Alignment of Homologous Protein Structures in the Presence of Domain Motions

Mosca R (1), Schneider RT (1,2)

(1) FIRC Institute of Molecular Oncology Foundation, Via Adamello 16, 20139 Milan, Italy
(2) European Institute of Oncology, Via Ripamonti 435, 20141 Milan, Italy

Motivation
Structural alignment is an important step in protein comparison. Well-established methods for solving this problem in the case of rigid structures already exist. Methods for three-dimensional alignment in the presence of domain motions have only been developed in recent years (Ye and Godzik 2003; Shatsky, Nussinov et al. 2004).
Methods
Here, we present a new method for the flexible alignment of homologous protein. In this method the space of all possible alignments is represented as a graph on a set of potential matching fragments in the two molecules, scored on the base of their structural similarity. Every path in the graph represents a possible alignment while the optimal alignment is considered to be the longest path in the graph, i.e. the path yielding to the maximum score. A scoring strategy based on difference distance matrices is used. The alignment algorithm is coupled with a genetic algorithm for the identification of conformationally invariant parts (Schneider 2002) to define subsets of atoms suitable for group-wise least-squares superposition The alignment is characterized by the number of superimposed atoms and a measure, RMSDflex, that summarizes the RMSD for the superposition of several rigid regions.
Results
The alignment tool based on this method is able to align homologous structures with large hinge motions, such as two structures of the molecular chaperon GroEL from two different species. When compared to a method like DaliLite (Holm and Sander 1993) by aligning a set of homologous kinases structures with both methods, the new algorithm reaches good performance in terms of coverage and RMSDflex, while requiring only about one fifth of the CPU-time used by Dali.

Contact email: roberto.mosca@ifom-ieo-campus.it

References
- Holm, L. and C. Sander (1993). "Protein structure comparison by alignment of distance matrices." J Mol Biol 233(1): 123-38.
- Schneider, T. R. (2002). "A genetic algorithm for the identification of conformationally invariant regions in protein molecules." Acta Crystallogr D Biol Crystallogr 58(Pt 2): 195-208.
- Shatsky, M., R. Nussinov, et al. (2004). "FlexProt: alignment of flexible protein structures without a predefinition of hinge regions." J Comput Biol 11(1): 83-106.
- Ye, Y. and A. Godzik (2003). "Flexible structure alignment by chaining aligned fragment pairs allowing twists." Bioinformatics 19 Suppl 2: II246-II255.