

## **CYBERINFRASTRUCTURE FOR GLYCOME RESEARCH**

Arun K. Datta

National University, San Diego (CA); contact: [adatta@nu.edu](mailto:adatta@nu.edu)

Advances in Cyberinfrastructure, particularly in grid computing, are changing the way research is conducted in all aspects of science and have led to the generation of seemingly limitless possibilities of national and international collaboration and sharing of data for research, education and training. One of the applications of this development has made it possible to glue together all the databases for collaboration that individual researchers develop and maintain for their research. This capability can be exploited for shared data utilization, data analysis and more robust data mining and visualization of data. Grid computing provides high performance computing to TeraGrid, an NSF funded program that supports a number of scientific projects. The cancer Biomedical Informatics Grid (caBIG) at the National Cancer Institute is enabling the research community to share data and knowledge on cancer and beyond. Our recent development on the grid enabled information network for childhood obesity surveillance is already making some progress. Grid computing is also enabling this author for high-throughput protein motif analysis in the glycosyltransferase family. Analysis of a protein motif provides a better understanding on many aspects of protein function, protein interaction, and gene/protein and organism evolution. It also reveals evolutionary relationships between protein sequences that are too distantly related. However, such analysis is inherently highly computationally intensive because of the exponential growth of the protein databases and the combinatorial number of ways in which protein motifs interact in protein-protein interaction network. We have used MotifNetwork environment, built on biologically grid-enabled workflows, to serve this purpose and show that there is no evolutionary relatedness between the mammalian and bacterial sialylmotifs. This and other such plausible roles of grid computing in glycome research will be discussed.