

タンパク質立体構造モデリングプログラム実行速度の高速化

研究代表者 岩館 満雄 研究員

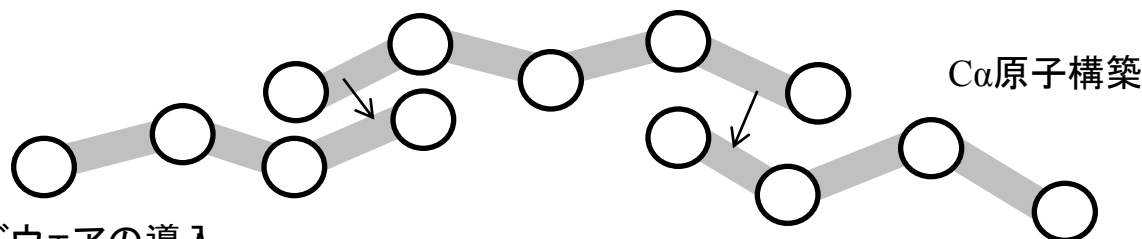
研究の目的

今年度の研究では高速化は電子部品ハードと近年公表されつつあるデータベースサーチに関する新アルゴリズムを効率的に取り入れることによる組み合わせによるアプローチを試みる。

データベース探索アルゴリズム

代表者の開発するタンパク質高次構造予測プログラムFAMS (Full Automatic Modeling System) は、データベースを高速に探索することを基軸においたプログラムであり、三次元座標の誤差の最小二乗を生み出す回転行列を求めるため、特異値分解の計算を実行することによって大量に且つ高速に解く必要がある。近年、この計算量を大幅に削減することを可能とするアルゴリズムが開発され (J Comput Biol. 2010 Mar;17(3):203-19. Searching protein 3-D structures in linear time. Shibuya T.) 理論的には確立し、実用的な有用性が求められている。

具体的には、FAMSは大きく4段階から成り立っているプログラムであり。①タンパク質のアミノ酸のC α 原子の構築、②タンパク質主鎖の構築、③構造保存部位の構築、④全原子の構築、の4段階である。特に①ではタンパク質のC α データベースから、②では主鎖のデータベースから誤差の最小二乗のデータを検出する必要があり、RMSDの閾値以下のタンパク質データベースを探索する方法の部分に導入することによって探索時間の縮小の効果が期待でき、適用を試験中である。



ハードウェアの導入

電子部品の活用の方角を模索することは継続するとともに、この有用な可能性の高いアルゴリズムの適用を試み、ソフト、ハードの両面からの高速化に取り組む予定である。

アルゴリズムの最適化によってFAMSの高速化に加えて、従来パソコンで処理していた論理を専門化した電子回路への導入を試みる。

ArduinoなどのC言語論理を導入することによってどの程度の高速化が望めるのかを検討しており、そのほかGPUなどのより効果の期待できるものへ移植することも検討中。