

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO
INSTITUTO DE QUÍMICA
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOQUÍMICA

GABRIELLE CARVALHO MIGUENS CABO

INVESTIGAÇÃO DO NEUROPEPTÍDEO F CURTO (sNPF) E DE SEU RECEPTOR
(sNPFR) EM TECIDOS DE *RHODNIUS PROLIXUS*: implicações no controle do vetor da
doença de Chagas

Rio de Janeiro

2025

Gabrielle Carvalho Miguens Cabo

INVESTIGAÇÃO DO NEUROPEPTÍDEO F CURTO (sNPF) E DE SEU RECEPTOR (sNPFR) EM TECIDOS DE *RHODNIUS PROLIXUS*: implicações no controle do vetor da doença de Chagas

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Bioquímica, Instituto de Química, Universidade Federal do Rio de Janeiro, como requisito parcial à obtenção do título de Mestre em Ciências.

Orientadora: Prof^a Ana Claudia do Amaral Melo

Coorientadora: Dra. Nathália Faro de Brito

Rio de Janeiro

2025

CIP - Catalogação na Publicação

C116i Cabo, Gabrielle Carvalho Miguens
Investigação do neuropeptídeo F curto (sNPF) e de seu receptor (sNPFR) em tecidos de *Rhodnius prolixus*: implicações no controle do vetor da doença de Chagas / Gabrielle Carvalho Miguens Cabo. -- Rio de Janeiro, 2025.
80 f.

Orientador: Ana Claudia do Amaral Melo.
Coorientador: Nathália Faro de Brito.
Dissertação (mestrado) - Universidade Federal do Rio de Janeiro, Instituto de Química, Programa de Pós Graduação em Bioquímica, 2025.

1. *Rhodnius prolixus*. 2. neuropeptídeo F curto. 3. expressão gênica. 4. trans splicing. I. do Amaral Melo, Ana Claudia, orient. II. Faro de Brito, Nathália, coorient. III. Título.

Elaborado pelo Sistema de Geração Automática da UFRJ com os dados fornecidos pelo(a) autor(a), sob a responsabilidade de Miguel Romeu Amorim Neto - CRB-7/6283.

Gabrielle Carvalho Miguens Cabo

INVESTIGAÇÃO DO NEUROPEPTÍDEO F CURTO (sNPF) E DE SEU RECEPTOR (sNPFR) EM TECIDOS DE *RHODNIUS PROLIXUS*: implicações no controle do vetor da doença de Chagas

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Bioquímica, Instituto de Química, Universidade Federal do Rio de Janeiro, como requisito parcial à obtenção do título de Mestre em Ciências.

Aprovado em: 26 de março de 2025.

Prof.^a Ana Claudia do Amaral Melo

(Instituto de Química, Universidade Federal do Rio de Janeiro – UFRJ)

Dr.^a Nathália Faro de Brito

(Programa de Terapia Celular e Gênica, Instituto Nacional de Câncer - INCA)

Prof. Marcos Dias Pereira

(Instituto de Química, Universidade Federal do Rio de Janeiro – UFRJ)

Prof. David Majerowicz

(Faculdade de Farmácia, Universidade Federal do Rio de Janeiro – UFRJ)

AGRADECIMENTOS

Primeiramente, agradeço aos meus pais, pois sem eles, eu não chegaria aqui. Minha querida mamãe Monique, companhia no tédio, nos sorrisos e nas choradeiras. É, sem dúvida, a pessoa mais verdadeira que eu conheço e que me ensina todos os dias a ser uma mulher completa e bem resolvida. Meu pai Alexandre me ensina a ser resiliente e não desistir. Além de ter puxado a cara dele, também puxei a vontade de sair e explorar o mundo. Obrigada por terem me priorizado, trabalho para um dia retribuir todo o esforço colocado em mim. Amo vocês!

Agradeço às minhas orientadoras, professora Ana Claudia Melo e dra. Nathalia Brito. Admiro muito as duas e sou grata por ser orientada por mulheres tão incríveis! Ana, obrigada pelos aprendizados, pela paciência e os conselhos que vão além da academia. Nathália, obrigada por me inspirar e me preparar enquanto eu fui sua IC, espero ainda aprender muito mais com você.

Agradeço também à professora Monica Moreira por todos os ensinamentos e boas conversas na hora do café.

Minhas meninas Mylena e Elisa, que me conheceram em todas as atualizações, desde a que eu comia meleca até a que eu estou me tornando mestre em ciências. Minha querida Nayara, que chegou e ficou, e por quem sou extremamente grata pela amizade que nós construímos. Minha fiel Luiza, que mesmo estando longe, eu consigo sentir todo o apoio e a admiração. Meu amigo Matheus (o Vianna), que é minha companhia de shows, bares e conversas sobre a vida. Muito obrigada a todos vocês que sempre estão ao meu lado e que me lembram sempre o que eu mereço de verdade.

Não posso esquecer dos meus amigos do LABBMOVE. As admiráveis ICs Emely, Jaqueline (JaquEmely!) e Luiza, e os fantásticos vizinhos de bancada Luís Guilherme e Beatriz. As minhas companheiras que entendem perfeitamente o que é viver a pós-graduação, Pamela, Brenda e Márcia. Vocês foram um alívio no meio de tanto caos e são capazes de qualquer coisa, obrigada pela parceria.

Agradeço aos colegas do LAMMP, que tantas vezes abriram as portas do laboratório para que eu pudesse usar os equipamentos. Aos professores do Instituto de Química que ministraram disciplinas incríveis. Aos professores que me incentivaram nos seminários de acompanhamento, Rodrigo Volcan, Gilberto Domont e, especialmente, o professor Anderson Pinheiro, que nos deixou tão cedo. Ciência se faz compartilhando!

Sem a colônia de *Rhodnius prolixus*, no CCS, esse trabalho também não seria possível. Agradeço às técnicas e à professora Katia Gondim, pela colaboração e paciência.

Agradeço à CAPES, pelo apoio financeiro.

Agradeço a todos os funcionários, servidores e terceirizados, por manterem a universidade de pé.

Obrigada!

*You got your passion, you got your pride
But don't you know that only fools are satisfied?
Dream on, but don't imagine they'll all come true
When will you realize, Vienna waits for you?*

(Vienna, Billy Joel)

RESUMO

CABO, Gabrielle Carvalho Miguens. Investigação do neuropeptídeo F curto (sNPF) e de seu receptor (sNPFR) em tecidos de *Rhodnius prolixus*: implicações no controle do vetor da doença de Chagas. Rio de Janeiro, 2025. Dissertação (Mestrado em Bioquímica) – Instituto de Química, Universidade Federal do Rio de Janeiro. Março, 2025.

Os neuropeptídeos regulam processos fisiológicos em artrópodes, como alimentação e comportamento. Produzidos principalmente no sistema nervoso central, eles interagem com receptores acoplados à proteína G (GPCR), desencadeando cascatas de sinalização celular. Entre eles, o *short neuropeptide F* (sNPF), e seu receptor sNPFR, destacam-se por regular alimentação e imunidade, possuindo conservação evolutiva e similaridade funcional com o neuropeptídeo Y (NPY) de vertebrados. Em *Rhodnius prolixus*, uma única proteína precursora gera uma forma ativa de sNPF. Contudo, a sequência codificante do receptor ainda não está completamente elucidada. Em insetos hematófagos, compreender a busca pela alimentação é crucial para estratégias de controle vetorial. Diferente do que ocorre em outros insetos, no genoma de *R. prolixus* estão anotados três genes contíguos parciais (RPRC002266, RPRC002268 e RPRC002269) associados ao receptor sNPFR. Assim, este estudo tem como objetivo identificar a sequência completa do sNPFR, realizar uma análise *in silico* das proteínas sNPF e sNPFR, incluindo a modelagem tridimensional de suas estruturas e a predição da energia de ligação proteína-ligante por *docking* molecular, além de investigar seus perfis de expressão em diferentes tecidos. RNA total foi extraído de tecidos de ninfas, tratado com DNase I, e transcrito reversamente em cDNA. Reações de PCR, com *primers* específicos para os genes do sNPFR, amplificaram a sequência-alvo, posteriormente clonada e sequenciada. A tradução para aminoácidos foi realizada pela ferramenta ExPASy. A topologia da proteína foi prevista pelo DeepTMHMM, enquanto a expressão gênica em diferentes tecidos foi avaliada por qPCR, utilizando *RproR18S* como gene de referência. O sequenciamento revelou três regiões codificantes: o exon 1 presente no gene RPRC002266, o exon 2 no gene RPRC002268 e o exon 2 no RPRC002269. A proteína resultante possui sete domínios transmembranares típicos de GPCRs, com posições conservadas. A maior parte dos transcritos do sNPFR foi observada no cérebro, seguido de intestino, túbulo de Malpighi e corpo gorduroso. Transcritos para o sNPF foram encontrados no cérebro e corpo gorduroso. Baseado na anotação do genoma para o sNPFR, que mostra 3 genes, podemos estar diante de um evento de *trans-splicing*, já previamente descrito em Hemiptera. A expressão ampla de sNPFR reforça seu papel fisiológico diversificado, enquanto o sNPF parece ter funções específicas no cérebro e no corpo gorduroso. Este estudo é o passo inicial para esclarecer a via de sinalização pelo sistema sNPF/sNPFR em *R. prolixus*, abrindo novas frentes para o estudo do comportamento deste inseto.

Palavras-chave: *Rhodnius prolixus*; neuropeptídeo F curto; expressão gênica; *trans-splicing*.

ABSTRACT

CABO, Gabrielle Carvalho Miguens. Investigation of the short neuropeptide F (sNPF) and its receptor (sNPFR) in *Rhodnius prolixus* tissues: implications for Chagas disease vector control. Rio de Janeiro, 2025. Dissertação (Mestrado em Bioquímica) – Instituto de Química, Universidade Federal do Rio de Janeiro. Março, 2025.

Neuropeptides regulate physiological processes in arthropods, such as feeding and behavior. Primarily produced in the central nervous system, they interact with G protein-coupled receptors (GPCRs), triggering intracellular signaling cascades. Among them, the short neuropeptide F (sNPF) and its receptor, sNPFR, stand out for their role in regulating feeding and immunity, exhibiting evolutionary conservation and functional similarity to the vertebrate neuropeptide Y (NPY). In *Rhodnius prolixus*, a single precursor protein generates an active form of sNPF. However, the coding sequence of its receptor remains incompletely elucidated. In hematophagous insects, understanding feeding behavior is crucial for vector control strategies. Unlike other insects, the *R. prolixus* genome contains three contiguous partial genes (*RPRC002266*, *RPRC002268*, and *RPRC002269*) associated with the sNPFR receptor. Thus, this study aims to identify the complete sequence of sNPFR, perform an *in silico* analysis of the sNPF and sNPFR proteins—including three-dimensional modeling of their structures and the prediction of protein-ligand binding energy via molecular docking—as well as investigate their expression profiles in different tissues. Total RNA was extracted from nymph tissues, treated with DNase I, and reverse-transcribed into cDNA. PCR reactions using specific primers for sNPFR genes amplified the target sequence, which was subsequently cloned and sequenced. Amino acid translation was performed using the ExPASy tool. Protein topology was predicted using DeepTMHMM, while gene expression in different tissues was evaluated by qPCR, with *RproR18S* as the reference gene. Sequencing revealed three coding regions: exon 1 in the *RPRC002266* gene, exon 2 in *RPRC002268*, and exon 2 in *RPRC002269*. The resulting protein contains seven transmembrane domains, characteristic of GPCRs, with conserved positions. The highest sNPFR transcript levels were observed in the brain, followed by the gut, Malpighian tubule, and fat body. Transcripts for sNPF were found in the brain and fat body. Based on the genome annotation showing three genes for sNPFR, this may represent a trans-splicing event, which has been previously described in Hemiptera. The broad expression of sNPFR reinforces its diverse physiological role, while sNPF appears to have more specific functions in the brain and fat body. This study represents an initial step toward elucidating the sNPF/sNPFR signaling pathway in *R. prolixus*, opening new avenues for studying the behavior of this insect.

Keywords: *Rhodnius prolixus*; short neuropeptide F; gene expression; trans splicing.

LISTA DE FIGURAS

Figura 1 – Ciclo biológico de <i>Rhodnius prolixus</i>	14
Figura 2 – Formas de transmissão do <i>T. cruzi</i>	15
Figura 3 – Desenho esquemático da olfação em insetos.	17
Figura 4 – Mapa do vetor plasmidial pGEM®-T Easy Vector com os pontos de referência.	28
Figura 5 – Perfil de expressão dos genes Rpro-R18S e Rpro-sNPF.	31
Figura 6 - Perfil de expressão dos genes Rpro-R18S e Rpro-sNPFR.	32
Figura 7 - Perfil de expressão relativa do gene codificante sNPF e sNPFR em <i>Rhodnius prolixus</i>	32
Figura 8 – Visualização do supercontig KQ034095.	35
Figura 9 – Alinhamento de sequências de aminoácidos das proteínas previstas RPRC002266 com sNPFR de <i>Anopheles gambiae</i> (ABD96049.1) e <i>Aedes aegypti</i> (AVR59278.1).	36
Figura 10 - Alinhamento de sequências de aminoácidos das proteínas previstas RPRC002268 com sNPFR de <i>Anopheles gambiae</i> (ABD96049.1) e <i>Aedes aegypti</i> (AVR59278.1).	38
Figura 11 - Alinhamento de sequências de aminoácidos das proteínas previstas RPRC002269 com sNPFR de <i>Anopheles gambiae</i> (ABD96049.1) e <i>Aedes aegypti</i> (AVR59278.1).	39
Figura 12 – Desenho esquemático que representa a amplificação do fragmento esperado.	40
Figura 13 – Alinhamento entre o resultado do sequenciamento (primeira linha) e a sequência do fragmento esperado (segunda linha).	42
Figura 14 - Alinhamento de sequências de nucleotídeos entre regiões do sequenciamento e os genes RPRC002266, RPRC002268 e RPRC002269, de <i>R. prolixus</i>	43
Figura 15 - Desenho esquemático representando uma possível via de transcrição do sNPFR.	46
Figura 16 – Sequência de RNA/mRNA previsto para sNPFR, obtida após sequenciamento.	48
Figura 17 – Sequência primária de Rpro-sNPFR e topologia prevista.	49
Figura 18 - Alinhamento múltiplo de sequências baseado na proteína Rpro-sNPFR.	50
Figura 19 - Análise filogenética dos genes de receptores de sNPF e NPY, e do receptor semelhante ao do peptídeo liberador de prolactina (PrRPR-like).	52
Figura 20 - Características da sequência da proteína precursora de sNPF.	54
Figura 21 – Teste de diferença de distância local previsto para cada posição dos aminoácidos nos modelos.	55
Figura 22 – Modelo de Rpro-sNPFR gerado pelo <i>AlphaFold2</i>	56
Figura 23 – Rpro-sNPFR gerado por modelagem por homologia.	57
Figura 24 – Comparação entre os modelos Rpro-sNPFR.	58
Figura 25 – Comparação estrutural de Rpro-sNPFR com NPY2-R de <i>H. sapiens</i>	60
Figura 26 – Sítio de interação Rpro-sNPFR (modelo SWISS-Model) com o ligante Rpro-sNPF.	62

LISTA DE TABELAS

Tabela 1 — Oligonucleotídeos iniciadores utilizados em reações de PCR.....	25
Tabela 2 – Protocolo de incubação das reações de amplificação em cadeia da polimerase.....	28
Tabela 3 — Ranking da afinidade de ligação obtida com <i>docking</i> molecular às cegas e direcionado.....	63

LISTA DE ABREVIATURAS E SIGLAS

AF2	AlphaFold2
CPI	Células produtoras de insulina
Cryo-EM	Crio-eletromicroscopia
ECL	Loop extracelular
GPCR	Receptor acoplado à proteína G
ICL	Loop intracelular
Lola	Lack of longitudinals
mod(mdg4)	Modificador do mdg4
OBPs	Proteínas ligadoras de odor
OK	Orcocinina
OR	Receptor olfativo
Orco	Correceptor olfativo
ORF	Quadro de leitura aberto
PAM	Peptidil-glicina- α -amidação monooxigenase
pb	Pares de base
PDB	Protein Data Bank
plDDT	Teste de diferença de distância local previsto
PrPR	Receptor do peptídeo liberador de prolactina
PrRPR-like	Receptor semelhante ao do peptídeo liberador de prolactina
qPCR	PCR quantitativo
RMN	Ressonância Magnética Nuclear
RNA _t .	RNA total
sNPF.	Neuropeptídeo F curto
sNPFR.	Receptor do neuropeptídeo F curto
SNPs.	Polimorfismo de nucleotídeo único
TM.	Domínio transmembrana
VB.	VectorBase

SUMÁRIO

1	INTRODUÇÃO.....	13
1.1	<i>RHODNIUS PROLIXUS</i>	13
1.2	DOENÇA DE CHAGAS.....	14
1.3	COMUNICAÇÃO QUÍMICA.....	16
1.4	SISTEMA NEUROENDÓCRINO	17
2	OBJETIVOS.....	21
2.1	OBJETIVO GERAL	21
2.2	OBJETIVOS ESPECÍFICOS	21
3	MATERIAIS E MÉTODOS.....	22
3.1	CRIAÇÃO E MANUTENÇÃO DE <i>RHODNIUS PROLIXUS</i>	22
3.2	DISSECÇÃO DOS INSETOS.....	22
3.3	DESENHO DOS OLIGONUCLEOTÍDEOS INICIADORES.....	23
3.4	EXTRAÇÃO DE RNA TOTAL.....	23
3.5	DETERMINAÇÃO DE CONCENTRAÇÃO DE ÁCIDOS NUCLEICOS.....	24
3.6	TRATAMENTO DE RNA COM DNASE I	24
3.7	SÍNTESE DA FITA DE DNA COMPLEMENTAR (cDNA).....	25
3.8	REAÇÃO DE AMPLIFICAÇÃO EM CADEIA DA POLIMERASE.....	25
3.9	ANÁLISE POR ELETROFORESE EM GEL DE AGAROSE	26
3.10	PURIFICAÇÃO DOS PRODUTOS DE PCR	26
3.11	CLONAGEM MOLECULAR E SEQUENCIAMENTO DO SNPFR	27
3.12	ANÁLISE FILOGENÉTICA	29
3.13	ANÁLISE <i>IN SILICO</i>	30
4	RESULTADOS E DISCUSSÃO.....	31
4.1	QUANTIFICAÇÃO DE TRANSCRITOS EM CONDIÇÃO DE JEJUM.....	31
4.2	SEQUÊNCIA CODIFICANTE DO RECEPTOR DE SNPF	34
4.3	ANÁLISE DE DOMÍNIOS CONSERVADOS DA PROTEÍNA RPRO-SNPFR.....	46
4.4	ANÁLISE FILOGENÉTICA	51
4.5	CARACTERIZAÇÃO ESTRUTURAL <i>IN SILICO</i> DA PROTEÍNA RPRO-SNPFR.....	55
5	CONCLUSÃO.....	63
6	PERSPECTIVAS.....	65
7	REFERÊNCIAS.....	66
	ANEXO I	72
	ANEXO II.....	80

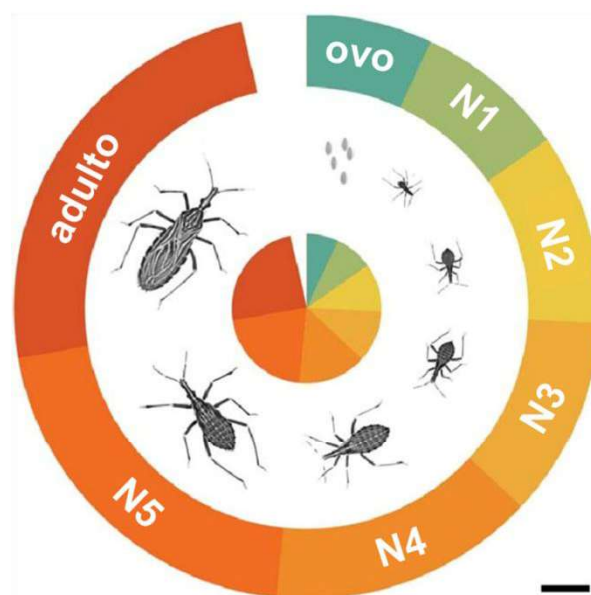
1 INTRODUÇÃO

1.1 *Rhodnius prolixus*

Rhodnius prolixus (Ordem: Hemiptera), popularmente conhecido como barbeiro, é um inseto triatomíneo e um dos principais vetores da Doença de Chagas (DC). Juntamente com *Triatoma infestans*, *T. brasiliensis* e *Panstrongylus megistus* (Almeida; Lima; Costa, 2014), é uma das espécies domiciliadas que ocorrem no Brasil. Em ecótopos artificiais, esses insetos se abrigam em galinheiros, celeiros, pocilgas e se alimentam do sangue dos hospedeiros vertebrados de sangue quente que ali residem, incluindo o ser humano.

O ciclo biológico de *R. prolixus* conta com ovo, cinco fases de ninfa e a fase adulta, quando ocorre a formação das asas e a maturação sexual (Figura 1). Para completar cada muda, o inseto necessita de um repasto sanguíneo, essencial para obter os nutrientes necessários para a ecdise e produção de ovos. É durante a alimentação que o inseto contaminado transmite o protozoário *Trypanosoma cruzi*, agente etiológico da DC. Assim, a alimentação desses insetos possui grande relevância sanitária, especialmente porque a maior incidência da doença é observada nas regiões Norte e Nordeste do Brasil, principalmente em locais onde há maior disparidade socioeconômica (Pereira *et al.*, 2015). Embora *R. prolixus* seja considerado hematófago estrito, observou-se que suas ninfas também são capazes de ingerir diferentes soluções de açúcar (Costa *et al.*, 2024). Esse achado sugere que, na natureza, esses insetos podem obter energia não apenas do sangue, mas também pela ingestão de açúcares.

Figura 1 – Ciclo biológico de *Rhodnius prolixus*. Inicialmente, uma fêmea adulta faz a postura dos ovos, que eclodem em ninfa de 1º estágio. Por se tratar de um inseto hemimetábolo, o inseto passa por mais quatro estágios ninfais, tendo características semelhantes ao estágio adulto, exceto a presença de asas e a maturação sexual. Após o último estágio ninfal, há a emergência em adultos.

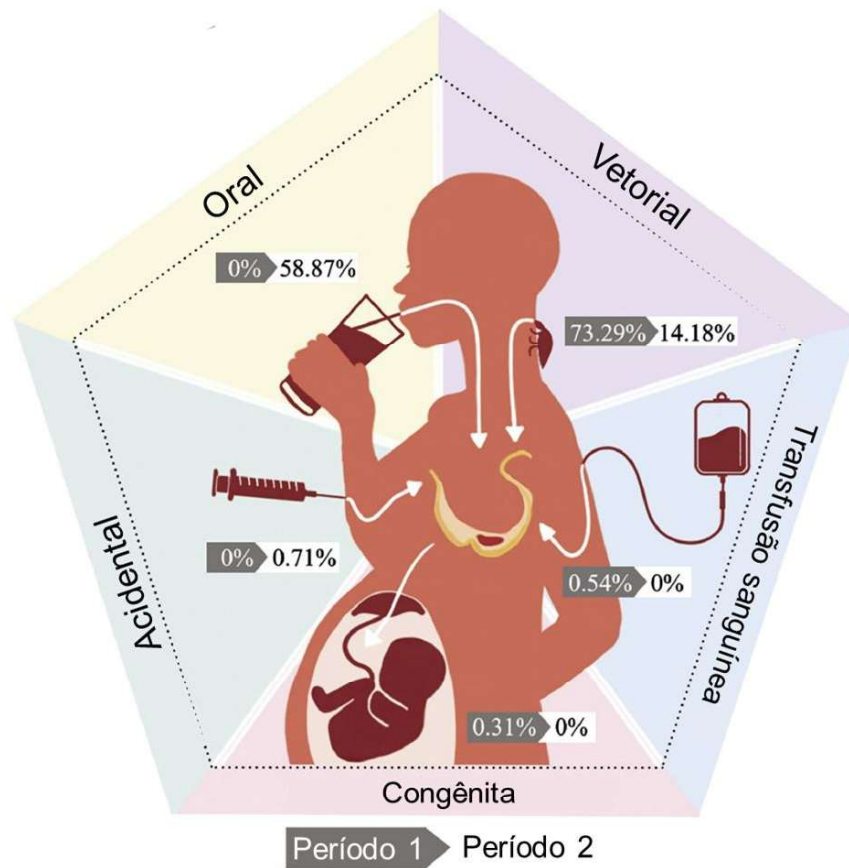


Fonte: Adaptado de Nunes-da-Fonseca *et al.*, 2017.

