Protein Structural Databases

Ramanathan Sowdhamini

National Center for Biological Science

Distant relationships between proteins are often reflected as high structural similarity despite poor sequence identity. Few such distant relationships extend to similarities in function and can be linked in evolutionary terms as superfamilies. We maintain structural databases to provide easy access to reliable alignments of structural members of protein superfamilies and putative members in genome databases. PASS2 is a superfamily alignment database containing 628 multi member superfamilies and 566 structure annotated sequences of single member superfamilies. Conserved interacting motifs and the hidden markov models for the different superfamilies are presented in PASS2 database. SMoS is a database of Structural Motifs of aligned protein domain superfamilies. These structural motifs, along with their sequence and spatial orientation, represent the conserved core structure of each superfamily fold. The database on Genomic Distribution of protein structural domain Superfamilies (GenDiS) provides record for the distribution of 4001 protein domains organized as 1194 structural superfamilies across 18,997 genomes at various levels of hierarchy in taxonomy by employing a three fold sequence search approach. DSDBASE is a database on disulphide bonds in proteins that provides information on native disulphides and also those that were stereochemically possible between pairs of residues for 19,612 entries in Protein Brookhaven Databank (PDB). This database has potential applications in protein engineering and in structural biology.