Chemistry Central Journal

Poster presentation

Open Access

A distance-based method for evaluating protein-structures

P-P Heym*, W Brandt and L Wessjohann

Address: Leibniz Institute of Plant Biochemistry, Weinberg 3, 06120 Halle/Saale, Germany
* Corresponding author

from 4th German Conference on Chemoinformatics
Goslar, Germany. 9–11 November 2008

Published: 5 June 2009

Chemistry Central Journal 2009, **3**(Suppl 1):P47 doi:10.1186/1752-153X-3-S1-P47

This abstract is available from: http://www.journal.chemistrycentral.com/content/3/S1/P47

The *ab initio* prediction of a protein's 3D-structure from its amino acid sequence remains an unsolved problem in bioinformatics. While the backbone conformation is dependent on phi- and psi-torsion angles, the formation of secondary and tertiary structure in proteins is based on non-covalent amino acid – amino acid-interactions and the conformations of their side chains. Every non-covalent interaction like van der Waals interaction or hydrogen bonding, between two arbitrary atoms leads to an optimal distance. One can assume that a folded protein is formed by maximizing either the number of non-covalent interactions or, as in this approach, the number of optimal distances.

A distance-based method was developed. Therefore the atom pair nonhydrogen distances including side chains of a protein and each distance were evaluated within intervals of 0.5 Å. Assigned probabilities are dependent on interacting atoms which themselves depend on the amino acid combination. These probabilities were derived from statistical analyses on 4187 PDB X-ray structures whose resolution was below 2.0 Å.

It is possible to evaluate given pdb-formatted protein files and produce an output similar to the Z-score [1] used in PROSAII. A graphical output showed the all-atom interactions within a defined range of amino acids analogous to the PROSAII-plot [2] which determines whether a protein structure is native-like folded or not.

The program, written in R language using the additional package bio3d [3], receives a set of protein structures as input and is able to evaluate more than a hundred structures in a few seconds.

References

1. Sippl MJ: *Current Opinion in Structural Biology* 1995, **5:**229-235.
2. Sippl MJ: *Proteins-Structure Function and Genetics* 1993, **17:**355-362.
3. Grant BJ, Rodrigues APC, Elsawy KM, McCammon JA, Caves LSD: *Bioinformatics* 2006, **22:**2695-2696.