

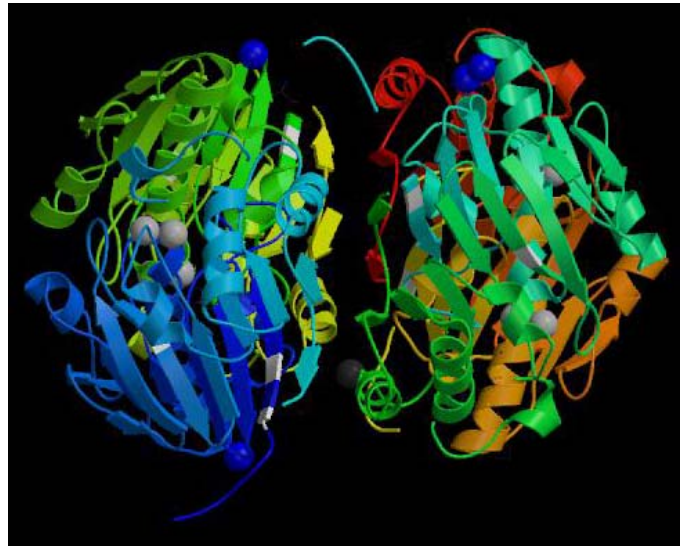
## *Biologists and computer scientists join forces to lead biology revolution*

Just as detectives use computers to compare fingerprints from crime scenes with millions in databases, Argonne scientists use computers to mine genetic information. This information is essential to progress in medicine and biotechnology.

Piecing together genetic information by computer is a relatively new field called bioinformatics. Argonne researchers are the first in this field to use hundreds of computers around the clock to analyze genomic information. Bioinformatics is a key component of Argonne's multi-million dollar, multi-disciplinary structural biology program, which provides guidance that can reduce the cost of identifying unique structures of medical and biotechnological significance.

The biology revolution came about with massive genome sequencing. More than 350 fully sequenced genomes are publicly available, and more than 780 are in the pipeline. But a known genome is alphabet soup – the human genome alone has 3 billion base pairs.

Using bioinformatics tools and advanced computers, Argonne researchers can take a small amount of known information from one genome – for example, genes involved in energy production – and search for them in all other genomes. If the same sequences are found in other genomes, then that suggests that those organisms have a similar physiology. In structural studies, comparison of protein sequences with those from proteins of known structure allows researchers to make conjectures about the presence of novel folds in proteins with no similarity to what is known. Comparative analysis of pathogenic and nonpathogenic *Mycobacterium* species implicated several genes in causing disease. Knowing these genes, researchers can seek them in other genomes; when found, their presence tells the researchers that organism is potentially pathogenic. Medical researchers can use this data to develop treatments.



*Bioinformaticists use computers to quickly perform comparative analysis to piece together the form and functions of enzymes produced by certain bacteria.*

Argonne's contributions to bioinformatics include databases and analytical tools that guide biological research. The main database, PUMA2, combines information from 22 databases. The team has also developed "Pathos" and "Chisel," software tools that work with PUMA to find specific genomic sequences. Pathos is a database for biodefense research. It contains all publicly available genomes of pathogens, including anthrax and plague. Chisel enables identification of eukaryotic and bacterial versions of the same enzyme functions.

The bioinformatics group coupled its extensive network of data and tools with the Grid. Spearheaded at Argonne, Grid technology enables supercomputers in different locations to work together seamlessly. With this kind of computing power, researchers can perform in one week comparisons that would take 18 months with a single computer.

October 2006

