ESTUDO BIOINFORMÁTICO DA METALOPROTEASE AAEL011292 EM *AEDES AEGYPTI*

Autores: Arthur Alves Neves, Beatriz Alves Azevedo, Bruno Regis Lyrio,Tatiane Sanches Soares

Universidade Federal do Espírito Santo/Departamento de Biologia, Alto Universitário, S/N, Guararema - 29500-000 - Alegre-ES, Brasil, arthur.neves.89@edu.ufes.br, beatriz.a.azevedo@edu.ufes.br, bruno.ferraz@ufes.br, tatiane.s.soares@ufes.br

Resumo

O AAEL011292 é um gene de *Aedes aegypti* responsável por codificar uma proteína que apresenta características sugestivas de uma possível função como metaloprotease, sendo relevante sua caracterização molecular. O objetivo deste trabalho foi caracterizar, por meio de ferramentas de bioinformática, os aspectos estruturais e funcionais da proteína codificada pelo gene para isso, a estrutura tridimensional da proteína foi predita utilizando o AlphaFold, permitindo a análise de sua conformação e dos níveis de confiança do modelo. A caracterização funcional e a identificação de domínios proteicos foram realizadas por meio do InterProScan. Foram identificados domínios relacionados às peptidases da família M1, incluindo Peptidase_M1, Peptidase_M1_N e ERAP1_C, além de um domínio catalítico de metaloproteases do tipo “zincin”. Também foram observados domínios associados ao Tricorn interacting factor 3 e à Neutral Protease Domain 2. Os resultados indicam que a proteína codificada pelo gene AAEL011292 apresenta características estruturais e funcionais compatíveis com uma metaloprotease da família M1. Dessa forma, a análise *in silico* contribui para a caracterização da proteína e fornece informações que podem auxiliar na realização de futuros estudos experimentais sobre sua função em *A. aegypti*.

Palavras-chave: *Aedes aegypti*. AAEL011292. Metaloprotease. Bioinformática. Gene.

Área do Conhecimento: Ciências Biológicas-Biologia Molecular.

Introdução

As arboviroses constituem um importante desafio para a saúde pública mundial, especialmente em regiões tropicais e subtropicais, onde as condições climáticas e ambientais favorecem a proliferação e a disseminação dos mosquitos vetores (Fernandes *et al,*2024). No Brasil, doenças como dengue, zika, chikungunya e febre amarela apresentam elevada relevância epidemiológica em razão da ocorrência de casos e do potencial de desencadear surtos e epidemias, ocasionando impactos significativos sobre a saúde da população (Lima-Camara, 2016).

Nesse contexto, o mosquito *A. aegypti* destaca-se como um dos principais vetores de arbovírus de importância médica, devido à sua ampla distribuição geográfica, elevada capacidade de adaptação ao ambiente urbano e estreita associação com áreas densamente povoadas. Além disso, a resistência do vetor a diferentes métodos de controle constitui um obstáculo para a implementação de estratégias eficazes de prevenção e controle (Brasil, 2023; Rahman *et al.*, 2021).

Diante da importância epidemiológica das arboviroses, diferentes estratégias têm sido empregadas para o controle de *A. aegypti*, incluindo o uso de larvicidas e inseticidas. Embora essas medidas possam contribuir para a redução das populações do vetor, a utilização recorrente de compostos químicos pode favorecer a seleção de indivíduos resistentes, reduzindo a eficácia dessas estratégias ao longo do tempo. Esse processo constitui um desafio para o controle do mosquito e evidencia a necessidade de ampliar a investigação de mecanismos biológicos e moleculares envolvidos em sua sobrevivência e adaptação, visando à identificação de novos alvos que possam contribuir para o desenvolvimento de estratégias de controle mais eficazes (Dias *et al.*, 2022).

Deste modo, a investigação dos mecanismos fisiológicos, metabólicos e moleculares de *A. aegypti* pode contribuir para a identificação de proteínas e vias biológicas essenciais à manutenção do vetor. A caracterização de moléculas envolvidas em processos fundamentais, como desenvolvimento,

metabolismo e reprodução, permite ampliar o conhecimento sobre a biologia da espécie e pode auxiliar na identificação de potenciais alvos moleculares para futuras abordagens de controle (Degner *et al.,* 2019).

Dessa forma, estudos direcionados à caracterização de proteínas associadas à funções fisiológicas importantes tornam-se relevantes para a compreensão dos mecanismos que sustentam a adaptação e a sobrevivência do mosquito, bem como para a identificação de novos alvos que possam ser explorados no desenvolvimento de estratégias de controle (Conway *et al*., 2023).

Entre as proteínas de interesse nesse contexto, destacam-se as metaloproteases, uma classe de enzimas proteolíticas que utiliza íons metálicos, frequentemente o zinco, para realizar a clivagem de ligações peptídicas. Essas enzimas participam de diversos processos biológicos e apresentam diferentes famílias e funções no organismo. Em *A. aegypti*, poucas metaloproteinases já foram identificadas, Kantor *et al* (2017) identificaram e caracterizaram nove metaloproteases de matriz codificadas pelo genoma de *A. aegypti* e observaram alterações na expressão de algumas dessas enzimas durante diferentes processos fisiológicos, evidenciando a diversidade dessas proteínas e sua relevância para a biologia do vetor

No contexto das metaloproteases presentes em *Aedes aegypti*, destacam-se as aminopeptidases N, pertencentes à família M1 de metalopeptidases dependentes de zinco. Essas proteínas apresentam regiões conservadas associadas à atividade catalítica e estão envolvidas na clivagem de peptídeos e proteínas, desempenhando diferentes funções fisiológicas nos organismos. Em *A. aegypti*, estudos de caracterização molecular identificaram algumas aminopeptidases N pertencentes à família M1, evidenciando a diversidade dessa classe de proteínas no mosquito (Aroonkesorn *et al.*, 2015).

Além disso, análises transcriptômicas identificaram o gene AAEL011292 entre os genes que codificam metaloproteases com elevada abundância de transcritos no trato reprodutivo de fêmeas após o acasalamento, sugerindo uma possível associação desse gene com processos fisiológicos relacionados à reprodução (Alfonso-Parra *et al.*, 2016). Dessa forma, a caracterização estrutural e funcional da proteína AAEL011292 por meio de ferramentas de bioinformática pode contribuir para ampliar o conhecimento sobre suas características moleculares e possíveis funções na fisiologia de *A. aegypti*. Nesse sentido, a utilização de ferramentas como o AlphaFold (Jumper, *et al*. 2021), para a predição de sua estrutura tridimensional e o InterProScan (Jones, *et al*. 2014) para a identificação e caracterização de domínios proteicos, pode fornecer informações relevantes para a compreensão de suas características estruturais e funcionais, além de auxiliar na formulação de hipóteses que possam ser avaliadas em futuros estudos experimentais.

Metodologia

A seleção da metaloprotease alvo baseou-se em uma pesquisa prévia na literatura científica, complementada por buscas em bases de dados genômicos, como o *National Center for Biotechnology Information* (NCBI) e o *Vector Base*.

O gene selecionado, identificado sob o código de acesso AAEL011292 no banco de dados *Vector Base*, Após a seleção, a sequência completa de aminoácidos correspondente ao transcrito da proteína foi obtida a partir do *Vector Base* no mês de agosto de 2026. Em seguida, essa sequência primária foi submetida a ferramentas computacionais para análises funcionais e estruturais *in sílico*.

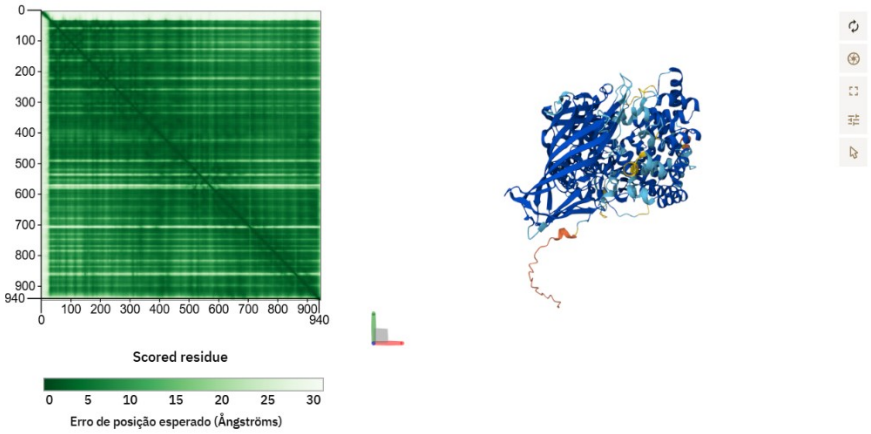
A identificação e a caracterização dos domínios conservados da proteína foram realizadas por meio da ferramenta InterProScan. Para a predição e obtenção da estrutura tridimensional, utilizou-se o banco de dados AlphaFold Protein Structure *Database*. A partir da modelagem obtida, foram avaliadas a conformação da molécula e as regiões de confiança da predição estrutural, utilizando como parâmetro a métrica técnica pLDDT (*Predicted Local Distance Difference Test*). Essa avaliação das regiões de confiança permitiu a descrição detalhada da composição global da proteína e da disposição espacial de seus resíduos de aminoácidos, assegurando a confiabilidade do modelo gerado.