1.2 DOENÇA DE CHAGAS

O primeiro caso da DC foi descrito em 1909, quando Carlos Chagas identificou a infecção em uma criança de 2 anos, em Minas Gerais (Lidani *et al.*, 2019). Apesar dos sintomas leves observados na fase aguda, a fase crônica - que costuma ocorrer entre 10 e 15 anos após a infecção - afeta principalmente o coração e o trato digestivo, impactando significativamente a qualidade de vida dos indivíduos acometidos (Gullo *et al.*, 2012; Simões *et al.*, 2018). Por se tratar de uma doença tropical negligenciada, ainda não há perspectiva de vacinas. Assim, o controle da DC é baseado exclusivamente na prevenção da transmissão, incluindo melhorias habitacionais em regiões endêmicas e o combate químico ao vetor (Dias, 2007). O ciclo dessa zoonose é mantido entre vertebrados e insetos triatomíneos que desempenham um papel crucial na transmissão do *T. cruzi*. Em áreas endêmicas, a forma mais comum de transmissão ocorre por via vetorial, frequentemente por meio da ingestão acidental de partes de triatomíneos contaminados (Figura 2) (Franco-Paredes *et al.*, 2020; Santos *et al.*, 2024).

Figura 2 – Formas de transmissão do *T. cruzi*. A figura compara as formas de transmissão oral, vetorial, acidental, congênita e por transfusão sanguínea em dois períodos diferentes (Período 1: 2001-2006 e Período 2: 2007-2021).



Fonte: Adaptado de Santos *et al.*, 2024.

Contudo, não existe uma estratégia eficaz de controle populacional dos triatomíneos, que poderia ser uma forte aliada na redução da incidência da doença. Diante desse cenário, torna-se essencial desenvolver estratégias de controle que levem em consideração a fisiologia do inseto-alvo, buscando minimizar o impacto ambiental do uso de inseticidas e contornar a baixa eficácia dos repelentes de proteção individual contra o *R. prolixus* (Franco *et al.*, 2018). Uma abordagem promissora envolve “confundir” a comunicação química desses insetos com o ambiente, de maneira mais específica e sustentável.

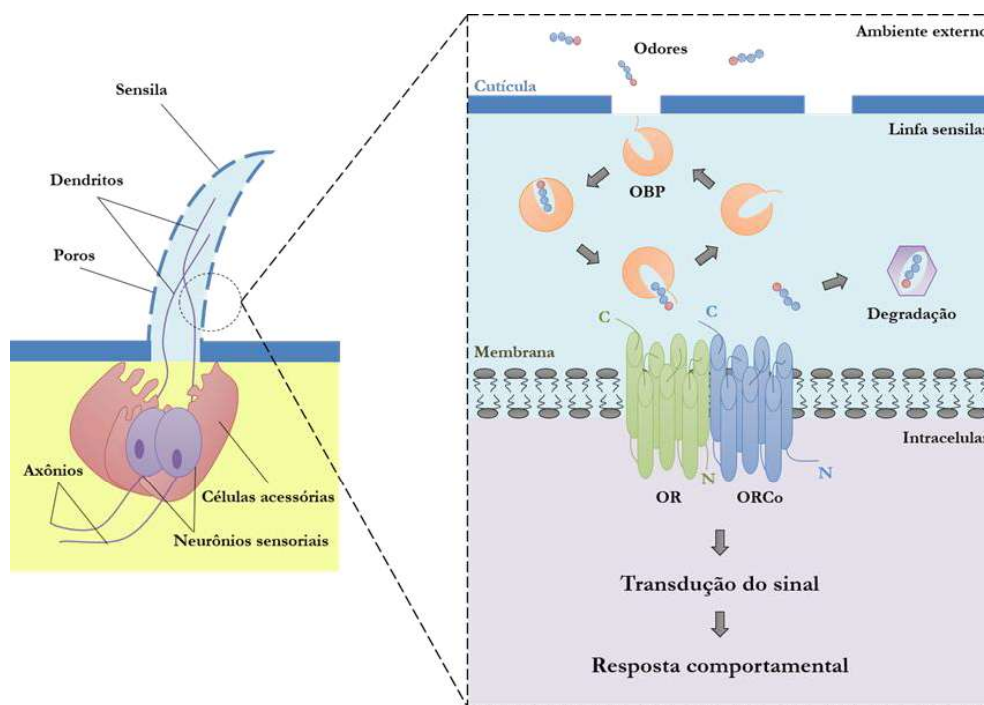
1.3 COMUNICAÇÃO QUÍMICA

A sobrevivência dos insetos depende basicamente de seu comportamento, o qual é mediado pela interpretação de estímulos (mecânicos, visuais, térmicos e químicos) detectados por células especializadas que compõem o sistema nervoso periférico (Gullan; Cranston, 2017). Entre esses estímulos, a quimiorrecepção de odores e feromônios, que ocorre principalmente nas antenas, envolve uma complexa cascata de eventos celulares (Hansson; Stensmyr, 2011).

A resposta comportamental dos insetos se inicia quando odores, que são compostos voláteis orgânicos (COVs) presentes no ambiente atravessam os poros das sensilas das antenas e são capturados pelas proteínas ligadoras de odor (OBP) ou de feromônio (PBPs) (Brito; Moreira; Melo, 2016). Essas OBPs/PBPs são pequenas proteínas solúveis responsáveis por transportar COVs até um complexo proteico heteromérico composto por um receptor olfativo (OR) e o correceptor olfativo (Orco) que estão expressos nos dendritos das membranas dos neurônios sensoriais olfativos, na proporção de 1:3, respectivamente (Zhao *et al.*, 2024). A ligação do odor ao OR desencadeia uma cascata de sinalização celular, culminando na propagação do impulso nervoso até o sistema nervoso central (Figura 3). A informação química, transformada em impulso elétrico, leva o inseto a exibir uma resposta comportamental, sendo que a correta interpretação dos estímulos ambientais é fundamental para a identificação de hospedeiros, parceiros sexuais, locais de postura de ovos, entre outros comportamentos essenciais para a sua sobrevivência (Brito; Moreira; Melo, 2016).

O papel do Orco nesse processo é bem estabelecido em diferentes insetos (Costa-da-Silva *et al.*, 2024; Paulo *et al.*, 2021; Tribble *et al.*, 2017). Em *R. prolixus*, por exemplo, o silenciamento do *Orco* resulta na incapacidade do inseto de localizar o hospedeiro, na redução do consumo de sangue e na diminuição da taxa de ecdise (Franco *et al.*, 2016). No entanto, a necessidade pela procura do alimento (sangue) ocorre por meio de outra via de sinalização, a via neuroendócrina.

Figura 3 – Desenho esquemático da olfação em insetos. O estímulo químico penetra a cutícula das sensilas olfativas, são transportados por uma proteína ligadora de odor (OBP) ou de feromônio (PBP) até o complexo proteico receptor olfativo (OR)-correceptor olfativo (Orco), onde o OR é ativado e ocorre a transdução do sinal, que provoca uma resposta comportamental.



Fonte: Brito; Moreira; Melo, 2016.

1.4 SISTEMA NEUROENDÓCRINO

Além da sinalização olfativa mediada pelo Orco, que desempenha um papel fundamental na detecção de COVs e na modulação do comportamento alimentar (Franco *et al.*, 2016), outros mecanismos neuroendócrinos também são fundamentais na regulação desse comportamento nos insetos. Um exemplo é o neuropeptídeo F curto (sNPF) e seu receptor (sNPFR), que pertence à família de receptores acoplados à proteína G (GPCR) (Caers *et al.*, 2012; Zhou *et al.*, 2024). Esses componentes estão envolvidos no estímulo da alimentação, influenciando diretamente a resposta comportamental do inseto diante de sinais alimentares (Gullan; Cranston, 2017; Zhao *et al.*, 2025).

O neuropeptídeo F curto (sNPF) se liga ao seu receptor sNPFR, juntos, participam de vários processos fisiológicos, incluindo a modulação da sinalização sensorial. A relação entre o

sNPFR e o Orco tem atraído crescente atenção no campo da fisiologia de insetos, especialmente no entendimento dos mecanismos de sinalização olfativa (Cholewiński *et al.*, 2024). Estudos recentes destacaram a localização e a interação funcional entre esses dois receptores nas antenas de diversas espécies de insetos (Li *et al.*, 2022).

No contexto do processamento olfativo, a co-localização de sNPFR e Orco nas antenas de insetos como *Bactrocera dorsalis*, sugere que o sNPFR pode desempenhar um papel importante na modulação da atividade dos NSOs que expressam Orco, o que é essencial para o funcionamento dos ORs (Cholewiński *et al.*, 2024). Essa interação entre sNPFR e Orco tem implicações funcionais significativas. Ela pode levar a respostas aumentadas ou diminuídas a estímulos olfativos, dependendo da sinalização mediada pelo sNPFR. Essa modulação é crucial para os comportamentos adaptativos dos insetos, como a busca por alimentos e o acasalamento, que, como discutido anteriormente, dependem fortemente da interpretação de sinais olfativos. Assim, a interação entre a sinalização olfativa mediada por Orco e os mecanismos neuroendócrinos como o sistema sNPF/sNPFR contribui para a complexidade do comportamento alimentar e reprodutivo dos insetos (Cholewiński *et al.*, 2024).

A primeira identificação do sNPF foi descrita no mosquito *Aedes aegypti*, onde foi designado como "peptídeos da cabeça de *Aedes*" (Nässel *et al.*, 2008). Essa descoberta marcou um avanço significativo no entendimento da sinalização de neuropeptídeos em insetos, pois destacou a presença do sNPF no SNC e seus papéis potenciais em vários processos fisiológicos. Desde então, a caracterização do sNPF se expandiu, revelando seu envolvimento em diversas funções essenciais, como a regulação da alimentação e os ritmos circadianos (Dillen *et al.*, 2013, 2014). Além disso, a ampla distribuição desse neuropeptídeo em várias espécies de insetos sugere uma conservação evolutiva e funcional em diferentes táxons (Fadda *et al.*, 2019b; Nagata *et al.*, 2012).

Por exemplo, no gafanhoto do deserto, *Schistocerca gregária*, foi demonstrado que o sNPF exerce um efeito inibitório sobre o comportamento alimentar (Dilln *et al.*, 2014). Em insetos sociais, como as abelhas melíferas, o aumento artificial de sNPF causou o aumento do apetite visual e sensibilidade para estímulos olfativos (Bestea *et al.*, 2021, 2022). Para além da função na alimentação, foi relatado que sNPF regula a função reprodutiva em órgãos reprodutivos masculinos da larva-da-farinha, *Tenebrio molitor* (Marciniak *et al.*, 2020).

Em *R. prolixus*, um único neuropeptídeo F curto (sNPF) é processado a partir de uma proteína precursora, diferentemente de *Aedes aegypti* (Liesch; Bellani; Vosshall, 2013),

Drosophila melanogaster (Broeck, 2001) e *Apis mellifera* (Hummon *et al.*, 2006), que geram múltiplos sNPFs a partir de um único proneuropeptídeo. No entanto, assim como em outras espécies de artrópodes, a sequência desse neuropeptídeo é altamente conservada, apresentando os seguintes aminoácidos: RxPxLRLRF-amida (Fadda *et al.*, 2019).

O processamento dos proneuropeptídeos ocorre no retículo endoplasmático rugoso (RER) ao longo das vias secretórias, onde as proteínas precursoras sofrem clivagem endoproteolítica em resíduos de aminoácidos básicos, geralmente organizados em pares ou trios. Após essa clivagem, uma carboxipeptidase remove os resíduos resultantes, levando à formação do peptídeo maduro. Caso a extremidade carboxílica contenha uma glicina, ocorre uma modificação pós-traducional na qual essa glicina é removida e o terminal convertido em uma amida pela ação da enzima peptidil-glicina- α -amidação monooxigenase (PAM) (Sossin; Fisher; Scheller, 1989). Essa modificação é essencial para a estabilidade e a atividade biológica do sNPF, garantindo sua funcionalidade nos processos de sinalização neuroendócrina e na regulação do comportamento alimentar.

A atividade fisiológica de neuropeptídeos como o sNPF em *R. prolixus* tem sido associada à regulação da motilidade intestinal e à absorção de nutrientes, destacando sua importância no ciclo de vida do inseto (Gonzalez; Orchard, 2009). Além disso, a identificação de precursores de neuropeptídeos em *R. prolixus* ampliou o conhecimento sobre seu sistema neuroendócrino. Estudos revelaram uma diversidade de neuropeptídeos, incluindo o sNPF, envolvidos na regulação de processos fisiológicos essenciais, como diurese e contração muscular (Ons *et al.*, 2011). A presença de sNPF e de seu receptor em *R. prolixus* sugere uma função conservada entre diferentes espécies de insetos, onde esse neuropeptídeo pode atuar na regulação da homeostase energética e de funções reprodutivas (Orchard *et al.*, 2023). Essa conservação funcional é particularmente relevante para insetos hematófagos, como o *R. prolixus*, devido às altas demandas energéticas associadas à alimentação sanguínea e à subsequente produção de ovos.

Portanto, compreender a interrelação entre os neuropeptídeos, seus receptores e o correceptor olfativo (Orco) é fundamental para elucidar as interações do inseto com o ambiente e seu hospedeiro. Estudos indicam que o sistema sNPF/sNPFR pode influenciar diretamente o comportamento alimentar (Amir *et al.*, 2022; Dillen *et al.*, 2014) e as interações entre o vetor e o hospedeiro (Christ *et al.*, 2017), tornando-se, assim, um alvo promissor para o desenvolvimento de novas estratégias de controle.

2 OBJETIVOS

2.1 OBJETIVO GERAL

Este estudo teve como objetivo estudar o neuropeptídeo F curto (sNPF) e seu receptor (sNPFR) em *Rhodnius prolixus*, analisando a expressão relativa de ambos, além de caracterizar a sequência completa do *sNPFR* e suas características físico-químicas determinando os aminoácidos que fazem parte do sítio de ligação com o sNPF por meio de análises *in silico*.

2.2 OBJETIVOS ESPECÍFICOS

- Determinar, por qPCR, o perfil de expressão dos genes *sNPFR* e *sNPF* em diferentes tecidos de *R. prolixus*;
- Obter a sequência completa do transcrito do receptor sNPFR de *R. prolixus*.
- Caracterizar a proteína codificada pelo *sNPFR*, incluindo predição de domínios funcionais, estrutura tridimensional e análise filogenética.
- Desenvolver um modelo tridimensional do sNPF para ancoramento molecular e análise de interação com o sNPFR;

3 MATERIAIS E MÉTODOS

3.1 CRIAÇÃO E MANUTENÇÃO DE *Rhodnius prolixus*

Os insetos foram mantidos em uma colônia de *R. prolixus*, a 28°C e umidade relativa de 75-85%, respeitando o fotoperíodo de 12 h em fase de luz e 12h em fase de escuridão. Todos os insetos foram alimentados com sangue de coelhos, em intervalos de três semanas. Os protocolos de experimentação animal foram realizados respeitando as diretrizes da Comissão de Avaliação do Uso de Animais em Pesquisa da Universidade Federal do Rio de Janeiro (CAUAP/UFRJ, Processo#01200.001568/2013-87), baseados no Guia do Instituto Nacional de Saúde e Uso de Animais de Laboratório (ISB#309-05377-3). Os funcionários técnicos do biotério dos coelhos seguiram estritas regras de manejo com os animais, respeitando a periodicidade de manipulação de cada indivíduo.

3.2 DISSECÇÃO DOS INSETOS

Foram dissecadas ninfas de 5º estágio (N5) no dia anterior ao repasto sanguíneo (3-5 semanas após a última alimentação). Inicialmente, os insetos foram mantidos por 3 min no congelador para permanecerem imóveis. Com o auxílio de pinças, foram retiradas probóscides, pares de antenas e patas. Com uma das pinças, o inseto foi imobilizado na posição ventral e, a partir de incisões nas laterais do abdômen com tesoura cirúrgica curva, o conteúdo abdominal foi exposto. Em seguida, foram retirados o intestino (anterior, médio e posterior), túbulos de Malpighi e corpo gorduroso. Todo o conteúdo do intestino foi retirado com a pinça. Todos os tecidos foram lavados em solução salina 0,9% gelada.

3.3 DESENHO DOS OLIGONUCLEOTÍDEOS INICIADORES

Os oligonucleotídeos iniciadores (Tabela 1) foram desenhados utilizando-se a ferramenta online Primer3Plus (<https://www.primer3plus.com/>), segundo critérios de qualidade descritos em THORTON & BASU (2011). A formação de estruturas secundárias, temperatura de *melting* (T_m) e conteúdo de GC (GC%) foram previstas com a ferramenta *OligoAnalyzer*TM Tool (<https://www.idtdna.com/pages/tools/oligoanalyzer>).

Tabela 1 - Oligonucleotídeos iniciadores utilizados em reações de PCR

Identificação	Iniciador	Sequência (5'-3')	Produto de PCR (pb)	T_m (°C)
Primer A	Senso	CGTAGTCGGTAGGAATCGTGCT	843	60
Primer B	Anti-senso	TCGTTTCAGCCAAGCGTATAA		
sNPF	Senso	CTCCTGGGCTCACAAAGTT	91	56
	Anti-senso	CTGAAGGAAGGTGGGATCAT		
sNPFR	Senso	CGTAGTCGGTAGGAATCGTGCT	105	60
	Anti-senso	GTGTAAATGGTACAGCCAATGTGC		
Rpro-R18S_F	Senso	TCGGCCAACAAAAGTACACA	104	60
	Anti-senso	TGTCGGTGTAAGTGGCATGT		

3.4 EXTRAÇÃO DE RNA TOTAL

O processo de extração do RNA total foi realizado a partir dos tecidos coletados durante a dissecação dos insetos. Inicialmente, todos os tubos passaram por uma etapa de congelamento em nitrogênio líquido e maceração com pistilos autoclavados. Em seguida, as amostras foram homogeneizadas em TRIzol (Invitrogen, Carlsbad, EUA) (100 μ L/mg de amostra) e incubadas por 5 min em temperatura ambiente para completa dissociação do complexo de nucleoproteínas. Logo após, foi adicionada uma alíquota de clorofórmio (20% do volume de TRIzol inicial), seguida de agitação em vórtex por 30 seg e incubação por 5 min em temperatura ambiente. As amostras foram centrifugadas por 15 min, a 12000 g, a 4 °C. Nessa etapa, ocorreu a separação das fases aquosa, menos densa; intermediária e compacta, contendo DNA; e orgânica, mais densa, contendo fenol-clorofórmio.

A fase aquosa, na qual o RNA total está disperso, foi coletada em um novo tubo de 1,5 mL, cuidadosamente, para evitar contaminação com as outras fases da mistura. A esse novo tubo foi adicionado isopropanol em volume correspondente a 1,5 vezes o volume recuperado da fase aquosa. Em seguida, as amostras foram incubadas por 10 min a -20 °C e, após esse período, centrifugadas por 10 min em velocidade máxima (g ou rpm), a 4 °C. O RNA total foi precipitado em forma de precipitado branco e gelatinoso no fundo dos tubos. O sobrenadante foi descartado invertendo-se o tubo sobre um pedaço de papel absorvente.

Para a lavagem do precipitado de RNA, foi adicionado etanol 75% gelado na proporção de 1:1 em relação ao volume de TRIzol utilizado na primeira etapa, seguido de breve agitação em vórtex. Em seguida, realizou-se uma centrifugação por 5 a 10 min a 12.000 g, a 4 °C, seguida do descarte do sobrenadante. Essa etapa foi repetida pelo menos duas vezes, seguida de uma última lavagem com etanol >85% gelado. Por fim, as amostras foram deixadas para secar em ar atmosférico dentro do fluxo laminar por até 10 min. Finalmente, o RNA foi reconstituído em 30 µL de água livre de RNase e DNase.

3.5 DETERMINAÇÃO DE CONCENTRAÇÃO DE ÁCIDOS NUCLEICOS

A quantificação de ácidos nucleicos foi realizada em espectrofotômetro *NanoDrop One* (ThermoFisher, Waltham, EUA), a 260 nm, utilizando-se 1 µL das amostras para a leitura do branco (água livre de nucleases) e da concentração das amostras. A pureza das amostras foi determinada pelas razões de absorbância em 260/280 nm e 260/230 nm, utilizadas para detectar a presença de proteínas e compostos fenólicos contaminantes, respectivamente. A razão A260/A280 para DNA puro foi considerada em torno de 1,8, enquanto para RNA puro, em torno de 2,0. Já a razão A260/230 foi mantida entre 2,0 e 2,2.

3.6 TRATAMENTO DE RNA COM DNASE I

A fim de eliminar possíveis resíduos de DNA genômico, a amostra de RNA foi tratada com *DNase I, RNase free* (1 U/µL) (Promega, Madison, EUA). Inicialmente, foi preparada uma

mistura contendo 1 μL de tampão de reação (10X Reaction Buffer com MgCl_2) e 1 μL da enzima DNase I para cada uma das amostras de RNA, seguido da distribuição de 2 μL dessa mistura em cada micro tubo. Posteriormente, foi adicionado 1 μg de RNA, respeitando-se o volume mínimo de 1 μL e máximo de 8 μL . Quando necessário, o volume da mistura foi ajustado até 10 μL com água livre de nucleases. Em seguida, os tubos foram incubados por 30 min a 37 °C. Para interromper a reação, foi adicionado 1 μL de EDTA 50 mM e os tubos incubados por 10 min a 65 °C.

3.7 SÍNTESE DA FITA DE DNA COMPLEMENTAR (cDNA)

A primeira fita de DNA complementar foi preparada utilizando-se o kit *GoScriptTM Reverse Transcription Mix, Random Primers* (Promega, Madison, EUA). A cada tubo contendo 1 μg de RNA tratado com DNase I, foram adicionados 4 μL da solução tampão (GoScript Buffer, Random primers), 2 μL de enzima (GoScript Enzyme Mix) e 4 μL de água livre de nucleases. Por fim, a reação foi incubada em termociclador segundo as etapas de anelamento a 25 °C/5 minutos; extensão a 42 °C/60 min; e inativação a 70 °C/15 min.

3.8 REAÇÃO DE AMPLIFICAÇÃO EM CADEIA DA POLIMERASE

A amplificação dos genes de interesse foi realizada a partir das amostras de cDNA preparadas anteriormente, utilizando o gene *Rpro-R18S* como referência (MAJEROWICZ *et al.*, 2011). Para as reações semiquantitativas, foi utilizado o kit *GoTaq[®] Hot Start Master Mix* (Promega, Madison, EUA), de acordo com o protocolo do fabricante, em 35 ciclos para genes de interesse e em 25 ciclos para o gene de referência (*Rpro-R18S*), em termociclador *Veriti[®] Thermal Cycler - 96 well* (Applied Biosystems, Waltham, EUA).

Para as reações de qPCR, foi utilizado o kit *SYBR[®] Green Master Mix* (Bio-Rad, Hercules, EUA), empregando triplicatas biológicas e técnicas, durante 35 ciclos e temperatura de 60 °C. O equipamento utilizado foi um termociclador *CFX96 Touch Real-Time PCR Detection System* (Bio-Rad, Hercules, EUA). As informações de incubação (tempo e

temperatura) estão expostas na Tabela 2. A expressão relativa dos genes foi determinada pelo método $2^{-\Delta\Delta C_t}$ (Livak; Schmittgen, 2001).

Tabela 2 – Protocolo de incubação das reações de amplificação em cadeia da polimerase

	Temperatura (°C)	Tempo (PCR)	Tempo (RT-qPCR)	#ciclos
Desnaturação inicial	95	2 min	10 min	1
Desnaturação	95	30 s	15 s	1
Anelamento	T_m	30 s	15 s	25 a 40*
Extensão	72	1 min	1 min	1

*Número de ciclos: 25 a 35 ciclos para PCR e 35 a 40 ciclos para RT-qPCR.

3.9 ANÁLISE POR ELETROFORESE EM GEL DE AGAROSE

Os produtos de PCR foram submetidos à eletroforese em gel de agarose 1% em tampão TAE pH 8 (Tris-Acetato 40 mM, EDTA 1 mM) e corados com *GelRed* (Biotium, Fremont, EUA). Cada gel foi registrado em fotodocumentador *DNR MiniBIS Pro-Bio-Imaging Systems* (Bio-America Inc., Fort Dodge, EUA).

3.10 PURIFICAÇÃO DOS PRODUTOS DE PCR

Após avaliação em eletroforese os produtos de PCR selecionados como inserto para a clonagem molecular foram purificados utilizando-se o kit *QIAquick PCR Purification* (QIAGEN, Hilden, Alemanha). O produto foi isolado a partir do gel de agarose, utilizando uma lâmina. A cada volume de gel, foram adicionados 5 volumes de solução tampão fornecida pelo fabricante e a mistura foi incubada a 50 °C por 10 min, até que todo o gel estivesse dissolvido. Em seguida, as etapas de ligação, lavagem e eluição foram realizadas em centrífuga a 13.000 rpm, em temperatura ambiente.

