A New Method for Matching Protein Secondary Structure Elements or C-alpha Atoms in Protein Structure Comparison

Edward S. C. Shih **Tsan-Fu Wang** **Ming-Jing Hwang**
mjhwang@ibms.sinica.edu.tw

Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan
128 Yen-chiou Yuan Rd., Sec. 2, Taipei 11529, Taiwan

Keywords: protein structure comparison, signal/noise amplification, pair-picking algorithm

1 Introduction

Recent years have seen a number of spectacular discoveries on surprisingly similar structures of proteins whose evolutionary kinship cannot be recognized based on sequence analysis alone [1]. These discoveries owe much to breakthroughs in mathematical algorithms by which automated structural comparisons are made possible [2,3]. However, while existing methods for comparing protein structures have been very satisfactory, further enhancement of their computational efficiency is still desired especially considering that the entries of Protein Data Bank are continuing to expand rapidly. The present work constitutes such an effort.

2 Method and Results

Existing methods of protein structure comparison usually work with pairwise relationships between two sets of structural entities such as secondary structure elements or simply C-alpha atoms. Key to the success of these methods is a heuristic algorithm, which differs from one method to another, that can efficiently identify an optimal pairing combination of all matchable structural entities by picking out 'true' pairings (signal) among numerous 'false' pairings (noise). It is this 'pair-picking' (i.e., finding equivalent structural entities) process that is most time consuming. In the present work, an intuitively derived simple formula having the ability to greatly amplify the signal while suppress the noise – thus making the 'picking' much more easier – has been developed. Our preliminary testing, results to be presented at the meeting, indicates the new method compares favorably with some of the widely used (e.g., DALI [4], VAST [5], and CE [6]).

References

[1] Gibrat, J.F., Madej, T. and Bryant, S.H., Surprising similarities in structure comparison, *Curr. Opin. Struct. Biol.*, 6(3):377–385, 1996.

[2] Holm, L. and Sander, C., Mapping the protein universe, *Science*, 273(5275):595–603, 1996.

[3] Brown, N. P., Orengo, C. A., and Taylor, W. R., A protein structure comparison methodology, *Computers Chem.*, 20:359–380, 1996.

[4] Holm, L. and Sander, C., Protein structure comparison by alignment of distance matrices, *J. Mol. Biol.*, 233(1):123–138, 1993.

[5] Madej, T., Gibrat, J.F. and Bryant, S.H., Threading a database of protein cores, *Proteins*, 23(3):356–369, 1995.

[6] Shindyalov, I.N. and Bourne, P.E., Protein structure alignment by incremental combinatorial extension (CE) of the optimal path, *Protein Eng.*, 11(9):739–747, 1998.