

# MODELAGEM COMPUTACIONAL DE PROTEÍNAS ASSOCIADAS À RESISTÊNCIA A INSETICIDAS EM *Anophelesdarlingi* ENCONTRADO NA REGIÃO AMAZÔNICA

Adonis de Melo LIMA

LIMA, Adonis de Melo. **Modelagem computacional de proteínas associadas à resistência a inseticidas em *anophelesdarlingi* encontrado na região amazônica.** Projeto de investigação científica do Curso de Biomedicina – Faculdade Integrada Brasil Amazônia (FIBRA), Belém, 2014.

A malária é uma doença infecciosa transmitida por mosquitos do gênero *Anopheles*. O principal vetor da malária no Brasil é o *Anophelesdarlingi*, presente no interior de todo o território, sendo responsável, provavelmente, pela transmissão da maioria dos casos em nosso país. O objetivo das ações de controle vetorial em saúde pública é diminuir a transmissão de doenças por meio da eliminação de parte da população de mosquitos infectados, que frequentemente está em contato com o homem. Essa eliminação pode ser feita com base no combate a larvas ou adultos por meio da

eliminação de criadouros (controle físico), da utilização de organismos naturalmente predadores dos mosquitos (controle biológico), ou de produtos capazes de matar os mosquitos (controle químico). Essa última é uma das metodologias mais utilizadas como parte de programas de combate aos vetores, principalmente por meio de ações como a borrifação intradomiciliar, nebulização espacial e, mais recentemente no Brasil, o uso de mosquiteiros impregnados. Atualmente, os inseticidas utilizados em saúde pública são derivados de piretroide, uma classe de inseticidas que surgiu em substituição aos organoclorados, representados principalmente pelo DDT, muito utilizado no combate à malária até a década de 1990. O uso inadvertido desse produto, tanto na saúde pública quanto na agricultura, pode ter sido um dos causadores de um efeito indesejado em um produto dessa classe: o desenvolvimento de resistência por parte da praga alvo. Um dos mecanismos envolvidos no processo resistência está relacionado à presença da família enzimática glutathione S transferases (GSTs), envolvida no processo de desintoxicação, desempenhando um mecanismo fundamental de defesa desses insetos a inseticidas. A modelagem por

homologia estrutural parte do princípio de que as proteínas são relacionadas evolutivamente e, portanto, compartilham uma estrutura similar. Com isso, o modelo de uma proteína com estrutura desconhecida (alvo) pode ser construído com base em um alinhamento de uma proteína com estrutura conhecida (modelo). Isso geralmente envolve quatro etapas: (1) identificação dos homólogos, que pode ser utilizada como modelo(s) para modelagem; (2) alinhamento da sequência-alvo para o(s) modelo(s); (3) construção de um modelo para o alvo com base nas informações a partir do(s) alinhamento(s); e (4) avaliação do modelo. Com isso, torna-se relevante conhecer a estrutura tridimensional da proteína da GST do mosquito *A. Darlingi*. O objetivo da investigação foi elucidar com mais acurácia suas prováveis funções biológicas não esclarecidas ainda, utilizando ferramentas de bioinformática para construir modelo teórico, usando modelagem por homologia da GST, obtida e anotada a partir do genoma de *Anopheles darlingi*, encontrado na região amazônica. Para isso foi feita a identificação das proteínas moldes apresentando sequências primárias similares, realizada por meio de alinhamentos locais da proteína-alvo contra as sequências proteicas do PDB; e

gerado modelo tridimensional por meio de modelagem homologia molecular da melhor sequência primária da proteína GST, obtida do genoma de *A. darlingi*; foram analisados os resultados obtidos entre as estrutura-alvo; e as estruturas-modelo e utilizados diversos programas de validação sobre o modelo proposto; e feita avaliação dos modelos construídos, analisando a qualidade estereoquímica, a energia livre do sistema e o mapeamento do potencial eletrostáticos molecular. O material e método usados para a obtenção das sequências-alvo foram o sequenciamento, a partir de *Anopheles darlingi* coletado no Brasil, usando a plataforma de sequenciamento Roche 454. A sequência foi depositada no *National Center for Biotechnology Information* (NCBI). Os números crescentes a respeito dos genes e das proteínas de muitos organismos proporcionaram o surgimento de inúmeros bancos de dados, com a tentativa de organizar e distribuí-los. Dentre os vários bancos existentes, o trabalho utilizou o GenBank, UniProt, Protein Data Bank (PDB) e o PDBSum. O alinhamento de sequências é um método de comparação que procura determinar o grau de similaridade entre duas ou mais sequências, ou a