3.11 CLONAGEM MOLECULAR E SEQUENCIAMENTO DO sNPFR

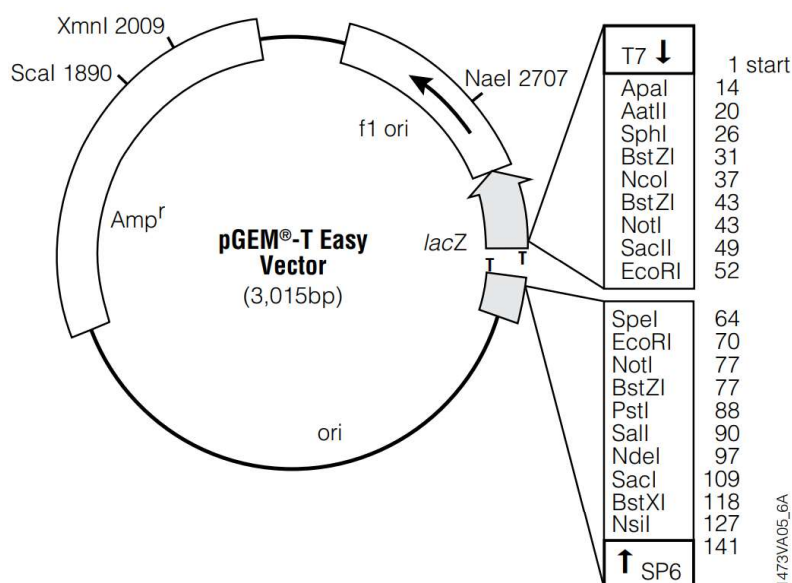
O produto de PCR purificado do sNPFR (RPRC002266 – RPRC002269, ver Tabela 1) foi utilizado como inserto na reação de ligação com o vetor plasmidial *pGEM®-T Easy* (Promega, Madison, EUA) (Figura 4), que confere resistência à ampicilina às células transformadas. A reação foi realizada misturando 5 µL da solução tampão da enzima T4 ligase, 3 µL de inserto (6,9 ng/µL), 0,6 µL do vetor (50 ng/µL) e 1 µL da T4 ligase (3 unidades Weiss/µL), respeitando a razão 2:1 (inserto:vetor), calculados a partir da equação 1. Em seguida, a reação foi incubada a 4° C/16 h, seguida de 16 °C/1 h.

Equação 1 – Cálculo da razão molar de inserto:vetor e da quantidade do inserto (DNA)

$$quantidade\ (ng)\ do\ inserto = \frac{quantidade\ (ng)\ do\ vetor \times tamanho\ (kb)\ do\ inserto}{tamanho\ (kb)\ do\ vetor} \times \frac{inserto}{vetor}$$

Células quimiocompetentes de *Escherichia coli* JM109 Competent Cells (Promega, Madison, EUA) foram transformadas com o produto da reação de ligação por meio do método do choque térmico. Inicialmente, 1 µL do produto da ligação foi misturado à 5 µL das células quimiocompetentes. Em seguida, a mistura foi incubada por 30 min no gelo, seguida de aquecimento a 42 °C por 30 seg, e, por fim, resfriada no gelo por mais 30 seg. Aos tubos, foram adicionados 250 µL de meio SOC em temperatura ambiente. As células foram, então, incubadas por 1 h 30 min sob agitação a 150 rpm. Após esse período, 250 µL das células transformadas foram espalhadas em placas contendo meio LB sólido com IPTG, ampicilina e X-gal, e incubadas em estufa bacteriológica por 16 h a 37 °C.

Figura 4 – Mapa do vetor plasmidial pGEM®-T Easy Vector com os pontos de referência.



Fonte: Promega.

A fim de confirmar o sucesso da ligação, as colônias brancas foram submetidas à reação de PCR. Para isso, treze colônias selecionadas foram tocadas com ponteiros descartáveis esterilizados, e misturadas em tubos individuais contendo 10 µL da mistura reacional de PCR, com movimentos de cima para baixo, utilizando uma pipeta. Os oligonucleotídeos utilizados foram T7 (5'–TAATACGACTCACTATAGGG–3') e SP6 (5'–ATTTAGGTGACACTATAGAA–3') por 35 ciclos, com temperatura de anelamento de 52 °C. Os produtos de PCR foram analisados por eletroforese em gel de agarose 1%, corado com *GelRed* (Biotium, Fremont, EUA), e visualizados no aparelho transiluminador UV *DNR MiniBIS Pro-Bio-Imaging Systems* (Bio-America Inc., Fort Dodge, EUA).

Em sequência, as ponteiros foram dispensadas em um tubo Falcon contendo 3 mL de meio LB líquido e ampicilina (100 mg/mL) e incubados a 37 °C por, aproximadamente, 19 h, para posterior purificação e isolamento do DNA plasmidial utilizando o kit *QIAprep Spin Miniprep* (QIAGEN, Hilden, Alemanha). Após o crescimento, o tubo que apresentou turbidez foi centrifugado por 3 minutos a 7000 x g, em temperatura ambiente. O sobrenadante foi descartado e o precipitado foi ressuspenso em 250 µL da solução tampão P1 contendo RNase A, transferido para tubo de microcentrifuga e homogeneizado com agitação vigorosa.

Em seguida, foram adicionados 250 µL do tampão P2, e a mistura invertida de 4 a 6 vezes, a fim de completar a lise celular. 350 µL de tampão N3 foram adicionados e a mistura foi invertida até que a solução se tornasse turva. Após essa etapa, a mistura foi centrifugada por

10 min a 13000 rpm, em temperatura ambiente, e 800 µL do sobrenadante foram transferidos para uma coluna fornecida pelo fabricante. Após centrifugação por 1 minuto, a fração eluída foi descartada. A coluna foi lavada com 500 µL de tampão BP para eliminar quaisquer traços de nucleases e de carboidratos, necessário quando utilizando células da série JM.

Logo após, 750 µL do tampão PE foram adicionados e a coluna foi centrifugada a 13000 rpm/1 minuto. Novamente, a fração eluída foi descartada e a coluna centrifugada em velocidade máxima por mais 1 minuto. Por fim, a coluna foi posicionada em um novo tubo limpo para a eluição do DNA plasmidial. Foram adicionados 50 µL de água livre de nucleases, deixada em repouso por 1 minuto e centrifugado por 1 minuto, em temperatura ambiente, a 13000 rpm. A dosagem da concentração foi realizada por espectrometria em *NanoDrop One* (ThermoFisher, Waltham, EUA).

O clone selecionado foi sequenciado na Plataforma de Sequenciamento de DNA (PSEQDNA-UFRJ) do Instituto de Biofísica Carlos Chagas Filho, UFRJ, utilizando o oligonucleotídeo iniciador anti-senso pUC/M13 (5' – CAGGGTTTTCCAGTCACGAC – 3'), pelo método da eletroforese capilar de análise de fragmentos (Sanger). O cromatograma foi analisado no programa UGENE 51.0 (Rússia).

3.12 ANÁLISE FILOGENÉTICA

Após a realização do blastp (banco de dados: non redundant), foram selecionadas sequências de aminoácidos similares, provenientes de outros insetos, com base em valores mínimos de 60% de identidade, 80% de cobertura e *e-value* de 1e-115. Mantiveram-se os critérios de identidade, com mínimo de 70% de cobertura e *e-value* 1e-15, para a realização do blastp com a proteína precursora de sNPF.

A identificação de possíveis peptídeos sinalizadores foi realizada com o software online *SignalP 6.0 Server* (<https://services.healthtech.dtu.dk/services/SignalP-6.0/>), e esses peptídeos foram retirados da sequência de Rpro-sNPF antes do alinhamento múltiplo de sequências (MSA), que foi realizado no *Clustal Omega* (<https://www.ebi.ac.uk/jdispatcher/msa/clustalo>). As relações filogenéticas foram feitas por meio do software MEGA 12 (<https://www.megasoftware.net/>), por meio do método *Neighbor-Joining* com 1000 repetições

bootstrap. As árvores filogenéticas foram editadas na ferramenta online *iTOL* (<https://itol.embl.de/>) e os alinhamentos de sequências editados na ferramenta Jalview (<https://www.jalview.org/>).

3.13 ANÁLISE *IN SILICO*

A sequência prevista para a proteína Rpro-sNPFR foi obtida pela ferramenta de tradução *in silico* ExPasy (<https://web.expasy.org/translate/>). As propriedades físico-químicas, como ponto isoelétrico e peso molecular, foram previstas pela ferramenta ProtParam (<https://web.expasy.org/protparam/>). O peptídeo maduro Rpro-sNPF foi modelado na ferramenta PEP-FOLD 3 (<https://bioserv.rpbs.univ-paris-diderot.fr/services/PEP-FOLD3/>). A topologia da proteína Rpro-sNPFR foi prevista na ferramenta online DeepTMHMM (<https://dtu.biolib.com/DeepTMHMM>). A modelagem de estrutura tridimensional de Rpro-sNPFR foi realizada pelo método *ab initio* utilizando o AlphaFold2 (ColabFold2) (MIRDITA *et al.*, 2022) e o modelo gerado pela modelagem por homologia foi realizado no SWISS-MODEL (<https://swissmodel.expasy.org/>) (WATERHOUSE *et al.*, 2018). As estruturas cristalizadas de NPY2-R ligado a NPY (PDB: 7YOO), de *Homo sapiens*, foi obtida no banco *Protein Data Bank* (PDB) (<https://www.rcsb.org/>). A ancoragem (*docking*) molecular foi realizada na ferramenta online DockThor (<https://dockthor.lncc.br/v2/>). As imagens das estruturas foram geradas no PyMol (<https://www.pymol.org/>).

4 RESULTADOS E DISCUSSÃO

4.1 QUANTIFICAÇÃO DE TRANSCRITOS EM CONDIÇÃO DE JEJUM

Com o objetivo de identificar os tecidos que apresentam expressão dos genes *sNPF* e *sNPFR*, foi traçado o perfil de expressão por meio de PCR semiquantitativo em antenas, cérebro, intestino, corpo gorduroso e túbulos de Malpighi. O gene ribossomal R18S (Majerowicz *et al.*, 2011) foi utilizado como controle para validação da integridade dos RNAs. Por estar presente em todos os tecidos (Figuras 5 e 6), o *Rpro-R18S* confirmou a qualidade das amostras e serviu como referência para a análise comparativa da expressão dos genes de interesse.

Esses resultados indicaram que o transcrito de *sNPF* foi detectado no cérebro e corpo gorduroso (Figura 5). Já o gene de *sNPFR* (iniciadores: RPRC002266) mostrou expressão no cérebro, intestino, corpo gorduroso e túbulos de Malpighi (Figura 6).

Figura 5 – Perfil de expressão dos genes *Rpro-R18S* e *Rpro-sNPF*. O RNA total foi extraído de ninfas de 5º estágio, tratados com DNase I e utilizado na reação de transcrição reversa para síntese de cDNA. As reações de PCR foram realizadas utilizando oligonucleotídeos iniciadores específicos para cada gene. O gene ribossomal R18S foi utilizado como controle da integridade de cDNA.

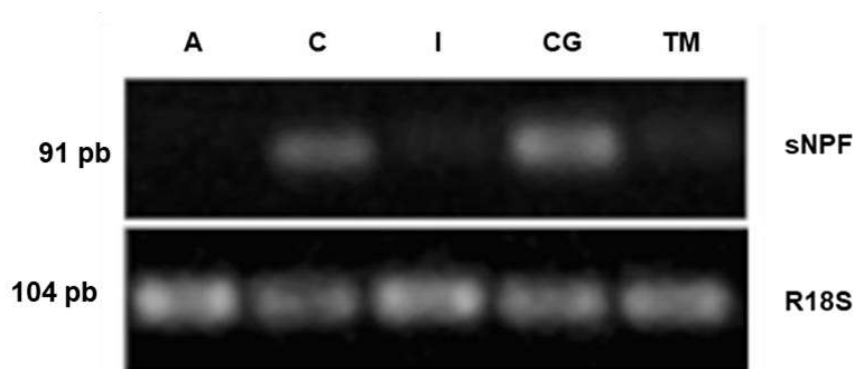
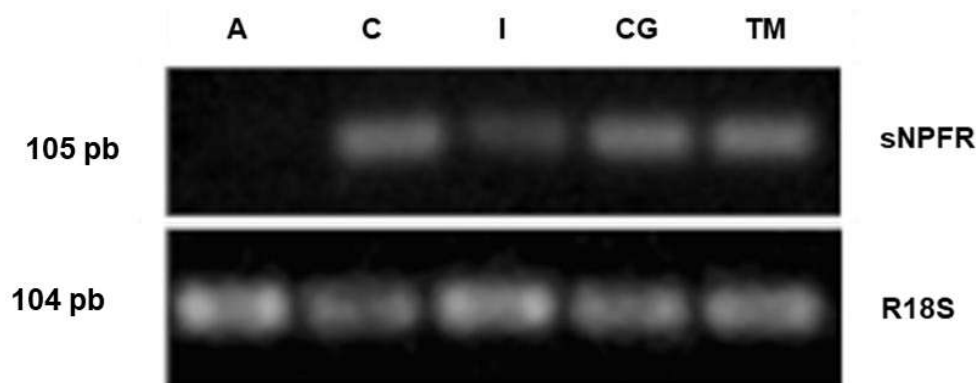
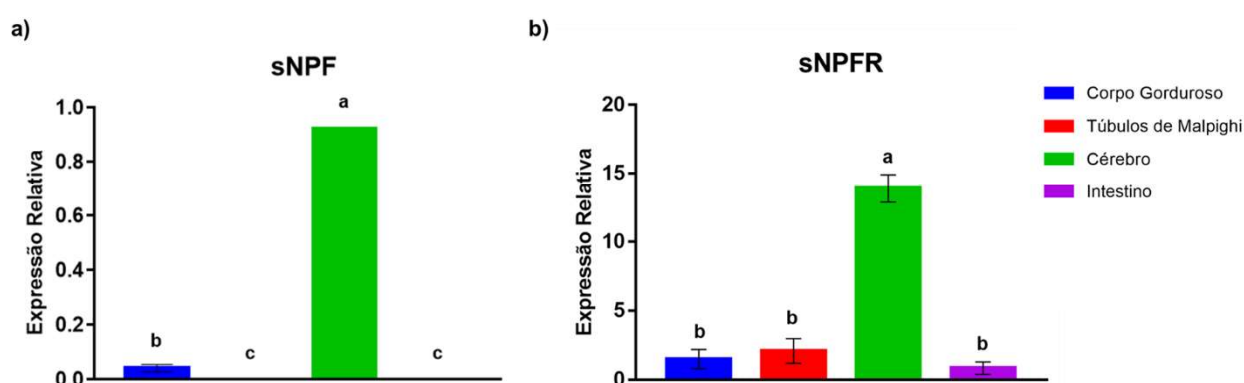


Figura 6 - Perfil de expressão dos genes *Rpro-R18S* e *Rpro-sNPFR*. O RNA total foi extraído de ninfas de 5º estágio, tratados com DNase I e utilizado na reação de transcrição reversa para síntese de cDNA. As reações de PCR foram realizadas utilizando oligonucleotídeos iniciadores específicos para cada gene. O gene ribossomal R18S foi utilizado como controle da integridade de cDNA.



Em seguida, foi realizada a análise quantitativa dos transcritos por qPCR, empregando o método do Ct comparativo ($2^{-\Delta\Delta C_t}$) (Livak; Schmittgen, 2001) utilizando o gene *Rpro-R18S* como referência. Como esperado, o cérebro apresentou a maior expressão relativa de sNPF (Figura 7a) e sNPFR (Figura 7b).

Figura 7 - Perfil de expressão relativa do gene codificante sNPF e sNPFR em *Rhodnius prolixus*. Os ensaios de qPCR foram realizados utilizando cDNA de insetos adultos alimentados. **a)** Expressão relativa de sNPF, representados em relação ao valor obtido para cérebro (1,0). **b)** Expressão relativa de sNPFR, representados em relação ao valor obtido para intestino (1,0). Os valores de Ct obtidos foram normalizados em relação ao gene de referência *Rpro-R18S* e quantificados pelo método comparativo ($2^{-\Delta\Delta C_t}$). Os resultados estão apresentados na forma de média $\pm$ desvio padrão para 3 determinações independentes. A comparação entre os grupos foi realizada utilizando ANOVA. Letras diferentes indicam diferenças estatisticamente significantes ($p < 0,05$).



Estudos anteriores indicam que o *sNPF* é expresso em diferentes populações neuronais no sistema nervoso central de *D. melanogaster* (Nässel *et al.*, 2008), sendo a maioria composta por interneurônios dos *mushroom bodies*, estruturas associadas ao processamento de

informações sensoriais (Rehan; Bulova; O'Donnell, 2015). O tamanho dessas estruturas reflete a capacidade cognitiva dos insetos, influenciando desde a retenção de memória até a resposta aos desafios ambientais e sociais (Watanabe; Kubo, 2015).

Além do seu papel no sistema nervoso central, o sNPF também está envolvido na regulação metabólica e comportamental. Um dos papéis principais de sNPF no corpo gorduroso, que é órgão principal para estoque de energia em insetos, está relacionado aos seus efeitos regulatórios na homeostase energética (Lee *et al.*, 2004).

No corpo gorduroso de *D. melanogaster*, a expressão de sNPFR nas células produtoras de insulina (CPI) regula a sinalização dos peptídeos semelhantes à insulina (ILPs), revelando mecanismos pelos quais o sNPF participa dos processos fisiológicos relacionados à alimentação, metabolismo e desenvolvimento (Fadda *et al.*, 2019). Os hormônios derivados desse órgão atuam em conjunto com neuropeptídeos como o sNPF para modular não apenas processos metabólicos, mas também respostas comportamentais, como a ingestão de alimentos (Lee *et al.*, 2004).

Em mamíferos, o neuropeptídeo Y sofre um decréscimo em neurônios orexigênicos no estado de saciedade via sinalização da insulina (Mayer; Belsham, 2009). De maneira análoga, em insetos, o neuropeptídeo F curto (sNPF) desempenha um papel fundamental na regulação do estado alimentar e na modulação da busca por alimento. A expressão de sNPF nesses neurônios sugere que esse peptídeo pode atuar como um modulador dos neurônios sensoriais olfativos (Nässel *et al.*, 2008).

Os neurônios sensoriais olfativos abrigam os receptores olfativos e o correceptor olfativo (Leal, 2013). Nesse contexto, foi observado que a sensibilidade olfativa é dependente da sinalização de sNPF, mediada pela produção de ILPs (Root *et al.*, 2011). Esses achados sugerem que, após o processamento da proteína precursora no sistema nervoso central, o neuropeptídeo maduro é transportado por meio da hemolinfa até diferentes tecidos do inseto, como o corpo gorduroso, onde o receptor sNPFR está presente, em células produtoras de insulina (Lee *et al.*, 2008).

4.2 SEQUÊNCIA CODIFICANTE DO RECEPTOR DE sNPF

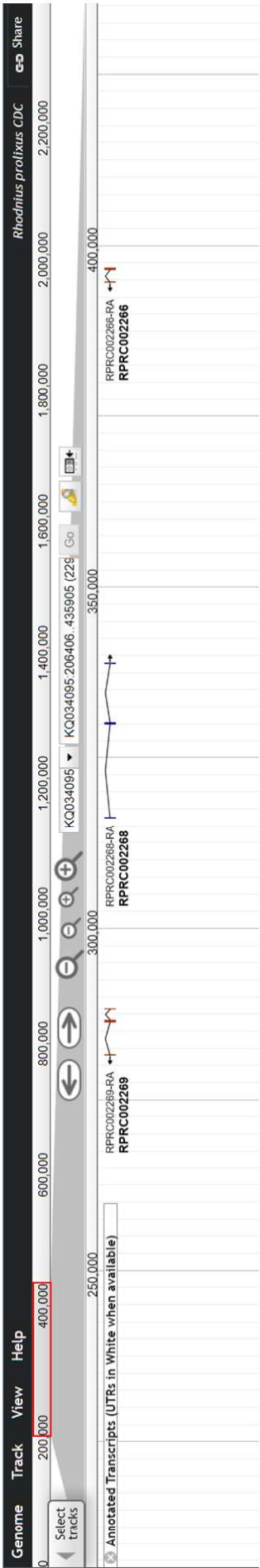
No contexto dos artrópodes, os receptores de neuropeptídeos são frequentemente identificados juntamente com seu neuropeptídeo correspondente por meio de análises genômicas e transcriptômicas comparativas (Caers *et al.*, 2012; Marciniak *et al.*, 2022).

Em *R. prolixus*, a análise transcriptômica do sistema nervoso central revelou a presença de receptores acoplados à proteína G (Ons *et al.*, 2016). Entre eles, destaca-se o receptor sNPFR, cujo provável ligante endógeno é o neuropeptídeo F curto (sNPF). No banco de dados do genoma de *R. prolixus*, três transcritos foram anotados como possíveis componentes do sNPFR: RPRC002266, RPRC002268 e RPRC002269 (Ons *et al.*, 2016). Diferentemente de outros insetos, que possuem apenas uma sequência codificante para o sNPFR (Mertens *et al.*, 2002), a presença de múltiplos transcritos em *R. prolixus* sugere uma complexidade única.

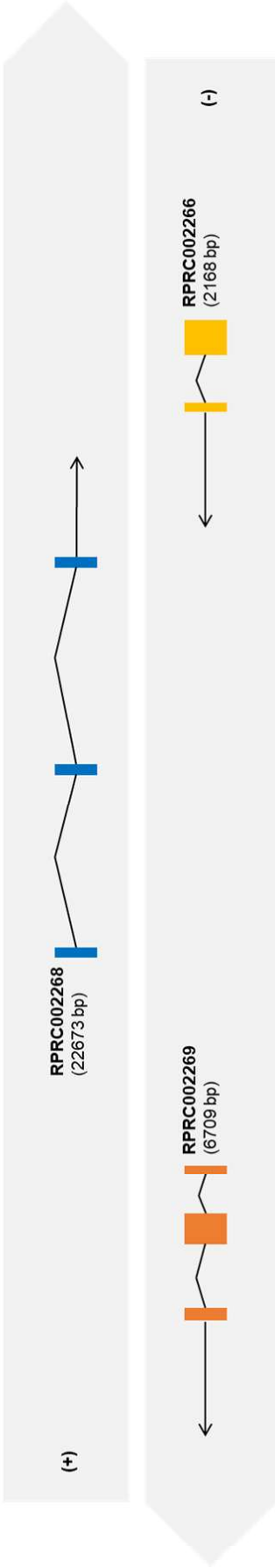
Os transcritos estão organizados em sequência na mesma região genômica, com RPRC002266 e RPRC002269 anotados na fita reversa e RPRC002268 na fita direta (Figura 8a). Dentre eles, apenas RPRC002266 possui dois exons, enquanto os demais apresentam três (Figura 8b). Esses transcritos codificam três proteínas previstas de 113, 118 e 130 aminoácidos, correspondentes a RPRC002266, RPRC002268 e RPRC002269, respectivamente.

Figura 8 – Visualização do supercontig KQ034095. a) Imagem do visualizador JBrowse, aproximada na localização de 391486 a 399080 (7.59 Kb), onde se encontram os transcritos anotados RPRC002266 (fita reversa), RPRC002268 (fita reversa), RPRC002269 (fita reversa) e RPRC002266 (fita reversa). b) Esquema ilustrado representando a posição dos transcritos no genoma. Sinal (+) indica fita direta, sinal (-) indica a fita reversa. Retângulos sólidos representam sequências codificantes (exons), linhas sólidas pretas indicam sequências não codificantes (introns).

a)



b)



Em todos dos insetos nos quais esse receptor foi estudado até o momento, há um único gene que codifica o sNPFR (Amir *et al.*, 2022; Li *et al.*, 2022; Liu *et al.*, 2021). No entanto, em *R. prolixus*, foram identificados três transcritos associados a esse receptor (Ons *et al.*, 2016). Para investigar a organização desse gene em *R. prolixus*, foi realizado um alinhamento entre as proteínas previstas (RPRC002266, RPRC002268 e RPRC002269) com os receptores sNPFR de *Anopheles gambiae* (575 aminoácidos) e *Aedes aegypti* (529 aminoácidos) (Christ *et al.*, 2018), considerando que ambos são vetores de doenças e possuem receptores com tamanho semelhante.

Ao comparar as sequências do sNPFR de *A. gambiae* e *A. aegypti* com o transcrito RPRC002266, observou-se uma maior identidade na região inicial das sequências dos mosquitos (Figura 9), correspondendo às posições 94 a 157 em *A. gambiae* e 63 a 126 em *A. aegypti*. Para o transcrito RPRC002268, a correspondência foi identificada na região intermediária, entre as posições 160 a 278 em *A. gambiae*, e 129 a 246 em *A. aegypti* (Figura 10). Por fim, os aminoácidos mais conservados do RPRC002269 foram localizados entre as posições 268 a 393 de *A. gambiae* e 236 a 261 em *A. aegypti* (Figura 11).

Figura 9 – Alinhamento de seqüências de aminoácidos das proteínas previstas RPRC002266 com sNPF de *Anopheles gambiae* (ABD96049.1) e *Aedes aegypti* (AVR59278.1). À esquerda, o identificador das espécies analisadas e os limites das seqüências que foram utilizados no alinhamento. À direita e à esquerda dos aminoácidos, está representado as suas posições. As barras amarelo/marrom sob as seqüências indicam o grau de conservação dos aminoácidos. O retângulo vermelho representa as porções onde se concentram o maior grau conservação.

