



## MODELAGEM ESTRUTURAL DE PROTEÍNA OsJGSTF3 COM POTENCIAL DE DESINTOXICAÇÃO DE HERBICIDAS EM *Oryza sativa japonica*

Ravenna Lins Rodrigues (1); Felipe França de Oliveira (2); Vinícius Costa Amador (3); Rafael Trindade Maia (4)

1 Discente do curso de Engenharia de Biotecnologia e Bioprocessos, Centro de Desenvolvimento Sustentável do Semiárido, Universidade Federal de Campina Grande, Sumé, PB, Sumé, Paraíba, Brasil. Email: ravennalinsrodrigues@gmail.com

2 Discente do curso de Engenharia de Biotecnologia e Bioprocessos, Centro de Desenvolvimento Sustentável do Semiárido, Universidade Federal de Campina Grande, Sumé, PB, Sumé, Paraíba, Brasil. Email: felipe\_ufcg2016@hotmail.com

3 Pós-graduando em Agronomia- Melhoramento Genético de Plantas. Universidade Federal Rural de Pernambuco - UFRPE, PPGAMGP, Recife, Pernambuco, Brasil. E-mail: vinicius.costa.amador@gmail.com

4 Professor Doutor. Centro de Desenvolvimento Sustentável do Semiárido, Universidade Federal de Campina Grande, Sumé, PB, Sumé, Paraíba, Brasil. E-mail: rafael.trindade@ufcg.edu.com.

### RESUMO

O arroz é uma planta herbácea da família das gramíneas, gênero *Oryza*, que alimenta mais da metade da população humana e portanto tem grande importância para pesquisa científica. Apesar da relevância, poucos são os estudos para minimizar os problemas intrínsecos ao cultivo destes cereais. O principal fator limitante no cultivo do arroz são as plantas daninhas, devido competirem por recursos inerentes a um desenvolvimento saudável, fazendo necessário o controle químico que se apresenta como método mais usado no manejo de plantas invasoras, onde em muitos casos a cultivar sofre perdas produtivas, pelo contato com herbicidas utilizados. A excreção de xenobióticos, incluindo desintoxicação de herbicidas está intrinsecamente relacionada a superfamília de enzimas glutationa s-transferases (GST's) que conferem em arroz (*Oryza sativa*) a proteção ao estresse biótico e abiótico, atuando na biotransformação de proteção contra estresse oxidativo. A bioinformática através do método de modelagem molecular por homologia, surge com uma ferramenta adequada para a predição teórica da estrutura de proteínas, sendo uma poderosa alternativa para a superação de alguns obstáculos envolvidos na elucidação de estruturas protéicas por técnicas experimentais, visto sua celeridade quanto ao processo elucidativo e de validação de estruturas protéicas em curto espaço de tempo e a custos reduzidos. Portanto, esse trabalho tem como objetivo elucidar a estrutura tridimensional/funcional de uma proteína de arroz da família das Glutaciona S-transferase e validá-la por metodologia *in-silico*, demonstrando sua importância biotecnológica e as vantagens voltadas para a agroindústria e indústria de defensivos agrícolas.

**Palavras chave:** Bioinformática, Herbicida, *Oryza sativa*, Glutaciona



## STRUCTURAL MODELING OF OSJGSTF3 PROTEIN WITH POTENTIAL DETOXIFICATION OF HERBICIDES IN *Oryza sativa japonica*

### ABSTRACT

Rice is an herbaceous plant in the grass family, from the genus *Oryza*, which feeds more than half the human population and therefore has great global importance. In spite of such importance, there are just a few studies to minimize the problems intrinsic to the cultivation of these cereals. One of the main limiting factor in the cultivation are rice weeds, due to competitor resources, inherent in a healthy development, the chemical control is still the most important form of weeds management, where in many cases the cultivar suffers productive losses, by contact with the herbicide. It is known that the superfamily of glutathione s-transferase (GSTs) enzymes confer on rice (*Oryza sativa*), a biotic and abiotic aesthetic, acts on biotransformation to protect against oxidative stress and excretion of xenobiotics, including herbicide detoxification. It is necessary to explore the mechanism of interaction, as well as the proteins with herbicides in question. A bioinformatics uses the molecular modeling method, for theoretical prediction of the structure of proteins, being a powerful alternative for overcoming some obstacles involved in the elucidation of protein structures by experimental techniques, since its celerity in the elucidation process and validation of protein structures in a short time and at reduced costs. This work aims to elucidate a three-dimensional/functional structure of a rice protein of a Glutathione S-transferase family and to validate it by in silico methodology, demonstrating its biotechnological importance and as advantages for agroindustry and agro-industry.