Resultados

Para a caracterização estrutural *in silico* da proteína AAEL011292, utilizou-se o algoritmo AlphaFold para a modelagem tridimensional da cadeia polipeptídica e determinação das margens de

erro relativas entre resíduos (Figura 1). Os resultados indicam a presença de um domínio central bem estruturado e regiões terminais com características de maior flexibilidade funcional.

Figura 1: Estrutura tridimensional da proteína codificada pelo gene AAEL011292, gerada pelo AlphaFold.



Fonte: Autores, 2026.

A predição da estrutura tridimensional da proteína (figura da direita) codificada pelo gene AAEL011292, gerada por meio do modelo computacional AlphaFold, apresentou uma proteína com o tamanho de ~940 aminoácidos e revelou uma conformação tridimensional com alta confiabilidade estrutural global. A análise visual do modelo demonstra que a vasta maioria da cadeia polipeptídica apresenta escores pLDDT superiores a 90 (representados pelas regiões em azul escuro), indicando um altíssimo grau de confiança na predição do dobramento de seu núcleo estrutural globular, onde presumivelmente se localizam os domínios funcionais e catalíticos da metaloprotease. Em contrapartida, observou-se uma região terminal em conformação de alça solta (representada em tons de laranja e vermelho), indicativa de um segmento com escore pLDDT inferior a 50.

As métricas de validação espacial corroboram a estabilidade do modelo gerado. O gráfico de Erro Posição Esperado (PAE) (figura da esquerda), que afere a precisão da orientação relativa entre os resíduos, apresentou uma matriz com predominância de zonas em verde-escuro, traduzindo um erro posicional inferior a 5 Ângströms para a maior parte das interações. Essa homogeneidade na matriz PAE evidencia que os domínios estruturais da enzima não foram apenas individualmente bem modelados, mas que o empacotamento tridimensional e a orientação espacial relativa entre eles possuem alta confiabilidade preditiva. As raras faixas de elevado erro posicional na matriz correspondem de forma exata ao segmento terminal desordenado evidenciado no modelo 3D.

A análise da sequência de aminoácidos da proteína AAEL011292, realizada por meio da ferramenta computacional InterProScan, revelou a presença de domínios conservados que confirmam sua classificação na família das metalopeptidases M1 (Figura 2). A arquitetura de domínios identificada demonstra uma organização estrutural definida. Na região N-terminal, observou-se a presença do domínio Peptidase_M1_N, sobreposto a assinaturas relacionadas a fatores de interação estrutural (tricorn interacting factor f3 domain) e regiões de homologia com a leucotrieno A4 hidrolase. Na porção central da proteína, identificou-se o domínio catalítico principal, anotado conjuntamente como Peptidase_M1 e domínio de metaloproteases dependentes de zinco ("*zincins*"), além do domínio de protease neutra 2 (Neutral Protease Domain 2).

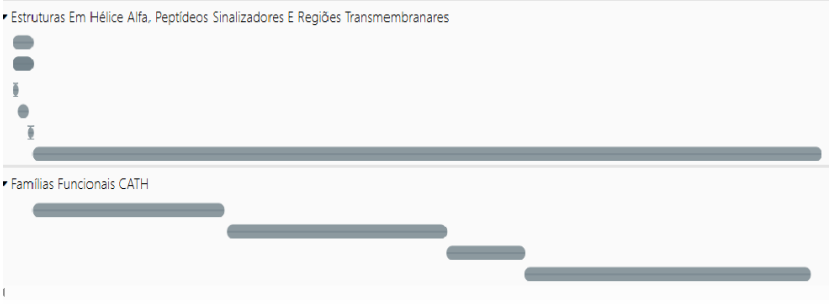
Figura 2: Análise dos domínios da proteína usando o Interproscan



Fonte: Interproscan.2026.

