;8(11):e1000547.
44. Santiago PB, Assumpção TC, de Araújo CN, Bastos IM, Neves D, da Silva IG, et al. A Deep Insight into the Sialome of *Rhodnius neglectus*, a Vector of Chagas Disease. *PLoS Negl Trop Dis*. 2016;10(4):e0004581.
45. Santiago PB, de Araújo CN, Charneau S, Bastos IMD, Assumpção TCF, Queiroz RML, et al. Exploring the molecular complexity of *Triatoma dimidiata* sialome. *J Proteomics*. 2018;174:47-60.
46. Praça YR, Santiago PB, Charneau S, Mandacaru SC, Bastos IMD, Bentes KLDS, et al. An Integrative Sialomic Analysis Reveals Molecules From. *Front Cell Infect Microbiol*. 2021;11:798924.
47. Nevoa JC, Mendes MT, da Silva MV, Soares SC, Oliveira CJF, Ribeiro JMC. An insight into the salivary gland and fat body transcriptome of *Panstrongylus lignarius* (Hemiptera: Heteroptera), the main vector of Chagas disease in Peru. *PLoS Negl Trop Dis*. 2018;12(2):e0006243.
48. Ribeiro JM, Andersen J, Silva-Neto MA, Pham VM, Garfield MK, Valenzuela JG. Exploring the sialome of the blood-sucking bug *Rhodnius prolixus*. *Insect Biochem Mol Biol*. 2004;34(1):61-79.
49. Assumpção TC, Francischetti IM, Andersen JF, Schwarz A, Santana JM, Ribeiro JM. An insight into the sialome of the blood-sucking bug *Triatoma infestans*, a vector of Chagas' disease. *Insect Biochem Mol Biol*. 2008;38(2):213-32.

50. Martin-Martin I, Paige A, Valenzuela Leon PC, Gittis AG, Kern O, Bonilla B, et al. ADP binding by the *Culex quinquefasciatus* mosquito D7 salivary protein enhances blood feeding on mammals. *Nat Commun.* 2020;11(1):2911.
51. Jablonka W, Kim IH, Alvarenga PH, Valenzuela JG, Ribeiro JMC, Andersen JF. Functional and structural similarities of D7 proteins in the independently-evolved salivary secretions of sand flies and mosquitoes. *Sci Rep.* 2019;9(1):5340.
52. Mans BJ, Calvo E, Ribeiro JM, Andersen JF. The crystal structure of D7r4, a salivary biogenic amine-binding protein from the malaria mosquito *Anopheles gambiae*. *J Biol Chem.* 2007;282(50):36626-33.
53. Calvo E, Mans BJ, Ribeiro JM, Andersen JF. Multifunctionality and mechanism of ligand binding in a mosquito antiinflammatory protein. *Proc Natl Acad Sci U S A.* 2009;106(10):3728-33.
54. Stollar EJ, Smith DP. Uncovering protein structure. *Essays Biochem.* 2020;64(4):649-80.
55. Walder M, Edelstein E, Carroll M, Lazarev S, Fajardo JE, Fiser A, et al. Integrated structure-based protein interface prediction. *BMC Bioinformatics.* 2022;23(1):301.
56. Jisna VA, Jayaraj PB. Protein Structure Prediction: Conventional and Deep Learning Perspectives. *Protein J.* 2021;40(4):522-44.
57. Kaczanowski S, Zielenkiewicz P. Why similar protein sequences encode similar three-dimensional structures? *Theoretical Chemistry Accounts.* 2010;125(3):643-50.
58. Abriata LA, Dal Peraro M. State-of-the-art web services for de novo protein structure prediction. *Brief Bioinform.* 2021;22(3).
59. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* 2018;46(W1):W296-W303.
60. Yang J, Zhang Y. Protein Structure and Function Prediction Using I-TASSER. *Curr Protoc Bioinformatics.* 2015;52:5.8.1-5.8.15.
61. Du Z, Su H, Wang W, Ye L, Wei H, Peng Z, et al. The trRosetta server for fast and accurate protein structure prediction. *Nat Protoc.* 2021;16(12):5634-51.
62. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. *Nature.* 2021;596(7873):583-9.