**Keywords:** Bioinformatic, Herbicide, *Oryza sativa*, Gluthatione

### INTRODUÇÃO

O *Oryza sativa*, também conhecida popularmente como “arroz”, é um dos cereais mais produzido e consumido do mundo, tornando o principal alimento dos países desenvolvidos sendo ultrapassado apenas pelo trigo (1). Pelo o seu excelente valor nutritivo, tornou-se parte da culinária brasileira e tornando assim o principal alimento de países asiáticos, fornecendo valor energético e proteico para os consumidores (2). No período de 2016 a 2017 foi previsto um aumento no consumo do arroz no qual foi constatado um total de 2,5 bilhões de toneladas do grão em escala mundial (3). Não obstante sua notável importância, verifica-se certa escassez de estudos para desenvolver técnicas com potencial de minimizar problemas intrínsecos ao cultivo desta cerealífera, como a competição com plantas daninhas, por exemplo (4). A maior parte dos custos na produção do arroz vem principalmente da utilização de fertilizantes, agrotóxicos, utilização de maquinários e sementes que representa em



torno de 60% do custo (3). Devido o decréscimo produtivo e de qualidade causado e a dificuldade na realização da colheita da cultura (5-6) faz-se necessário o uso de herbicidas. Herbicidas são seletores químicos capazes de classificar ervas daninhas de plantas produtivas provocando assim a morte desses organismos. A classificação de herbicidas se dá através de agrupamentos de acordo com seu mecanismo de aplicação nas plantas e sua estrutura química básica. Geralmente herbicidas que fazem parte do mesmo grupo químico exibem efeitos similares nas plantas susceptíveis, apesar de existir exceções a regras (7).

AS Glutathione S-Transferase (GSTs) são enzimas que tem como principal função catalisar o ataque nucleofílico da sua forma reduzida em compostos orgânicos que possuem em suas estruturas químicas um átomo de carbono, nitrogênio ou enxofre eletrofilico, e se encontram dessa forma em meio biológico como homo ou heterodímero possuindo assim o papel importante na biotransformação e excreção de xenobióticos além de proteção contra estresse oxidativo (8). Sendo enzimas sintetizadoras que podem adicionar ou substituir tripeptídios Glutathione (GSH;  $\gamma$ -Glu-Cys-Gly) sem auxílio de ribossomo por um determinado núcleo eletrofilico que recebe essas determinadas moléculas (9). Por fazer parte de um conjunto de super enzimas, as glutathione são encontradas intracelularmente em altas concentrações, essencialmente em todos os organismos aeróbicos. Devido ao potencial inibitório das enzimas relacionadas à GSH tendo como alvo para desenvolver substâncias que apresentam características farmacológicas (8). O processo de geração atua no meio intracelular, a partir da atuação de duas enzimas que fazem a ligação peptídica entre os aminoácidos envolvidos, classificadas em duas fases, na seguinte sequência: inicialmente a biossíntese, que começa com a catalisação da ligação peptídica dos aminoácidos ácido glutâmico e cisteína, produzida através da enzima  $\gamma$ -glutamylcisteína sintetase; conseqüentemente, este dipeptídeo, une-se a glicina pela reação da enzima glutathione sintetase. (10)

