Poster I-102

Classifiers of the protein interactions with their substrate/inhibitors

Authors:
Vera S. Garcia (*Unicamp/Embrapa*)
Paula Kuser Falcão (*Embrapa*)

Short Abstract: Basically all biological processes require protein-ligand identifications. One major challenge to biophysics-chemistry is try to predict these interactions. One step towards solving this mystery is analyzing the structural features involved in the complex formation. A good biological model to start these studies is the cystein-protease family.

Long Abstract:
Many fundamental biological processes require protein/ligand identifications to fulfill their biological function. This process is termed biological recognition. One major challenge to biophysics-chemistry is try to predict these interactions (recognition) between proteins and their ligands. Being able to predict the protein recognition methods using structural data would be valuable in understanding the structural-functional relationship and also would facilitate the drug design procedure. One step toward acquiring this knowledge is to analyze all the structural features involved in the complex formation, aiming to understand the structural interactions with their complex pattern of recognition and specificity, and trying to map those features responsible for the enzyme interactions with a specific substrate.
For this work, it was established that a good biological model should have: a well described catalytic mechanism; a good number of structurally solved representatives in the Protein Data Bank, including the isolated enzyme and in complex with different substrates; being best if found in all live kingdoms and the possibility to raise a grater interest (for example being a drug target). One of the families meeting the criteria mentioned above is the cystein-protease family.
Cystein-protease constitutes an important class of enzymes involved in the formation and hydrolysis of peptide bonds. Besides its obvious vital role in cell homeostasis, it is involved in apoptosis, Parkinson disease, muscle dystrophy, osteoporosis, is a potential target for parasitic treatment, and more. It includes plant proteases as papain and actinidin, mammalian lysosomal proteases (cathepsins B, C, H, K, X, etc.) cytosolic calcium activated proteases (calpain), parasitic proteases (cruzain), viral proteases (picornain and adenain) and many others. This large family of enzymes has a characteristic molecular topology, and all have in common the catalytic triad Cys, His, Asn (or Asp). Briefly the catalytic mechanism involves the deprotonation of the cystein sulfhydryl by the adjacent histidine, followed by the nucleophilic attack of the sulfur on the peptide carbonyl carbon.
Initially, we selected the protein structures of the members of this family using different data bases. Since the study requires protein structures solved with good resolution, only experimentally determined structures with resolution 2.9 Å or higher were selected. Both, search methods and manual analysis were used to exclude the “false positive” family members. Since the proteins grouped into families have in common evolutionary relationship and enzymatic activities and the investigation is based, primarily, in structural features, structural analysis were done to create a data bank of similar structures for these family

members, forming a structural cluster inside the family. This data bank excluded proteins which the overall structure were easily identified as being different from every other members.
With those structural clusters the alignment experiments were performed and the differences analyzed. In the next step we will perform docking experiments with the structural available substrates. With these we hope to be able to infer some structural parameters that are crucial for the interaction and complex formation.