Prospecting novel proteins from *Deinococcus radiodurans*: a model for putative heat shock proteins

Ricardo Valle Ladewig Zappala[1,2], Pedro Geraldo Pascutti[1,2], Manuela Leal da Silva[2,1] e Claudia A. S. Lage[1]

[1]Instituto de Biofísica Carlos Chagas Filho, Universidade Federal do Rio de Janeiro (RJ), [2]Diretoria de Metrologia Aplicada às Ciências da Vida, Instituto Nacional de Metrologia, Qualidade e Tecnologia (RJ)

The Deinococcaceae group comprises some of the robust known extremophilic bacteria. Attempts have specially focused on responses against extreme doses of gamma radiation, to explain survival mechanisms of *Deinococcus radiodurans* against simultaneous stresses, as desiccation and heat. *D. radiodurans* has many defensive mechanisms, and transcriptomes already made in response to gamma radiation and desiccation revealed that some genes were transcribed to proteins of undefined functions, while others have never been expressed under those conditions. Therefore, it is expected that such genes with obscure function can code for novel resistance proteins to these extremophilic conditions. The present study aims to identify and perform function prediction for hypothetical, unique proteins of *D. radiodurans*, without similarity to any other known protein. A group of proteins expressed in *D. radiodurans* after gamma radiation was retrieved, which hypothetical functions were predicted by the best scores after Psi-BLAST alignments and CD-search. Information about the proteins was gathered through alignments against Uniprot and PDB databases. Using molecular modeling tools as I-TASSER, SWISS MODEL and MODELLER, 3D models were successfully built for 20 out of 26 hypothetic proteins and they were initially evaluated by Ramachandran's Plot and RMSD. The best models from were then submitted to structural classification on SCOP and CATH servers. That way, we were able to speculate about the function of some candidates, and generate other models whenever structural and sequential annotations disagree. Among the 20 analyzed proteins, one of the most interesting was the *DR0491* gene product, showing 25% identity and 41% similarity covering 90% of the sequence correspondent to the *Escherichia coli* heat shock protein Hsp31. This may represent an essential role on catalysis of damaged proteins, as well as proper folding assistance on other unstable proteins. Moreover, the catalytic essential residues of the *E. coli* protein were also found at the correct position in our *D. radiodurans* modeled protein. Other 19 proteins seem to bind to nucleic acids, acting on metabolic regulation under severe conditions. This particular resistance toolbox with novel and exclusive proteins was referred as the "Black Box Genome of *D. radiodurans*". Additionally, our results reveal promising candidates for future biotechnological approaches.

Funding support – CAPES, CNPq