Automation of polyproteins GAG and GAG-POL-1 cleavage site mapping: A study of large scale molecular docking

Fernando Limoeiro Lara de Oliveira[1 2], Maria Fernanda Ribeiro Dias.[2] , Manuela Leal da Silva[2]

[1]Universidade Federal do Rio de Janeiro; [2]Instituto Nacional de Metrologia, Qualidade e Tecnologia – Diretório de Metrologia Aplicada às Ciências da Vida (DIMAV)

Viral infections affect populations over the globe. The most of the treatments available for the population is based on prevention of infection rather than extermination of the virus. The difficulty for the design and development of antiviral drugs results from different viruses and their mutational capacity. The drugs in generally acts on biomolecular mechanisms common to both viruses and hosts, making specificity for drugs a challenge, which will culminate in losses on health and life quality. In this context, the use of bioinformatics tools may provide a time and cost efficient approach for the study and development of antivirals, since such techniques can deal with the screening of huge amounts of biological and experimental data very rapidly. In this work, we are showing the early stages of the development of large-scale workflow for docking and mapping of the Gag and Gag-Pol poliproteins cleaved by the HIV-1 Protease. This is a well-studied mechanism of the HIV, in which GAG and GAG-pol are clived into active, smaller proteíns. These poliproteins will act on the maturation of the HIV virion. Our goal is to map the clivage sites of the poliproteins based on a more conformational approach, using the molecular docking. The first step of the workflow consists on the fragmentation of both gag and gag-pol into smaller peptides of different lengths (4, 6 and 8 acid amino). Each fragment size characterizes a different assay. We build a tridimensional structure of each peptide using the software Modeller, for both modeling and structure optimization. Each group of structures containing is automatically prepared for the molecular docking procedures using the OpenBabel and Autodock Tools softwares for a series of operations. Our preliminary tests on the workflow are being made on groups of fragments containing four acid amino. The peptides sequences were created using a Python script. The construction of the models resulted in satisfactory scores by Modeller, yielding a final number of 1445 structures, docking-ready. The next step is to perform the docking (AutodockVina) on the 4-residues GAG and GAG-pol peptide groups, in order to assess the optimal parameters for the mapping process. References and preview studies showed that peptide size, location and composition are crucial to determine substrate recognition by the enzyme. Therefore we believe that such methods, applied on natural substrates of enzymes, may provide useful information for the development of new maturation inhibitory drugs.

Financial support: CAPES and CNPq.