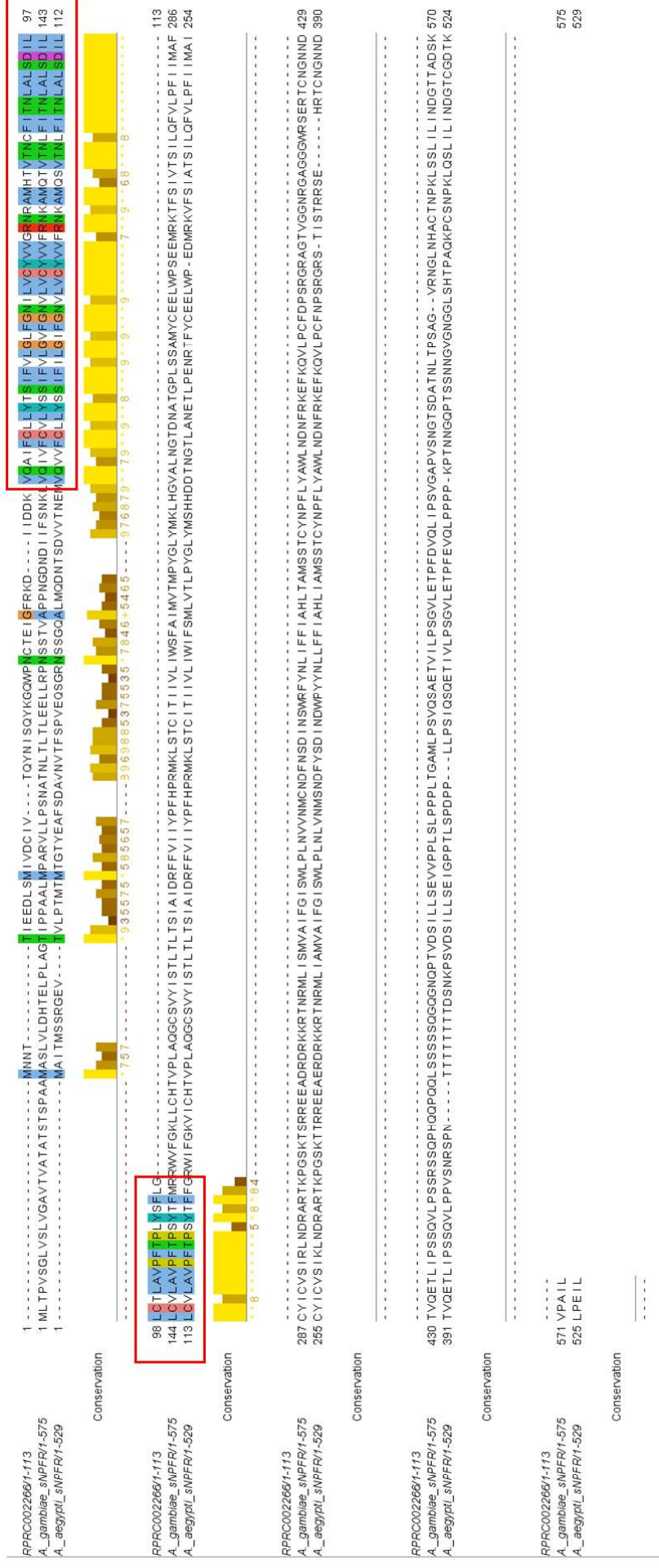


Figura 10 - Alinhamento de seqüências de aminoácidos das proteínas previstas RPRC002268 com sNPFR de *Anopheles gambiae* (ABD96049.1) e *Aedes aegypti* (AVR59278.1). À esquerda, o identificador das espécies analisadas e os limites das seqüências que foram utilizados no alinhamento. À direita e à esquerda dos aminoácidos, está representado as suas posições. As barras amarelo/marrom sob as seqüências indicam o grau de conservação dos aminoácidos. O retângulo vermelho representa as porções onde se concentram o maior grau conservação

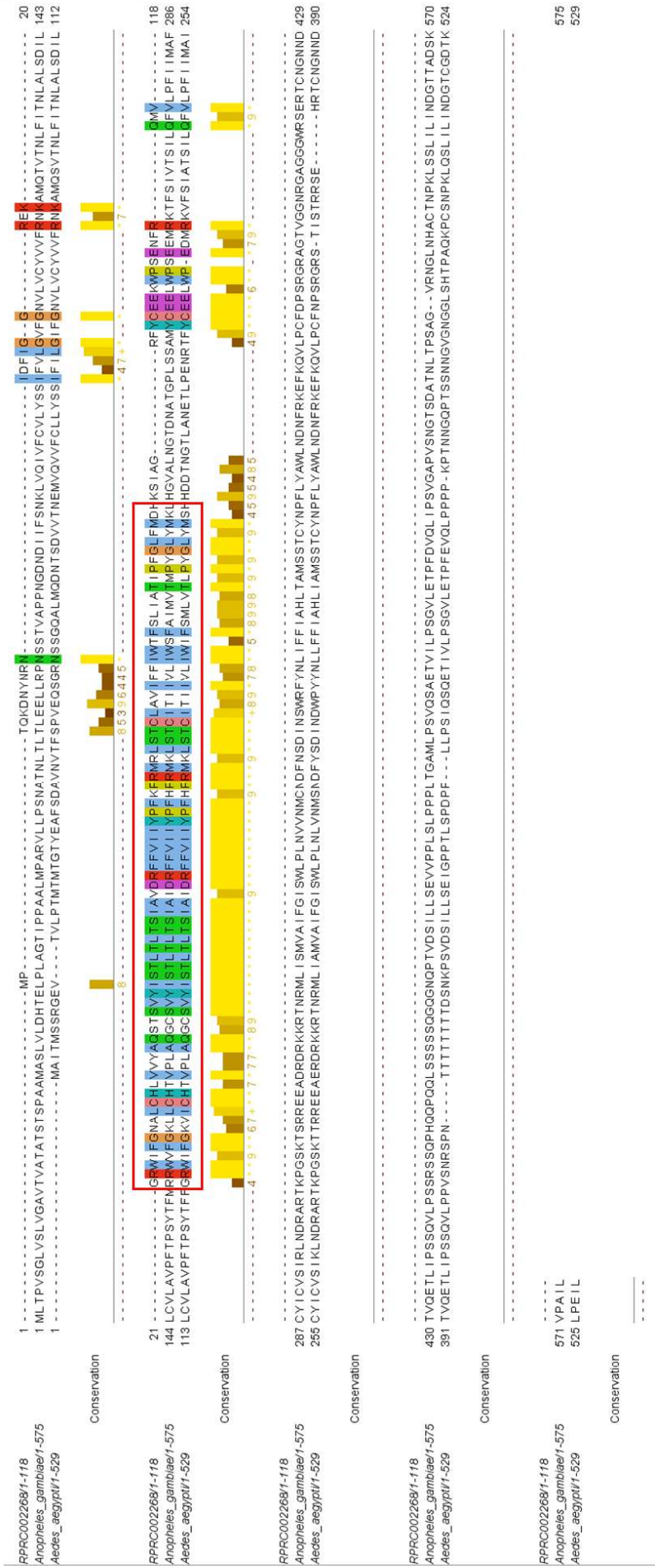
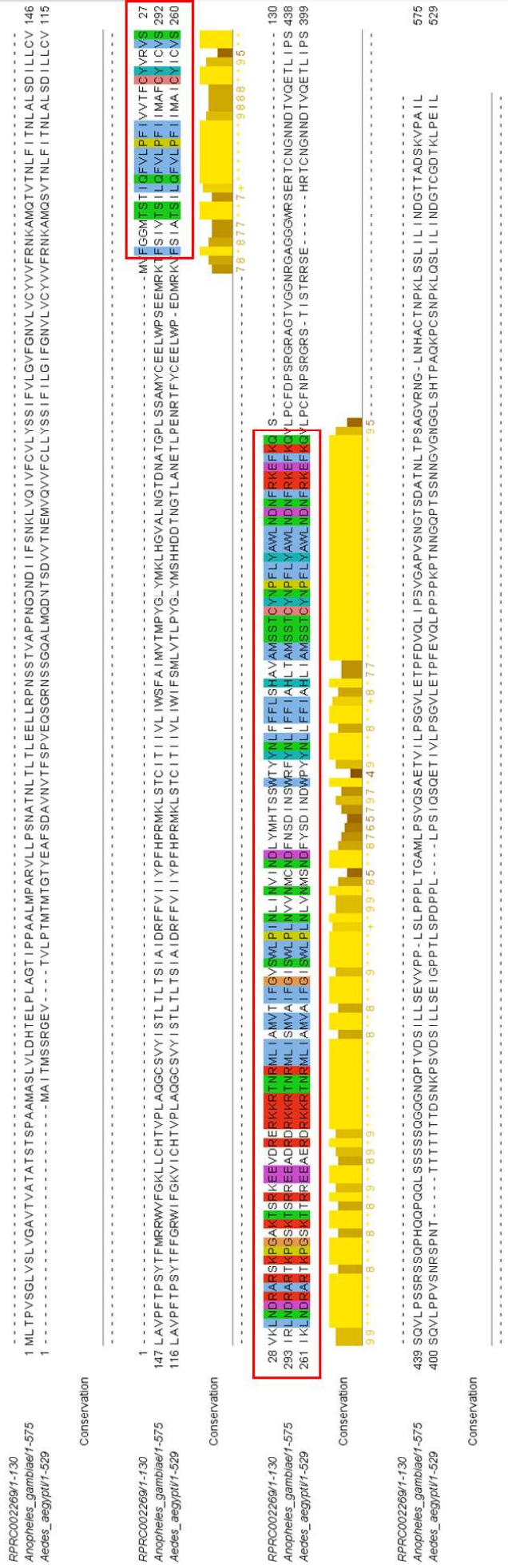


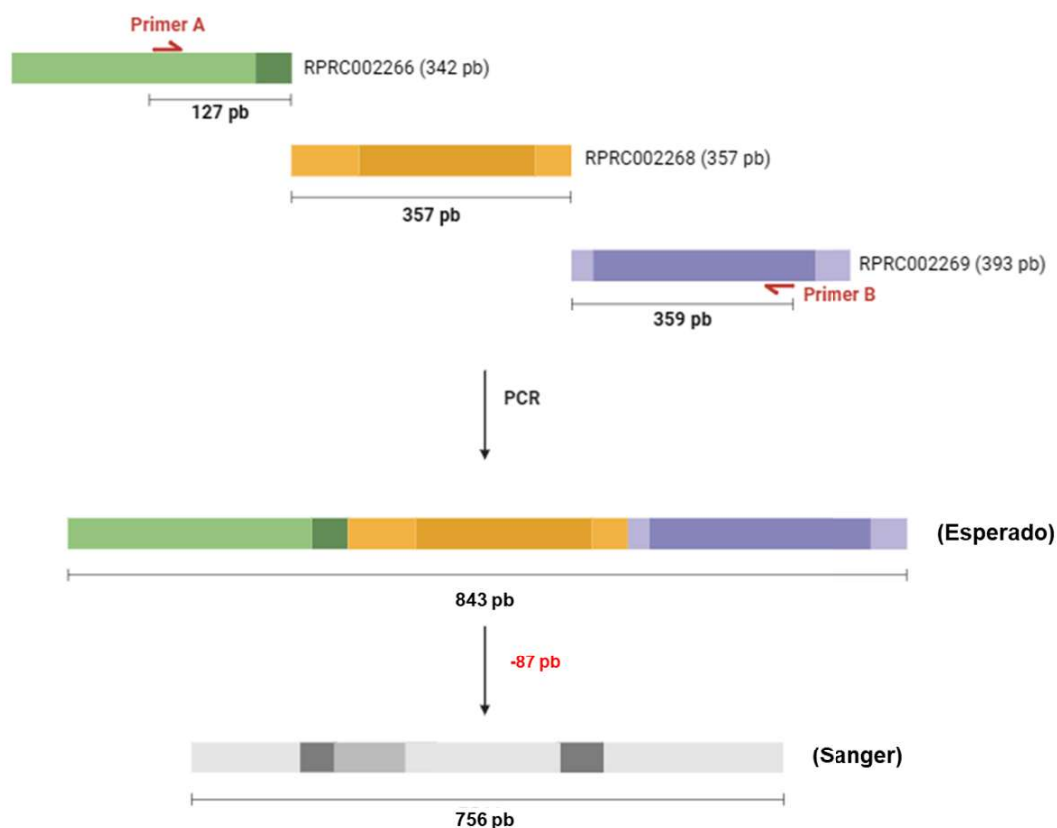
Figura 11 - Alinhamento de seqüências de aminoácidos das proteínas previstas RPRC002269 com sNPFR de *Anopheles gambiae* (ABD96049.1) e *Aedes aegypti* (AVR59278.1). À esquerda, o identificador das espécies analisadas e os limites das seqüências que foram utilizados no alinhamento. À direita e à esquerda dos aminoácidos, está representado as suas posições. As barras amarelo/marrom sob as seqüências indicam o grau de conservação dos aminoácidos. O retângulo vermelho representa as porções onde se concentram o maior grau conservação



Apesar de serem três transcritos, a comparação com as sequências dos mosquitos sugere que eles se organizam como parte de um único gene. Assim, a proteína sNPFR em *R. prolixus* parece ser composta pelo transcrito RPRC002266 na região inicial, RPRC002268 na porção intermediária e RPRC002269 na região final, indicando que esses três transcritos podem ser processados em uma única proteína.

A partir dessa análise, foi possível definir os oligonucleotídeos iniciadores utilizados para elucidar a sequência completa do quadro de leitura aberto (ORF) da proteína sNPFR. Com base na posição desses oligonucleotídeos (Figura 12) e na organização prevista dos transcritos, estimava-se a amplificação de um fragmento de 843 pb, contendo os exons correspondentes a 127 pb do RPRC002266, 357 pb de RPRC002268 e 359 pb do RPRC002269 (Figura 12). No entanto, o sequenciamento revelou um fragmento de 756 pb, indicando que 87 pb não estavam presentes no RNA maduro.

Figura 12 – Desenho esquemático que representa a amplificação do fragmento esperado. Cada transcrito está representado por uma cor e suas extensões (pb – pares de base) e identificadores VectorBase estão anotados à direita dos *exons*, que são representados por retângulos sólidos coloridos. Setas vermelhas indicam a posição dos oligonucleotídeos iniciadores utilizados na reação de amplificação (Primer A e Primer B). Abaixo, o fragmento esperado de 843 pb e o fragmento de 756 pb obtido no sequenciamento.

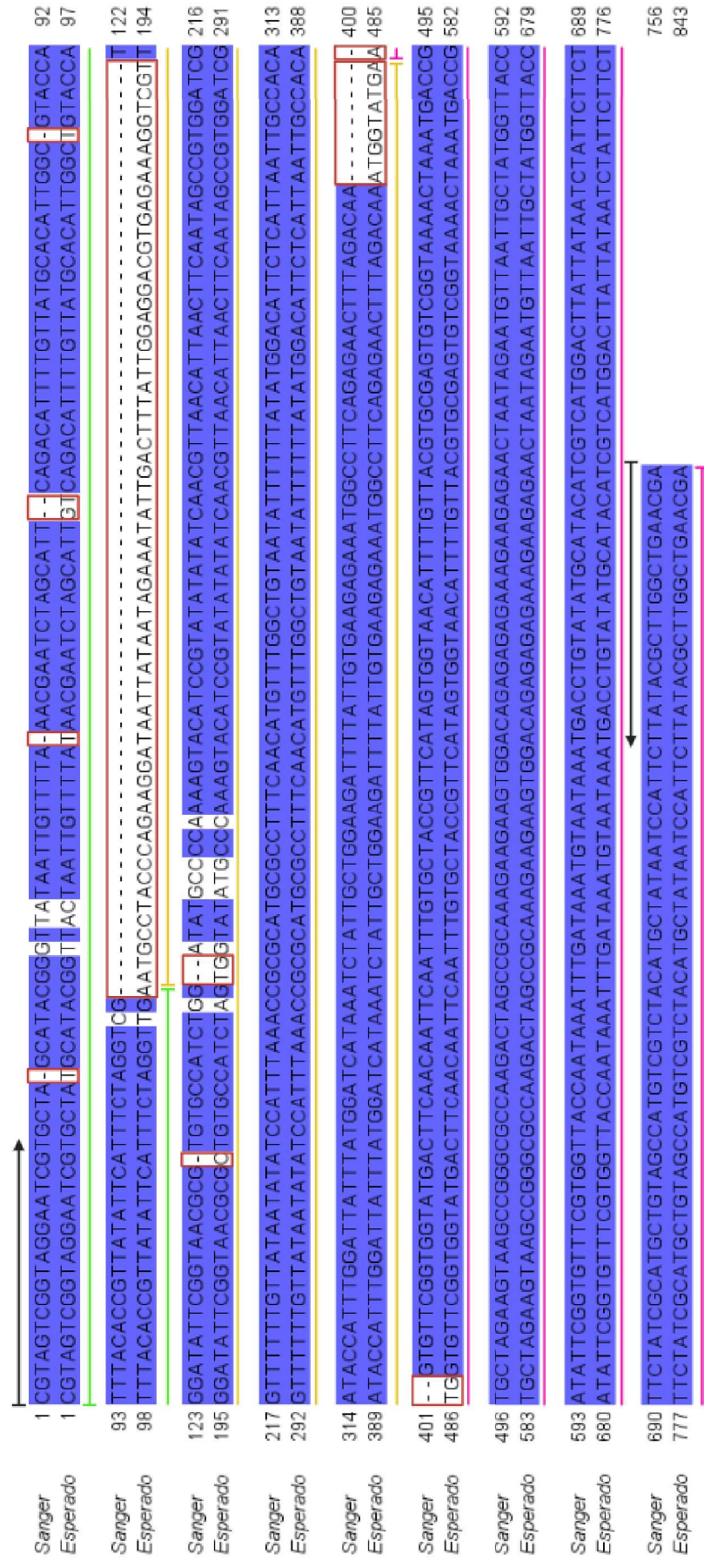


Para identificar as bases ausentes, foi realizado um alinhamento entre a sequência obtida pelo sequenciamento Sanger (756 pb) e o fragmento esperado (843 pb) (Figura 12). As regiões

cujos nucleotídeos possuíam 100% de identidade estão coloridas de azul na Figura 13. Dessa forma, foi possível identificar a região de divergência entre as sequências.

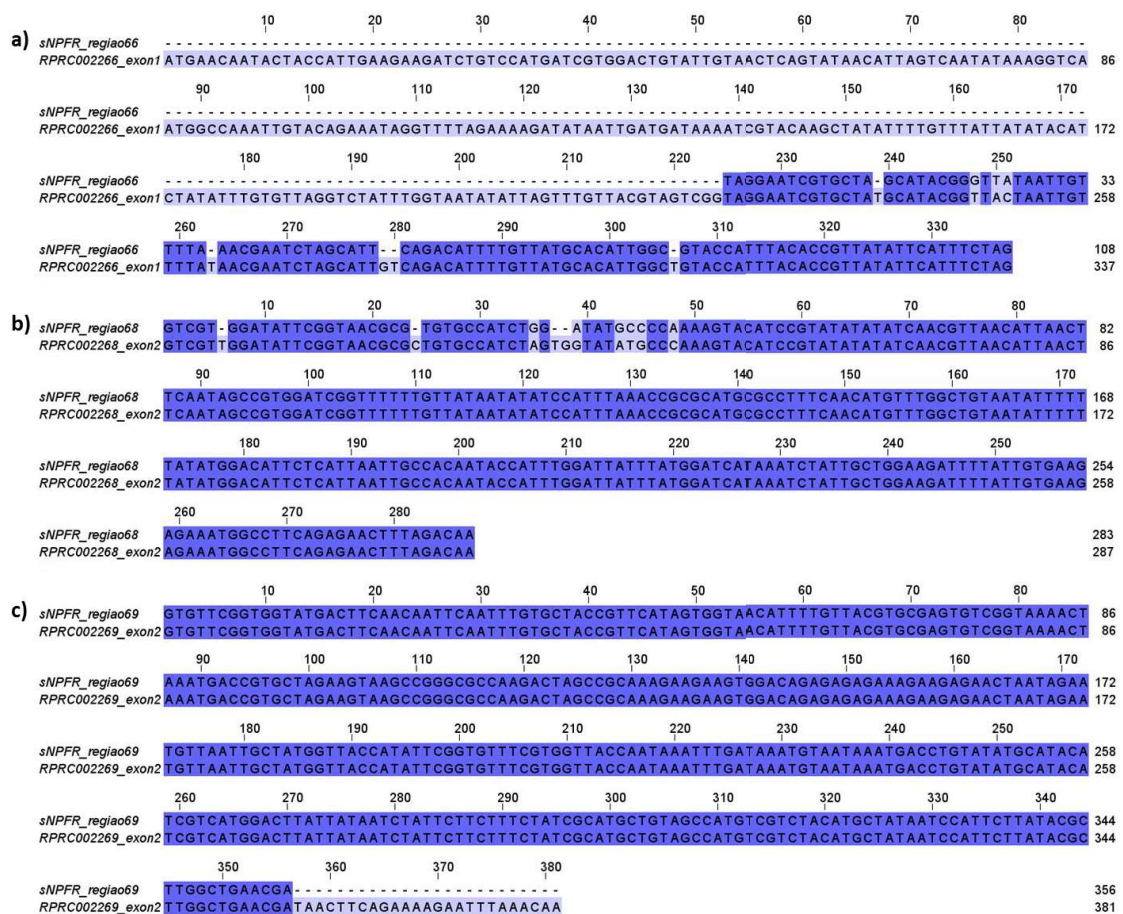
Como o fragmento esperado foi construído com base nas sequências anotadas dos transcritos, foi possível delimitar precisamente as regiões completas de cada um no alinhamento (Figura 13). No caso do transcrito RPRC002266, verificou-se que 5 pares de bases (pb) não foram identificados na sequência obtida por Sanger. Para RPRC002268, a ausência foi de 79 pb, enquanto para RPRC002269, apenas 3 pb não foram detectados.

Figura 13 – Alinhamento entre o resultado do sequenciamento (primeira linha) e a sequência do fragmento esperado (segunda linha). As linhas sólidas sobre as sequências destacam os aminoácidos referentes a cada transcrito, **verde**: RPRC002266, **amarela**: RPRC002268, **rosa**: RPRC002269. Os nucleotídeos ausentes estão destacados por retângulos **vermelhos**. Os oligonucleotídeos utilizados para a amplificação estão destacados por duas setas. Os aminoácidos coloridos em azul apresentam 100% de identidade.



Com base nas regiões correspondentes a cada transcrito presentes na sequência Sanger, foi realizada uma busca por BLASTn contra o genoma de *R. prolixus*, com o objetivo de identificar os exons aos quais pertencem os nucleotídeos sequenciados e ausentes. Os resultados indicaram que a região sequenciada correspondente ao transcrito RPRC002266 alinhou-se exclusivamente ao exon 1 desse transcrito (*sNPFR_região66*, Figura 14a). De maneira semelhante, a região associada ao transcrito RPRC002268 (*sNPFR_região68*, Figura 14b) alinhou-se ao exon 2 do respectivo transcrito. Por fim, a região referente ao transcrito RPRC002269 (*sNPFR_região69*, Figura 14c) também apresentou alinhamento com o exon 2 de seu transcrito correspondente. Esses achados sugerem a ocorrência de um processamento do mRNA.

Figura 14 - Alinhamento de sequências de nucleotídeos entre regiões do sequenciamento e os genes RPRC002266, RPRC002268 e RPRC002269, de *R. prolixus*. As sequências identificadas como *sNPFR_região66*, *sNPFR_região68* e *sNPFR_região69* (primeira linha), que correspondem ao resultado do sequenciamento, são comparadas aos transcritos RPRC002266, RPRC002268 e RPRC002269 (segunda linha), anotados no banco de dados. As regiões 100% conservadas estão destacadas em azul escuro. **a)** Alinhamento entre a sequência experimental com o exon 1 do gene RPRC002266. **b)** Alinhamento entre a sequência experimental com o exon 2 do gene RPRC002268. **c)** Alinhamento entre a sequência experimental com o exon 2 do gene RPRC002269.



A presença de uma sequência única contendo partes dos transcritos anotados no genoma pode indicar a ocorrência de *splicing* alternativo, um mecanismo comum em eucariotos que permite a geração de múltiplas variantes proteicas a partir de um único gene. Nesse caso, os três transcritos, RPRC002266, RPRC002268 e RPRC002269, poderiam representar variantes de *splicing* que, isoladamente ou em conjunto, contribuem para a formação do receptor funcional.

O *splicing* alternativo é um mecanismo fundamental que contribui para a diversidade de isoformas de proteínas, permitindo especialização e adaptação funcional em organismos eucariotos (Marasco; Kornblihtt, 2023). Atualmente, sabe-se que existem dois tipos principais de *splicing* alternativo: *cis-splicing* e *trans-splicing*.

