IDENTIFICATION, FUNCTIONAL ANNOTATION & MOLECULAR MODELING OF GLYCOXYL HYDROLASES FROM ACHATINA FULICA METAGENOME

Guimarães, M.¹,²; da Silva, M.L.¹; Gomes, D.E.B.¹

¹Grupo de Biologia Computacional, Diretoria de Metrologia Aplicada às Ciências da Vida – DIMAV, Instituto Nacional de Metrologia, Qualidade e Tecnologia – INMETRO, Duque de Caxias, Rio de Janeiro, Brazil; ²Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil.

E-mail: maguimaraes0@gmail.com

There is a major international effort to develop renewable alternatives to fossil fuels. Bagasse, a cellulose-rich biomass and product of sugarcane industrialization, is a great source for second-generation ethanol, not competing with the food industry. Cellulases (glycosyl hydrolases – GH) are enzymes that degrade cellulose. Extensive research is being performed in attempt to obtain cellulases with higher activity on pretreated biomass substrates by screening and sequencing from cellulose degrading microorganisms. Here, we probed the metagenome of the african snail's gastrointestinal content and identified two potential GH candidates through sequential, functional & structural annotation. Also, we characterized their active sites and generated computational models to predict interactions with substrates. Both GH were identified by searching the NCBI database and their functions were annotated by similarity to UniProt and KEGG records. The Protein Data Bank (PDB) was probed for suitable templates while models were built with Swiss-Model and validated by QMEAN scoring. The interaction with substrates, 1-6 glucose residues (GR) or xylose residues (XR), was predicted by molecular docking using AutoDock. The sequential and functional annotations revealed the sequences to be a β-D-xylosidase (GH43, EC:3.2.1.37) and an endoglucanase (GH8, EC:3.2.1.4) from Enterobacteriaceae sp. with a coverage 99% and 93%, and identity 87% and 81%, respectively. Structural templates, PDB codes 2EXJ and 1WZZ, covered 99% and 96% of sequences, with 55% and 38% identity, allowing the generation of reliable models, the best values have QMEAN < -3.53 and -4.98. Molecular docking indicate that 2XR and 3,4GR are better substrates to GH43, while extensive chains bind improperly. For GH8, substrates were found to be D-glucose and cellobiose (1,2GR), all others showed unstable interactions at the site. As a perspective molecular dynamics is being performed to refine models & confirm stability of binding modes, while GH were cloned, produced, with good crystals for GH8.

Keywords: Biofuels, cellulase, metagenome.

Acknowledgements: Capes-BioComputacional