O método de modelagem comparativa, também chamada de modelagem por homologia, baseia-se no conceito de evolução molecular. As estruturas de proteínas tridimensionais (3D) apresentam grande importância e interesse para desenhos racionais de experimentos biológicos, como a mutagênese voltado para a evidênciação de inibidores específicos. Sendo assim, verifica-se que a modelagem por



homologia é um método confiável para gerar a estrutura tridimensional da proteína através da sequência de aminoácidos (11). Duas condições são, geralmente realizáveis: A primeira é conhecer a família proteica a que pertence a proteína-problema selecionando a proteína-molde diretamente do PDB e a segunda condição é procurar de forma sistemática um ou mais moldes em um banco de dados de estruturas primárias derivadas de proteínas armazenadas no PDB para identificação da família na qual a proteína-problema faz parte. Por fim, a validação tem como objetivo assegurar que as etapas atende às necessidades do processo, ou seja, a efetivação de que estas cumpram suas determinações. A validação é uma etapa fundamental que pode ser efetuada em diferentes níveis de conformação estrutural.

## METODOLOGIA

**Data Mining:** A mineração de dados da proteína-problema foi feita no servidor NCBI, onde foi encontrado e obtida a sequência (FASTA).

**Escolha do molde, modelagem:** O servidor SWISS-MODEL (12) usa um do banco de dados de proteína remanescentes (PDB), tem como pretensão eliminar algumas instabilidades(11). Os prováveis modelos são classificados conforme seu valor, identidade de sequência, resolução e qualidade da estrutura(11). Essa construção de um modelo de homologia implicam basicamente em quatro etapas sucessivas: [1] identificação e seleção de modelo(s) estrutural e seleção de proteínas-molde; [2] alinhamento das sequências-alvo de resíduos e estrutura(s) do modelo; [3] construção das coordenadas do modelo; [4] validação do modelo. Para essa busca, podem ser utilizados bancos de dados de sequência, ferramentas de busca, como o BLAST. Após a definição do molde, ocorre o alinhamento da sequência-problema com a sequência-molde. O alinhamento tem como propósito alinhar resíduos estruturalmente similares levando em conta elementos de estrutura secundária(13). Inicialmente se faz a identificação de no mínimo uma proteína de estrutura tridimensional trivial, que será utilizada como molde para definição da estrutura da proteína-problema(14).

**Validação estrutural:** O processo de validação da estrutura tridimensional foi feito na plataforma de modelagem do SWISS-MODEL baseia-se em avaliação da qualidade do empacotamento global da proteína, os possíveis erros estruturais em



regiões localizadas e os parâmetros estereoquímicos. (15). Este processo é feito através do programa PROCHECK (16-17) para verificar a qualidade estereoquímica da estrutura. Foi calculado para valores para o ANOLEA (18) e o GROMOS (<http://www.gromos.net/>).

ANOLEA – Avalia a qualidade de estruturas de proteínas com uma energia de interação atômica não-local. O processo é feito por meio da análise de ambiente dos átomos pesados por meio de cálculos de energia, sendo possível zonear regiões de alta energia na proteína, as quais estão relacionadas a erros pontuais ou regiões de interação. O ANOLEA sumariza os resultados, quantificando aminoácidos situados em zonas de maior energia (em número e porcentagem) e quais são eles.(19).

GROMOS: O programa GROMOS (GRONIGEN MOLECULAR SIMULATION) descreve no eixo Y a representação do campo de forças empíricos para cada aminoácido da cadeia de proteínas. (20)

A visualização da molécula foi através do VMD *Visual Molecular Dynamics* (<http://www.ks.uiuc.edu/Research/vmd/>).

## RESULTADOS E DISCUSSÃO

**Data Mining:** A mineração de dados da proteína-problema foi feita no servidor NCBI, onde foi encontrado e obtida a sequência FASTA (linear) de uma proteína putativa da super-família Glutathione S-Transferase, classe “Phi” de arroz (*Oryza sativa japonica*), com 224 aminoácidos de comprimento código de acesso AAG32477.1.