63. Moulton J, Pedersen JT, Judson R, Fidelis K. A large-scale experiment to assess protein structure prediction methods. *Proteins.* 1995;23(3):ii-v.
64. Pereira J, Simpkin AJ, Hartmann MD, Rigden DJ, Keegan RM, Lupas AN. High-accuracy protein structure prediction in CASP14. *Proteins: Structure, Function, and Bioinformatics.* 2021;89(12):1687-99.
65. Hou J, Wu T, Cao R, Cheng J. Protein tertiary structure modeling driven by deep learning and contact distance prediction in CASP13. *Proteins: Structure, Function, and Bioinformatics.* 2019;87(12):1165-78.
66. Verli H. *Bioinformática : da biologia à flexibilidade molecular.* <http://hdl.handle.net/10183/166105>: Sociedade Brasileira de Bioquímica e Biologia Molecular; 2014.
67. Croitoru A, Park SJ, Kumar A, Lee J, Im W, MacKerell AD, et al. Additive CHARMM36 Force Field for Nonstandard Amino Acids. *J Chem Theory Comput.* 2021;17(6):3554-70.
68. Hollingsworth SA, Dror RO. Molecular Dynamics Simulation for All. *Neuron.* 2018;99(6):1129-43.

69. Kitchen DB, Decornez H, Furr JR, Bajorath J. Docking and scoring in virtual screening for drug discovery: methods and applications. *Nat Rev Drug Discov.* 2004;3(11):935-49.
70. Brooijmans N, Kuntz ID. Molecular Recognition and Docking Algorithms. *Annual Review of Biophysics and Biomolecular Structure.* 2003;32(1):335-73.
71. Gohlke H, Klebe G. Approaches to the Description and Prediction of the Binding Affinity of Small-Molecule Ligands to Macromolecular Receptors. *Angewandte Chemie International Edition.* 2002;41(15):2644-76.
72. Sutherland JJ, Nandigam RK, Erickson JA, Vieth M. Lessons in Molecular Recognition. 2. Assessing and Improving Cross-Docking Accuracy. *Journal of Chemical Information and Modeling.* 2007;47(6):2293-302.
73. Grünberg R, Leckner J, Nilges M. Complementarity of structure ensembles in protein-protein binding. *Structure.* 2004;12(12):2125-36.
74. Valle A, Pérez-Socas LB, Canet L, Hervis YP, de Armas-Guitart G, Martins-de-Sa D, et al. Self-homodimerization of an actinoporin by disulfide bridging reveals implications for their structure and pore formation. *Sci Rep.* 2018;8(1):6614.
75. Svedberg T, Fåhræus R. A NEW METHOD FOR THE DETERMINATION OF THE MOLECULAR WEIGHT OF THE PROTEINS. *Journal of the American Chemical Society.* 1926;48(2):430-8.
76. Lorber B, Fischer F, Bailly M, Roy H, Kern D. Protein analysis by dynamic light scattering: methods and techniques for students. *Biochem Mol Biol Educ.* 2012;40(6):372-82.
77. Bhattacharjee S. DLS and zeta potential - What they are and what they are not? *J Control Release.* 2016;235:337-51.
78. Gasteiger E, Hoogland C, Gattiker A, Duvaud Se, Wilkins MR, Appel RD, et al. Protein Identification and Analysis Tools on the ExPASy Server. In: Walker JM, editor. *The Proteomics Protocols Handbook.* Totowa, NJ: Humana Press; 2005. p. 571-607.
79. Almagro Armenteros JJ, Tsirigos KD, Sønderby CK, Petersen TN, Winther O, Brunak S, et al. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat Biotechnol.* 2019;37(4):420-3.
80. Blom N, Sicheritz-Pontén T, Gupta R, Gammeltoft S, Brunak S. Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. *Proteomics.* 2004;4(6):1633-49.
81. Gupta R, Brunak S. Prediction of glycosylation across the human proteome and the correlation to protein function. *Pac Symp Biocomput.* 2002:310-22.
82. Steentoft C, Vakhrushev SY, Joshi HJ, Kong Y, Vester-Christensen MB, Schjoldager KT, et al. Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. *EMBO J.* 2013;32(10):1478-88.
