

STUDYING STRUCTURE FUNCTION RELATIONSHIP OF PROTEINS USING “REMEDIATED” PDB FILES

Mazoni I., Borro L.C., Alvarenga D., Jardine J.G., NeshichG.

Embrapa Informática Agropecuária, Nucleo de Bioinformatica, Campinas, SP

In August of 2007, PDB/RCSB started to offer a remediated PDB files. The major difference introduced by this format is related to the number of chains present in the PDB file. Namely, files such as the 1cho.pdb, containing the alpha-chymotrypsin with its inhibitor turkey ovomucoid third domain, have suffered a change in number of chains present in the enzyme structure: previously there was only one chain, named “E”. Now, there are 3 chains: “E”, “F” and “G”. The inhibitor chain remained unchanged. The reason why the RCSB decided to “break-up” one into the three chains in this particular case is based on the fact that the enzyme structure contains two gaps, which are treated with certain degree of difficulty by the software producers.

We have invested a significant amount of time and CPU resources in order to curate and analyze what are the consequences for in-silico analysis of the active site 3D environment after implementing remediated PDB files to STING_DB and STING interface.

For this purpose, we analyzed 2 different proteins, chosen to represent those files which suffered a change in total number of chains. Results show that the ensemble of protein structure descriptors did not change with respect to range of values used for selecting amino acids into the restricted group of active site residues.