**Escolha do molde, modelagem:** O algoritmo de alinhamento global do SWISSMODEL, encontrou um molde (PDBID: 4ri6) de *Populus tremula* em estado homo-dimérico, com 56.60 identidade de alinhamento, elucidado por meio de difração de raio-X, com resolução de 1.52 Å [2].



```

Model_01:A  MAAPVTVYGPMISSPAVARVAACLLEKDVFPQVEPVDMSKGEHKSPSFLKLQPPGQVPAFK
Model_01:B  MAAPVTVYGPMISSPAVARVAACLLEKDVFPQVEPVDMSKGEHKSPSFLKLQPPGQVPAFK
4ri6.1.A    MATPVTIYGPPLSTAVSRVLATLIEKDVPFHLLIPIDLKSGEQKKPEYLKIQPPGQVPAFK

Model_01:A  DSLTTVFESRAICRYICDQYADSGNK-TLMGRKEDGAVGRAAIEKWIEAEGQSFNPPSLA
Model_01:B  DSLTTVFESRAICRYICDQYADSGNK-TLMGRKEDGAVGRAAIEKWIEAEGQSFNPPSLA
4ri6.1.A    DESITLTFESRAICRYICDKYADKGNKSLY----GTDILSKANIDQWVETDGTGQTFGPPSGD

Model_01:A  MAFQLAFAPFMGRATDMAVVEQNEAKLVKVLVDVYEQWLGENQYFAGDEFSLADLVHMPNT
Model_01:B  MAFQLAFAPFMGRATDMAVVEQNEAKLVKVLVDVYEQWLGENQYFAGDEFSLADLVHMPNT
4ri6.1.A    LVHDLFLFSS---VPVDEALIKKNVDKLAKVLDIYEQKLGQTRFLAGDEFSFADLSHLPNG

Model_01:A  DLLVRKTNKAG-LFTERKNLAKWWDEVSARPSWKKVVELQNVPRPS
Model_01:B  DLLVRKTNKAG-LFTERKNLAKWWDEVSARPSWKKVVELQNVPRPS
4ri6.1.A    DYLVNS-TDKGYLFTSRKNVNRWWTEISNRESWKKVLEMRKN----

```

Figura 1. Alinhamento feito pelo SWISSMODEL, no formato Clustal-W, para ambas subunidades da proteína-problema homodimérica GSTF3 de *Oryza sativa Japonica*, e a proteína-molde de *Populos tremula*.

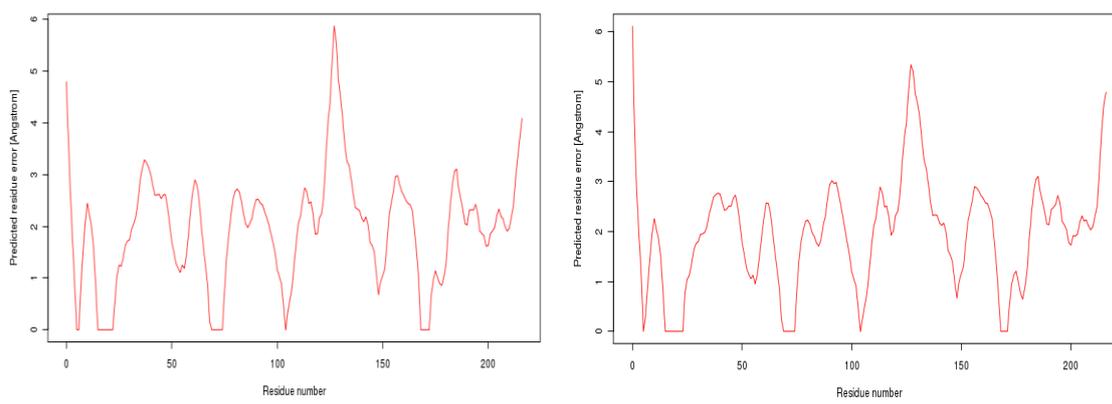
As coordenadas atômicas foram portanto salvas em formato/extensão .pdb, para que se seguisse o processo de validação estrutural do modelo teórico, no servidor SWISSMODEL.

**Validação estrutural:** O processo de validação da estrutura tridimensional baseou-se em avaliação da qualidade do empacotamento global da proteína, os possíveis erros estruturais em regiões localizadas e os parâmetros estereoquímicos. (14).