O *cis-splicing*, a forma mais comum em insetos, permite a geração de isoformas de mRNA a partir do transcrito primário (pre-mRNA) de um único gene, por meio da exclusão de íntrons e exons (Marasco; Kornblihtt, 2023). Em *D. melanogaster*, estima-se que 31% dos seus genes passem por *splicing* alternativo, o que aumenta consideravelmente a variedade de proteínas produzidas (Daines *et al.*, 2011). Em *A. gambiae*, a resposta imune é regulada pela diversidade de isoformas geradas a partir do gene *Dscam*, em resposta à infecção por diferentes patógenos (Dong; Taylor; Dimopoulos, 2006), permitindo um alcance maior na resposta imune adaptativa.

Os eventos de *cis-splicing* alternativo são comuns em vias neuropeptidérgicas, onde genes precursores geram diferentes transcritos de neuropeptídeos. Um exemplo é o gene da orcocinina (OK), regulador da ecdise, que produz dois precursores dos neuropeptídeos OKA e OKB, por meio de *splicing* alternativo de um único gene (Sterkel *et al.*, 2012). No caso do hemimetábolo *R. prolixus*, OKA é expresso no sistema nervoso central, enquanto OKB e OKC, uma terceira variante, tanto no sistema nervoso central quanto no periférico (Wulff *et al.*, 2017).

Por outro lado, os eventos de *trans-splicing* envolvem a junção de exons de diferentes transcritos primários (pre-mRNA) para a formação de isoformas (Lasda; Blumenthal, 2011). Nos tripanossomatídeos, todos os mRNAs são processados por eventos de *trans-splicing* (Bernhard *et al.*, 2025). Embora seja menos comum em insetos, esse processo também foi documentado, sendo o gene *mod(mdg4)* de *D. melanogaster* o mais estudado (Yu *et al.*, 2014).

O primeiro registro desse fenômeno em insetos ocorreu na mosca da fruta (*D. melanogaster*), quando Labrador *et al.* (2001) investigaram a existência de um mRNA originado da fusão de dois RNAs precursores, transcritos a partir de duas unidades

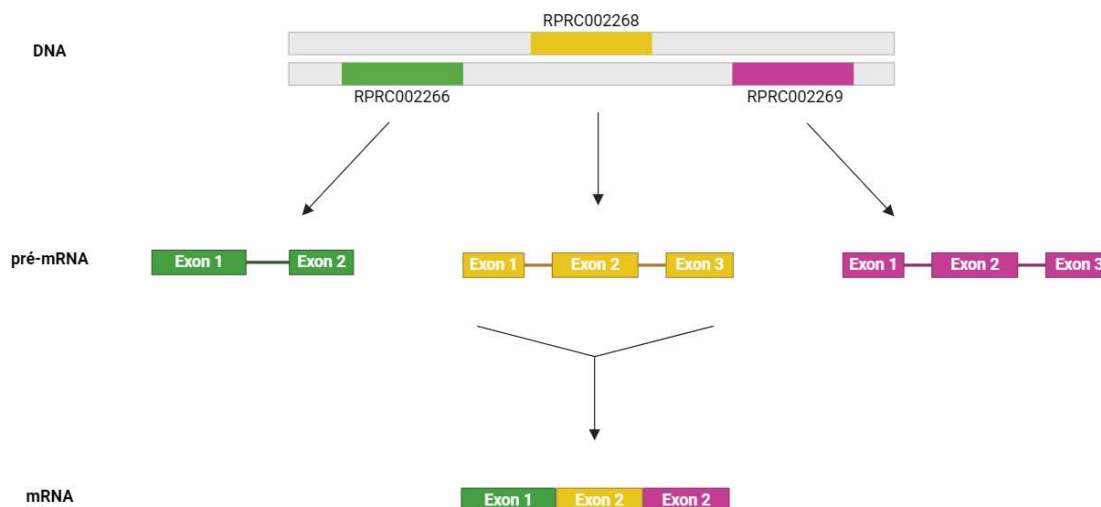
transcricionais opostas do gene *mod(mdg4)*. Dorn, Reuter e Loewendorf (2001) demonstraram que múltiplos elementos iniciadores de transcrição no locus desse gene permitem que exons colineares e antiparalelos atuem como precursores do evento de *trans-splicing*.

Outros genes, como *lola* em *D. melanogaster*, também passam por *trans-splicing*. Assim como *mod(mdg4)*, os exons alternativos de *lola* correspondem à região terminal dos mRNAs (Horiuchi; Aigaki, 2006). No entanto, ao contrário de *mod(mdg4)*, os exons do gene *lola* são codificados de forma colinear.

Esse mecanismo regulatório também já foi documentado no bicho-da-seda (*Bombyx mori*) por meio de uma análise transcriptômica abrangendo todas as etapas de desenvolvimento do inseto. Utilizando parâmetros diferenciados de mapeamento de RNA-seq contra o genoma disponível, Shao *et al.* (2012) identificaram 2648 novos transcritos, incluindo 1923 possíveis eventos de *trans-splicing*. Esses eventos foram classificados em cinco grupos, confirmados por análise de polimorfismo de nucleotídeos únicos (SNPs) quiméricos e validados por RT-PCR. Além disso, um estudo filogenético mapeou um sistema evolutivo de aproximadamente 400 milhões de anos em oito espécies para identificar eventos de *trans-splicing*. Das espécies analisadas, *Acyrtosiphon pisum*, hemíptero como *R. prolixus*, apresentou a maior frequência de eventos, seguido de *D. melanogaster* e *A. aegypti* (Kong *et al.*, 2015).

Portanto, a identificação de transcrito do receptor sNPFR em *R. prolixus* como estando associado a múltiplos genes podem estar diretamente relacionada a eventos de *trans-splicing*, mecanismos já identificados em insetos da mesma ordem. O *trans-splicing* permitiria a fusão de exons provenientes de diferentes precursores transcricionais (Figura 15), o que corrobora com nossos resultados.

Figura 15 - Desenho esquemático representando uma possível via de transcrição do sNPFR.



4.3 ANÁLISE DE DOMÍNIOS CONSERVADOS DA PROTEÍNA RPRO-SNPFR

Sabe-se que, assim como outros receptores neuropeptidérgicos, o receptor do neuropeptídeo F (sNPFR) pertence à família A dos GPCR (Ons *et al.*, 2016). Com base nisso, foram realizadas análises estruturais e de motivos conservados no receptor sNPFR de *R. prolixus*, com o objetivo de confirmar que a sequência de aminoácidos prevista exibe as características típicas de um GPCR da família A.

Devido à posição dos oligonucleotídeos iniciadores (Figura 12) na sequência experimental, não foi possível obter a sequência completa do Rpro-sNPFR (ORF) que correspondesse integralmente às características esperadas para um GPCR, principalmente devido ao fato de gerar uma proteína com apenas quatro domínios transmembranares. Dessa forma, os exons incompletos do *RPRC002266*, que correspondem à região aminoterminal da proteína, e do *RPRC002269*, que correspondem à região carboxiterminal (Figura 14a e 14c), foram completados manualmente com base nas sequências anotadas no VectorBase, resultando em uma sequência final de 1004 nucleotídeos (Figura 16).

A sequência de nucleotídeos montada (Figura 16) codificou uma proteína com 334 aminoácidos (Figura 17a), com peso molecular de 38,87 kDa e ponto isoelétrico teórico de 9,28. A predição de topologia dessa proteína sugere uma estrutura com sete domínios α -hélice

transmembranares conservados, cujos aminoácidos estão localizados nas posições 50-73 (TM1), 94-109 (TM2), 121-142 (TM3), 161-182 (TM4), 206-228 (TM5), 263-284 (TM6) e 299-322 (TM7) (Figura 17b). Foi observada a presença de resíduos de asparagina (N) no TM1, de prolina (P) nos TM4 a TM7, e de cisteína (C) no ECL1. Além disso, foi identificada uma substituição na tirosina do motivo conservado DRY, que foi trocada por fenilalanina, formando o motivo DRF. Essa substituição já foi observada no receptor de allostatinas de *D. melanogaster* (Verlinden *et al.*, 2015; Wang; Chin-Sang; Bendena, 2012).

Os receptores acoplados à proteína G (GPCR) são grupos de proteínas transmembranares fundamentais em diferentes processos de sinalização celular, pois interagem com diferentes ligantes, como hormônios, neuropeptídeos, neurotransmissores e estímulos químicos. Por conta disso, representam alvos interessantes para estudo e desenvolvimento de estratégias terapêuticas. São caracterizados por suas estruturas únicas, que consistem em sete domínios transmembranares em α -hélices, além de N-terminal extracelular e C-terminal intracelular (Peng; Yang; Chen, 2010). Além disso, os receptores da família A apresentam, geralmente, resíduos conservados, incluindo um resíduo de asparagina no domínio transmembranar (TM) 1, aspartato no TM2 e prolina em TM4 ao TM7. Nos *loops* extracelulares (ECL1 e ECL2), são encontrados resíduos de cisteína, importantes para a estabilização do receptor. Também é observado um motivo conservado DRY (Asp-Arg-Tyr) logo ao final do TM3, e sítios de fosforilação e palmitoilação no carboxiterminal importantes para a transdução do sinal (Verlinden *et al.*, 2015).

Figura 16 – Sequência de RNA/mRNA previsto para sNPFR, obtida após sequenciamento. Oligonucleotídeos iniciadores utilizados para a obtenção do produto sequenciado estão identificados por setas pretas. As regiões do produto sequenciado estão destacadas de acordo com os transcritos identificados, **verde**: região 66; **amarelo**: região 68; **rosa**: região 69. As porções destacadas em **vermelho** representam as sequências parciais dos exon 1 de RPRC002266 (VB) e exon 2 RPRC002269 (VB) que foram inseridas manualmente.

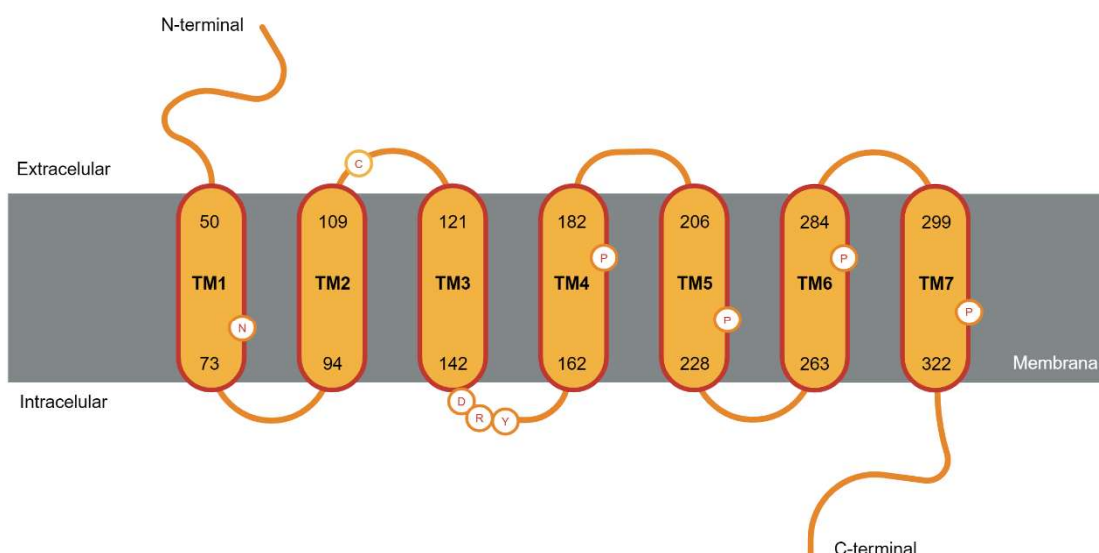
sNPFR	1	ATGAACAA	TACTACCA	TTGAAGA	AGATCTGTCC	ATGATCGTGG	ACTGTATTGTAA	CTCAGTATAACA	TTAGTCAATATAA	AGGTCAAT	88	
sNPFR	89	GGCCAAAT	TGTACAGA	AAATAGG	TTTTAGAAA	AGATATAA	TTGATGATAA	AAATCGTACA	AGCTATATTT	TGTTTATTATACATCTAT	176	
sNPFR	177	ATTTGTGT	TAGGTCTA	TTTGGTA	ATATATTAG	TTTGTAC	CGTAGTCGG	TAGGAATCGT	GCTAGCATA	CGGGTTATAA	TGTTTAAAC	264
sNPFR	265	GAATCTAG	CATT	CAGACAT	TTTGT	TATGCACAT	TGGCGTACCA	TTTACACCG	TTATATTCAT	TTCTAGG	TCGTGGATATTCGGTAACG	352
sNPFR	353	CGTGTGCC	ATCTGG	ATATGCCCC	AAAGTACAT	CCGTATATAT	ATCAACGTTAAC	ATTCAACTTCAAT	AGCCGTGGAT	CGGTTTTTTGT	440	
sNPFR	441	TATAATAT	ATCCATT	TAAACCG	CGCATGCG	CCTTTCAAC	ATGTTGGCTGTAA	TATTTT	TATATGGACAT	TTCTCAT	TAAATTGCCACA	528
sNPFR	529	ATACCA	TTTGGAT	TATTTATGG	ATCATAA	ATCTATTGCT	GGAAGATTTT	ATTGTGAAG	AGAAATGG	CCTTCAGAG	AACTTTAGACAAG	616
sNPFR	617	TGTT	CGGTGG	TATGACT	TTCAACA	ATTCAATTTGT	GCTACCGTT	CATAGTGGT	AACATTTTGT	TACGTGCGAG	TGTCGGTAAAACTAAA	704
sNPFR	705	TGACCG	TGCTAGA	AGTAAG	CCGGCGCC	AAGACTAG	CCGCAAGA	AGAAGTGG	ACAGAGAG	AAAGAA	GAAGAACTAATAGAATGTTA	792
sNPFR	793	ATTGCT	ATGGTTAC	CATATT	CGGTGTTTCGT	GGTTACCA	ATAAATTTGATA	AAATGTAA	TAAATGACCT	GTATATGC	ATACATCGTCAT	880
sNPFR	881	GGACT	TATTATA	ATCTATTCT	TTCTAT	ATCGCAT	GTCTGTAG	CCCATGTCTAC	ATGCTATA	AATCCATTCT	TATACGCTTGGCTGAA	968
sNPFR	969	CGA	AACTTC	AGAAA	AGAAATTT	AAACAA	ATTTTCGT	GTA				1004

Figura 17 – Sequência primária de Rpro-sNPFR e topologia prevista. a) Sequência de aminoácidos da proteína Rpro-sNPFR. Em **negrito**, destacados os domínios transmembranares; em **vermelho**, aminoácidos conservados para GPCRs; sublinhado, motivo conservado DRF. **b)** Predição da topologia da proteína. As posições do início dos sete domínios transmembranares (TM) estão indicados. Três loops intracelulares (ICL) e três loops extracelulares (ECL). Porção amino-terminal na região extracelular e carboxi-terminal na região intracelular.

a)

```
>Rpro-sNPFR
MNNTTIEEDLSMIVDCIVTQYNISQYKGQWPNCTEIGFRKDIIDDKIVQAIFCLLYTSIFVLGL
FGNILVCYVVGRRNRASIRVIVLNESSIQTCYAHWRTIYTVIFISRSWIFGNACAIWICPKST
SVYISTLTLTSIAVDREFVIIYPFKPRMRLSTCLAVIFFIWTFSLIATIPFGLFMDHKSIAGRF
YCEEKWPSENFRQVFGGMTSTIQFVLPFIVVTECYVRVSVKLNDRARSKPGAKTSRKEEVDRER
KKRTNRMLIAMVTIFGVSWLPINLINVLINDLYMHTSSWTYYNLFFFLSHAVAMSSTCYNFLYA
WLNNETSEKNLNKFR
```

b)



A realização de um BLASTp, seguida de um alinhamento múltiplo de sequências, revelou que os resíduos transmembranares são os mais conservados entre as espécies analisadas, o que reforça a importância estrutural desses domínios (Figura 18).

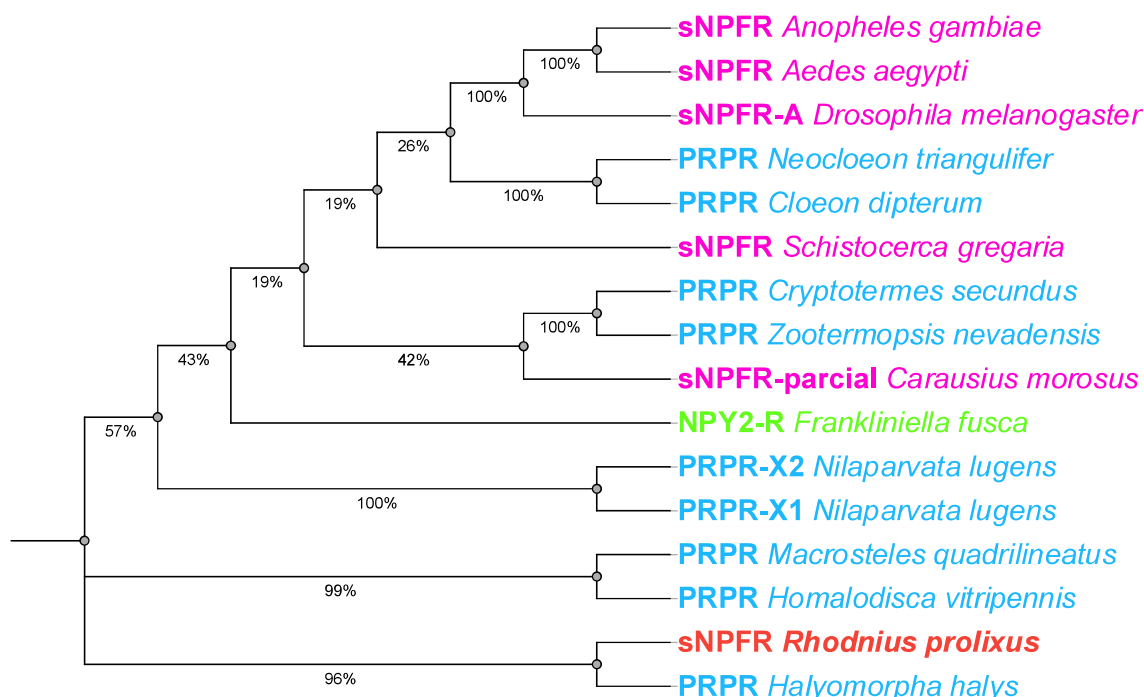
4.4 ANÁLISE FILOGENÉTICA

A análise filogenética permite inferir a função de proteínas putativos ao compará-las evolutivamente com proteínas conhecidas e funcionalmente caracterizadas. Por meio de alinhamentos de sequências, é possível identificar regiões que se mantiveram conservadas ao longo do tempo e em diversas espécies. Além disso, a construção de uma árvore filogenética possibilita a formação de grupos (*clusters*) de proteínas semelhantes, com base em modelos evolutivos.

O método do *neighbor-joining*, utilizado nesta análise, é uma abordagem quantitativa empregada na construção de árvores filogenéticas, baseada em matrizes de distância derivadas de dados genéticos (Evans; Sheneman; Foster, 2006; Li, 2015). A técnica *bootstrap* fornece a base estatística para validar quantas vezes um determinado clado é observado em cada árvore gerada, sendo crucial para confirmar as relações evolutivas (Bland; Altman, 2015).

A análise filogenética da proteína Rpro-sNPFR foi realizada pelo método NJ, com 1000 repetições de *bootstrap*. A análise revelou que a sequência Rpro-sNPFR apresenta 96% de homologia com o receptor semelhante ao do peptídeo liberador de prolactina (PrRPR-*like*) de *Halyomorpha halys* (XP_066907729.1) (Figura 19).

Figura 19 - Análise filogenética dos genes de receptores de sNPF e NPY, e do receptor semelhante ao do peptídeo liberador de prolactina (PrRPR-like). A árvore consenso foi construída utilizando o método de *neighbor-joining*, com 1000 repetições *bootstrap*.



Os receptores de peptídeos liberadores de prolactina (PrRPR), que também são GPCRs da família A (Lagerström *et al.*, 2005), desempenham um papel crucial na comunicação neuroendócrina dos mamíferos, podendo atuar como estimuladores do apetite (Yamashita *et al.*, 2013). A redução da expressão desses receptores em ratos está associada a alterações metabólicas e respostas anormais ao estresse (Takayanagi; Onaka, 2010).

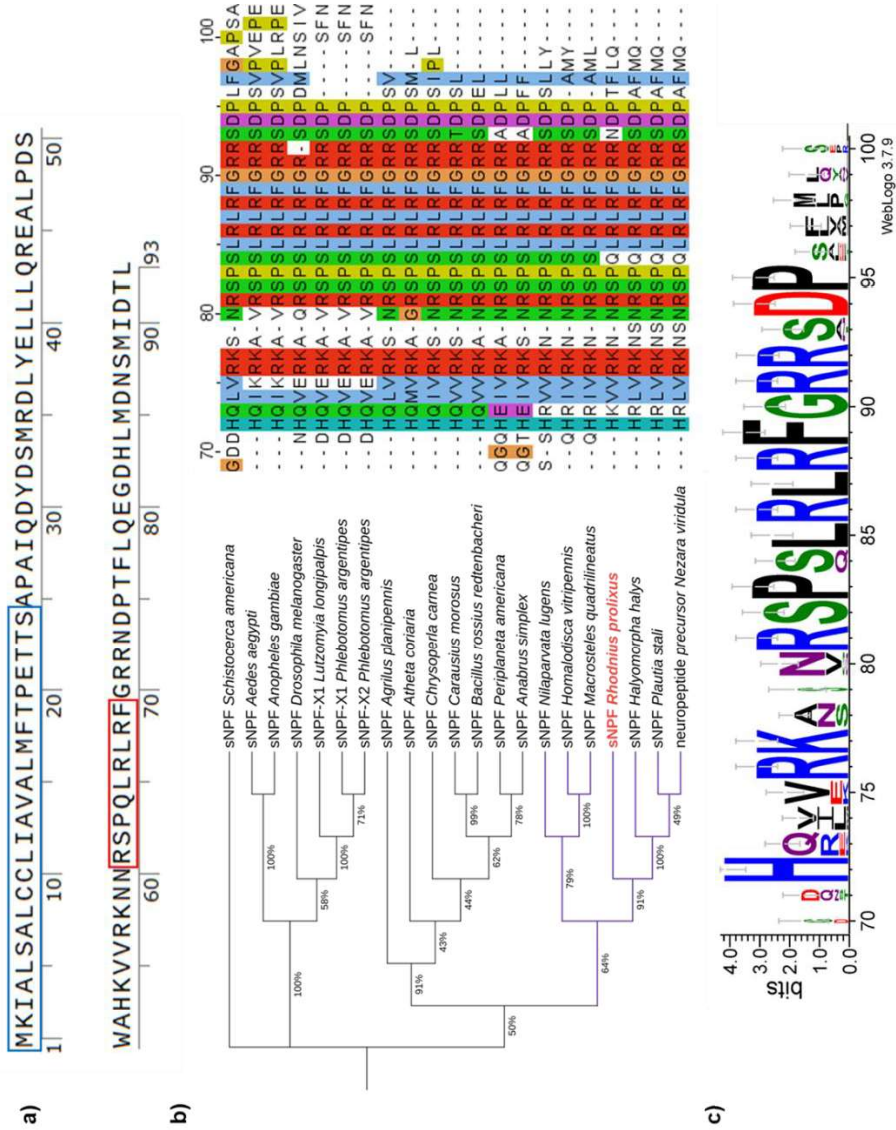
A natureza conservada desses receptores sugere que os PrRPR-like em insetos podem ter evoluído a partir de um GPCR ancestral comum, associado à regulação da prolactina e vias hormonais (Lawrence *et al.*, 2000). Além disso, sua semelhança funcional com receptores de neuropeptídeos envolvidos em funções neuroendócrinas indica a preservação de mecanismos endócrinos ao longo da evolução. Assim, a de-orfanização desses receptores em insetos pode contribuir para uma reclassificação mais precisa de sua função, ampliando a compreensão do seu papel fisiológico.

Na proteína precursora de sNPF em *R. prolixus*, o peptídeo sinal é composto por 24 aminoácidos, sendo clivado entre as sequências Thr-Thr-Ser e Ala-Pro-Ala (Figura 20a). O alinhamento múltiplo de sequências revelou que os aminoácidos do neuropeptídeo maduro estão em posições conservadas, assim como em outras espécies (Figura 20b). Além disso, a

análise indica que, em *R. prolixus*, a proteína precursora é processada para gerar apenas um peptídeo sNPF.

Como esperado, a proteína precursora de sNPF de *R. prolixus* agrupou-se com outras sequências de hemípteros, reforçando a conservação filogenética dessa proteína (Figura 20b). Dentre essas, 91% de homologia foi observada com a proteína de *H. halys*, *Plautia stali* e *Nezara viridula*. Além disso, o alinhamento múltiplo demonstrou que o motivo característico do neuropeptídeo (RxPxLRLRF) está em posição conservada.