similaridade entre fragmentos dessas sequências (MUNIZ, 2003). O valor aceitável de similaridade na modelagem por homologia é acima de 25% da sequência entre proteína-molde e proteína-alvo (SALI, 1993; D'ALFONSO *et al.*, 2001; VITUKUP *et al.*, 2001). O “e-value”, parâmetro estimado pelo BLAST, expressa a dificuldade para encontrar uma sequência perfeitamente idêntica, ou seja, quanto menor o valor, menor a chance de tal comparação ter sido encontrada por pura coincidência (GIBAS & JAMBECK, 2001). Foram levados em consideração os valores de “score” e de resolução em Angstrom da difração por raio-X. Na modelagem por homologia, os valores significativos de “score” devem estar maiores de 200 para esse programa. No estudo utilizamos programas como o Clustalw, BLAST e o BLASTP para realizar comparações entre as sequências escolhidas. A construção de um modelo tridimensional (alvo) é feita por meio da utilização de uma estrutura molde, gerada depois de um alinhamento consistente. Foi utilizado o *software* MODELLER 9.10 (MARTI-RENOM *et al.*, 2000) para as construções dos modelos. Os *softwares* usados no trabalho para avaliação dos modelos foram: PROCHECK

(LASKOWSKI *et al.*, 1993) e o VERIFY3D (LUTHY *et al.*, 1992). Para analisar erros estruturais, utilizou-se o VERIFY3D. Foi observado que os 210 aminoácidos (AA) presentes no modelo teórico dividiram-se em duas zonas distintas. No domínio N-terminal (correspondente aos 78 ácidos aminados iniciais), foi possível identificar a conservação de vários resíduos, tais como SER9, HIS50, VAL52, GLU64, SER65, ARG66, que são os locais de interação com o cofator GSH. Quanto ao domínio C-terminal, a outra parte restante composta por 132 aminoácidos, verificou-se a conservação de cinco  $\alpha$  hélices também presentes no modelo. O gráfico de Ramachandram plotado mostrou 97,7% de todos os resíduos em regiões quimicamente favoráveis. O desvio calculado pelo RMSD foi de 0.097Å, mostrando que há pequenas diferenças de conformação entre as estruturas-alvo e do molde. Usando o *Verify 3D*, foi constatado que 86,94% dos resíduos têm compatibilidade 3D-1D. O ANOLEA foi usado para calcular a energia conformacional de cada cadeia lateral da proteína, mostrando a correspondência entre três resíduos do sítio ativo, SER9 e HIS50 VAL52, localizados em regiões de elevada energia e em loops.

De acordo com os valores obtidos nas validações, foi possível confirmar a boa qualidade do modelo gerado. As regiões do sítio ativo mostraram-se conservadas, ajudando, assim, a compreensão do mecanismo de desintoxicação utilizado pelo mosquito. Torna-se plausível, portanto, a elaboração bloquear a operação da classe delta, tornando-se novamente eficaz DDT em populações de mosquitos já resistentes. Conclui-se que a resistência a inseticidas se tornou uma grande barreira ao combate ao anofelino. Diante dos bons resultados encontrados na validação do modelo, pode-se afirmar que a modelagem molecular por homologia é uma alternativa confiável para predição de estruturas tridimensionais, além de acelerar o processo de elucidação de proteínas utilizadas como fármacos em curto espaço de tempo e a custos reduzidos. Os resultados permitiram visualizar a estrutura completa da enzima GST de *Anopheles darlingi*, incluindo seu sítio de catálise, que se mostrou conservado durante o processo evolutivo, sugerindo funções semelhantes e expectativas quanto aos estudos de interações com o substrato GSH, possível apenas por meio de estudo de docagem e dinâmica molecular.

**Palavras-chave:** Modelagem. Proteína. Inseticidas. Região Amazônica.

## REFERÊNCIAS

D'ALFONSO G, TRAMONTANO A, LAHM A (2001) Structural conservation in single- domain proteins: implications for homology modeling. *Journal of structural biology* 134 (2-3):246-256.

GIBAS, C.; JAMBECK, P. Desenvolvendo Bioinformática: ferramentas de software para aplicações em biologia. Tradução Milarepa Ltda.- Rio de Janeiro: Campus, 2001.

LASKOWSKI, R. a., MACARTHUR, M. W., MOSS, D. S., & THORNTON, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography*, 26(C), 283–291. doi:10.1107/S0021889892009944.

LUTHY, R., BOWIE, J. U., & EISENBERG, D. (1992). Assessment of protein models with three-dimensional profiles. *Nature*, 356(6364), 83–85.

MARTÍ-RENOM, M.A.; STUART, A.C.; FISER, A.; SÁNCHEZ, R.; MELO, F.; SALI, A. Comparative protein structure modeling of genes and genomes. *Annual Review of Biophysics and Biomolecular Structure*, Palo Alto, v. 29, p. 291-325, 2000.

MUNIZ, J. R. C. Aplicação da bioinformática nos estudos dos genes e enzimas envolvidos na síntese da gomafastidiana produzida pela *Xylela fastidiosa*. 124p. Dissertação (Mestrado) - Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2003.

SALI, A., & BLUNDELL, T. L. (1993). Comparative protein modelling by satisfaction of spatial restraints. *Journal of Molecular Biology*, 234(3), 779–815. doi:10.1006/jmbi.1993.1626.

VITKUP, D., MELAMUD, E., MOULT, J., and SANDER, C. 2001. Completeness in structural genomics. *Nat. Struct. Biol.* 8: 59–66.