Mutation Analysis for AgrC from Staphylococcus aureus

Samuel da Silva Guimarães, Danielle Mendes Silva, Mônica Pacheco Silva, Pedro Marcus Vidigal, Andréa de Oliveira Barros Ribon, Sabrina de Azevedo Silveira

Universidade Federal de Viçosa

Staphylococcus aureus is one of the main pathogens of bovine mastitis. The analysis of single nucleotide polymorphisms (SNPs) on the genomes of four strains of *S. aureus* associated with mastitis showed the presence of variations in the sequence of the agrC protein, which can influence in the expression of genes related to the illness. Thus, it was hypothesized that these variations could be related with different manifestations of the disease. We searched for the reference sequence of agrC in Protein Data Bank (PDB) and found the 4BXI.A entry (Crystal structure of ATP binding domain of AgrC from *Staphylococcus aureus*), which comprehends just part of the sequence, with 153 residues (278-430). From the 4BXI.A, a new search was made in the PDB to find similar structures, resulting in a set of one representative chain of 82 different structures, grouped by 40% of sequence similarity. A pairwise structural alignment was performed using the MultiProt software, aligning each of the structures with the 4BXI.A, and a visual representation was generated with the CINEMA color scheme so that each sequence is shown as a line and the each column corresponds to the alignment position. Also, we used the Expectation Maximization algorithm (EM) to make similar sequences appear next to each other. Subsequently, the interactions were modeled as graphs in which atoms represent nodes and edges represent interactions between atoms. To calculate the interactions we used the Voronoi diagram followed by the Delaunay triangulation and we labeled nodes as positively charged, negatively charged, aromatic, hydrophobic, donor or acceptor. The edges were labeled according to a distance criteria and the type of edges as hydrogen bond, aromatic stacking, hydrophobic, repulsive and salt bridge. The computed interactions were mapped from the atomic level to the residue level, so that nodes represent residues and edges represent interactions between residues. From the set of the protein modeled as graphs, some centrality measures (degree, closeness, betweenness), that are commonly used in complex networks, were calculated using the iGraph, each of them providing a distinct perspective of centrality. This approach enabled us to identify the varying positions of the agrC sequence that are potentially important to protein structure and should be studied further to evaluate the prevalence of this variation in a larger number of strains and also to determinate its effect on the ability of bacteria to cause disease.

This work has been supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG).