Functional prediction of stress-modulated proteins of Deinococcus radiodurans

Ricardo Valle Ladewig Zappala[1], Manuela Leal da Silva[2], Pedro Geraldo Pascutti[1], Claudia de Alencar Santos Lage[1],

1 Universidade Federal do Rio de Janeiro
2 Instituto Nacional de Metrologia Qualidade e Tecnologia

Abstract

The Deinococcaceae group comprises some of the robust known extremophilic bacteria. Attempts have specially focused on responses against extreme doses of gamma radiation or desiccation to explain survival of Deinococcus radiodurans against them. Many defensive mechanisms were shown to exist in D. radiodurans, and transcriptomes already performed 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 unknown functions could code for novel resistance proteins to those extreme 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. Sequences of a group of 26 proteins, with 23 expressed in D. radiodurans after radiation or desiccation stresses were retrieved, which hypothetical functions were predicted by the best scores after 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 built for all hypothetic proteins and they were mainly evaluated by Ramachandran's Plot, RMSD and DOPE score. The best models were then submitted to structural classification on SCOP and CATH servers. This approach enabled us to better infer about the function of twenty candidates for new extremophilic proteins, of which thirteen went through comparative modeling with multiple templates. Three seemed to belong to the group of intrinsically disordered proteins, and three have not aligned to any proper templates by comparative modeling. Among the seven predicted proteins using structures from SWISS MODEL are: DR0438, a DNA binding protein; DR1263, a N-glycosidase; DR1314, a photosystem-like transmembrane protein; DR1370, a structural lipoprotein; DR2073, a kinase; and DR2441, an acetyl-transferase. Among the 26 analyzed proteins, the most interesting one appears to be the DR0491 gene product, showing 25% identity, 41% similarity, and 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. After several steps of investigation, modeling and structural analyses, complementary tools such as phylogeny, molecular dynamics and molecular docking were also performed to strengthen the significance of the observed results, and this particular resistance toolbox with novel and exclusive proteins was referred as the "Black Box Genome of D. radiodurans".

Funding: CNPq