A análise apontou a presença de uma provável estrutura correspondente a um peptídeo sinalizador na região N-terminal, seguida por um extenso domínio classificado como não citoplasmático (*Non-cytoplasmic domain*), que abrange a quase totalidade da cadeia polipeptídica. Quanto à organização estrutural, foram identificadas diferentes regiões de estrutura secundária, com destaque para estruturas em hélice alfa, além da classificação de quatro regiões pertencentes a famílias funcionais CATH, distribuídas ao longo da sequência proteica e apresentando diferentes extensões (Figura 3). Esses resultados indicam uma proteína com organização estrutural complexa e predominantemente não citoplasmática, características que podem estar relacionadas à sua função molecular.

Figura 3: Análise das estruturas, sequências e famílias funcionais pelo Interproscan.



Fonte: Autores, 2026.

Discussão

Após a análise realizada pelo programa InterProScan, foram identificados domínios que podem desempenhar papéis importantes nas funções metabólicas da proteína codificada pelo gene AAEL011292, destacando-se a presença de uma região homóloga à leucotrieno A4 hidrolase (LTA4H), localizada aproximadamente entre os aminoácidos 24 e 245. Essa região pode estar relacionada a processos fisiológicos importantes, considerando que estudos anteriores demonstraram aumento da expressão de AAEL011292 no trato reprodutivo de fêmeas de *Aedes aegypti* após o acasalamento, com elevada expressão observada 24 horas após a cópula (Alfonso-Parra *et al.*, 2016). Esses resultados sugerem uma possível participação da proteína em processos fisiológicos associados ao período pós-acasalamento. Além disso, foi identificada uma região associada ao fator de interação Tricorn F3, relacionado a processos de degradação de proteínas (TAMURA *et al.,* 1996). Essa característica sugere uma possível participação de AAEL011292 em processos de processamento e degradação proteica, podendo contribuir para funções metabólicas importantes no organismo. A análise pelo CATH (Class, Architecture, Topology and Homologous superfamily), apresentada na Figura 3, também identificou regiões relacionadas a famílias de peptidases, as aminopeptidases e glutamil peptidases. As aminopeptidases estão envolvidas na remoção de aminoácidos das extremidades de peptídeos N-terminal, enquanto as glutamil peptidases participam da clivagem de ligações peptídicas envolvendo resíduos de glutamato **(**RAWLINGS *et al.,* 2016). Dessa forma, os domínios identificados pelo InterProScan, associados aos dados de expressão previamente descritos na literatura, fornecem indícios de que AAEL011292 pode desempenhar funções relacionadas ao metabolismo proteico e à fisiologia pós-acasalamento em fêmeas de *A. aegypti.*

A predição da estrutura tridimensional da proteína AAEL011292, realizada por meio do AlphaFold, revelou uma conformação predominantemente globular e elevada confiança estrutural na maior parte da cadeia polipeptídica. A predominância de regiões com valores de pLDDT superiores a 90 indica alta confiança na predição local da estrutura, sendo compatível com regiões globulares bem definidas. Em contrapartida, a região terminal apresentou baixos valores de pLDDT, indicando maior incerteza na predição estrutural. Regiões com baixo pLDDT são frequentemente associadas a segmentos estruturalmente flexíveis ou intrinsecamente desordenados, podendo apresentar maior acessibilidade ao solvente. Estudos que avaliaram as predições do AlphaFold demonstraram que regiões de baixa confiança apresentam enriquecimento de regiões intrinsecamente desordenadas e relação com características de acessibilidade ao solvente (Akdel *et al.,* 2022). Assim, a baixa confiança observada na região terminal de AAEL011292 pode estar relacionada à maior flexibilidade conformacional dessa porção da proteína e à possível exposição ao solvente. Entretanto, essa interpretação deve ser considerada como uma hipótese estrutural, uma vez que o pLDDT, isoladamente, não permite determinar diretamente a exposição ao solvente.

Conclusão

A análise bioinformática da proteína AAEL011292 permitiu caracterizar sua arquitetura de domínios e sua organização estrutural, evidenciando características compatíveis com proteínas da família das metaloproteases M1. A identificação de regiões relacionadas à LTA4H e ao fator de interação Tricorn F3, associada ao aumento da expressão de AAEL011292 após o acasalamento descrito na literatura, sugere uma possível participação da proteína em processos de metabolismo e processamento proteico durante a fisiologia pós-acasalamento de *Aedes aegypti*. A predição realizada pelo AlphaFold complementou esses resultados ao demonstrar uma estrutura tridimensional predominantemente bem definida, com alta confiança na maior parte da proteína. Dessa forma, este estudo contribui para a caracterização inicial de uma proteína ainda pouco explorada, fornecendo informações estruturais e funcionais que podem servir de base para futuros estudos experimentais sobre seu papel na fisiologia reprodutiva e no metabolismo de *A. Aegypti*.

