Mutation Operator and its Effects on Protein Structure Prediction in Genetic Algorithms

Trent Higgs[a], Bela Stantic[a], Md Tamjidul Hoque[b] and Abdul Sattar[a]

[a] *Institute for Integrated and Intelligent Systems Griffith University, Queensland, Australia*
[b] *Discovery Biology, Eskitis Institute for Cell & Molecular Therapies Griffith University, Queensland, Australia*

Abstract and Objective

We have used Genetic Algorithms (GAs) for the Protein Structure Prediction (PSP) problem, which utilise two main search operators: crossover, and mutation. However, in this paper, we have focused particularly on identifying optimal mutation that can be used to avoid local minima effectively. This is done by altering the chromosomes randomly within the identified range of rotational degree of freedom to provide better diversity to the search algorithms. We have identified the optimal range of rotational moves by empirically analyzing the instantiates of the conformational effects using the mutation operator applied on couple of proteins from the PDB.

Introduction

GAs have proven to be quite useful in the PSP field [1, 2]. They consist of two main search operators: crossover and mutation. Mutation can be used to escape local minima, it achieves this by randomly altering a select amount of chromosomes in the population to introduce diversity.

In this research we have looked at the mutation operator in detail and analysed how it effected the free energy calculations, and the overall structure similarity with the native conformation. This was to identify the optimal range of rotational moves to be applied to the PSP problem when using a GA approach.

Methodology and Results

To test the mutation operator in GAs for PSP we have used a single point pivot rotation move [1]. This is done by performing random mutations on a set of predicted protein structures for a specific target protein. For our experiments we have used a number of proteins, each of which consist of exactly 200 predicted structures generated by Rosetta using *ab initio* protocols. Rotation was performed on the *x*, *y*, and *z* axis separately. Due to space limitations, we will only show the results for protein 2ptl using *y* rotation, as it obtained results very similar to every other protein and axis we examined. For energy calculations we applied Rosetta's fitness function, and for structural similarity we used root mean square deviation (RMSD). Our results can be seen in Figure 1.

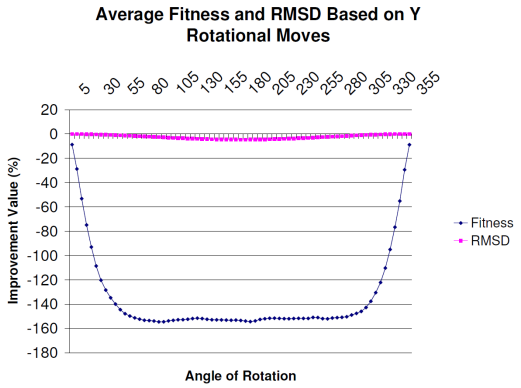

Figure 1-Average improvements for 2ptl y rotational moves.

Discussion and Conclusion

In this research we analysed the effects the mutation operator had on the PSP process in GAs. From our experiments (see above) two main points can be concluded: (1) The rotational moves that provided us with the best results were those that had smaller displacement or magnitude (i.e. 5º to 40º) compared to any further larger magnitude (see Figure 1), and (2) The fitness value of a protein appears to more effected by rotational moves, when compared to RMSD (notice in Figure 1 the fluctuations of the fitness values compared to the RMSD values).

References

[1] Unger, R. & Moult, J. Genetic algorithms for 3d protein folding simulations, Journal of Molecular Biology, 1993, 231, 75-81.

[2] Hoque, T., Chetty, M. & Dooley, L. A guided genetic algorithm for protein folding prediction using 3D hydrophobic-hydrophilic model, in IEEE Congress on Evolutionary Computation, 2006, pp. 8103-8110.

Address for correspondence

Emails: {T.Higgs, B.Stantic, T.Hoque, A.Sattar}@griffith.edu.au