Figura 20 - Características da sequência da proteína precursora de sNPF. a) Sequência primária com peptídeo sinal destacado em azul e a sequência do neuropeptídeo destacado em vermelho. b) À esquerda: Árvore filogenética da proteína precursora de sNPF em diferentes espécies. A árvore consenso foi construída utilizando o método de *neighbor-joining*, com 1000 repetições *bootstrap*. Os clados contendo hemípteros estão destacados em roxo. À direita: Alinhamento múltiplo das sequências de aminoácidos da proteína sNPF em diferentes insetos, correspondente às espécies expostas na árvore filogenética. c) Logo representando os aminoácidos em posições conservadas da proteína precursora, representando a sequência do peptídeo maduro.

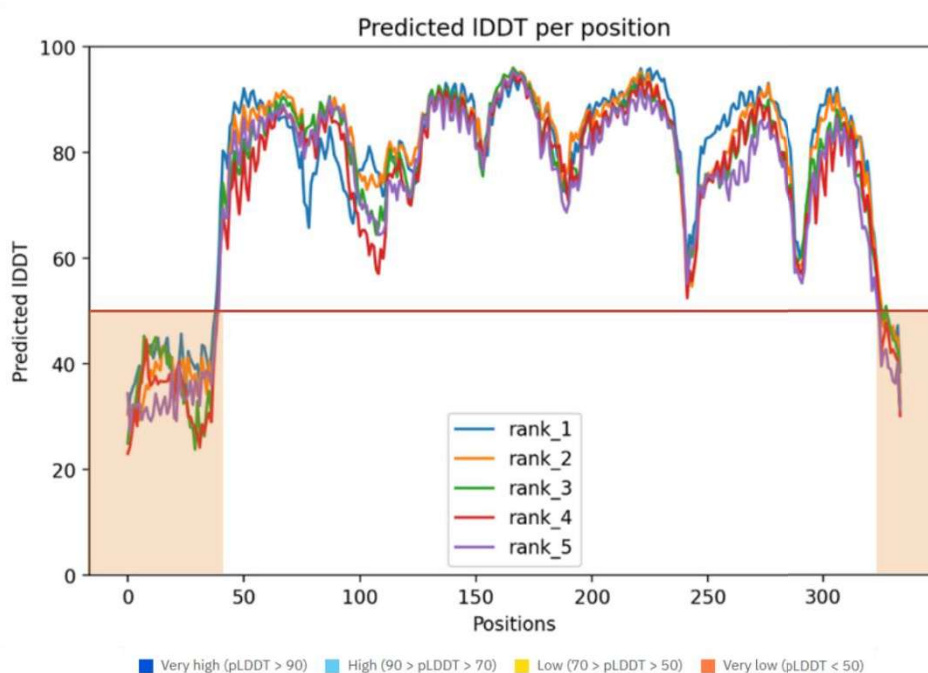


4.5 CARACTERIZAÇÃO ESTRUTURAL *IN SILICO* DA PROTEÍNA Rpro-sNPFR

Para reforçar a hipótese de que se trata de um GPCR, foi realizada a predição da estrutura tridimensional da proteína utilizando a modelagem *ab initio* disponível no *AlphaFold2* (AF2). Esse método apresenta uma acurácia comparável à de técnicas experimentais, como cristalografia de raios-X, crio-eletromicroscopia (Cryo-EM) e ressonância magnética nuclear (RMN), devido ao treinamento de sua rede neural com um vasto banco de dados de estruturas de proteínas e suas respectivas sequências (Pak *et al.*, 2023).

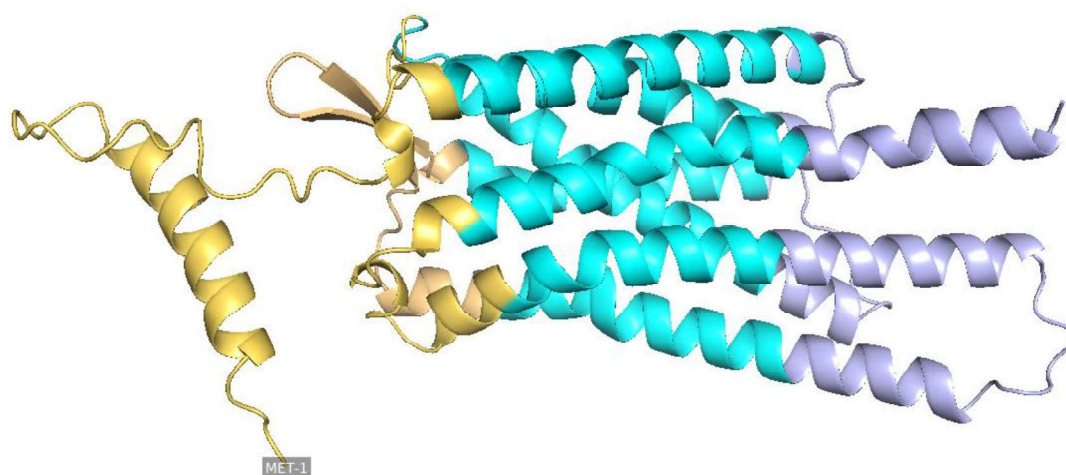
A análise dos modelos gerados para o Rpro-sNPFR revelou que os valores de pLDDT (teste de diferença de distância local prevista), que indicam a confiabilidade da predição estrutural, foram mais baixos na porção amino e carboxiterminal (Figura 21). O AF2 classifica regiões de baixa confiança as regiões que são naturalmente desordenadas ou para as quais há informações insuficientes para uma predição mais precisa (Mariani *et al.*, 2013).

Figura 21 – Teste de diferença de distância local previsto para cada posição dos aminoácidos nos modelos. Confiabilidade é considerada muito alta se pLDDT > 90; alta, se 90 > pLDDT > 70; baixa, se 70 > pLDDT > 50; muito baixa, se pLDDT < 50. Posições com valor de pLDDT abaixo de 50 estão indicados em laranja.



Na figura 22, o modelo Rpro-sNPFR está colorido de acordo com a topologia da proteína. Os domínios transmembranares estão indicados em azul, a região extracelular em amarelo, e a região intracelular em lilás. A α -hélice observada na região extracelular, representando a porção amino-terminal, apresentou um valor de pLDDT previsto inferior a 50.

Figura 22 – Modelo de Rpro-sNPFR gerado pelo AlphaFold2. Os domínios transmembranares estão indicados em azul ciano; a região extracelular, indicada em amarelo; e a região intracelular indicada em lilás.



O AF2 não busca *templates* diretamente em estruturas experimentais, como faz a modelagem por homologia tradicional. No entanto, ele pode incorporar informações de alinhamento de sequência e estruturas conhecidas de forma indireta, pois é treinado com um grande conjunto de dados de estruturas experimentais disponíveis em bancos como o PDB (*Protein Data Bank*) (Urvás *et al.*, 2024). Dessa forma, embora não dependa diretamente de *templates*, seu treinamento permite que ele reconheça padrões estruturais comuns.

Com isso em mente, foi realizada uma segunda modelagem tridimensional pelo método da homologia, utilizando a ferramenta *SWISS-Model*, com o objetivo de comparar os modelos gerados para Rpro-sNPFR nas duas técnicas. O modelo por homologia foi gerado com base na estrutura do receptor GPCR putativo de *T. infestans*, que apresentou 80,21% de identidade de sequência e uma cobertura de 86%. A qualidade do modelo foi avaliada por meio de um gráfico de Ramachandran, com 95,86% dos aminoácidos em regiões favoráveis, indicando baixa frequência de impedimentos estéricos (Figura 23). A principal diferença observada entre os modelos foi a ausência da α -hélice na região extracelular do Rpro-sNPFR no modelo por homologia. A sobreposição dos dois modelos gerou um valor de RMSD (desvio quadrático médio da raiz), que mede a diferença média entre as posições atômicas de duas estruturas, de 0,935 Å (Figura 24c).

Figura 23 – Rpro-sNPFR gerado por homologia. a) Estrutura tridimensional gerada no SWISS-Model. Colorido de acordo com valores de pLDDT. **b, c, d)** Gráficos de Ramachandran. **b)** Resíduos gerais. **c)** Resíduos de glicina. **d)** Resíduos de prolina. **e)** Resíduos pré-prolina.

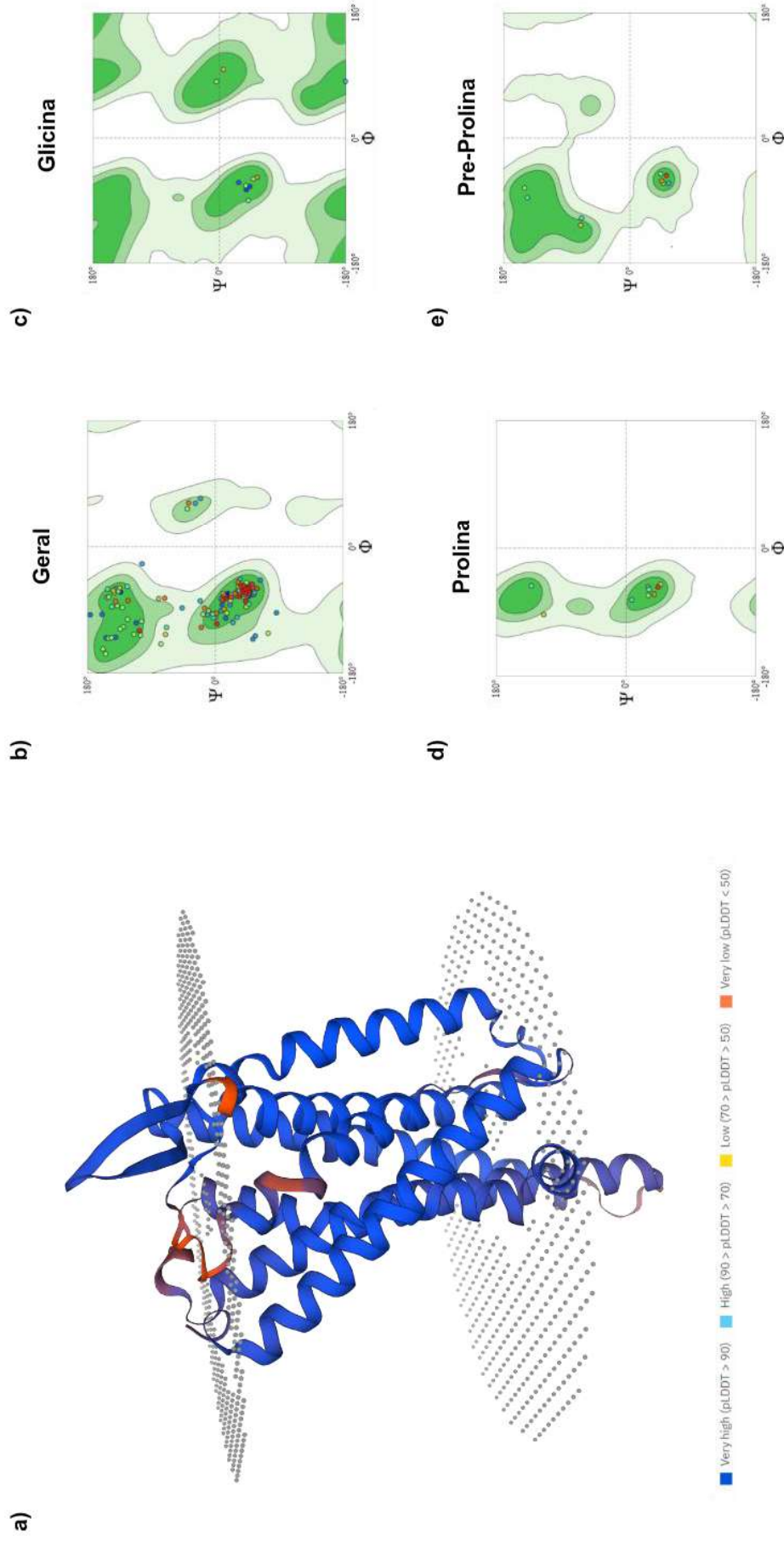
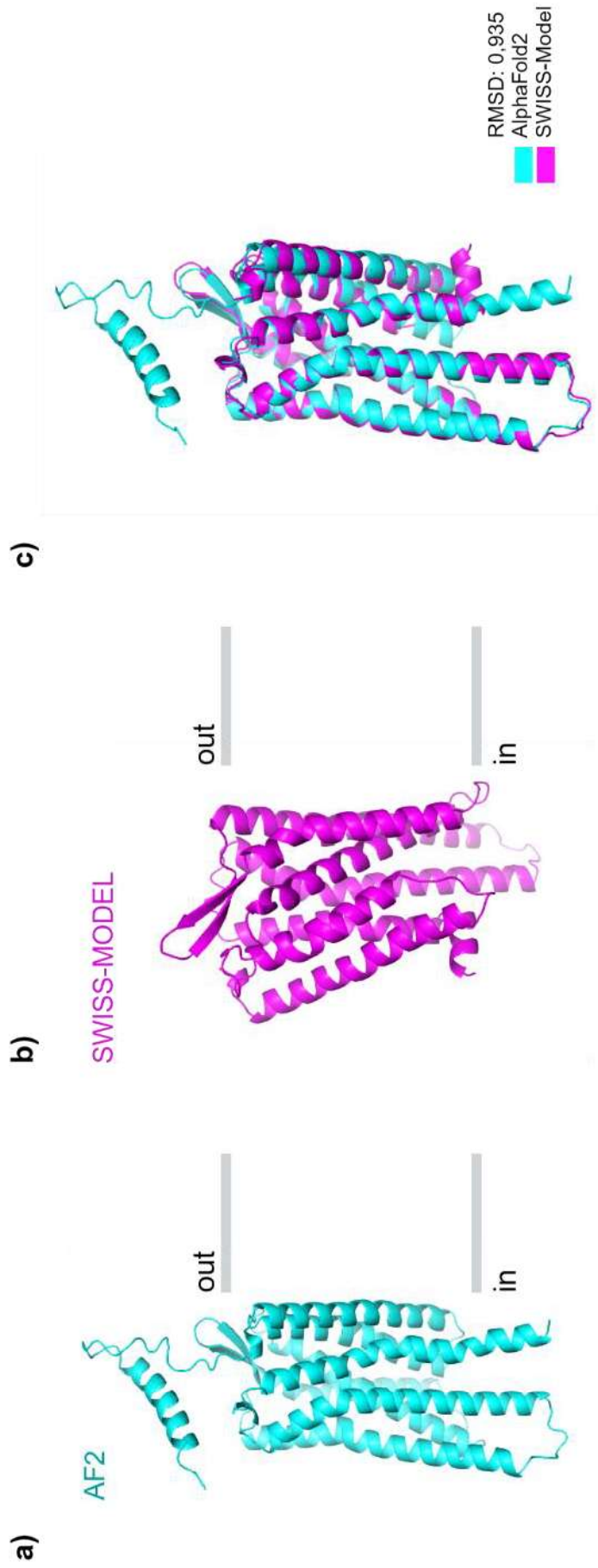


Figura 24 – Comparação entre os modelos Rpro-sNPF_R. a) Modelo gerado pelo AlphaFold2. b) Modelo gerado pelo SWISS-MODEL. c) Sobreposição estrutural. RMSD: 0,935.



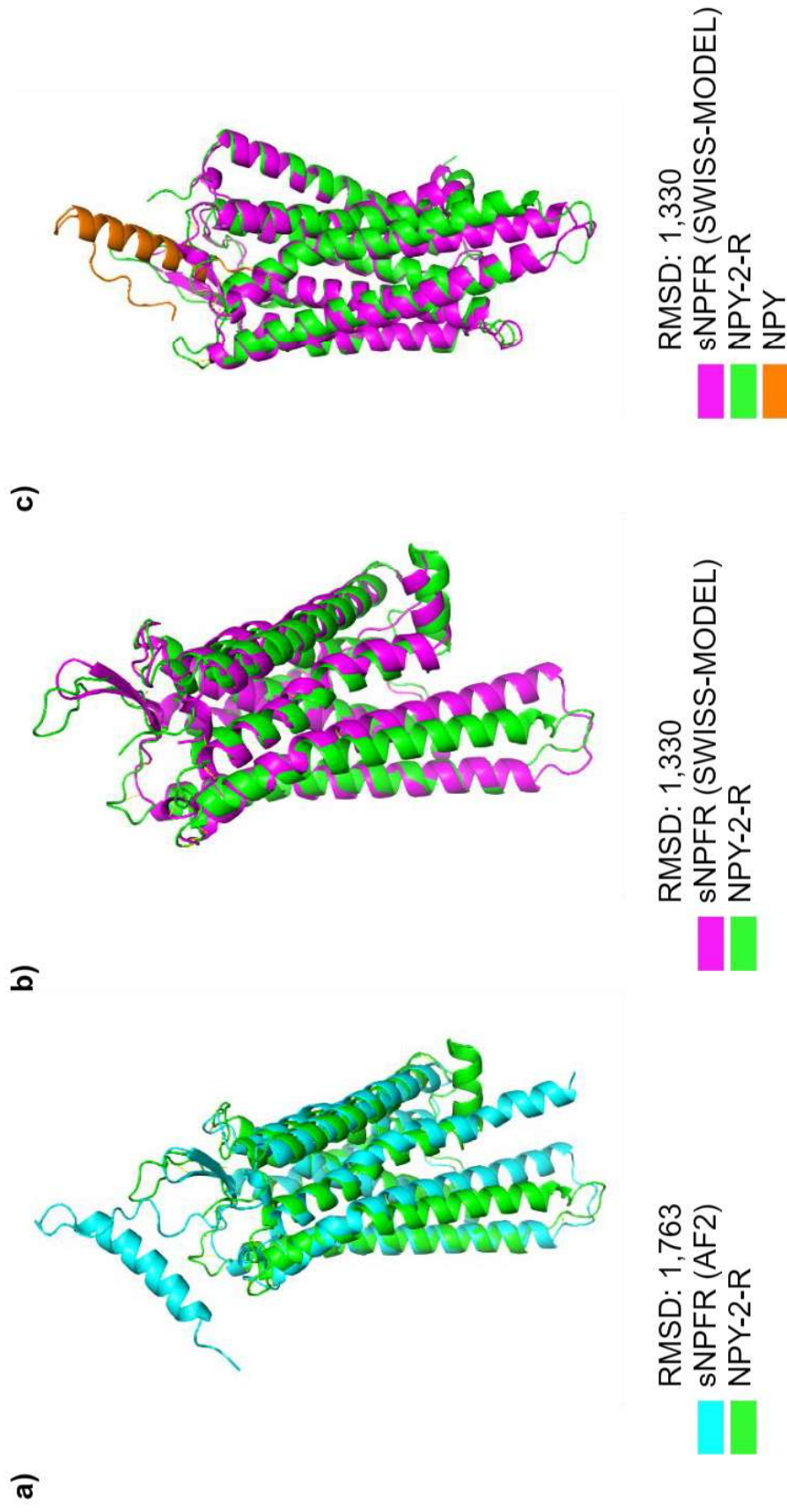
Analisando as estruturas obtidas por crio-eletromicroscopia de NPY2R e dos receptores de prolactina de *Homo sapiens* (Tang *et al.*, 2022; Li *et al.*, 2024), ambos ortólogos de Rpro-sNPFR, nota-se a ausência dos aminoácidos da porção amino-terminal nessas estruturas. Isso ocorre devido às limitações da técnica empregada, uma vez que regiões naturalmente desordenadas e mais flexíveis, como as extremidades das proteínas, apresentam menor densidade eletrônica, o que diminui a resolução da estrutura (Merk *et al.*, 2016). Outra evidência foi a baixa definição na porção amino-terminal do receptor de orexina (OX2R) de *H. sapiens*, onde está localizada uma pequena α -hélice extracelular que precede o primeiro domínio transmembrana na proteína (Hong *et al.*, 2021), situação semelhante à observada no modelo gerado pelo AF2. Por essa razão, na modelagem por homologia, a porção amino-terminal do Rpro-sNPFR foi retirada, resultando em uma conformação estrutural sem a α -hélice extracelular (Figura 24b).

Portanto, a baixa confiabilidade nas extremidades do modelo gerado pelo AF2 pode ser explicada pela maior flexibilidade estrutural dessas regiões, tornando-as mais propensas à ausência de sinal em técnicas experimentais, o que, conseqüentemente, afeta a quantidade de informações que beneficiariam a rede neural do AF2. A diferença observada entre as duas modelagens pode refletir não apenas as diferenças metodológicas, mas também a possibilidade de que a α -hélice extracelular seja uma estrutura dinâmica, como observado por Hong *et al.* (2021).

Para a validação e previsão do sítio de ligação nos modelos, foram realizadas sobreposições estruturais entre a estrutura do NPYR humano na presença do NPY (Tang *et al.*, 2022) e os modelos previstos de Rpro-sNPFR, com o objetivo de identificar o possível sítio de ligação do sNPF. O receptor NPYR do tipo 2 possui 29,12% de identidade de sequência e 95% de cobertura comparado ao Rpro-sNPFR. Apesar de diferentes sequências primárias, ambas as proteínas pertencem à família dos GPCRs e desempenham funções semelhantes nos dois organismos.

A sobreposição gerou um valor de RMSD de 1,763 para o modelo AF2 (Figura 25a) e RMSD de 1,330 para o modelo SWISS-Model (Figura 25b). Valores de RMSD menores ou iguais a 2,0 Å indicam uma alta semelhança estrutural entre as proteínas. Dessa forma, foi possível inferir um provável sítio de ligação para o ligante endógeno do Rpro-sNPFR, a partir da posição do NPY na estrutura do NPYR humano (Figura 25c). Considerando que RMSD foi menor na comparação entre NPY2-R e o modelo gerado por homologia (SWISS-Model), esse foi escolhido para o ensaio de *docking* molecular.

Figura 25 – Comparação estrutural de Rpro-sNPFR com NPY2-R de *H. sapiens*. a) Sobreposição da estrutura de NPY2-R de *H. sapiens* obtida por crio-eletromicroscopia com o modelo para Rpro-sNPFR gerado pelo *AlphaFold2*. b) Sobreposição da estrutura de NPY2-R de *H. sapiens* obtida por crio-eletromicroscopia com o modelo para Rpro-sNPFR gerado por homologia (SWISS-Model). c) Sobreposição do modelo obtido por homologia com a estrutura experimental de NPY2-R na qual NPY está localizado.



O *docking* molecular é uma abordagem computacional essencial para a identificação de sítios de interação proteína-ligante (Pinzi; Rastelli, 2019). A principal maneira de quantificar a energia de interação em ensaios de *docking* molecular é por meio da variação da energia livre de Gibbs (ΔG), que indica a estabilidade do ligante no sítio ativo da proteína. Valores mais negativos de ΔG indicam uma interação mais favorável entre proteína e ligante. Esse método permitiu identificar os resíduos envolvidos na interação do receptor com seu ligante endógeno.

O ensaio de *docking* molecular foi realizado com o modelo tridimensional do peptídeo sNPF maduro (RSPQLRLR-amida) de *R. prolixus* (Rpro-sNPF), gerado com PEP-FOLD3.

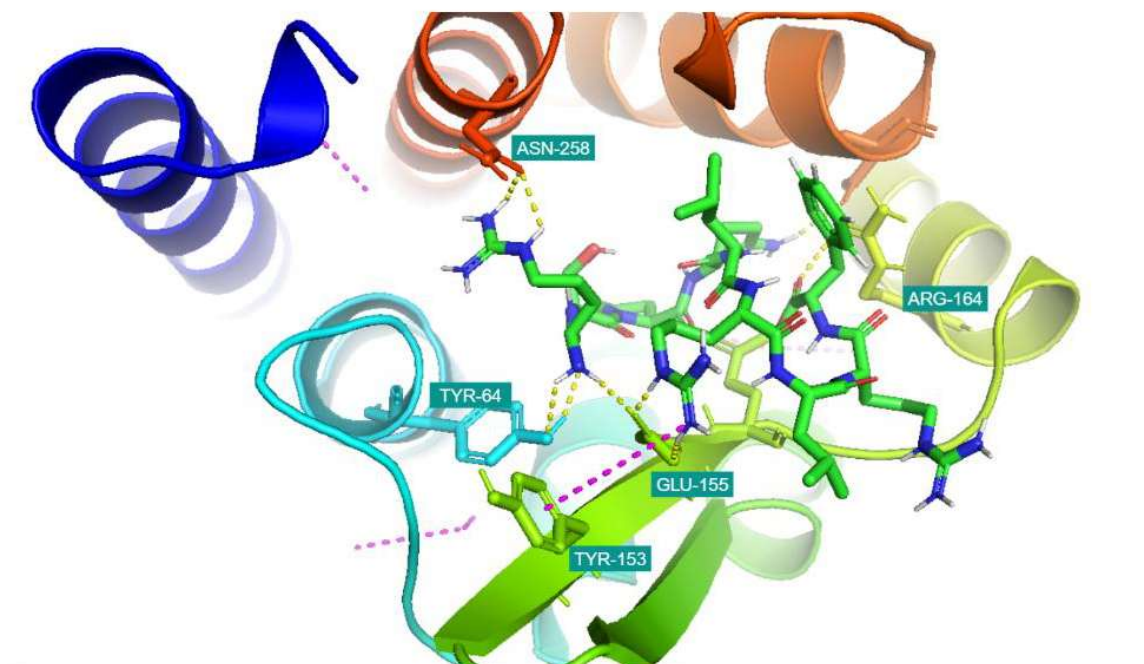
No *docking* às cegas (*blind*), a superfície da proteína é explorada em busca do sítio de interação com o ligante, enquanto no *docking* direcionado, o usuário fornece a provável posição do sítio de interação. Com base na sobreposição estrutural realizada entre o modelo de Rpro-sNPF e NPY2-R, o sítio de ligação foi posicionado próximo aos *loops* extracelulares.