O gráfico do erro residual exhibe o erro do resíduo ao longo da cadeia em função de seu desvio, medido em ângstroms (Figura 2A e Figura 2B). Gráficos de qualidade local () em função do resíduo de aminoácido também conseguem descrever o padrão de comportamento homodimérico da GST estudada, como descreve a Figura 2C.

A

B



C

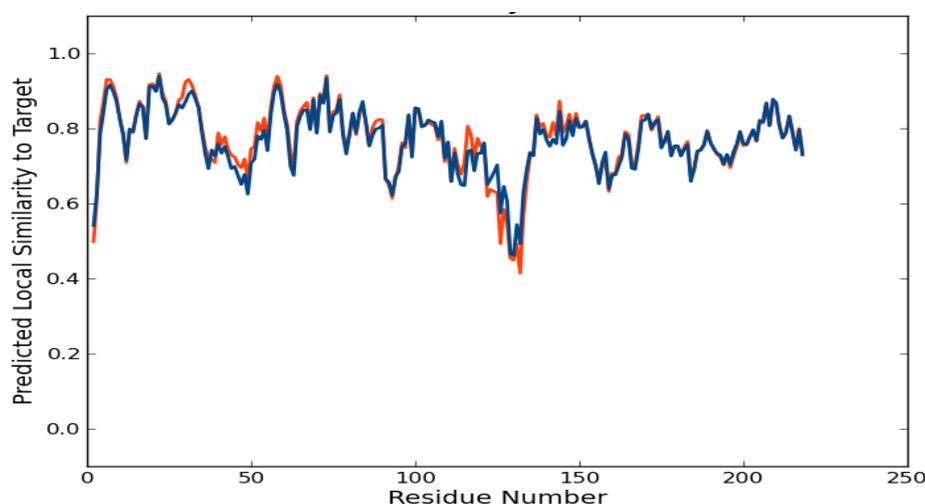


Figura 2. Gráficos de qualidade residual-local do modelo teórico

O QMEAN (*Quality Mean Score*) para seis parâmetros avaliou a interação de parâmetros de pseudo-energias em conjunto com o Z-score, respeitando assim, parâmetros existentes em estruturas obtidas experimentalmente com tamanho proporcional ao modelo teórico em estudo (21).

O Quality Mean Score para 6 parâmetros em conjunto com o Z-Score, Figura 3 A, foi de 0,73, o que pode ser considerado bom, já que este valor tem que ser próximo a 1, para um modelo de melhor qualidade. (21)

Na Figura 3 B os círculos em escala de cinza representam scores do QMEAN para outras estruturas de tamanho similares, obtidas com metodologias experimentais, depositadas no PDB. Dessa forma, o Z-score do QMEAN indica quantos desvios padrões o score difere dos valores esperados das estruturas experimentais. Nesse caso, quanto mais próximo da parte preta melhor, pois menor Z-score, menor desvio padrão. Logo, o Z-score do QMEAN foi de 0,23 o qual pode ser considerado um bom



valor, também, como podemos observar encontra-se próximo à parte preta. (21) Os parâmetros de qualidade aferidos para qualidade estereoquímica total foram considerados bons (Figura 3C)

