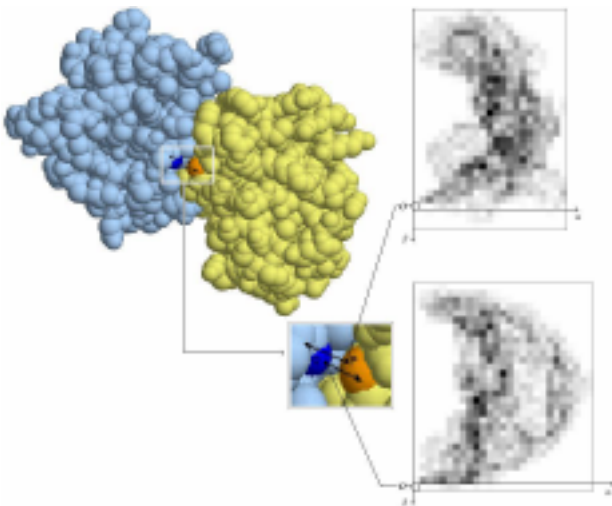


Proteomics Applications in Grid

The main tasks of proteomics are to understand protein functionality, both identifying the key role amino acids in the protein sequence and describing how interactions with other biological components take place. A typical approach to accomplish the identification of functional residues in proteins is to use methods which consist in finding similarities between well annotated proteins and those under examination. Thanks to the BioinfoGRID project a number of Bioinformatics tools can now be used to analyse sequences patterns to discover similarities by using domain profiling and Hidden Markov models.

However, sometimes these similarities are limited to a small number of amino acids in the protein sequence, such as the case of a particular functional site where the residues are sparse inside the sequence. In this context structural methods based on Grid technology can be used to compare the three-dimensional structures. In particular methods based on molecular surface similarities can give interesting information when protein function can be difficult to infer from the primary sequence.



Example of protein docking prediction identified using surface correlation.

Both the sequence based approach and the surface analysis for identifying protein functionality are computationally very expensive because they rely on a huge amount of comparison between specific protein patterns and reference databases. Moreover, modern high throughput analysis techniques have enhanced the problem of handling great quantities of data in short time, thus the possibility to use Grid computing represents an ideal solution to give researchers scalable solutions for Proteomics analysis. In the frame of the BioinfoGRID project a well-structured infrastructure has been created to perform functional protein analysis based on the resources of the EGEE project Grid platform.

The results achieved in the Proteomics applications are very promising. A number of sequence-based applications have been successfully ported on the Grid platform and thanks to the efficient system implemented for database replication, they present scalable performance. Web interfaces have been developed to enable user-friendly access to these resources by the Bioinformatics community.