A afinidade de ligação foi maior no *docking* direcionado, com valor de -10,045 kcal/mol, enquanto no *docking* às cegas foi de -9,731 kcal/mol. Os aminoácidos envolvidos e os valores de afinidade de ligação previstos estão apresentados na Tabela 3. Portanto, o sítio ativo proposto no *docking* direcionado parece ser mais adequado, considerando os aminoácidos que participam da interação. As tirosinas podem auxiliar no reconhecimento e estabilização do ligante, como evidenciado pela interação π - π que entre o resíduo Tyr-153 e a arginina presente no sNPF (Figura 26).

Tabela 3 – Ranking da afinidade de ligação obtida com *docking* molecular às cegas e direcionado.

Tipo de <i>docking</i>	Afinidade de ligação (kcal/mol)	Resíduos envolvidos	Interações principais
Direcionado	-10,045	Tyr-64, Tyr-153*, Glu-155, Glu-156, Arg-164, Asp-246, Asn-258	Interações polares e interação π - π *
<i>Blind</i>	-9,731	Glu-154, Val-226	Interações polares

Figura 26 – Sítio de interação Rpro-sNPFR (modelo SWISS-Model) com o ligante Rpro-sNPF. Linha pontilhada amarela representa as interações polares; linha pontilhada rosa representa as interações π - π .



5 CONCLUSÃO

Por fim, a sequência codificante de Rpro-sNPFR foi elucidada, mas a hipótese de que se tratar de um evento de *splicing* alternativo entre transcritos de genes diferentes ainda requer investigação mais aprofundada. A replicação de métodos computacionais, que já se mostraram eficientes anteriormente, pode trazer novas descobertas para os estudos envolvendo esse inseto. Aliado a isso, a caracterização funcional do gene é interessante para a total compreensão dos seu papel sobre o comportamento desse inseto.

A análise *in silico* da proteína prevista se mostrou condizente com as características esperadas para essa classe de receptores. Além disso, a ancoragem molecular entre a proteína e o peptídeo maduro reforça que se trata de um receptor cujo ligante endógeno é o sNPF. Esse achado, contudo, não exclui a necessidade de realização de uma validação estrutural por técnicas experimentais.

Em resumo, o mecanismo regulatório da fome por sNPF/sNPFR está intrinsicamente associado ao sistema olfativo de *R. prolixus*. Ao final da digestão, ocorre um aumento na expressão do gene sNPF em células neurosecretoras e da presença do neuropeptídeo na região glomerular que conecta neurônios do SNC com neurônios sensoriais olfativos, onde colocalizam-se o receptor de sNPF (sNPFR) e o correceptor olfativo (Orco). Com a ativação do sNPFR, ocorre um aumento na expressão de Orco, ocasionando no aumento da sensibilidade olfativa, que facilita a busca por alimento (hospedeiro).

6 PERSPECTIVAS

No contexto do controle vetorial de *Rhodnius prolixus* e outros insetos-praga, a técnica de RNA de interferência (RNAi) se destaca como uma abordagem capaz de silenciar genes de interesse para modular o comportamento ou reduzir a sobrevivência dos insetos. A aplicação dessa técnica, aliada à realização de bioensaios, possibilita investigar a relação entre sNPF/sNPFR e o correceptor olfativo (Orco), fornecendo informações sobre o papel desse sistema neuropeptidérgico no comportamento e suas possíveis interações com o sistema olfativo.

Além disso, a análise proteômica antes e após o silenciamento gênico permitiria traçar um perfil proteico detalhado dos insetos, possibilitando a identificação de proteínas diferencialmente expressas. Isso permitiria avaliar se as alterações observadas em nível translacional impactam a síntese proteica, revelando mecanismos compensatórios e potenciais novos alvos para estudo.

Por fim, compreender a função fisiológica desse neuropeptídeo contribuirá para o desenvolvimento de estratégias inovadoras de controle vetorial, incluindo a identificação de alvos para RNAi e o aprimoramento de armadilhas para monitoramento populacional.

7 REFERÊNCIAS

- ALMEIDA, C. E.; LIMA, M. M.; COSTA, J. Ecologia dos vetores. Em: **Vetores da doença de chagas no Brasil**. SciELO Books, 2014. p. 210–235.
- AMIR, M. B. *et al.* Short Neuropeptide F and Its Receptor Regulate Feeding Behavior in Pea Aphid (*Acyrtosiphon pisum*). **Insects**, v. 13, n. 3, p. 282, 13 mar. 2022.
- BERNHARD, H. *et al.* Structural basis of Spliced Leader RNA recognition by the *Trypanosoma brucei* cap-binding complex. **Nature Communications** **2025 16:1**, v. 16, n. 1, p. 1–17, 15 jan. 2025.
- BESTEIA, L. *et al.* The short neuropeptide F regulates appetitive but not aversive responsiveness in a social insect. **iScience**, v. 25, n. 1, 21 jan. 2021.
- BESTEIA, L. *et al.* The short neuropeptide F (sNPF) promotes the formation of appetitive visual memories in honey bees. **Biology letters**, v. 18, n. 2, 2022.
- BLAND, J. M.; ALTMAN, D. G. Statistics Notes: Bootstrap resampling methods. **BMJ**, v. 350, p. 2622, 2 jun. 2015.
- BRITO, N. F.; MOREIRA, M. F.; MELO, A. C. A. A look inside odorant-binding proteins in insect chemoreception. **Journal of Insect Physiology**, v. 95, p. 51–65, dez. 2016.
- BROECK, J. VANDEN. Neuropeptides and their precursors in the fruitfly, *Drosophila melanogaster* ☆. **Peptides**, v. 22, n. 2, p. 241–254, fev. 2001.
- CAERS, J. *et al.* More than two decades of research on insect neuropeptide GPCRs: an overview. **Frontiers in Endocrinology**, v. 3, 2012.
- CHOLEWIŃSKI, M. *et al.* Short neuropeptide F in integrated insect physiology. **Journal of Zhejiang University-SCIENCE B**, v. 25, n. 5, p. 389–409, 17 maio 2024.
- CHRIST, P. *et al.* Feeding-induced changes in allatostatin-A and short neuropeptide F in the antennal lobes affect odor-mediated host seeking in the yellow fever mosquito, *Aedes aegypti*. **PLOS ONE**, v. 12, n. 11, p. e0188243, 22 nov. 2017.
- CHRIST, P. *et al.* Functional characterization of mosquito short neuropeptide F receptors. **Peptides**, v. 103, p. 31–39, maio 2018.
- COSTA, M. C. *et al.* Sugar feeding in triatomines: a new perspective for controlling the transmission of Chagas disease. **Frontiers in Physiology**, v. 15, 15 out. 2024.
- COSTA-DA-SILVA, A. L. *et al.* Female *Aedes aegypti* mosquitoes use communal cues to manage population density at breeding sites. **Communications Biology**, v. 7, n. 1, p. 143, 31 jan. 2024.

DAINES, B. *et al.* The *Drosophila melanogaster* transcriptome by paired-end RNA sequencing. **Genome research**, v. 21, n. 2, p. 315–324, fev. 2011.

DIAS, J. C. P. Globalização, iniquidade e doença de Chagas. **Cadernos de Saúde Pública**, v. 23, n. 1, p. 13–22, 8 fev. 2007.

DILLEN, S. *et al.* Functional Characterization of the Short Neuropeptide F Receptor in the Desert Locust, *Schistocerca gregaria*. **PLoS ONE**, v. 8, n. 1, p. e53604, 4 jan. 2013.

DILLEN, S. *et al.* Identification of the short neuropeptide F precursor in the desert locust: Evidence for an inhibitory role of sNPF in the control of feeding. **Peptides**, v. 53, p. 134–139, mar. 2014.

DONG, Y.; TAYLOR, H. E.; DIMOPOULOS, G. AgDscam, a Hypervariable Immunoglobulin Domain-Containing Receptor of the *Anopheles gambiae* Innate Immune System. **PLOS Biology**, v. 4, n. 7, p. e229, 2006.

DORN, R.; REUTER, G.; LOEWENDORF, A. Transgene analysis proves mRNA trans-splicing at the complex mod(mdg4) locus in *Drosophila*. **Proceedings of the National Academy of Sciences of the United States of America**, v. 98, n. 17, p. 9724–9729, 14 ago. 2001.

EVANS, J.; SHENEMAN, L.; FOSTER, J. Relaxed neighbor joining: A fast distance-based phylogenetic tree construction method. **Journal of Molecular Evolution**, v. 62, n. 6, p. 785–792, 28 jun. 2006.

FADDA, M. *et al.* Regulation of Feeding and Metabolism by Neuropeptide F and Short Neuropeptide F in Invertebrates. **Frontiers in Endocrinology**, v. 10, n. FEB, 2019a.

FRANCO, T. A. *et al.* Silencing the odorant receptor co-receptor RproOrco affects the physiology and behavior of the Chagas disease vector *Rhodnius prolixus*. **Insect Biochemistry and Molecular Biology**, v. 69, p. 82–90, 1 fev. 2016.

FRANCO, T. A. *et al.* Reverse chemical ecology-based approach leading to the accidental discovery of repellents for *Rhodnius prolixus*, a vector of Chagas diseases refractory to DEET. **Insect Biochemistry and Molecular Biology**, v. 103, p. 46–52, dez. 2018.

FRANCO-PAREDES, C. *et al.* A deadly feast: Elucidating the burden of orally acquired acute Chagas disease in Latin America – Public health and travel medicine importance. **Travel Medicine and Infectious Disease**, v. 36, p. 101565, jul. 2020.

GONZALEZ, R.; ORCHARD, I. Physiological Activity of Neuropeptide F on the Hindgut of the Blood-Feeding Hemipteran, *Rhodnius prolixus*. **Journal of Insect Science**, v. 9, n. 57, p. 1–14, jul. 2009.

GULLAN, P. J.; CRANSTON, P. S. **Insetos: Fundamentos da Entomologia**. 5. ed. Rio de Janeiro: Roca, 2017.

GULLO, C. E. *et al.* Formas digestivas da doença de Chagas e carcinogênese: um estudo de associação. **Revista do Colégio Brasileiro de Cirurgiões**, v. 39, n. 2, p. 146–150, abr. 2012.

- HANSSON, B. S.; STENSMYR, M. C. Evolution of insect olfaction. **Neuron**, v. 72, n. 5, p. 698–711, 8 dez. 2011.
- HONG, C. *et al.* Structures of active-state orexin receptor 2 rationalize peptide and small-molecule agonist recognition and receptor activation. **Nature Communications** **2021 12:1**, v. 12, n. 1, p. 1–11, 5 fev. 2021.
- HORIUCHI, T.; AIGAKI, T. Alternative trans-splicing: a novel mode of pre-mRNA processing. **Biology of the Cell**, v. 98, n. 2, p. 135–140, 1 fev. 2006.
- HUMMON, A. B. *et al.* From the Genome to the Proteome: Uncovering Peptides in the Apis Brain. **Science**, v. 314, n. 5799, p. 647–649, 27 out. 2006.
- KONG, Y. *et al.* The evolutionary landscape of intergenic trans-splicing events in insects. **Nature communications**, v. 6, 2 nov. 2015.
- LABRADOR, M. *et al.* Protein encoding by both DNA strands. **Nature**, v. 409, n. 1000, p. 100, 22 fev. 2001.
- LAGERSTRÖM, M. C. *et al.* The Ancestry of the Prolactin-Releasing Hormone Precursor. **Annals of the New York Academy of Sciences**, v. 1040, n. 1, p. 368–370, 1 abr. 2005.
- LASDA, E. L.; BLUMENTHAL, T. Trans-splicing. **Wiley interdisciplinary reviews. RNA**, v. 2, n. 3, p. 417–434, maio 2011.
- LAWRENCE, C. B. *et al.* Alternative role for prolactin-releasing peptide in the regulation of food intake. **Nature Neuroscience** **2000 3:7**, v. 3, n. 7, p. 645–646, jul. 2000.
- LEAL, W. S. Odorant reception in insects: roles of receptors, binding proteins, and degrading enzymes. **Annual review of entomology**, v. 58, p. 373–391, 7 jan. 2013.
- LEE, K. S. *et al.* *Drosophila* short neuropeptide F regulates food intake and body size. **Journal of Biological Chemistry**, v. 279, n. 49, p. 50781–50789, 3 dez. 2004.
- LEE, K.-S. *et al.* *Drosophila* short neuropeptide F signalling regulates growth by ERK-mediated insulin signalling. **Nature Cell Biology**, v. 10, n. 4, p. 468–475, 16 abr. 2008.
- LI, H. *et al.* The short neuropeptide F receptor regulates olfaction-mediated foraging behavior in the oriental fruit fly *Bactrocera dorsalis* (Hendel). **Insect Biochemistry and Molecular Biology**, v. 140, p. 103697, 1 jan. 2022.
- LI, J. A fast neighbor joining method. **Genetics and Molecular Research**, v. 14, n. 3, p. 8733–8743, 2015.
- LI, Y. *et al.* Molecular mechanism of prolactin-releasing peptide recognition and signaling via its G protein-coupled receptor. **Cell Discovery** **2024 10:1**, v. 10, n. 1, p. 1–12, 3 set. 2024.
- LIDANI, K. C. F. *et al.* Chagas Disease: From Discovery to a Worldwide Health Problem. **Frontiers in Public Health**, v. 7, 2 jul. 2019.

- LIESCH, J.; BELLANI, L. L.; VOSSHALL, L. B. Functional and Genetic Characterization of Neuropeptide Y-Like Receptors in *Aedes aegypti*. **PLoS Neglected Tropical Diseases**, v. 7, n. 10, p. e2486, 10 out. 2013.
- LIU, B. *et al.* Identification of the Short Neuropeptide F and Short Neuropeptide F Receptor Genes and Their Roles of Food Intake in *Dendroctonus armandi*. **Insects**, v. 12, n. 9, p. 844, 19 set. 2021.
- LIVAK, K. J.; SCHMITTGEN, T. D. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the $2^{-\Delta\Delta CT}$ Method. **Methods**, v. 25, n. 4, p. 402–408, dez. 2001.
- MAJEROWICZ, D. *et al.* Looking for reference genes for real-time quantitative PCR experiments in *Rhodnius prolixus* (Hemiptera: Reduviidae). **Insect Molecular Biology**, v. 20, n. 6, p. 713–722, 19 dez. 2011.
- MARASCO, L. E.; KORNBLIHTT, A. R. The physiology of alternative splicing. **Nature reviews. Molecular cell biology**, v. 24, n. 4, p. 242–254, 1 abr. 2023a.
- MARCINIAK, P. *et al.* Short neuropeptide F signaling regulates functioning of male reproductive system in *Tenebrio molitor* beetle. **Journal of comparative physiology. B, Biochemical, systemic, and environmental physiology**, v. 190, n. 5, p. 521–534, 1 set. 2020.
- MARIANI, V. *et al.* IDDT: a local superposition-free score for comparing protein structures and models using distance difference tests. **Bioinformatics**, v. 29, n. 21, p. 2722–2728, 1 nov. 2013.
- MAYER, C. M.; BELSHAM, D. D. Insulin directly regulates NPY and AgRP gene expression via the MAPK MEK/ERK signal transduction pathway in mHypoE-46 hypothalamic neurons. **Molecular and cellular endocrinology**, v. 307, n. 1–2, p. 99–108, 13 ago. 2009.
- MERK, A. *et al.* Breaking Cryo-EM Resolution Barriers to Facilitate Drug Discovery. **Cell**, v. 165, n. 7, p. 1698–1707, 16 jun. 2016.
- MERTENS, I. *et al.* Characterization of the short neuropeptide F receptor from *Drosophila melanogaster*. **Biochemical and Biophysical Research Communications**, v. 297, n. 5, p. 1140–1148, 2002.
- MIRDITA, M. *et al.* ColabFold: making protein folding accessible to all. **Nature Methods** **2022 19:6**, v. 19, n. 6, p. 679–682, 30 maio 2022.
- NAGATA, S. *et al.* Effects of Starvation on Brain Short Neuropeptide F-1, -2, and -3 Levels and Short Neuropeptide F Receptor Expression Levels of the Silkworm, *Bombyx mori*. **Frontiers in Endocrinology**, v. 3, 2012.
- NÄSSEL, D. R. *et al.* A large population of diverse neurons in the *Drosophila* central nervous system expresses short neuropeptide F, suggesting multiple distributed peptide functions. **BMC neuroscience**, v. 9, 19 set. 2008.
- NUNES-DA-FONSECA, R. *et al.* *Rhodnius prolixus*: From classical physiology to modern developmental biology. **genesis**, v. 55, n. 5, 22 maio 2017.

ONS, S. *et al.* Neuropeptide precursor gene discovery in the Chagas disease vector *Rhodnius prolixus*. **Insect Molecular Biology**, v. 20, n. 1, p. 29–44, 19 fev. 2011.

ONS, S. *et al.* Identification of G protein coupled receptors for opsins and neurohormones in *Rhodnius prolixus*. Genomic and transcriptomic analysis. **Insect Biochemistry and Molecular Biology**, v. 69, p. 34–50, 2016.

ORCHARD, I. *et al.* Malpighian tubules of *Rhodnius prolixus*: More than post-prandial diuresis. **Frontiers in Insect Science**, v. 3, 30 mar. 2023.

PAK, M. A. *et al.* Using AlphaFold to predict the impact of single mutations on protein stability and function. **PLOS ONE**, v. 18, n. 3, p. e0282689, 1 mar. 2023.

PAULO, D. F. *et al.* Disruption of the odorant coreceptor Orco impairs foraging and host finding behaviors in the New World screwworm fly. **Scientific Reports**, v. 11, n. 1, p. 11379, 31 maio 2021.

PENG, Z. L.; YANG, J. Y.; CHEN, X. An improved classification of G-protein-coupled receptors using sequence-derived features. **BMC Bioinformatics**, v. 11, n. 1, p. 1–13, 9 ago. 2010.

PEREIRA, L. DOS S. *et al.* Clinical And Epidemiological Profile Of Elderly Patients With Chagas Disease Followed Between 2005-2013 By Pharmaceutical Care Service In Ceará State, Northeastern Brazil. **Revista do Instituto de Medicina Tropical de São Paulo**, v. 57, n. 2, p. 145–152, abr. 2015.

PINZI, L.; RASTELLI, G. Molecular Docking: Shifting Paradigms in Drug Discovery. **International journal of molecular sciences**, v. 20, n. 18, 1 set. 2019.

REHAN, S. M.; BULOVA, S. J.; O'DONNELL, S. Cumulative Effects of Foraging Behavior and Social Dominance on Brain Development in a Facultatively Social Bee (*Ceratina australensis*). **Brain Behavior and Evolution**, v. 85, n. 2, p. 117–124, 1 maio 2015.

ROOT, C. M. *et al.* Presynaptic facilitation by neuropeptide signaling mediates odor-driven food search. **Cell**, v. 145, n. 1, p. 133–144, 1 abr. 2011.

SANTOS, L. B. *et al.* Drastic reduction in the notification of acute cases of Chagas disease in the Northeast region of Brazil. Epidemiological evaluation in the period 2001–2021. **Acta Tropica**, v. 256, p. 107267, 1 ago. 2024.

SHAO, W. *et al.* Alternative splicing and trans-splicing events revealed by analysis of the *Bombyx mori* transcriptome. **RNA**, v. 18, n. 7, p. 1395, jul. 2012.

SIMÕES, M. V. *et al.* Chagas Disease Cardiomyopathy. **International Journal of Cardiovascular Sciences**, 2018.

SOSSIN, W. S.; FISHER, J. M.; SCHELLER, R. H. Cellular and molecular biology of neuropeptide processing and packaging. **Neuron**, v. 2, n. 5, p. 1407–1417, maio 1989.

STERKEL, M. *et al.* OKB, a novel family of brain-gut neuropeptides from insects. **Insect Biochemistry and Molecular Biology**, v. 42, n. 7, p. 466–473, 1 jul. 2012.

TAKAYANAGI, Y.; ONAKA, T. Roles of prolactin-releasing peptide and RFamide related peptides in the control of stress and food intake. **FEBS Journal**, v. 277, n. 24, p. 4998–5005, dez. 2010.

TANG, T. *et al.* Receptor-specific recognition of NPY peptides revealed by structures of NPY receptors. **Science Advances**, v. 8, n. 18, p. 1232, 1 maio 2022.

TRIBLE, W. *et al.* orco Mutagenesis Causes Loss of Antennal Lobe Glomeruli and Impaired Social Behavior in Ants. **Cell**, v. 170, n. 4, p. 727–735.e10, ago. 2017.

URVAS, L. *et al.* Benchmarking AlphaFold-Generated Structures of Chemokine-Chemokine Receptor Complexes. **Journal of Chemical Information and Modeling**, v. 64, n. 11, p. 4587–4600, 10 jun. 2024.

VERLINDEN, H. *et al.* The pleiotropic allatoregulatory neuropeptides and their receptors: A mini-review. **Journal of Insect Physiology**, v. 80, p. 2–14, 1 set. 2015.

WANG, C.; CHIN-SANG, I.; BENDENA, W. G. The FGLamide-allatostatins influence foraging behavior in *Drosophila melanogaster*. **PloS one**, v. 7, n. 4, 27 abr. 2012.

WATANABE, T.; KUBO, T. A new antigenic marker specifically labels a subpopulation of the class II Kenyon cells in the brain of the European honeybee *Apis mellifera*. **BIOPHYSICS**, v. 11, p. 73–77, 2015.

WATERHOUSE, A. *et al.* SWISS-MODEL: homology modelling of protein structures and complexes. **Nucleic Acids Research**, v. 46, n. W1, p. W296–W303, 2 jul. 2018.

WULFF, J. P. *et al.* Orcokinin neuropeptides regulate ecdysis in the hemimetabolous insect *Rhodnius prolixus*. **Insect Biochemistry and Molecular Biology**, v. 81, p. 91–102, 1 fev. 2017.

YAMASHITA, M. *et al.* Involvement of Prolactin-Releasing Peptide in the Activation of Oxytocin Neurones in Response to Food Intake. **Journal of Neuroendocrinology**, v. 25, n. 5, p. 455–465, 1 maio 2013.

YU, S. *et al.* Brahma regulates a specific trans-splicing event at the mod(mdg4) locus of *Drosophila melanogaster*. **RNA biology**, v. 11, n. 2, p. 134–145, 2014.

ZHAO, J. *et al.* Structural basis of odor sensing by insect heteromeric odorant receptors. **Science**, v. 384, n. 6703, p. 1460–1467, 28 jun. 2024.

ZHAO, J. *et al.* Complicated gene network for regulating feeding behavior: novel efficient target for pest management. **Pest Management Science**, v. 81, n. 1, p. 10–21, 10 jan. 2025.

ZHOU, Y. *et al.* Research progress on insect short neuropeptide and its potential applications in pest control. **Chinese Journal of Pesticide Science**, v. 26, n. 6, p. 1025–1036, 2024.

