

Protein Information Resource: A Community Resource for Expert Annotation of Protein Data

Cathy H. Wu and Winona C. Barker

Protein Information Resource, National Biomedical Research Foundation,
Georgetown University Medical Center, 3900 Reservoir Road, NW, Washington, DC
20007-2195

The Protein Information Resource (PIR) provides protein databases and analysis tools to support research on molecular evolution, functional genomics, and computational biology. PIR, along with the Munich Information Center for Protein Sequences and the Japan International Protein Information Database, maintains the PIR-International Protein Sequence Database, the most comprehensive, well-annotated, and non-redundant protein sequence database in the public domain. We have designed a bioinformatics infrastructure that can provide (i) high quality protein sequence data and annotation, while keeping pace with the influx of data from genome sequencing projects, (ii) an integrated system of databases and analytical tools for expert annotation and knowledge discovery, and (iii) easy database access and interoperability. Advances include: highly automated protein sequence classification and annotation, an enhanced Web site with new search tools for protein data mining and analysis, a new integrated classification database that provides comprehensive descriptions of family relationships and functional/structural annotations, database migration into an object-relational database system, and database distribution in XML format. [Supported by NIH grant P41-LM05798 and NSF grant DBI-9974855].

Presented at the Biophysical Society 45th Annual Meeting, Boston, MA, 2001.

[Back to Publications Page](#)