# ANÁLISE *IN SÍLICO* DE UM GENE (AAEL012774) DA MATRIZ EXTRACELULAR DE *AEDES AEGYPTI*.

**Eduardo Mendes de Oliveira, Eduarda Rodrigues Araujo Alves, Laís Marins Menezes, Bruno Regis Lyrio Ferraz, Tatiane Sanches Soares.**

Universidade Federal do Espírito Santo – UFES, Campus Alegre, Alto Universitário, S/N, Guararema – 29500-000 – Alegre - ES, Brasil, eduardo.oliveira.75@edu.ufes.br, eduardarodriguesaraujoalves5@gmail.com, laismarins28@gmail.com, bruno.ferraz@ufes.br, tatiane.s.soares@ufes.br.

**Resumo**

As arboviroses representam um importante problema de saúde pública e são transmitidas principalmente por mosquitos vetores, entre os quais se destaca *Aedes aegypti*. Durante o ciclo de infecção, fatores moleculares do próprio vetor podem influenciar a replicação e a disseminação viral, atuando de maneira favorável ou restritiva. Nesse contexto, a matriz extracelular (ECM), responsável pela manutenção da integridade e organização dos tecidos, desempenha também papel relevante na interação entre células e agentes patogênicos. Entre os componentes envolvidos na remodelação da ECM, destacam-se as metaloproteinases, enzimas que podem participar de processos relacionados à modulação dos tecidos, e potencialmente, à disseminação viral do vetor. Diante disso, a caracterização utilizando abordagens bioinformáticas de metaloproteinases de matriz extracelular, como a codificada pelo gene AAEL012774, podem ajudar a compreender sua possível participação na interação do vetor com os arbovírus.

**Palavras-chave**: *Aedes aegypti*. Metaloproteinases. Arbovírus. Bioinformática.

**Área do Conhecimento:** Ciências Biológicas; Biologia Molecular.

## Introdução

A ampla distribuição geográfica e a capacidade de adaptação do mosquito *Aedes aegypti* aos ambientes urbanos tornam necessário o aprofundamento do conhecimento sobre os mecanismos moleculares associados à sua biologia, principalmente devido aos arbovírus associados a esse vetor (Braga; Valle, 2007; WHO, 2023). As arboviroses representam um importante problema de saúde pública nas regiões tropicais e subtropicais do mundo, sendo o mosquito *Aedes aegypti* um dos principais vetores envolvidos em sua transmissão (*World Health Organization*, 2022). Entre as proteínas presentes no organismo do vetor, destacam-se as metaloproteinases, enzimas dependentes de íons metálicos, especialmente o zinco, para desempenhar suas funções catalíticas (Nagase; Woessner, 1999). Em insetos, essas enzimas estão relacionadas a diferentes processos fisiológicos, incluindo digestão de proteínas, desenvolvimento e remodelação tecidual e metamorfose (Kantor *et al.*, 2017; Gomes *et al.*, 2012). Embora metaloproteinases representantes dos gêneros *Anopheles* e *Culex* sejam bem caracterizadas, as funções moleculares de diversas dessas proteínas em *A. aegypti* ainda permanecem pouco caracterizadas (Gomes *et al.*, 2012).

A caracterização de proteínas por meio de abordagens bioinformáticas permite investigar suas propriedades moleculares a partir de informações disponíveis em bancos de dados genômicos e proteômicos. Essas análises possibilitam a identificação de sequências, domínios conservados, motivos funcionais e características estruturais, contribuindo para a formulação de hipóteses sobre a função de proteínas ainda não estudadas experimentalmente (Jumper *et al.*, 2021; Waterhouse *et al.*, 2018). Nesse contexto, a análise *in silico* de metaloproteinases de *A. aegypti* pode fornecer informações relevantes sobre sua composição molecular e possíveis características funcionais, além de permitir a comparação dessas proteínas com homólogas presentes em outras em outras espécies. A integração

dessas informações constitui uma estratégia importante para ampliar a compreensão da diversidade e da evolução das metaloproteinases presentes no mosquito vetor.

