GAIA: Generation of alternative alignments by an inverse approach

Tosatto SCE, Albiero A

Dept. of Biology and CRIBI Biotechnology Centre, University of Padova, Padova

Motivation

State-of-the-art methods for protein structure prediction based on the modelling of an unknown structure from a known template have been recently shown to achieve greater accuracy with the simultaneous modelling of ensembles of slightly different models, which have to be subsequently ranked using model quality assessment programs (MQAPs). One way to construct large ensembles of alternative models is from alternative alignments, generated either from varying alignment parameters or using different alignment methods. These approaches however do not guarantee a systematic and uniform coverage of conformational space and are therefore not guaranteed to find the most accurate solution, while spending much time exploring redundant solutions.

Methods

Here we propose a novel combinatorial optimization method for the systematic exploration of the alignment space defined by a limited set of initial alternatives. The method uses the initial alternatives to construct an inverse alignment matrix which can be traversed with regular dynamic programming to produce new solutions from suboptimal alignments. A divide and conquer step ensures the generation of truly different solutions ranked by relative frequency in all regions of alternative alignment.

Results

The inverse alignment approach was benchmarked on the comparative modelling and fold recognition homologous targets from the recent CASP-6 blind test and on a test set of 20 difficult alignments with little sequence identity. The results indicate that the inverse alignment approach is capable of ensembles of up to thousands of alternative alignments in minutes of computer time on a single desktop processor. The generated ensemble is enriched in accurate solutions and facilitates the MQAP selection of near-native models.

Contact email: silvio@cribi.unipd.it