

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)¹⁾ によるタンパク質の構造予測を行う。また、理工研執筆のために、GA を用いたタンパク質の構造予測について文献の調査を行う。

2 研究の進捗状況

2.1 Dual DGA によるタンパク質の構造予測

前回の報告で、Met-enkephalin を対象問題としたとき、Dual DGA は有効な解探索を行うことを確認した。今回は Met-enkephalin よりも複雑な構造を持つタンパク質、(Ala)₁₀ を対象問題とし、Dual DGA による構造予測を行った。(Ala)₁₀ において、存在する 30 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 -9.7kcal/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 1, 2, 3, 4, 5, 6 の 6 つのパターンを用意し、実験を行った。試行回数は 12 回である。

Table 1 パラメータ

| | | | |
|---------|-------------|-----------|----------|
| 個体数/島 | 2 | 設計変数 | 30 |
| 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 | 1,2,3,4,5,6 | 評価計算回数 | 30008000 |

Fig. 1 に実行結果を示す。Fig. 1 において、12 試行での median (中央値), average (平均値), worst (最悪値), best (最良値) の 4 つの値は全て終了世代の値である。移住間隔 2, 3 において、4 つの値がそれぞれほぼ等しいことが分かった。また、移住間隔 4, 5 においても同様のことを確認することができた。

中央値、平均値、最悪値において、移住間隔 1 のとき他の移住間隔と比較して、最も悪い値を示していることが分かった。

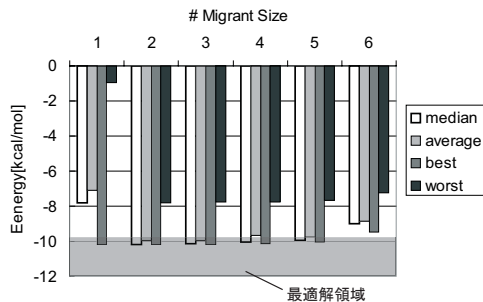


Fig. 1 実行結果

Fig. 2 に 12 回試行における終了世代の最適解発見率を示す。移住間隔 2, 3, 5 のとき、最適解発見率が 91% で

あり、移住間隔 4 のとき、最適解発見率が 83% であった。よって、これらの移住間隔における解探索は、最適解発見率が高く、有効であることが分かる。しかし、移住間隔 1, 6 における解探索は、最適解発見率が低く、有効でないことが分かった。

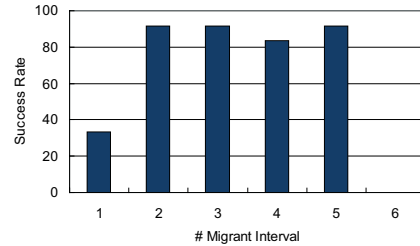


Fig. 2 最適解発見率

Dual DGA による探索で得た最適解の立体構造を Fig. 3 に示す。Fig. 3 より、ヘリックスが確認でき、良好な立体構造をしていることが分かった。

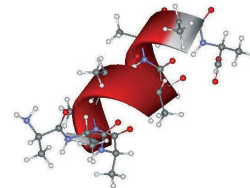


Fig. 3 (Ala)₁₀ の立体構造: エネルギー値 10.2[kcal/mol]

(Ala)₁₀ を対象問題とした時、パラメータチューニングを行うことにより、Dual DGA が有効な解探索を行うことを確認することができた。よって、今後はより複雑なタンパク質を対象問題し、Dual DGA で解くことを課題とする。

3 今後の課題

Dual DGA を用いて、より複雑なタンパク質を解く。また、理工研執筆を行う。

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