Evolutionary Approach to Protein Structure Prediction with Hydrophobic Interactions

T. W. de Lima, P. H. R. Gabriel, and A. C. B. Delbem
Institute of Mathematics and Computer Sciences
Sao Carlos, Sao Paulo, Brazil
telma@icmc.usp.br,
phrg@grad.icmc.usp.br,
acbd@icmc.usp.br

R. A. Faccioli, and I. N. da Silva
Sao Carlos Engineering School
Sao Carlos, Sao Paulo, Brazil
faccioli@sel.eesc.usp.br,
insilva@sel.eesc.usp.br

ABSTRACT

The Protein Structure Prediction (PSP) is to determine the protein tertiary structure from its amino acids. This paper presents the ProtPred and investigates its application. The first results showed that ProtPred is a consistent approach.

Categories and Subject Descriptors: I.2.1 Artificial Intelligence: Applications and Expert Systems

General Terms: Algorithms.

Keywords: Computational Biology, Bioinformatics, Protein Structure Prediction, Evolutionary Algorithm.

1. INTRODUCTION

The Protein Structure Prediction (PSP) problem aims to determine the protein tertiary structure from its sequence of amino acids. This paper presents an Evolutionary Algorithm (EA) to PSP problem using an *Ab initio* approach called ProtPred. We use a distinct fitness function to calculate hydrophobic interactions witch is, in general, not used by EAs with full-atom model.

2. PROTPRED

The proposed ProtPred is based on a classic EA. The algorithm starts initializing a random conformation. The torsion angles (ϕ, ψ, χ_i) are generated at random from the constrained regions. Afterwards, the energy of the conformation is evaluated. First, the protein's structure in internal coordinates (backbone and side-chain torsion angles) is transformed into Cartesian coordinates.

We proposed three kinds of crossover operators: one based on BLX-α operator; one using uniform crossover and another is a two-point crossover. Three kinds of mutation operators were proposed: the first acts only on angles of the peptide chain; the second and the third use an uniform mutation, modifying all the values of the backbone and side-chain torsion angles.

3. RESULTS

The ProtPred was tested using the population size 500 chromosomes and the maximum number of generations 100. The cost function have dielectric constant equal to 4.0.

GECCO'07, July 7–11, 2007, London, England, United Kingdom.
ACM 978-1-59593-697-4/07/0007.

We used four proteins from the Protein Data Bank (PDB): 1A11, 1ALE, 1DEP and 1DU1. Table 1 presents the lowest energy of each protein tested and their distance matrix error (DME). The ProtPred found adequate PSP for the tested proteins.

Table 1: Minimum fitness value and DME, with dielectric constant equal to 4.0, the maximum distance between atom i and atom j to compute van der Waals interactions 8Å, for electrostatic interaction 13Å and for hydrophobic interactions 8.5Å

Protein (PDB Id)	Minimum Energy (C)	DME(Å)
$1A11$	-262.9420	7.15
$1ALE$	-227.3685	4.54
$1DEP$	-237.1484	2.48
$1DU1$	-314.9491	3.77

4. CONCLUSIONS

Purposed of ProtPred is to develop a pure Ab initio algorithm that does not use any heuristics in the prediction process, in contrast with Ab initio approaches like Rosetta [1]. Thus, the proposed approach can explore any region of the search space to find an adequate structure conformation even though there is no similar protein structure previously known.

Another innovation of ProtPred is the use only of van der Waals and electrostatic energies and the addition of hydrophobic energy to a full-atom model. The other approaches to PSP with full-atom model do not use hydrophobic energy to evaluate their solutions and uses another potential energy.

5. ACKNOWLEDGMENTS

FAPESP and CNPq (Brazilian research foundations) to the financial support.

6. REFERENCES

[1] R. Bonneau, J. Tsai, I. Ruczinski, D. Chivian, C. Rohl, C. Strauss, and D. Baker. Rosetta in CASP4: progress in Ab initio protein structure prediction. *Proteins*, 5:119–126, 2001.