

Protein Modeling and Bioinformatics

- Protein Structure
- Family Databases
- Fold Identification
- Residue Mutation
- Rotamer Exploration
- Multiple Alignment
- Structural Family Analysis
- Homology Modeling
- Structural Quality Assessment

MOE™'s CASP validated applications for protein structure prediction are powerful, intuitive and easy to use, both for experts and occasional users. Powerful homologue identification, alignment technology and refinement methodology make high quality sequence to structure predictions routinely possible.

Protein Structure and Family Databases. The structures deposited in the Protein Data Bank are often problematic to use because of chronic format errors. CCG has processed all the PDB structures, corrected many common errors and produced a cleaned version of the PDB. Search the database using Code, Header, Compound, Title, HET groups, resolution, etc. The cleaned database is subjected to an exhaustive and iterative structural clustering procedure to generate the Structural Family Database. The result is a database of structural families in excellent agreement with expert hand curated family databases.

Fold Identification. Search the Structural Family Database with a fold detection methodology to identify relevant protein families. The search uses a FastA-type local alignment followed by a family membership test based upon full multiple alignment and Z-score significance testing. As a result, the folds of even distantly related homologues can be reliably identified with few false positives (unlike pairwise searches). Run the search in parallel with MOE/smp compute cluster technology to perform timely whole-genome identifications.

Structural Family Analysis. Understand the conserved features and differences between related protein structures (and homologous sequences). A 3D structural family analysis provides insight into conserved geometry, conserved water molecules, salt bridges, hydrogen bonds, hydrophobic contacts and disulfide bonds that are often undetected in sequence alignments. Use structural or sequence dendrograms to eliminate outliers and improve alignments.

Mutation and Rotamer Exploration. Discover accessible amino acid side chain conformations with MOE's Rotamer Explorer. This predicts the structure of amino acid mutations in a 3D protein structure. Rank candidate rotamers using an energy-based scoring function and visually analyze them using MOE's graphical interface.

Multiple Alignment. Find optimal alignments of protein sequences, given both sequence-only and structural data using CCG's unique technology. The number of protein structures is not limited. Use arbitrary constraints and weightings of secondary structures for the alignments. Simultaneously use sequence and 3D structure information to enhance the quality of the resulting alignment, especially in the problem areas.

Homology Modeling. Build a homology model from an amino-acid sequence by statistically assembling fragments of experimentally determined backbone structures from one or more templates, selection of sidechain conformations from a rotamer library, followed by a refinement protocol based on energy minimization in the AMBER '89/'94/'99, CHARMM22/27 or Engh-Huber forcefields. A ranked ensemble of protein structures is produced and written to a MOE molecular database for subsequent analysis or further refinement. Include environment units such as bound ligands and conserved waters in the structural template.

Structural Quality Assessment. Assess the reliability of predicted structures with statistical measures of quality derived from X-ray crystallographic data. Diagnostic measures such as Ramachandran Plots, Chi Plots and a Stereochemical Report make it easy to identify and isolate regions of predicted structures that require further treatment.

MOE
Molecular Operating Environment

Cheminformatics and QSAR
Pharmacophore Modeling
Structure-Based Design

Method Development and Deployment
Molecular Modeling and Simulations
High Throughput Discovery

www.chemcomp.com

