

PSA/GAc における交叉間隔と MCsweep 数の数値実験
青井 桂子

1 今月の課題

- PSA/GAc における 1 個体あたりの MCsweep 数と交叉間隔の検討
- 上記の実験の実行時間の測定

2 MCsweep 数と交叉間隔の検討と実行時間

2.1 実験概要

これまでの実験より, タンパク質立体構造予測問題において PSA/GAc は SA よりも探索性能が良いことが明らかである. PSA/GAc は一定間隔で交叉処理を行い良い個体の形質を伝達するためと考えられる. このため今回は特に交叉オペレータにおける交叉間隔と 1 個体あたりの MCsweep 数の検討を行う. 実験では PSA/GAc の総評価計算回数を凡そ一定とし, 1 個体あたりの MCsweep 数の検討に関しては 16 個体 × 6000MCsweep, 32 個体 × 3000MCsweep, 64 個体 × 1500MCsweep の 3 パターンを検討する. また, 交叉処理における交叉間隔をそれぞれ 1, 2, 4, 8, 16, 32, 64, 交叉処理なし (PSA) の 8 パターンについて実験を行い, この時の解探索性能を比較する. 実験で用いた対象タンパク質は Met-enkephalin, C-peptide, PTH(1-34) である.

2.2 交叉間隔と MCsweep 数の検討

Fig. 1 に得られたエネルギーの平均値を示す.

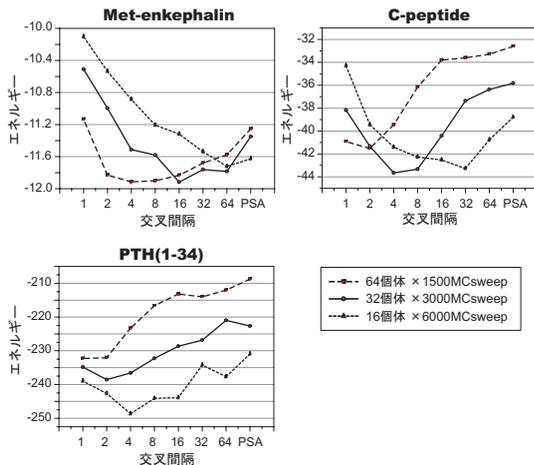


Fig. 1 エネルギーの平均値

実験結果より, 各タンパク質で適正な探索ステップや 1 個体あたりの MCsweep 数によって適正な交叉間隔は異なることが示された. またより原子量の大きなタンパク質に対しては, PSA/GAc の交叉間隔を頻繁に行うほど良い結果が得られた.

2.3 実行時間

交叉間隔と MCsweep 数の実験ではパラメータによって評価計算回数や実行時間が異なる. 本実験で用いた PSA/GAc は Fig. 2 に示すように並列可能部分と逐次部分に分けられる. 交叉部分は逐次処理になるため交叉間隔が短い場合には並列化効率が悪くなると考えられる.

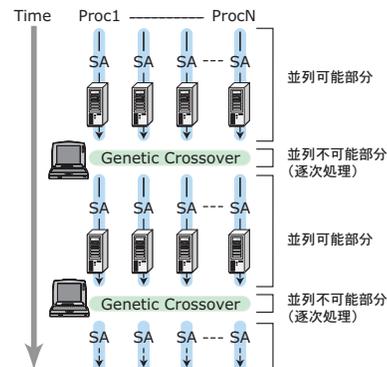


Fig. 2 PSA/GAc における並列可能部分と逐次部分

Fig. 3 に各パラメータでの実行時間を示す.

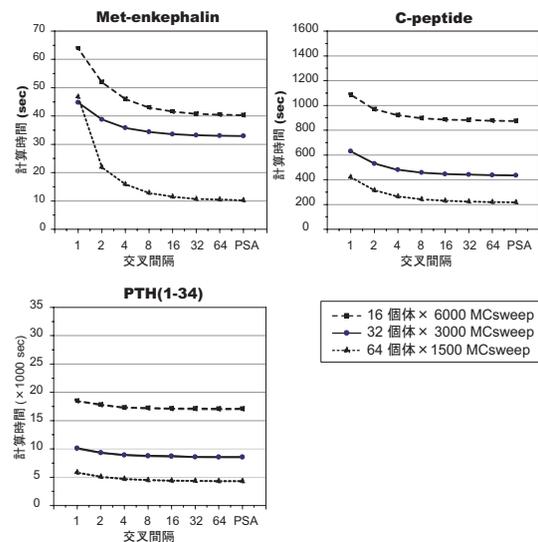


Fig. 3 実行時間

実験結果より交叉間隔が短い場合, Met-enkephalin では PSA の倍以上の時間を要していることから並列化効率が悪い. しかし, C-peptide や PTH(1-34) では交叉間隔が短い場合でも, 並列化効率はほとんど低下しない.

3 今後の研究課題

NetSolve を用いた PSA/GAc の非同期モデルの実装