UniProt: New website and latest developments

UniProt is the central resource for storing and interconnecting information from large and disparate sources, and the most comprehensive catalog of protein sequence and functional annotation. UniProt is built upon the extensive bioinformatics infrastructure and scientific expertise at European Bioinformatics Institute (EBI), Protein Information Resource (PIR) and Swiss Institute of Bioinformatics (SIB). UniProt is updated and distributed every four weeks and can be accessed online for searches or download at http://www.uniprot.org. UniProt has three protein databases optimized for different uses.

The UniProt Knowledgebase (UniProtKB) is a central access point for integrated protein information with cross-references to multiple sources. The UniProtKB contains two sections. UniProtKB/Swiss-Prot contains records with full manual annotation (or computer-assisted, manually-verified annotation) performed by biologists. UniProtKB/TrEMBL contains computationally generated records enriched with automatic classification and annotation.

The UniProt Archive (UniParc) is a comprehensive sequence repository, reflecting the history of all protein sequences, and linking identical sequences from other sources.

UniProt Reference Clusters (UniRef) merge closely related sequences based on sequence identity to speed up searches.

Recent UniProt efforts have focused on:
- Building the reference proteome set
- Releasing the new UniProt beta site
- Enriched flat files with evidence tags
- Exploring proteomics directories to identify unique peptides for experimental proteins in UniProtKB