Referências

AKDEL, M. et al. A structural biology community assessment of AlphaFold2 applications. **Nature Structural & Molecular Biology**, v. 29, n. 11, p. 1056-1067, nov. 2022. Disponível em: https://doi.org/10.1038/s41594-022-00849-w. Acesso em: 7 set. 2026.

ALFONSO-PARRA, C. et al. Mating-Induced Transcriptome Changes in the Reproductive Tract of Female Aedes aegypti. **PLoS Neglected Tropical Diseases**, v. 10, n. 2, e0004451, 2016. Disponível em: https://doi.org/10.1371/journal.pntd.0004451. Acesso em: 1 set. 2026.

AROONKESORN, A. et al. Two specific membrane-bound aminopeptidase N isoforms from Aedes aegypti larvae serve as functional receptors for the Bacillus thuringiensis Cry4Ba toxin implicating counterpart specificity. **Biochemical and Biophysical Research Communications**, v. 461, n. 2, p. 300-306, 2015. Disponível em: https://doi.org/10.1016/j.bbrc.2015.04.026. Acesso em: 2 set. 2026.

CONWAY, M. J.; HASLITT, D. P.; SWARTS, B. M. Targeting Aedes aegypti metabolism with next-generation insecticides. **Viruses**, v. 15, n. 2, p. 469, 2023. Disponível em: https://doi.org/10.3390/v15020469. Acesso em: 3 set. 2026.

DEGNER, E. C.; AHMED-BRAIMAH, Y. H.; BORZIAK, K.; WOLFNER, M. F.; HARRINGTON, L. C.; DORUS, S. Proteins, Transcripts, and Genetic Architecture of Seminal Fluid and Sperm in the Mosquito Aedes aegypti. **Molecular & Cellular Proteomics**, v. 18, p. S6-S22, 2019. Disponível em: https://doi.org/10.1074/mcp.RA118.001067. Acesso em: 1 set. 2026.

DIAS, Í. K. R. et al. Ações educativas de enfrentamento ao Aedes aegypti: revisão integrativa. **Ciência & Saúde Coletiva**, Rio de Janeiro, v. 27, n. 1, p. 231-242, 2022. Disponível em: https://www.scielo.br/j/csc/a/F3BmRZ5bBDSJLxYNpWjTwzp/?lang=pt. Acesso em: 30 ago. 2026.

FERNANDES, C. O. S.; FERNANDES, D. R. A. S.; BARACAT, R. V. M.; SILVEIRA, P. T. M.; BRAGA, G. O. Arboviroses emergentes e reemergentes no Brasil: dengue, chikungunya e zika. **Brazilian Journal of Implantology and Health Sciences**, v. 6, n. 8, p. 5036-5048, 2024. Disponível em: https://doi.org/10.36557/2674-8169.2024v6n8p5036-5048. Acesso em: 1 set. 2026.

KANTOR, A. M. et al. Identification and initial characterization of matrix metalloproteinases in the yellow fever mosquito, Aedes aegypti. **Insect Molecular Biology**, v. 26, n. 1, p. 113-126, 2017. Disponível em: https://doi.org/10.1111/imb.12275. Acesso em: 2 set. 2026.

LIMA-CAMARA, T. N. Arboviroses emergentes e novos desafios para a saúde pública no Brasil. **Revista de Saúde Pública**, São Paulo, v. 50, p. 36, 2016. Disponível em: https://doi.org/10.1590/S1518-8787.2016050006791. Acesso em: 2 set. 2026.

RAHMAN, R. U. et al. Insecticide resistance and genetic structure of Aedes aegypti populations from Rio de Janeiro State, Brazil. **PLoS Neglected Tropical Diseases**, v. 15, n. 2, e0008492, 2021. Disponível em: https://doi.org/10.1371/journal.pntd.0008492. Acesso em: 2 set. 2026.

RAWLINGS, N. D. et al. Twenty years of the MEROPS database of proteolytic enzymes, their substrates and inhibitors. **Nucleic Acids Research**, v. 44, n. D1, p. D343-D350, jan. 2016. Disponível em: https://doi.org/10.1093/nar/gkv1118. Acesso em: 5 set. 2026.

TAMURA, T. et al. Tricorn protease and its interacting factors in Thermoplasma acidophilum. **FEBS Letters**, v. 435, n. 2-3, p. 251-254, set. 1998. Disponível em: https://doi.org/10.1016/S0014-5793(98)01082-6. Acesso em: 1 set. 2026.