

2 個体分散遺伝的アルゴリズムによるタンパク質の構造予測  
岩橋 崇史

1 目標課題

2 個体分散遺伝的アルゴリズム ( Dual Individual Distributed Genetic Algorithm : Dual DGA )<sup>1)</sup> によるタンパク質の構造予測を行う。

2 研究の進捗状況

2.1 Dual DGA

Dual DGA<sup>1)</sup> は, DGA において 1 島内の個体数を 2 とし, その個体数に適した遺伝的オペレータを適用したものである。本手法は DGA と比較してアルゴリズムが単純化しており, かつ解探索能力に優れている。

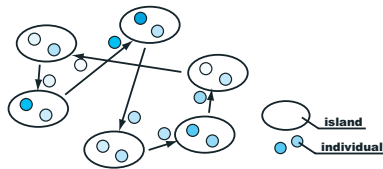


Fig. 1 Dual DGA

2.2 Dual DGA によるタンパク質の構造予測

構造予測を行うタンパク質は Met-enkephalin を用いた。存在する 19 個の二面角が設計変数に相当する。二面角の値から求まるエネルギー値が -11kcal/mol 以下のとき最適解とみなす。実験で使用したパラメータを Table 1 に示す。移住間隔を 2, 3, 4 の 3 つのパターンを用意し, 実験を行った。試行回数は 30 回である。

Table 1 パラメータ

個体数/島	2	設計変数	19
島数	3200	ビット数/設計変数	9
移住個体数/島	1	交叉率	1.0
移住間隔	2,3,4	評価計算回数	1900000

Fig. 2 に実行結果を示す。Fig. 2 において, 30 回試行における median ( 中央値 ), worst ( 最悪値 ), best ( 最良値 ) の 3 つの値は全て終了世代の値である。すべての移住間隔において, 中央値, 最悪値に差はないことが分かった。最悪値については, 移住間隔 3 のとき最も低い値を示していることが分かった。

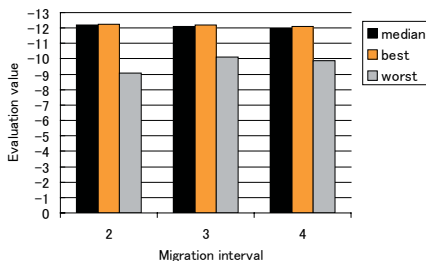


Fig. 2 実行結果

Fig. 3 に 30 回試行における終了世代の最適解および -10kcal/mol 以下の解の発見率を示す。移住間隔 3 のとき, 最適解発見率が 93% であり, 他の移住間隔の場合の最適解発見率よりも高いことが分かった。そして, -10 以下の解の発見率についても同様のことが言える。これより, 移住間隔 3 のとき, 他の移住間隔と比べて有効な解探索を行っていることが分かった。

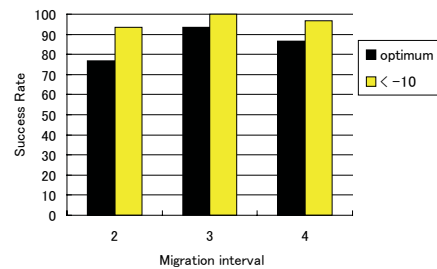


Fig. 3 最適解, -10 以下の解の発見率

先月の研究により, DGA による Met-enkephalin の構造予測において, 十分な個体数が必要であることが確認された。Dual DGA についても同様であり, 今回の実験では, 個体数を 6400 とした。そして, 移住間隔の大きさをいくつか用意し, 実験したところ, 移住間隔の大きさによって, 最適解発見率に違いが生じることが分かった。よって, Dual DGA を用いた Met-enkephalin の構造予測では, 十分な個体数とともに終了世代に対しての移住間隔の大きさ, すなわち移住回数が重要であることが分かった。

3 今後の課題

EVOLVE/G システム上で Dual DGA を実装し, タンパク質の構造予測を行う。

参考文献

1) Tomoyuki HIROYASU, Mitsunori MIKI, Masahiro HAMASAKI, Yusuke TANIMURA. A New Model of Distributed Genetic Algorithm for Cluster Systems: Dual Individual DGA. Proceedings of the International Conference on Parallel and Distributed Processing Techniques and Applications, Vol.1 (2000), pp.477-483