

ORAL PRESENTATION

Open Access

Exploiting physical properties in protein structure alignment

Shweta Shah*, Nick Sahinidis

From Sixth International Society for Computational Biology (ISCB) Student Council Symposium
Boston, MA, USA. 9 July 2010

Introduction

The functional properties of a protein are strongly dependent on its structural conformation. The primary question addressed in this work is how to determine structural and therefore functional similarity from 3D protein structures. The approach we take relies on protein structure alignment, which elucidates functional protein relationships that are not depicted by the sequence.

Most current structural alignment tools are based on geometric properties of protein structures, and do not incorporate the effect of physical properties responsible for protein shape and similarity. The effect of physical features, such as hydrogen bonding, bond and torsion angles, hydrophathy, and solvent accessibility, are not well understood and hence not used while developing structure alignment tools. In this work, we exploit physical properties in protein structural alignment algorithms. In particular, we incorporate dominant physical features in one of the exact protein structure alignment tools, CMOS [1], which is based on the contact map overlap maximization (MAX-CMO) formulation of protein structure alignment. CMOS is a state-of-the-art tool, providing fast structural alignments that are in excellent agreement with the SCOP fold family classification. An implementation of CMOS is accessible at <http://eudoxus.cheme.cmu.edu/cmos/cmos.html>.

Computational methodology and results

The dominant physical features were incorporated in CMOS and their impact on algorithm speed and quality of results was studied. Our computational studies indicate that a search space reduction based on secondary

Table 1 Effect of physical properties on protein structure alignment with CMOS

Physical Property	% matches eliminated	% deviation from optimal	Average speed up in CMOS
Secondary Structures	24	0.8	5x
Torsion angles	28	14	N.A.
Hydrophathy	12	7	N.A.
Solvent accessibility	42	16	N.A.

structures leads to an over five-fold reduction in CPU requirements while maintaining solution quality [Table 1]. Other features, such as hydrophathy, torsion angles, and solvent accessibility, result in alignments with a large deviation from the optimal alignment.

Conclusions

The physical features responsible for protein structural similarities are not well understood. While hydrophathy, torsion angles, and solvent accessibility were found to be of little use in this context, our computational study suggests that hydrogen bonding properties can be meaningfully exploited to reduce the search space and reduce the computational requirements of structure alignment with CMOS.

Published: 7 December 2010

Reference

1. Xie W, Sahinidis NV: A reduction-based exact algorithm for the contact map overlap problem. *Journal of Computational Biology* 2007, **14**:637-654.

doi:10.1186/1471-2105-11-S10-O3

Cite this article as: Shah and Sahinidis: Exploiting physical properties in protein structure alignment. *BMC Bioinformatics* 2010 **11**(Suppl 10):O3.

* Correspondence: shweta@cmu.edu
Department of Chemical Engineering, Carnegie Mellon University, Pittsburgh, PA 15213, USA