

逐次 SA の性能評価
青井 桂子

1 今月の課題

- 逐次 SA の性能評価
- Grid 環境の PSA/GAc の単位時間当たりの計算量

2 逐次 SA の性能評価

これまでタンパク質の立体構造予測においては SA がよく使用されてきた。岡本らは、小規模なタンパク質 (Met-enkephalin) を対象として立体構造予測における SA の有効性を確かめている。

逐次 SA による小規模なタンパク質 (Met-enkephalin, (Ala)₁₀) において、立体構造予測を行い、岡本らと同一のパラメータを用いることで、性能の比較を行う。

Table 1 に、本実験で用いたパラメータを示す。MC-sweep 数、最高温度、最低温度を岡本らが用いた値とし、近傍に関しては、岡本らの実験で経験的に良いとされてきた、 $180^\circ \rightarrow 54^\circ$ のものと、 180° 一定の場合の 2 パターンに関して実験を行った。

Table 1 パラメータ

target	Met-enkephalin	(Ala) ₁₀
amino acids	5	10
atoms	75	103
deheadral angles	19	30
total MCSweep	100000	11000
trial		50
max temperature		2.0
min temperature		0.1
cooling rate	0.999970042	0.999973
neighborhood range	180° constant, 180° → 54°	

Table 2 に 50 試行の最小エネルギーの平均値、最小値、最大値、中央値、また最適解発見率を示す。このとき、ECEPP/2 エネルギー関数に基づいた気相中において、Met-enkephalin は $E \leq -11kcal/mol$ の領域で、(Ala)₁₀ は $E \leq -9.7kcal/mol$ の領域で、最小エネルギー構造をとるため、この領域のエネルギーが得られた場合に最適解が発見できたものとする。

実験結果より、Met-enkephalin と (Ala)₁₀ の立体構造予測において、岡本らの実験で経験的に良いとされてきた $180^\circ \rightarrow 54^\circ$ のものよりも、 180° 一定の場合の方が解探索能力が優れていた。

Table 2 逐次 SA の性能

	Met-enkephalin		(Ala) ₁₀	
	180to54	180const	180to54	180const
average	-10.68	-10.88	-8.00	-9.46
minimum	-12.02	-12.08	-9.76	-9.83
maximum	-7.94	-7.53	2.23	-4.60
median	-10.76	-10.91	-9.52	-9.56
success rate	18/50	24/50	4/50	4/50

3 Grid 環境の PSA/GAc の計算量

個体数を増やした場合の、Grid 環境の PSA/GAc と通常の PSA/GAc の単位時間あたりの評価計算回数の比較を行う。

Grid 環境の PSA/GAc では、Grid RPC システムの一つである NetSolve を用いる。また、ユーザは 1 台の PC しか持たないものとし、NetSolve Server としては、CambriaSystem の 100 ノードを利用可能な資源とする。

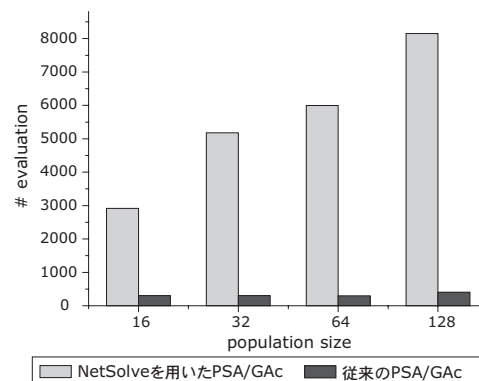


Fig. 1 単位時間 (1hour) あたりの評価計算回数の比較

Fig. 1 に単位時間あたりの評価計算回数の比較を示す。実験結果より、Grid 環境の PSA/GAc では個体数が増えると逐次 SA の並列数が増えるため、外部 Server で利用する計算資源が増える。ユーザが PC 1 台しか持たない場合でも、16 個体で 10 倍近く、128 個体の時には 20 倍もの計算資源を手に入れられた。

4 今後の課題

- NetSolve システムのオーバーヘッドの検討
- 理工研の執筆