

SA によるタンパク質立体構造予測  
米田真純

1 前月からの課題

1. MCsweep の検証
2. 一定温度 SA の解探索能力の検証

2 課題の達成状況および研究成果

2.1 MCsweep の検証

SA をタンパク質立体構造予測に適用する際、我々の研究グループでは各設計変数に対して先頭から順に生成処理・受理判定を行う Monte Carlo Sweep(MCsweep)を用いている。しかし、これまでに MCsweep を用いない SA の性能検証、および生成処理・受理判定を行う順番については検証を行っていない。そこで、対象問題を Met-enkephalin とし、SA をタンパク質立体構造予測に適用する際の生成処理・受理判定について検証を行った。

本実験では以下の 3 パターンについて比較を行った。

- 生成処理・受理判定を設計変数毎に先頭から順に行う MCsweep(Sequential)
- 生成処理・受理判定を行う順番をランダムにした MCsweep(Random)
- すべての設計変数を生成した後、受理判定を行う (Non-MCsweep)

評価計算回数は岡本らの実験と同じく 1900000 回に設定し、試行回数は 50 回とした。実験結果を Table 1 および Fig. 1 に示す。Fig. 1 はエネルギー推移を示しており、縦軸はエネルギー、横軸は評価計算回数である。

Table 1 実験結果

	Sequential	Random	Non-MCsweep
最良値	-12.10	-12.08	2.49
平均値	-10.40	-10.48	7.87
最悪値	-7.23	-7.60	16.12

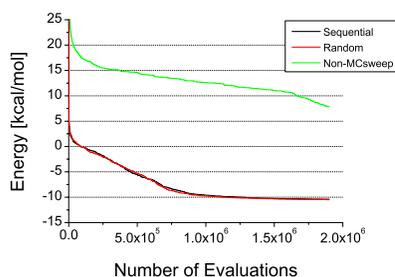


Fig. 1 エネルギー推移

Table 1 および Fig. 1 より、SA を用いたタンパク質立体構造予測における MCsweep の有効性が確認された。また、MCsweep において各設計変数の生成処理・受理判定を行う順番は解探索能力に大きな影響を与えなかった。

2.2 一定温度 SA による解探索能力

これまでの研究で SA におけるタンパク質立体構造予測には重要温度領域が存在することを確認した。しかし、この重要温度領域の検証実験では、逐次 SA に用いていた総探索数よりも少ない値を用いていた。そこで、本実験では逐次 SA と一定温度 SA の解探索能力を比較するために、一定温度 SA の総探索数をこれまでの逐次 SA と同じ 100000MCsweep に設定し、Met-enkephalin の立体構造予測に適用した。なお、一定温度 SA の温度は最高温度 10.0 と最低温度 0.01 の間を等比的に 32 分割した値を割り当てた。また、試行回数は 30 回とした。

Fig. 2 に一定温度 SA を Met-enkephalin の立体構造予測に適用した際に得られた最小エネルギーの平均値を示す。縦軸はエネルギーで横軸は温度である。

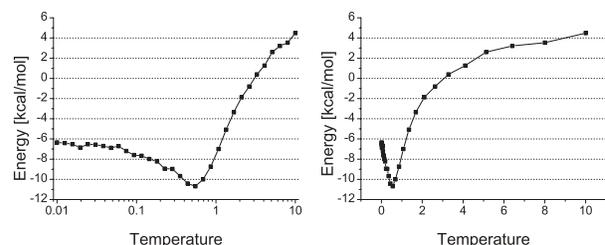


Fig. 2 Met-enkephalin における重要温度領域

Fig. 2 より、これまでの結果と同様に温度 0.3~0.7 の領域が良好な結果を得ていることを確認した。そこで、この温度領域での一定温度 SA と従来の温度スケジュールを持つ逐次 SA の性能比較を行った。

Table 2 に比較結果を示す。なお、逐次 SA は 50 回試行、一定温度 SA は 30 回試行の結果である。

Table 2 実験結果

温度	逐次 SA	一定温度 SA			
		2.0-0.1	0.69	0.55	0.44
最良値	-12.10	-10.95	-11.27	-11.58	11.63
平均値	-10.40	-10.00	-10.68	-10.41	-9.68
最悪値	-7.23	-9.08	-9.50	-8.22	-7.03
最適解発見率	30 %	0 %	30 %	50 %	30 %

Table 2 より、重要温度領域内の一定温度 SA が逐次 SA よりも高い解探索能力を持っていることがわかる。また、温度 0.44 の一定温度 SA は、温度パラメータチューニングを行った逐次 SA よりも高い解探索能力を得た。この結果より、Met-enkephalin を対象問題とした場合には重要温度領域内の一定温度 SA が最も高い性能を示していることが確認された。