

タンパク質立体構造予測への並列 SA の適用
吉田 武史

1 研究活動報告

今月行った研究活動を示す .

- チュニジアにて IEEE SMC 2002 の参加
- 並列 SA(PSA) をタンパク質に有効性を検証
- 遺伝的アルゴリズムを用いた適応的温度調節を行う並列 SA(PSA/AT(GA)) のタンパク質への適用

2 IEEE SMC 2002 への参加

10 月 6 ~ 9 日にチュニジアで開催された IEEE System and Man Cybernetics 2002 に参加した . 本学会はを目的としており , 情報系以外の多くの研究者が参加する大規模な学会であった . その中で私は「 Parallel SA with Adaptive temperature determined by GA 」の発表を行った . 国際学会への参加が 2 度目であったため , 発表準備の要領もわかり , 満足の行く発表をおこなえた . なお , 我々がチュニジアで撮影した写真を <http://mikilab.doshisha.ac.jp/~undry/tomo/smc2002/> に掲載する .

3 タンパク質への並列 SA の適用

3.1 実験概要

本論文では , タンパク質立体構造予測に並列 SA を適用し , 有効性を検証する . 本実験では , 解探索中に解交換を行わない並列 SA と同期を取りその時の最良の解を全プロセスに配布する並列 SA を用いた . 対象問題として Met-enkephalin と (Ala)10 を用い , Table 1 に用いたパラメータを示す .

Table 1 実験で用いたパラメータ

プロセス数	8
Mcsweep	6000
最高温度	2
最低温度	0.1
冷却率	0.9995
近傍	180 - 54
解交換周期	なし, 20, 100, 500
試行回数	30

3.2 実験結果

Met-enkephalin における結果を Table 2 に , (Ala)10 における結果を Table 3 に示す . またエネルギー推移の

平均値を Fig. 1 に示す . Fig. 1 では横軸に MCsweep 数 , 縦軸にエネルギーを示す .

Table 2 Met-enkephalin の実験結果

解交換周期	なし	20	100	500
最良値	-12.01	-11.72	-12.13	-12.08
平均値	-10.97	-7.46	-8.95	-10.52
中央値	-11.33	-7.12	-8.95	-10.86
最悪値	-9.11	-4.18	-3.51	-8.10
最適解発見率	57%	3%	10%	37%

Table 3 (Ala)10 の実験結果

解交換周期	なし	20	100	500
最良値	-9.60	-0.51	-9.98	-9.86
平均値	-8.35	5.51	-1.93	-7.38
中央値	-9.41	5.40	-0.56	-9.61
最悪値	-1.81	11.81	6.39	1.55
最適解発見率	0%	0%	20%	26%

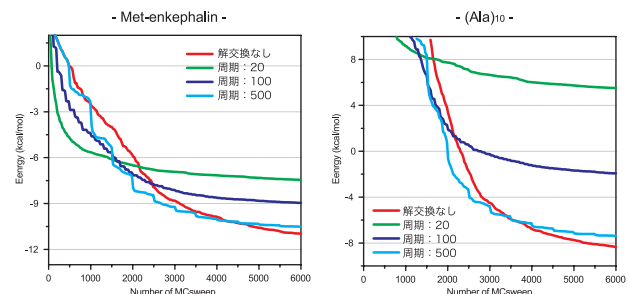


Fig. 1 エネルギー推移

この実験結果より , 両タンパク質において解交換を行わない並列 SA (独立型並列 SA) の解探索能力が高いことがわかる . また , 解交換を行う並列 SA では , 解交換周期が短い場合に解探索初期で局所解に収束しているのに対して , 解交換周期が長い場合には局所解に陥らず解探索を行っていることがわかる .

以上の結果より , 独立型並列 SA がタンパク質に適していると考えられる . しかし , 本実験では 8 プロセスしか使用しておらず , 初期解に依存が強いタンパク質では , より多いプロセス数で実験する必要がある .

4 今後の課題

- PSA/AT(GA) のタンパク質への適用
- PSA/AT(GA) の論文投稿
- タンパク質における PSA の有効性を検証