83. Ko J, Park H, Heo L, Seok C. GalaxyWEB server for protein structure prediction and refinement. *Nucleic Acids Res.* 2012;40(Web Server issue):W294-7.
84. Williams CJ, Headd JJ, Moriarty NW, Prisant MG, Videau LL, Deis LN, et al. MolProbity: More and better reference data for improved all-atom structure validation. *Protein Sci.* 2018;27(1):293-315.
85. Jurrus E, Engel D, Star K, Monson K, Brandi J, Felberg LE, et al. Improvements to the APBS biomolecular solvation software suite. *Protein Sci.* 2018;27(1):112-28.
86. Tian W, Chen C, Lei X, Zhao J, Liang J. CASTp 3.0: computed atlas of surface topography of proteins. *Nucleic Acids Res.* 2018;46(W1):W363-W7.
87. Dundas J, Ouyang Z, Tseng J, Binkowski A, Turpaz Y, Liang J. CASTp: computed atlas of surface topography of proteins with structural and topographical

- mapping of functionally annotated residues. *Nucleic Acids Res.* 2006;34(Web Server issue):W116-8.
88. Jo S, Kim T, Iyer VG, Im W. CHARMM-GUI: a web-based graphical user interface for CHARMM. *J Comput Chem.* 2008;29(11):1859-65.
 89. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, et al. Scalable molecular dynamics with NAMD. *J Comput Chem.* 2005;26(16):1781-802.
 90. Klauda JB, Venable RM, Freites JA, O'Connor JW, Tobias DJ, Mondragon-Ramirez C, et al. Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. *J Phys Chem B.* 2010;114(23):7830-43.
 91. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem.* 2004;25(13):1605-12.
 92. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. *J Mol Graph.* 1996;14(1):33-8, 27-8.
 93. Tubiana T, Carvaille JC, Boulard Y, Bressanelli S. TTClust: A Versatile Molecular Simulation Trajectory Clustering Program with Graphical Summaries. *J Chem Inf Model.* 2018;58(11):2178-82.
 94. Ward JH. Hierarchical Grouping to Optimize an Objective Function. *Journal of the American Statistical Association.* 1963;58(301):236-44.
 95. Hanwell MD, Curtis DE, Lonie DC, Vandermeersch T, Zurek E, Hutchison GR. Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *J Cheminform.* 2012;4(1):17.
 96. O'Boyle NM, Banck M, James CA, Morley C, Vandermeersch T, Hutchison GR. Open Babel: An open chemical toolbox. *J Cheminform.* 2011;3:33.
 97. Best RB, Zhu X, Shim J, Lopes PE, Mittal J, Feig M, et al. Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone ϕ , ψ and side-chain $\chi(1)$ and $\chi(2)$ dihedral angles. *J Chem Theory Comput.* 2012;8(9):3257-73.
 98. Rostkowski M, Olsson MH, Søndergaard CR, Jensen JH. Graphical analysis of pH-dependent properties of proteins predicted using PROPKA. *BMC Struct Biol.* 2011;11:6.
 99. Guedes IA, Barreto AMS, Marinho D, Krempser E, Kuenemann MA, Sperandio O, et al. New machine learning and physics-based scoring functions for drug discovery. *Sci Rep.* 2021;11(1):3198.
 100. Adasme MF, Linnemann KL, Bolz SN, Kaiser F, Salentin S, Haupt V J, et al. PLIP 2021: expanding the scope of the protein–ligand interaction profiler to DNA and RNA. *Nucleic Acids Research.* 2021.
 101. Sambrook J. *Molecular cloning : a laboratory manual.* In: Russell DW, editor. 3rd. ed. ed. Cold Spring Harbor, N.Y. :: Cold Spring Harbor Laboratory; 2001.
 102. Laemmli UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature.* 1970;227(5259):680-5.
 103. Kabsch W, Sander C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers.* 1983;22(12):2577-637.
 104. Micsonai A, Wien F, Bulyáki É, Kun J, Moussong É, Lee YH, et al. BeStSel: a web server for accurate protein secondary structure prediction and fold recognition from the circular dichroism spectra. *Nucleic Acids Res.* 2018;46(W1):W315-W22.
