

PDGA によるタンパク質の構造予測
岩橋 崇史

1 目標課題

PDGA によるタンパク質の構造予測を行い、パラメータが解探索に与える影響を調査する。

2 研究の進捗状況

2.1 PDGA 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。

実験で共通するパラメータを Table 1 に示す。

Table 1 パラメータ

設計変数	19	交叉率	1.0
ビット数/設計変数	9	突然変異率	0.01
評価計算回数	1900000	エリート保存	1

Table 2 で示すように、島サイズ、総個体数、移住間隔の 3 つのパラメータを変更し、構造予測を行った。

Table 2 変更したパラメータ

島サイズ	4			10			20		
島数	50	100	200	20	40	80	10	20	40
移住個体	1			1			2		
移住間隔	10	50	100	10	50	100	10	50	100

Fig. 1, Fig. 2 に 30 回試行における終了世代の最適解および -10kcal/mol 以下の解の発見回数を示す。Fig. 1 は、島サイズ 4 の場合を示す。Fig. 2 は、総個体数 800 の場合を示す。

Fig. 1 より、移住間隔 10, 50 のとき、島数が多くなるほど最適解および -10kcal/mol 以下となる解の発見率が高くなることが分かった。

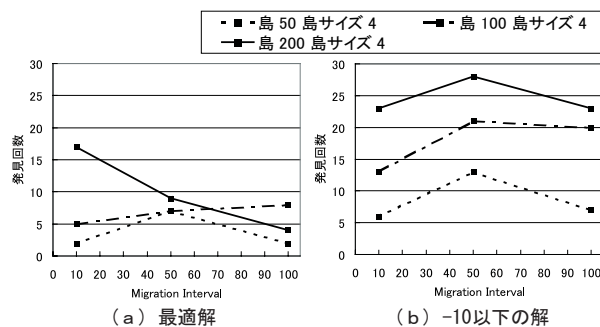


Fig. 1 島サイズ 4

Fig. 2 より、移住間隔 10, 50 のとき、島数が多くなるほど最適解および -10 以下の解の発見率が高くなるこ

とが分かった。移住間隔 100 のとき、島数に関係なく最適解発見率が下がる。しかし、島数が多くなるほど -10kcal/mol 以下の解の発見率が高くなる。これより、島サイズの大きさよりも、島数が解探索に大きく影響すると考えられる。

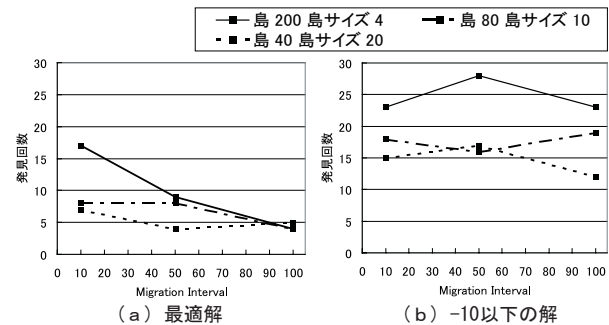


Fig. 2 総個体数 800

今回の実験より、島サイズの大きさ、移住間隔に関係なく、島数が多いほど多数の局所的探索を行い、有効な解探索を行うことが分かった。また、移住間隔を長くすることによって、大域的な探索が可能となる。しかし、移住間隔 100 のとき、終了世代で最適値に到達せず、結果的に有効な解探索を行えないと考えられる。よって、移住間隔は終了世代までの移住回数とのバランスが必要であると考えられる。

以上より、有効な Met-enkephalin の構造予測を行うためには、多くの島数が必要であることが分かった。よって、Met-enkephalin の構造予測に、2 個体 GA¹⁾ が有効であると考えられる。

3 今後の課題

タンパク質グループとしては、2 個体 GA¹⁾ によるタンパク質の構造予測を行う。

また、Grid グループとしては、oasis に Globus のインストールを行う。

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