Entre as principais ferramentas disponíveis para esse tipo de investigação, destacam-se aquelas destinadas à predição e análise da estrutura tridimensional e à caracterização de domínios funcionais das proteínas. Ferramentas como *AlphaFold* (Jumper *et al.*, 2021) permitem a predição da estrutura tridimensional de proteínas a partir de suas sequências de aminoácidos, enquanto SWISS-MODEL (Waterhouse *et al.*, 2018) possibilita a construção de modelos estruturais por modelagem comparativa. A identificação de domínios e regiões funcionais pode ser realizada por meio de plataformas como o *InterProScan* (Jones *et al.*, 2014), que integra informações provenientes de diferentes bancos de dados para a caracterização funcional de proteínas, permitindo a identificação de famílias, domínios, regiões conservadas e possíveis funções moleculares.

Diante disso, este estudo tem como objetivo realizar a caracterização *in silico* de uma metaloproteinase de *Aedes aegypti* (AAEL012774), utilizando diferentes abordagens de bioinformática. Para isso, serão realizadas análises da sequência proteica, identificação de características estruturais e funcionais da proteína estudada, contribuindo para ampliar o conhecimento sobre as metaloproteinases de *A. aegypti* e fornecer subsídios para futuras investigações experimentais sobre suas funções biológicas.

**Metodologia**

Para a efetuação da modelagem preditiva da estrutura tridimensional da metaloproteinase AAEL012774, primeiramente foi obtida a sequência de aminoácidos da proteína através do banco de dados *VectorBase (VectorBase,* 2024) e *National Center for Biotechnology Information* (NCBI) *Gene/Protein* (NCBI, 2024), a partir do identificador gênico correspondente. A sequência em formato FASTA foi utilizada na análise estrutural e funcional. A modelagem estrutural tridimensional foi realizada utilizando a plataforma *SWISS-MODEL*, utilizando a abordagem por homologia. A plataforma realiza a busca de estruturas ou modelos relacionados à sequência-alvo para sua utilização como *templates*, incluindo estruturas experimentais disponíveis na *SWISS-MODEL Template Library* (SMTL) e modelos vindos do *AlphaFold Database* (AFDB). Os modelos gerados foram avaliados com base nos parâmetros de qualidade disponibilizado pela plataforma, incluindo a identidade de sequência, cobertura, GMQE (*Global Quality Estimation*) e QMEANDisCo (*Qualitative Model Energy Analysis Distance Constraints*) quando disponível. Os *templates* utilizados na modelagem também foram analisados quanto à sua origem e às características estruturais.

A identificação dos domínios funcionais, regiões conservadas e sítios catalíticos da proteína AAEL012774 foi realizada por meio da ferramenta *InterProScan*, que integra dados de múltiplos bancos de domínios como *Pfam*, *SMART* e Prosite (Jones *et al.*, 2014).

**Resultados**

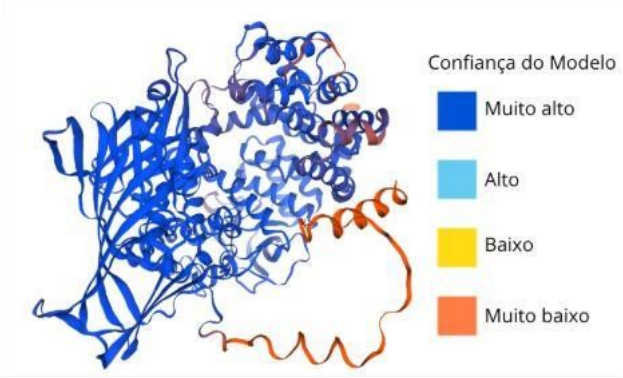
A sequência proteica da AAEL012774 apresentou 900 aminoácidos. A partir dessa sequência foi realizada a modelagem estrutural da proteína AAEL012774 utilizando o *SWISS-MODEL* resultando em dois modelos preditivos, que foram comparados quanto aos parâmetros estabelecidos (Tabela 1). O Modelo 01 (Figura 1) apresentou identidade de sequência de 98,44%, cobertura de 1,00 e GMQE (*Global Model Quality Estimate*) de 0,90, utilizando como *template* um modelo estrutural proveniente do *AlphaFold Database* (Q16L34.1.A). O Modelo 02 (Figura 2) apresentou identidade de sequência de 35,17%, cobertura de 0,93 e GMQE de 0,71, tendo como *template* a estrutura 4WZ9.1.A, correspondente à proteína APN1 de *Anopheles gambiae*.