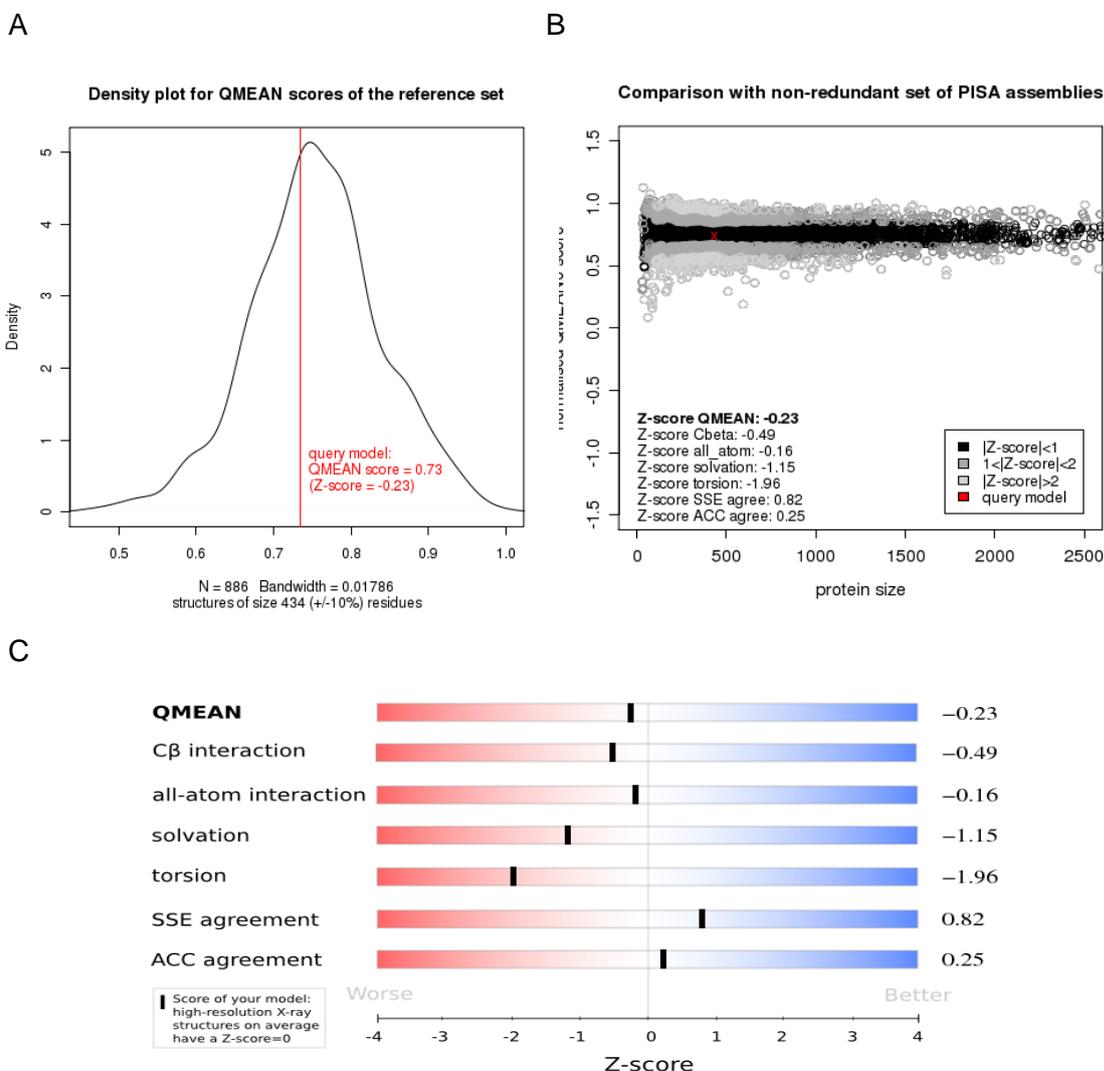


Figura 3, Descritores de qualidade média em função do índice de Z-Score (Figura 1A). Qualidade do modelo teórico comparado à modelos reais, de mesmo tamanho (Figura 3B). Descrição de qualidade do modelo em função das médias de qualidade estereoquímicas (Figura 3C).

O gráfico de Ramachandran mostrou 94,2% dos resíduos situavam-se em regiões energeticamente mais favoráveis (vermelho), 5,0% em regiões permitidas (amarelo) , 0,8% em regiões generosamente permitidas (creme) e 0,0% em regiões proibidas (branco) como mostra a Figura 9.0. Um bom modelo, requer mais de 90% de



seus resíduos de aminoácidos situados em regiões favoráveis do gráfico de Ramachandram (A, B e L), sendo desconsiderados resíduos como glicina, prolina, e os que estiverem situados nas extremidades, pois estes apresentam padrões estereoquímicos diferentes dos demais resíduos, apresentando uma natural perturbação termodinâmica (16).