 105. Evans R, O'Neill M, Pritzel A, Antropova N, Senior A, Green T, et al. Protein complex prediction with AlphaFold-Multimer. *bioRxiv.* 2021:2021.10.04.463034.

106. Kelly SM, Jess TJ, Price NC. How to study proteins by circular dichroism. *Biochim Biophys Acta*. 2005;1751(2):119-39.
107. Champagne DE. Antihemostatic strategies of blood-feeding arthropods. *Curr Drug Targets Cardiovasc Haematol Disord*. 2004;4(4):375-96.
108. Deeb O, Zhou J. *In silico* Methods for Drug Design. *Current Topics in Medicinal Chemistry*. 2023;23(3):155-7.
109. Katsila T, Spyroulias GA, Patrinos GP, Matsoukas MT. Computational approaches in target identification and drug discovery. *Comput Struct Biotechnol J*. 2016;14:177-84.
110. Ribeiro JM, Martin-Martin I, Arcà B, Calvo E. A Deep Insight into the Sialome of Male and Female *Aedes aegypti* Mosquitoes. *PLoS One*. 2016;11(3):e0151400.
111. Bussacos AC, Nakayasu ES, Hecht MM, Parente JA, Soares CM, Teixeira AR, et al. Diversity of anti-haemostatic proteins in the salivary glands of *Rhodnius* species transmitters of Chagas disease in the greater Amazon. *J Proteomics*. 2011;74(9):1664-72.
112. Ononamadu CJ, Abdalla M, Ihegboro GO, Li J, Owolarafe TA, John TD, et al. identification and study of potential anti-mosquito juvenile hormone binding protein (MJHBP) compounds as candidates for dengue virus - Vector insecticides. *Biochem Biophys Rep*. 2021;28:101178.
113. de Brito GA, Rocha de Oliveira PF, de Andrade Silva CM, de Araújo Neto MF, Leite FHA, Mesquita PRR, et al. Identification of Bioactive Compounds against *Aedes aegypti* (Diptera: Culicidae) by Bioassays and in Silico Assays. *Chem Biodivers*. 2021;18(9):e2100242.
114. Kritsi E, Liggri PGV, Stamati ECV, Tsitsanou KE, Zographos SE, Michaelakis A, et al. A Combined Computational Methodology for the Discovery of Hit Compounds with Putative Insect Repellency Properties. *ChemMedChem*. 2022;17(16):e202200271.
115. Xu C, Ng DT. Glycosylation-directed quality control of protein folding. *Nat Rev Mol Cell Biol*. 2015;16(12):742-52.
116. Moghadam M, Ganji A, Varasteh A, Falak R, Sankian M. Refolding process of cysteine-rich proteins:Chitinase as a model. *Rep Biochem Mol Biol*. 2015;4(1):19-24.
117. HABER E, ANFINSEN CB. Side-chain interactions governing the pairing of half-cystine residues in ribonuclease. *J Biol Chem*. 1962;237:1839-44.
118. Singh SM, Panda AK. Solubilization and refolding of bacterial inclusion body proteins. *J Biosci Bioeng*. 2005;99(4):303-10.
119. Martin-Martin I, Smith LB, Chagas AC, Sá-Nunes A, Shrivastava G, Valenzuela-Leon PC, et al. D7 Salivary Protein Prevents Host Hemostasis and Inflammation. *Biomolecules*. 2020;10(10).
120. Andronopoulou E, Labropoulou V, Douris V, Woods DF, Biessmann H, Iatrou K. Specific interactions among odorant-binding proteins of the African malaria vector *Anopheles gambiae*. *Insect Mol Biol*. 2006;15(6):797-811.
121. Mascarenhas NM, Gosavi S. Understanding protein domain-swapping using structure-based models of protein folding. *Prog Biophys Mol Biol*. 2017;128:113-20.
122. Valenzuela JG, Pham VM, Garfield MK, Francischetti IM, Ribeiro JM. Toward a description of the sialome of the adult female mosquito *Aedes aegypti*. *Insect Biochem Mol Biol*. 2002;32(9):1101-22.