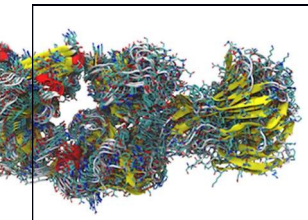

IV ASBioSim – Advanced School on Biomolecular Simulation: Protein Engineering with Rosetta, from fundamental principles to tutorials
05 to 10/May, 2019, Recife/PE, Brazil

GriDoMol: Integrating a computational grid and Rosetta framework for reverse vaccinology studies

Luiz Felipe Gomes Rebello Ferreira[1]; Marcelo Zaldini Hernandes[1]

[1]*Pharmaceutical Sciences Department, Federal University of Pernambuco*

Abstract: The reverse vaccinology is a computational (in silico) approach that, with a pathogen genome in hands, seeks for antigens capable of triggering immunologic response and, thus, develop a vaccine that can be used against specific diseases. By using the reverse vaccinology methods, one can abbreviate the development time involved on the development of a new vaccine to 1 or 2 years, significantly faster when compared to the conventional methods, which usually ranges from 5 to 15 years. This speedup happens because the reverse vaccinology can identify and test (computationally) practically all proteic antigens that may be expressed by the pathogen, building a comprehensive "catalog" of antigen candidates, increasing the chances of finding antigens with the potential of becoming a vaccine. Furthermore, as only the best candidates are experimentally tested, the reverse vaccinology also incurs the reduction of financial costs associated on the discovery of new vaccines. Therefore, a reverse vaccinology method has been elaborated and implemented on our laboratory's in-house software known as GriDoMol in order to prepare, distribute and analyze a large quantity of molecular docking involving epitopes candidates and several human Major Histocompatibility Complex (MHC) I and II alleles using several modules of the Rosetta framework distributed among the computers that integrates our computational grid infrastructure. The GriDoMol's reverse vaccinology module possess a curated database composed of 21 MHC I and 12 MHC II allele's structure and takes as input FASTA files containing the amino acids sequence that represents each of epitope candidates that may be addressed to be docked into the alleles group of one of the MHC's classes. Our reverse vaccinology methodology has been divided into three phases: i) Complex creation (FixBB), stabilization (Relax) and scoring (FlexPepDock); ii) Filtering, by windows and frequencies occurrences, and default flexible peptide docking and iii) Refinement steps by performing the flexible peptide docking using higher number of solutions. This methodology has been applied in a study case that aimed to discover potential epitopes candidates that may trigger an immunologic response against the *Leishmania braziliensis*, which has been publish under the title "Combination of In Silico Methods in the Search for Potential $CD4^+$ and $CD8^+$ T Cell Epitopes in the Proteome of *Leishmania braziliensis*" on the "Frontiers in Immunology" journal. A total of 32.658 complexes (epitope + allele) have been computationally (in silico) tested and the top 10 epitope candidates have been selected for the experimental tests. It is important to note that each individual molecular docking involving a single complex, that is, one epitope candidate being docked into one allele, has no logic or temporal dependency on docking calculations involving other complexes, representing an overall "bag-of-tasks" type of approach, which is the ideal scenario for a task distribution involving computational grids. In this way, by having both the speedup provided by the MPI implementation on Rosetta's source code and the computational grid speedup provided by the integrated GriDoMol's modules, the necessary time to execute a massive quantity of molecular docking calculations is greatly reduced when compared to the execution on a single computer. The capacity of inducing the proliferation of peripheral blood mononuclear cell (PBMC), derivate of humans that have been cured after treatments, has been measured and all these epitope candidates have stimulated the proliferation of PBMC and on five of these epitope candidates the proliferation capacity have been statistically significant when comparing the patients and the control volunteers. Additional studies are being conducted by the experimental group, using both in vitro and in vivo models, in order to elucidate the immunologic behavior of these epitopes candidates that has been selected by our GriDoMol's program.