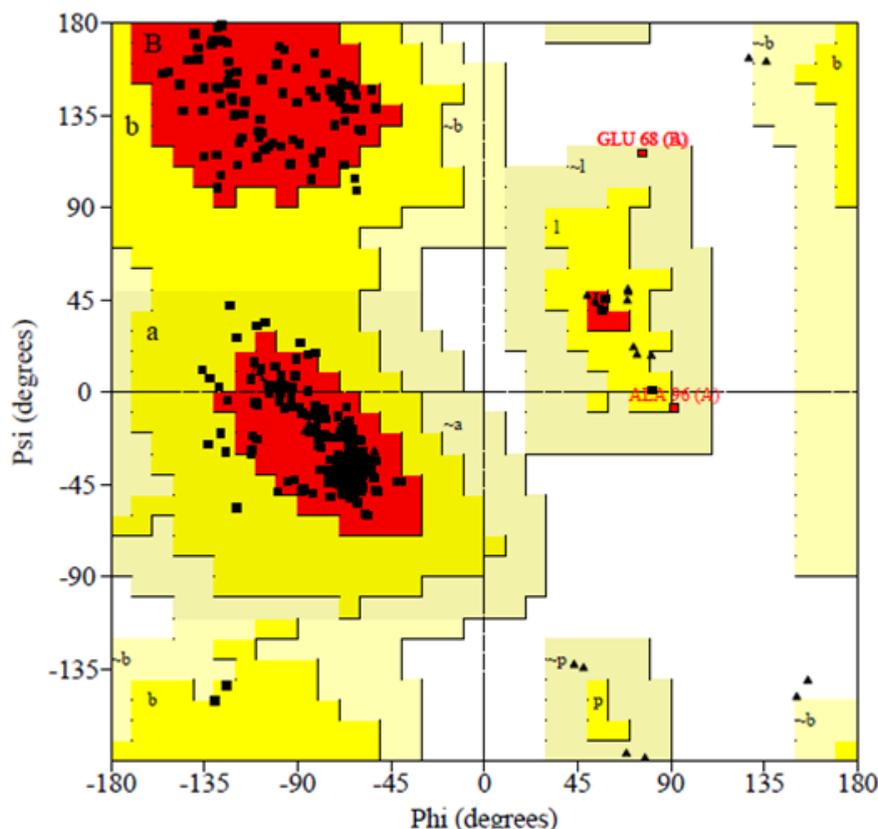


Figura 9. Descritores de estabilidade estereoquímica dos resíduos de aminoácido em função dos ângulos permitidos Phi-Psi, regiões com gradiente de cor mais quente são mais favoráveis termodinamicamente.

## CONCLUSÃO

Os indicadores de estabilidade estereoquímica e termodinâmica da molécula, indicam que o modelo teórico desenvolvido e estudado apresenta potencial para representar apropriadamente um modelo real em estudos teóricos. Apresentando resultados satisfatórios quanto a confirmação de sua estabilidade físico-química estrutural.



A bioinformática se mostrou uma ferramenta célere e eficiente para elucidação estrutural, e para estudos bioquímicos-estruturais de proteínas. Este tipo de ferramenta pode auxiliar a exploração do mecanismo de ação metabólica das proteínas.

## REFERÊNCIAS

1. Costa JA, Silva MC, Oliveira IP, Costa FR, Lima AF. Respostas de Aplicações de Diferentes Doses de Zinco na Cultura do Arroz em Solos do Cerrado. Rev Focu Mont Bel (FMB). 2015; 9 (1): 59 - 139.
2. Cordeiro ACC, Suhre E, Medeiros RD, Vilarinho AA. Sistemas de Cultivo e Manejo de Água na Produção de Diferentes Genótipos de Arroz em Várzea, no Estado de Roraima. Pesq Agropec Trop. 2010; 40(3):362-369.
3. CONAB. 2017 [cited 11 October 2017]. Available from: <http://www.conab.gov.br/conteúdos.php?a=1252>
4. Silva M, Durigan J. Períodos de Interferência das Plantas Daninhas na Cultura do Arroz de Terras Altas: I – Cultura do IAC 202. Planta Daninha. 2006;24(4):685-694.
5. Cardoso G, Alves P, Beltão N, Vale L. Períodos de interferência das plantas daninhas em Algodoeiro de Fibra Colorida 'BRS Safira'. Rev Ciê Agron. 2010; 41(3):456-462.
6. Freitas R, Berger P, Ferreira L, Cardoso A, Freitas T, Pereira C. Interferência de Plantas Daninhas na Cultura de Algodão em Sistema de Plantio Direto. Plant Dan. 2002; 20(2): 197-205.
7. Oliveira RS, Constantin J, Inoue MH. Biologia e Manejo de Plantas Daninhas. 22<sup>a</sup> ed. Curitiba, PR: Omnipax; 2011.
8. Huber PC, Almeida WP. Glutathione e Enzimas Relacionadas: Papel Biológico e Importância em Processos Patológicos. Quim Nova. 2008 abr; 31(5): 1170-1179.
9. Dixon DP, Edwards R. Glutathione Transferases. The Arab Book/Ameri Soci of pln biolog. Mai 08 2010; 8:e0131. 1-15. doi: 10.1199/tab.0131. PMC3244946.
10. Amador VC, Maia RT, de Sousa AA. Bioinformática na Resistência de Plantas a Herbicidas. 1st ed. Novas Edições Acadêmicas; 2017.
11. Filho OAS, de Alencar RB. Modelagem de Proteínas por homologia. Quim Nova. 2003; 26(2): 253-295.
12. Arnold K, Bordoli L, Koop J, Schwede T. (2016). The SWISS-MODEL Workspace: A web-based environment for protein structure homology modelling. Bioinformatics, 22, 195-201.
13. Ramalho TC, Freitas MP, da Cunha EFF. Chemoinformatics: Directions Toward Combating Neglected Diseases. Rio de Janeiro: Bent e Book; 2012.
14. Lima EB. Detecção de Subfamílias Proteicas Isofuncionais utilizando integração de Dados e Agrupamento Espectral. Belo Horizonte [Internet]. 2015 [acesso em 2017 out 18]. Disponível em: [http://vulcano.grude.ufmg.br/dspace/bitstream/handle/1843/BUOS-APTNC/tese\\_elisa\\_boari\\_de\\_lima\\_2011695257.pdf?sequence=1](http://vulcano.grude.ufmg.br/dspace/bitstream/handle/1843/BUOS-APTNC/tese_elisa_boari_de_lima_2011695257.pdf?sequence=1).
15. Reizer J, Reizer A, Saier MH. The MIP Family of Integral Membrane Channel Proteins: Sequence Comparisons, Evolutionary Relationships, Reconstructed Pathway



of Evolution, and Proposed Functional Differentiation of the Two Repeated Halves of the Proteins. *Crit Revi in Bioch and Mole Biol.* 2008 set 26; 28 (3): 235-237.

16. Laskowski RA, MacArthur MW, Moss D, Thornton JM . PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Cryst.* 1993; 26(2):283-291.

17. Laskowski RA, Rullman JA, MacArthur MW, Kaptein R, Thornton JM. AQUA and PROCHECK-NMR: Programs for Checking the Quality of Protein Structures Solved by NMR. *J Biomol NMR.* 1996; 8(4): 477-86.

18. Melo F, Feytmans E. Assessing Protein Structures With a non-local Atomic Interaction Energy. *Jourl of Mole Biol.* 1998; 277(5): 1141-1152.

19. Melo F, Feytmans E. Assessing Protein Structures With a Non-local atomic interaction energy. *Jorn of Mol Biol.* 1998;277(5):1141-1152.

20. Gusteren W. *Biomolecular Simulation.* Zürich:Vdf, HochsChulverlag an der ETH; 1996.

21. Benkert P, Biasini M, Schwede T. (2011). "Toward the Estimation of the absolute quality of individual protein structure models". *Bioinformatics*, 27(3):343-50.