ANEXO I

SEQUÊNCIAS GENÔMICAS RPRC002266, RPRC002268, RPRC002269
(íntons destacados em amarelo)

>RPRC002266 | tamanho = 2168 pb

ATGAACAATACTACCATTGAAGAAGATCTGTCCATGATCGTGGACTGTATTGTAACCTCAGTATAACATTAGTCAA
TATAAAGGTCAATGGCCAAATTGTACAGAAATAGGTTTTAGAAAAGATATAATTGATGATAAAATCGTACAAGCT
ATATTTTGTATTATTATATACATCTATATTTGTGTTAGGTCTATTTGGTAATATATTAGTTTGTACGTAGTCCGT
AGGAATCGTGCTATGCATACGGTTACTAATTGTTTTATAACGAATCTAGCATTGTCAGACATTTTGTATGCACA
TTGGCTGTACCATTACACCGTTATATTCATTTCTAGGtaagtattatttccatttttattgatttaatttatat
tttattgaataagtctagactagccagtaagcgtagtgatgtaaccatacacaaatcaaatataattgagcggtt
cactgcaaaagtcgattaaactacataaaaacaaaaagtagagaaatgcaatgaaataatgtacatttttttaaaa
aaattcttgcttataataaaattcaataaaattgttaccattcagcgcctcctacttgaatttgagttgaatta
ttgctgagcacggagttaatcaaaatggccttctgaaataaattatttaataatttaataaccaaataaattttgtc
atttaaatggcgtaatactttattattttgggaaattttattacaaaatataaaaaagctatgaactgaaccaagtgc
tttgacacttaaatttaagataactaatttttaaaaaggtggagttaatcgattttgtgtagtgaacggctccctata
taactaacctaagtgtcgataaaaattatagacgtatgggaaatcggaagcgaagaattaaataggagaagaag
gtaggaagttaaaagtgaagaatcagtttgagttgggcagttgttatcaagttgtaccagttgaagacgttcttt
agttttaagtagtacacagctggtttaaaggagtttggttttctaaagcgttgctaggttacggatagggtttttgg
ttattggctaattaaagcagttgacgttgagaggttacaaaaaacattcatttttgatcgaatgttttttaggtt
atttatttttttaaaaaaaaggtgttaaatgtggtgtttttaaagttgaataaagtagtgatttctgcattaat
ttttgtgggtaaggggtcattgtggtataaaaagagaatgaaattgttgataaacttgcaaaaggaagccaccaat
agtgggtaaatggctaattataaaattaccacacaccgatatcaagtgtataattgcaaaagcagcactacaaaaa
atggccagacagattacagtaaatggtaataaaggagattctataaaactatgtgtagatcagttacctagacaa
ccatggtttaaggaactgacactaataaatgggtgtttacgacaataagcaggctgagagcaaaccaggcggtg
tgcgatcgtatcttcacaggatcaacaaaaaggtcctctctcctctttgcgtagattgcaatgaggaggagaat
ttcaaacatataataatgatatgtccacgttatgtagtagaaaggaaacgaatgttgacgacatgtacaggtat
ttagatgcaccatttcgctaggaagacataatattctccagcaacatctatgagcttaaatctgtagcagagcta
gcaatgaaatatgtatgtatttaaaagttattgccgattagaatgaatgcaagaaaatgattagagaagaattac
gtgatttttaatttcggttaagctagaagaatggcaggagcacagagtactgatccatccattagctgtagtttat
gtaaatagttgtatatagtagagggatgaatgtgtgaagcgaatgtattataagtaagaatggagagtgagtga
tattctagtgtacgtacgtatgctggctgtatggcgaaagtcaatgccaaaaaagcaaaaaaagtagg
gacttataaagttattttcaataaaaaacctctattttgtcaaatcctcgacgctgtacctatacttgagttcgtag
aatttgtagtttaaaagtgaatctttggatcttgaagaattacttatgtaaatttttacttttagGTTGA

>RPRC002268 | tamanho = 22673 pb

ATGCCTACCCAGAAGGATAATTATAATAGAAATATTGACTTTATTGGAGGACGTGAGAAAGgtaagaatagaaaa
aagaatggataaataaccaggccggaatatgggaataacggatgcgttagaagagttaatagggaacggatggt
aagtagcttaggacaactaaagagtggtgttattgctgagaaatgggttaacaacaaaacaaatacagagaaaa
taaccaattaatatcatgcgatacattattcttgtaacaataatcttcactgcctaacagaattaacatactaa
actttatttaccaaagaatttggttttttagttggaagtgaagatgaggtagtataaagaaaacggtagtaa
aattccagttattttattaaactcagccataaattattattagcagtcctaacacgcttttcaaaacggatttggtg
tttatttcttacaagtttgtagcgttaagatgccttaattcttcttcgacgttagaagctaaagtcttaagccgta
cttagtaagttactttaaatagtttaataacaaagtttctaggttaacaaaaacatgattattgttttaaaacg
ctgtagcgttttgtaataaaaaaagtagaaataggaaaaagggaaaagttatacataattaaaaaagaaaccttac
aaaaatatctactattatccgtattaaatagcattattatgccatttgatgatatttaatttacacaaaatctta
ttaatttaattaaaagttactttatttttaattaatgataaagtaaaatggaaaagtaaccctcaaaataaataaa
taataataataaagcttaacatgtctagctttcgaaagtaataataaaaataactacttttaatttttaaaaaa
aattaaattttataataataaaatgtaataatttttaataattttgtttctgggataaacgtgatatttta
aggttaaaaaaaagtagtagaactatattttttgtcctaaggtcactaggtcatggaaatttaataaatttaaat
actgttctttattaaatattttacaaatatttaaatgtaaaattcttttttggtacttttaataattttatttagtc

tagagagtaatttcaagggtttcagaattttttttcacattacttaggtaattgaaagttttgaaaaataatttt
 ttcgatgttaaggttgcggattttctaaaaaatttttaacttgataaaacaagagagaccgagtgtaaacgttga
 atggatttttcttctttttttctgagatttttctcatgatttctttaaattctgagtatatgtagataaaat
 ttaatggattgattccaaacacaagcaattcttttaaaagttaatcaattcaaaaacttgtaaaaaacaaagttga
 aaataagtcaataccattaactaaacaattttatgagcagggcctttcttatcacctaatttaagcaaggatttag
 gtgagagagaactgttcttatcacgttcagcctatttaagctgatacgagctttaaatatttaagaaatagaatc
 aacactctaaattgatctaaaaatgttcggaatagtaaaatcaattttcttctcattctgtcttttaaatcaatg
 ttcacagtgaaatagtgatttaacagaattttaaaatttttaaaagctcttttgagcttctgtgttcggattgtttg
 gaatgcaaaaaaatggtgacgaacttattcacctacaacagtagagctgaattttaaaattttcaaagctcttttg
 aacttccgtgttctcggattgtttggaatgcaaaaaaatggtgacgaacttattcaccttcaacaatagagctga
 atttaaaattttcaaagctcttttgaaacttccgtgttctcggattgtttggaatgcaaaaaaaagtgacgaact
 tattcaccttcaacaatagagctgaattttaaaattttcaaagctcttttgaaacttccgtgttctcggattgtttg
 gaatgcaaaaaaaagtgacgaacttattcaccttcaacagtagagctgaattttaaaattttcaaagctcttttg
 aacttccgtgttctcggattgtttggaatgcaaaaaaaagtgacgaacttattcaccttcaacagtagagctgaa
 ttttaaaatttttaaagctcttttgagcttctgtgttcggattgtttggaatgcaaaaaaatggtgacgaacttat
 tcacctacaacagtagagctgaattttaaaatttttaaagctcttttgagcttctgtgttcggattgtttggaatg
 caaaaaaatggtgacgaacttattcacctacaacagtagagctgaattttaaaatttttaaagctcttttgaaactt
 ccgtgttctcggattgtttggaatgcaaaaaaaagtgacgaacttattcacctacaacagtagagctgaatttaa
 aatttttaaagctcttttgagcttctgtgttcggattgtttggaatgcaaaaaaatggtgacgaacttattcacc
 tacaacagtagagctgaattttaaaattttcaaagctcttttgaaacttccgtgttctcggattgtttggaatgcaa
 aaaaaaagtgacgaacttattcaccttcaacaatagagctgaattttaaaattttcaaagctcttttgaaacttccg
 tgttctcggattgtttggaatgcaaaaaaatggtgacgaacttattcaccttcaacaatagagctgaatttaaaa
 ttttcaaagctcttttgaaacttccgtgttctcggattgtttggaatgcaaaaaaaagtgacgaacttattcacc
 tcaacaatagagctgaattttaaaattttcaaagctcttttgaaacttccgtgttctcggattgtttggaatgcaa
 aaaaaaagtgacgaacttattcaccttcaacagtagagctgaattttaaaattttcaaagctcttttgaaactttt
 gtgttctcggattgtttggaaggccaaaaaattgtcgaacttattcaccttatttgatctttcttggttggttaa
 atctgaagaaaagaaggaaactagataattcctaggttaacatgtcggcaatgaagtgaacgtggcagagttgaga
 ttattttatctgcggggcggtgtattcgtcagacatggacgaatgaggaccttgggaagaagcaaggaatgaa
 tttaaaagaaaacaatgaagaaaaaatgtttaaaatggtacagatgcttgtggagaatggtcccaggaaatgtg
 gctaaccgaaccagaaaatagaaccctcctagggaaagaaaagatccacaaagattttggacgaggaatttagac
 aaccaaatggtggaaagtgtatttagagatgtagattggtgaagtacctgaagggtaacctagaacccgtaaat
 tcaataaataaataaggaaatttaactaaactcaattcagcgggtgaacttttatatacaaaaaaatagtaattatct
 gattaaaatcataatgttcagatgggatgtatcataatgatgtatttagttggcttacattacttagtaacatct
 ctgacaaagagatactccgtaagaaagccactaaacataaccttacattgttacattccagataacaactctttt
 aatctattttatttaaaaataaaaaatggcacaaatcaatccagaacgcacaatttagaagttaccctagttctacc
 aaaactatgcagttggagattttactaagttacgaaccagaatctgttggctcagattgcaccagttcttattgaat
 gctctagatccgtctgtttgttgcttagcctgcaattccctaaatatctgcagctcaactcactgtaactaata
 agctttcatctgaatggtctgaaaattatctgcttattcaattttgacttgcattaaaaataattataaaaaataat
 agcgtttctaagtttattcttacgtacgtctcttaagcatttttagagctatatataaaatgtttacacattaat
 tacttgttcggctcgtgaccgagaacattcggccagtttggttagcgtccgactctcgcgagtcagcctcgcgcc
 gcgcgcgtactaatgtacgcattcgccttaatttctttttctacatatcttctatcgcttattactaagttat
 tattttataaccttttttattctttaaactttccatctatctcaacttgtttccgtttcttatgccaaatgcgtc
 actacctatagcttctatttcttttccccgttgtttgtgtgtctatattctagacctttgccccactacttctt
 tttggaccatgtggtcactcctccagctctatgcagttccgtgctaggcattgcttcagacgcttctttgctaatt
 aagcttttagctacctttaaccatccccgcctgctacgcacacttccgctgacgcctagtgttctctgtggggc
 attgcatcctacgcttctgtgtgtgttttgggtgttcttctaaacgcacatctcaactcgctacgcttctactgccc
 ctgacacctagcagtcgcgcatcatccgcatctattgcaactataaacttccctcaactgcatctagtgcattatt
 gcactattcccccttactcttccgcttacctcacttacacttccgatacgttctctgagcctgagcccaaatat
 tgctaccttttctcagagctcaacacacacacacacttacacaaacatatacaccgcttccgctccacagcttctt
 tttctcccatcgcttctcggcgcttctgtcaacatacctcacgcacatcgttccttgaccaaagggttaagtac
 ctctctctttttctctattttttttgtttactattacttcttacagcgccacttgggcaccttagcatctggg
 tgtaggcggttctcgcgcatctacattactatttttaaaattcatcaaggtaaacacattttttattaaagctcttta
 tactaattcagtcataatagtagaacttatagggaatggaatggaatgccacttctttagaaaactttagaaaatc
 taattcataagattcataatgaccactggttttgacaatgattatgaaaactacataaaaaattgtgaaaaatga
 cgtaaacccctttgogatatccctaatatttgcgagatatttaagatcaaaattacaattttcatcttattcg
 agaacgggtttcaatattaattttccaccagaaatacaccagttttgggaggtcctctgggtgcaacaaaaatgt
 tattggtgacgcaagaacaaaaataataaataatgaagattttaagaatgtagtgcaatgtaatggctgtggg
 tgaatcaggaaaaaagttgttttatcatcaataaaaagcaaccatcacatcatttttaaggtaccagtagtgc
 agacttcttgacctatgattacggttaagaatatcttgccaaggaggactgggtgttggctcatgagcgtggtttt
 taagtgggagatccccagtttgaatccgcgagtgacgtcattcacaatacttaatacatcataacactttttta
 aataaaatgtagtcataagtggtgattcgaaactggggatctccacttaaaaagccacgctcttgatcactagacc

[illegible]

tatgtatcagaaatcttgcatagaagtgaagcaatatattacatgaaagaacgcctgcatccaagctcagaattg
 ttgccaatctcggttagttagaattgtttaaaaatggttagacattataattaacaaagcgacaatgtaagc
 caggtatgtaatacgcgcgaccaggacagtctatatctctttctgattatatgtcctaattatatgatttaataa
 tattgtatataagaagttcatagactgcgtaggagttatttaagaatttaaagtgtaataagtgtaaaatacag
 ttgtaagtttgctattaagtaaaagcaagtaaaatataagtaactagaagaatgcttgattttaacgaggtagaat
 ggaaatagatctgagcacaagagatgagggcactagtcccaaacttgaaagatgaaagaaatatgtaaaacagttt
 caaggtttatgaaatgcataacctaaaaatcttaaaaaacaaatgaaacacatttatggtaatggaaaggtctaaa
 ttaaatatattagatggaatcacatgtaaaaaaaatcaaagtaatttaagatattaagtaattggcaaaaaatcc
 cgtcaaaaaactaagtaaaatgcaacttttgatcagtaggacatactagtataggagcttattgtagaatctgt
 gtgtgaaggtggggcagcagcaagagtaatatcagccaatgcgtattaagttcccatatgtaaaacatatggcc
 agtgtacttctttctgattggctaaaaatatcttgatactctgccacacctatttactccgattctacagtttagc
 tcctacagcaacatgttaccagatacaaaatttgaaattttgttgacttttaagaggatttttttttaaaatt
 ataaatgcaacttttatatcagttggaaatactggtataggagcttattgcagaatctgtgtgtgatgggtggaa
 cagtggaaggtgaaatcagccaatgcgtattaagagtcctatctgtacataaacacgcccactgtacttcattct
 gattggctgaaaccaacttgatattcgccacacacttttactgggattctaccttaactcctgcaggattatgt
 taccagatacaaaacttctattttttgtagcatttttaaggggattgtttttaattgtgtactctttttcgtat
 ttacctagtttctgtaattaatagacaccaaactcattatgttttaattagtaaaatttaaaagatgcattagtata
 tagaatttgaaattgtaaaattttcaattttttacccattgtaaatgaaatcaaaagtgtattgtaaaattttgtg
 aggaagcgttgtaataaatgtatgtaataataaagcgttgtaaaaaatccaaaccgttgtagagttttacatcatt
 cctttgaaatcaaattaaattaaaaataaattactagtttttgctattcatttactgttaataaataatgttttaa
 atagcaacttgaaataatccattactgtataatttttaattgtaaacttctatttttttttaagctttataaag
 ggaatcaatgaaacactgttataataacataaaaagaaagtgttataataaatgtatgcaataataaagtgtgtg
 ataattttattaccctttattgagttttacatcattctttataagtttataatcattgtttcttgctcggattag
 atagtgtgcttacatctgagaaacgtcagaatagtacgaaacaatgaagactgagatgaaatatcttaatatataca
 ctacgcttacaagataataactaagtaatttctagatggtatgaagtttagagaacatattttattgagagattcg
 taaaacagtttaaaaaattatgaaatgcattacctaaaagtcttataaacaatgaaacacatatatgcttatat
 aaaggtctaaattaaatatattttgactccatttaagtaaaagcaagtttagataaatgactagaagaatacttat
 ttttgacgagattgtgcggaatggatcaaagcaccgattaggtggaagtctgaaacctagaacatgacttcnn
 nnnnnnnaatatgtaaaacagtttcaaggtttatgaaatgcataacctaaaaatcttaaaaaacaaatgaaacac
 atttatggtaatggaaaggtctaaattaaatatattagatggaatcacatgtaaaaaaatcaaagtaatttaaga
 tattaagtaattggcaaaaaatcttgtaaaaaactaagtaaaatgcaacttttgatcagtaagacatactagt
 ataggagcttattgtagaatctgtgtgtgaaggtggggcagcagcaagagtaatatcagccaatgagtaataag
 tcccatctgtaaaacatatggccagtgacttctttctgattggctaaaaatatcttgatactctgccacacct
 atttactctgattctacagtttagctcctacagcaacatgttaccagttttgttgacttttatggggattttt
 tattaaatttataaatgcaacttttgatcagttggaaatactggtataggagcttattgcagaatctgtgtgtg
 atgggtggaacagtggaaggtgaaatcagccaatgcgtattaagagcctatctgtacataaacacgcccactgt
 acttcattctgattggctgaaaccatcttgatattcgccacacatcttactgtgattctaccttaactcctgc
 aggattatgttaccagatacaaaacttctattttttgtagcatttttaaggggattgtttttaattgtgtactc
 tttttcgtattttacctagtttctgtaattaatagacaccaaataattatgttttaattagtaaaatttaaaagatg
 cattagtatatagaatttgaaattgtaaaattttcaattttttacccacattgtaaatgaaatcaaaagtgtattg
 taaaaattttgtgaggaagcgttgtaataaatgtatgcaataataaagtgttgatgataattttatttaccctttatt
 gagttttacatcattcttttttaagtttataatcattgtttcttgctcggattagatagtggttacatctgaga
 aacgtcagaatagtatgaaacaatgaagactgaagtgaatatcttaatatatacactagccttacaagataataac
 taattaattttctagATGGTATGA

>RPRC002269 | tamanho = 6709 bp

ATGgctgggttactcctggtattttatctagtagccagtgagacactagcgtatttcagtaattggaattgttcct
 aggataggagttcacgcggccaatgttggttttagatgtggtggtgcgattatcaggatgtggattatagaaggag
 gaattccaacggttggttggtggatggcagaataaggggtttgttctgtgataaagggaatatttttaaatcttc
 gtgaaagcgtgtggttaaaagataataagggagcattaacaagacggcgaggacttttgattggaaggattgga
 tgaaggagctattagaagatttaggtaagccacacagctctgattagcaagtccagatgggtttaagtttaagcca
 tatagaggggttagttgttatgtgtgtttagtgtagtttccaaaacttaaaattgaattaaatcgaatttgaaaa
 aaggtgtgagagtggtatcatctgattttttccagtaaaacttactattttatgtttctagtataattcattatta
 tttttctttcctaattatgtttttctccccaaaaaataattaaattgaggttaattttgattataattttgtacc
 agttacgatgaaaaggaagatcttcattttgtgtgactagtaaaaagaaatcccatcataatttaataattaac
 ctaaaaacaaagaggtgtgctggtatgctgctttattagataaaccttctcaaagttaagatattgcacagtaaa
 agccctcagatattatgggaaaacacaaagcacgatggagaaaaactatgaagaaaatgtactaaacatagtttt
 tcttttgagactttttttttttgtaaagtacagagttttttacaggccgggattaactaccgattaggtatcgt
 ggacggaagggcggaagctcacatactttactataaacaatttttttctaacaataaaaaagtagtcagacgcgg

catccaaaattaaattagaccattggatagccaaaggatcaaagaattaattacgataagcgagcactgtgaagc
atcgcgcgcggtggagggggaagcgccgccatTTTTgagatcattttgctgctatcgagatttgcaattagtata
acttttgactcggagaagctacgatgttcgtttaagttttaatcgtcaactaatctatcgattgtatataaaaca
ttagtgtaaaatagtgacagtgaaatataatcaggtgttagaatcaaaaaatattataaaaaaatatataatgtcg
gcctgtacgcgtcgaaaagccaaaaccaaatacagagttcgtatatggttgaaagggcttctcaagaacaatcg
aatgtaattaataaaacaccatatgaccgagtcagctgtgctgccgggaacgtaaaagaatgaatgtggccgaa
cggatcaacgacggtgtccagtacatctggcagctgggtgtgttgaaattatggaagtggatgatgaaatcgctt
atacatcgacaacggactgtataggtatagattgtgttcgatagtcgatttgTTTTcttgatttcacctaata
cataaataaaaaatTTTTatacaaaaaatataaaccgttacgtaaacgtattcacgtacaacacacggccgtgtc
catgacacagtcacagttttttttttttttcttcctaggagctctaacgagagaacactgtccggaggcgctta
aatcaacaggaaaataaaaagtatatatatttggttaaggagaaataatattgtccaattcaaccacgaataaaa
attctcattttattatacagtttagtacaaaaagagtgggacagaaggaaattgatttaaattattcatcaaaaca
gaaaaaagaaaaaggaaaatagaagaaaaactgcgacaaaaaatctgaaaattttatacagcgccacaatttac
gtcggctattaaatgaaaacaaacaatgtctctgctctctgccaacaatggaaatggtagaaaccattctgagtt
tgtaagtttatttggttaaccaactagaggtacttttctgctctagcaaatcatttaaaatttagtcaaccttaa
attatttaattttaactgttttcagATTTCGTGA

ANEXO II

>sequência obtida com primer M13/pUC_anti-senso

5'

AMSGAMCGGGGCSRGTGATTGTAATACGACTCACTATAGGGCGAMKTGGGCCCCGACGTCGCATGCTCCCGGCCGC
CWTGGCGGCCGCGGGGAATTCGATCGTTCAGCCAAGCGTATAAGAATGGATTATAGCATGTAGACGACATGGCTAC
AGCATGCGATAGAAAAGAAGAATAGATTATAATAAGTCCATGACGATGTATGCATATACAGGTCATTTATTACATT
TATCAAATTTATTGGTAACCACGAAACACCGAATATGGTAACCATAGCAATTAACATTCTATTAGTTCTCTTCTT
TCTCTCTCTGTCCACTTCTTCTTTGCGGCTAGTCTTGGCGCCCGGCTTACTTCTAGCACGGTCATTTAGTTTTAC
CGACACTCGCACGTAACAAAATGTTACCACTATGAACGGTAGCACAAATTGAATTGTTGAAGTCATACCACCGAA
CACTTGTCTAAAGTTCTCTGAAGGCCATTTCTCTTCCACAATAAAATCTTCCAGCAATAGATTTATGATCCATAAA
TAATCCAAATGGTATTGTGGCAATTAATGAGAATGTCCATATAAAAAATATTACAGCCAAACATGTTGAAAGGCG
CATGCGCGGTTTTAAATGGATATATTATAACAAAAAACCGATCCACGGCTATTGAAGTTAATGTTAACGTTGATAT
ATATACGGATGTACTTTTGGGGCATATMCCACTAGATGGCACASC GCGTTACCGAATATCCMACGACCTAGAAAT
GAATATAACGGTGTAATGGTACMGCCAATGTGCATAACAAAATGTCTGAMMATGCTAGATTCGTTWTAAAACAA
TTATAACCCGTATGCMTAGCMCGATTYCC TAMCGACTACSAATCACCTAGTGAAATCSSGGCCGCTGCCAGYCS
ACATATGGGARAGCTCCCAAMSSCTTGGATGCATARCCTTGAGAWTCCWATATGCCCCCWACATASCTTGSCTA
ATCATGGGCATASSTGCTTCTGKG YAAAATTGCTAYCCGCTCMCAATYMCCCMCACWWTMSARCCCGRAMASCTA
ARATGRYAAAAGCCTAKGGGGTCCCTCAKTGAKSGAASCYRAMCTWCMATTAMATTSTASTWGTGCTWMCCTYGYCC
YMTTTTCCCAACKCGAGSRAACAACSATTGTSMGKRCKGTCTMCKAGYWCTGRTCAGYAGC 3'

>sequência reversa complementar

5'

GCTCTGACAGCTGGACGGCACAATGTTGTTCTCGGTTGGGAAAAAGGCAGGAGCACATAAATTAATGAGTGTTCT
CATGAGGACCCCTAGGCTTTTCATTTAGTTTCGGGTAGTGGGGATTGGAGCGGTAGCAATTTTCCAGAAGCATATGC
CCATGATTAGCAAGTATGTGGGGGGCATATGGATCTCAAGGTATGCATCCAAGTTGGGAGCTTCCCATATGTGCT
GGCAGGCGGCCGATTTCACTAGGTGATTGTAGTCGTAGGAATCGGCTAGCATACGGGTTATAATTGTTTTAAACG
AATCTAGCATTACAGACATTTTGTATGCACATTGGCGTACCATTTACACCGTTATATTTCATTTCTAGGTCTGTTGA
TATTCGGTAACGCGTGTGCCATCTAGTGGATATGCCCCAAAAGTACATCCGTATATATATCAACGTTAACATTAA
CTTCAATAGCCGTGGATCGGTTTTTTGTTATAATATATCCATTTAAACCGCGCATGCGCCTTTCAACATGTTTGG
CTGTAATATTTTTTATATGGACATTCTCATTAATTGCCACAATACCATTTGGATTATTTATGGATCATAAATCTA
TTGCTGGAAGATTTTATTGTGAAGAGAAATGGCCTTCAGAGAACTTTAGACAAGTGTTTCGGTGGTATGACTTCAA
CAATTCAATTTGTGCTACCGTTCATAGTGGTAACATTTTGTACGTGCGAGTGTTCGGTAAACTAAATGACCGTG
CTAGAAGTAAGCCGGGCGCCAAGACTAGCCGCAAAGAAGAAGTGGACAGAGAGAGAAAGAAGAGAACTAATAGAA
TGTTAATTGCTATGGTTACCATATTCGGTGTTCGTGGTTACCAATAAATTTGATAAATGTAATAAATGACCTGT
ATATGCATACATCGTCATGGACTTATTATAATCTATTCTTCTTCTATCGCATGCTGTAGCCATGTCGTCTACAT
GCTATAATCCATTCTTATACGCTTGGCTGAACGATCGAATTCGCGCGGCCGCCAGGCGGCCGGGAGCATGCGACG
TCGGGCCCATCGCCCTATAGTGAGTCGTATTACAATCACGCCCGTCT 3'