Tabela 1 - Comparação dos modelos estruturais preditivos da proteína AAEL012774 obtidos no *SWISSMODEL*.

| **Parâmetro** | **Modelo 01** | **Modelo 02** |
|---|---|---|
| *Template* | Q16L34.1.A | 4WZ9.1.A |
| Identidade da sequência (%) | 98,44 | 35,17 |
| Cobertura (%) | 100 | 93 |
| GMQE | 0,90 | 0,71 |
| Região modelada (aa) | 1-900 | 44-890 |
| Estado oligomérico | Monômero | Homodímero |
| QMEANDisco Global | -- | 0,74 ± 0,05 |

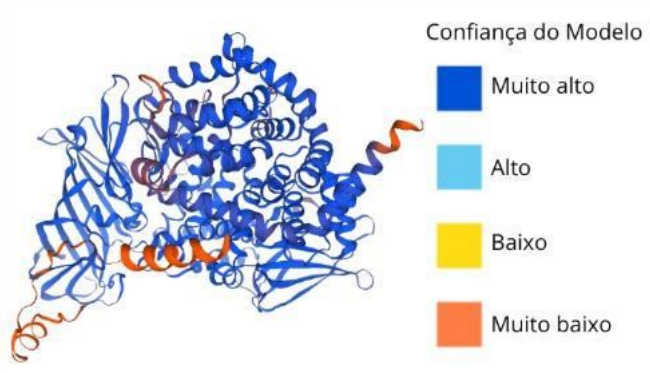
Fonte: Autores, 2026.

Figura 1 - Modelo estrutural dimensional 01.




Fonte: *SWISS-MODEL* (2026), adaptado pelos autores, 2026.

Figura 2 - Modelo estrutural dimensional 02.




Fonte: *SWISS-MODEL* (2026) adaptado pelos autores, 2026.

A análise da estrutura predita da proteína AAEL012774 no *AlphaFold Database* indicou elevada confiança na estrutura modelada. O valor médio de pLDDT (*Predicted Local Distance Difference Test*) foi de 90,06, sendo 76,4% dos resíduos classificados com confiança muito alta e 15,9% com confiança alta. Apenas 2,0% dos resíduos apresentaram baixa confiança e 5,7% apresentaram confiança muito

baixa. A distribuição dos valores de pLDDT indicou, portanto, maior confiabilidade na maior parte da estrutura predita, enquanto as regiões com menores valores de confiança apresentaram maior incerteza estrutural. Não foram identificadas estruturas experimentais correspondentes disponíveis no *Protein Data Bank* (PDB).

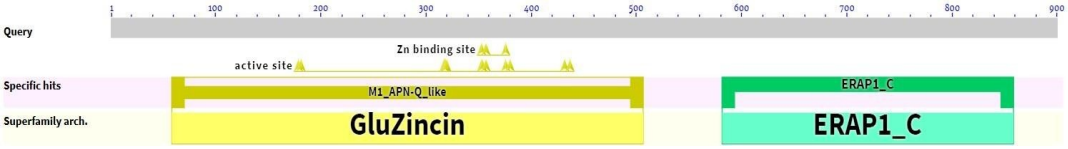
A análise funcional *in silico* da proteína AAEL012774 realizada pelo *InterProScan* indicou sua associação à família das peptidases M1, caracterizada por aminopeptidases dependentes de zinco (Figura 3). Foram identificados domínios relacionados à atividade de peptidases, incluindo Peptidase_M1_N e *Neutral Protease Domain 2* (Figura 3) A anotação por *Gene Ontology* (GO) indicou, na categoria de função molecular, os termos *zinc ion binding* (GO:0008270) e *metallopeptidase activity* (GO:0008237), enquanto na categoria de processo biológico foi identificado o termo *proteolysis* (GO:0006508). A análise complementar pelo PANTHER também associou a proteína aos termos *metalloaminopeptidase activity* (GO:0070006), *peptide binding* (GO:0042277) e *zinc ion binding* (GO:0008270), além dos processos de proteólise e catabolismo de peptídeos (GO:0043171) (Tabela 2). Quanto ao componente celular, foram atribuídas as categorias membrana (GO:0016020) e citoplasma (GO:0005737), enquanto o termo *extracellular space* (GO:0005615) foi identificado como obsoleto na anotação.

Tabela 2 – Termos de *Gene Ontology* associados à proteína AAEL012774 identificados pelo InterProScan e PANTHER.

| **Fonte** | ***Molecular function*** | ***Biological process*** | ***Cellular component*** |
|---|---|---|---|
| *InterPro* | *Zinc ion binding (GO:0008270; Metallopeptidase activity (GO:0008237)* | *Proteolysis* (GO:0006508) | *None* |
| PANTHER | *Metalloaminopeptidase activity* (GO:0070006); *peptide binding* (GO:0042277); *zinc ion binding* (GO:0008270) | *Proteolysis* (GO:0006508); *peptide catabolic process* (GO:0043171) | *Membrane* (GO:0016020); *Obsolete extracellular space* (GO:0005615); *cytoplasm* (GO:0005737) |

Fonte: Autores, 2026.

Figura 3 – Representação dos domínios e sítios funcionais identificados na sequência da proteína AAEL012774.




Fonte: NCBI, 2026.

## Discussão

A caracterização *in silico* da proteína AAEL012774 revelou características estruturais e funcionais compatíveis com uma metaloaminopeptidase da família M1. A modelagem realizada pelo *SWISSMODEL* apresentou dois modelos, sendo o Modelo 01 o que apresentou melhores parâmetros de qualidade, com 98,44% de identidade, cobertura de 1,00 e GMQE de 0,90. A análise complementar pelo *AlphaFold* apresentou pLDDT médio de 90,06, indicando elevada confiança na maior parte da estrutura predita. O pLDDT é uma medida de confiança atribuída aos resíduos individuais da estrutura, permitindo avaliar quais regiões apresentam maior ou menor confiabilidade na predição (Jumper *et al*., 2021). A elevada confiança observada para a AAEL012774 reforça a utilidade do modelo estrutural obtido para a investigação de características da proteína. Dessa forma, a integração dos resultados do *SWISS-MODEL* e *AlphaFold* fornece maior suporte à caracterização estrutural da proteína,

especialmente considerando a ausência de uma estrutura experimental correspondente disponível no PDB.

Os resultados do *InterProScan* e PANTHER reforçaram a possível função da AAEL012774 como uma metaloaminopeptidase, devido à identificação da família Peptidase M1, de domínios característicos e dos termos de *Gene Ontology* relacionados à ligação a íons zinco e à atividade de metalopeptidase. A identificação de termos associados à proteólise e ao catabolismo de peptídeos também é consistente com essa anotação funcional. Entretanto, por se tratar de predições *in silico*, esses resultados não confirmam experimentalmente a atividade enzimática da proteína, sendo necessários estudos bioquímicos para validar sua função e investigar seu possível papel na remodelação da matriz extracelular e na interação vetor-vírus em *A. aegypti*.

Esses resultados são particularmente relevantes diante da ausência de uma estrutura experimental correspondente disponível no PDB, evidenciando o potencial das ferramentas de modelagem estrutural *in silico* para auxiliar na caracterização de proteínas ainda não estudadas experimentalmente e fornecer informações estruturais que podem orientar futuras investigações sobre sua função biológica.

## Conclusão

A análise *in silico* da proteína AAEL012774 revelou características estruturais e funcionais compatíveis com uma metaloaminopeptidase da família M1, com elevada confiança na estrutura tridimensional predita e evidências de ligação a zinco e atividade de metalopeptidase. A modelagem tridimensional realizada pelo *SWISS*-MODEL apresentou modelos com diferentes níveis de similaridade, sendo o modelo 01 aquele que apresentou os melhores parâmetros de identidade, cobertura e GMQE.

As análises funcionais realizadas pelo *InterProScan* e PANTHER identificaram domínios e regiões associados à atividade de metalopeptidases dependentes de zinco, além de termos relacionados à ligação ao zinco e à proteólise. Esses resultados contribuem para a caracterização molecular da proteína AAEL012774 e fornecem subsídios para futuras análises experimentais de sua função em *Aedes aegypti*.

## Referências

BRAGA, I. A.; VALLE, D. *Aedes aegypti*: histórico do controle no Brasil. **Epidemiologia e Serviços de Saúde,** v. 16, n. 2, p. 113–118, 2007. Disponível em http://scielo.iec.gov.br/scielo.phpscript=sci_arttext&pid=S167949742007000200006&lng=pt&nrm=iso. Acesso em: 18 agosto 2026.

GOMES, B.; TOLLEFSEN, D. M.; PAES, A. M. A.; *et al.* Effect of metal ion chelators on proteolytic activity in midgut homogenates of *Aedes aegypti* larvae. **Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology**, v. 163, n. 3-4, p. 205–211, 2012. DOI: https://doi.org/10.1016/j.cbpb.2012.07.002.

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. DOI: 10.1111/imb.12275.

JONES, P. *et al.* InterProScan 5: genome-scale protein function classification. **Bioinformatics**, v. 30, n. 9, p. 1236– 1240, 2014. DOI: https://doi.org/10.1093/bioinformatics/btu031.

JUMPER, J. *et al.* Highly accurate protein structure prediction with AlphaFold. **Nature,** v. 596, p. 583–589, 2021. DOI: https://doi.org/10.1038/s41586-021-03819-2.

NAGASE, H.; WOESSNER, J. F. Matrix metalloproteinases. J**ournal of Biological Chemistry**, v. 274, n. 31, p. 21491–21494, 1999. DOI: https://doi.org/10.1074/jbc.274.31.21491.

NCBI – NATIONAL CENTER FOR BIOTECHNOLOGY INFORMATION. Gene and Protein Databases. Bethesda (MD): **NCBI**, 2024. Disponível em: https://www.ncbi.nlm.nih.gov/. Acesso em: 19 agosto 2026.

VECTORBASE. Bioinformatics Resource for Invertebrate Vectors of Human **Pathogens.** 2024. Disponível em: https://vectorbase.org/. Acesso em: 19 agosto 2026.

WATERHOUSE, A. *et al.* SWISS-MODEL: homology modelling of protein structures and complexes. **Nucleic Acids Research**, v. 46, n. W1, p. W296–W303, 2018. DOI: https://doi.org/10.1093/nar/gky427.

WORLD HEALTH ORGANIZATION. Dengue and severe dengue: fact sheet. Geneva: **WHO**, 2022. Disponível em: https://www.who.int/news-room/fact-sheets/detail/dengue-and-severe-dengue. Acesso em: 15 agosto 2026.

WORLD HEALTH ORGANIZATION. Global arbovirus initiative: tackling the increasing burden of arboviruses worldwide. Operational guidance. Geneva: **WHO**, 2023. Disponível em: https://www.who.int/initiatives/global-arbovirus-initiative. Acesso em: 16 agosto 2026.