

遺伝的交叉を用いた並列シミュレーテッドアニーリングによる タンパク質立体構造予測

Protein Tertiary Structure Predictions by Parallel Simulated Annealing using Genetic Crossover

小椋 真貴 (日本電気) 正 三木 光範 (同志社大工)
正 廣安 知之 (同志社大工) 岡本 祐幸 (岡崎国立共同研究機構)
青井 桂子 (同志社大院) 小椋 信弥 (同志社大院)

Maki OGURA, NEC Corporation
Mitsunori MIKI, Doshisha University, Tatara Miyakodani 1-3, Kyo-Tanabe, Kyoto
Tomoyuki HIROYASU, Doshisha University, tomo@is.doshisha.ac.jp
Yuko OKAMOTO, Institute for Molecular Science
Keiko AOI, Graduate School of Engineering, Doshisha University
Shinya OGURA, Graduate School of Engineering, Doshisha University

Key word: Simulated Annealing, Genetic Algorithms, Crossover, Protein Tertiary Structure Predictions

1 はじめに

自然に存在するタンパク質の立体構造は、そのタンパク質が持つエネルギーの最小状態に対応している。そのため、エネルギーを最小とするような立体構造を最適化手法を用いて求めることで立体構造予測が可能であると考えられる。これまでタンパク質立体構造予測においては、シミュレーテッドアニーリング (Simulated Annealing : SA) がよく使用されてきた。しかしタンパク質のエネルギー関数は複雑であり、従来の SA では容易にタンパク質立体構造予測を行うことができない。そのため、立体構造予測のための新たな手法の開発が必要とされている。

そこで本研究では、これまでに我々が提案している遺伝的交叉を用いた並列シミュレーテッドアニーリング (Parallel Simulated Annealing using Genetic Crossover : PSA/GAc) によってタンパク質立体構造予測を行い、従来の立体構造予測に用いられている SA よりも有効なアルゴリズムであることを検証する。

PSA/GAc は、並列 SA と遺伝的アルゴリズム (Genetic Algorithm : GA) のハイブリッドアルゴリズムである。タンパク質の持つエネルギー関数の特徴から、局所的な探索が得意かつ計算の高速化を図ることができる並列 SA に、大域的な探索が得意かつ部分解の組み合わせで最適解が得られる問題に有効な GA のオペレータを取り入れた本アルゴリズムは、タンパク質立体構造予測にさらに有効であると考えられる。

2 遺伝的交叉を用いた 並列シミュレーテッドアニーリング

遺伝的交叉を用いた並列シミュレーテッドアニーリング (PSA/GAc) は Fig. 1 に示すように、複数の SA を並列に実行する。さらに、一定期間ごとの SA の解の伝達時に、GA のオペレータである遺伝的交叉を用いたものである。遺伝的交叉の処理では、もとの親と生成した子との 4 個体間のうち評価値の高い 2 個体を選択して、選択された 2 個体から次の探索を続ける。ある設計変数の最適値が求まっている場合、遺伝的交叉によってその設計変数の最適値を他の SA 探索に伝達することができるため、アニーリングの収束を早めることができる。これらの処理を繰り返し、終了条件を満たすまで探索を続ける。

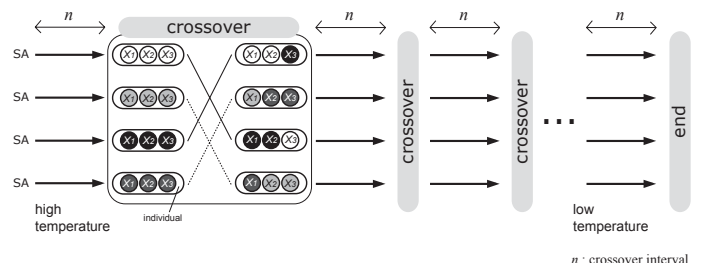


Fig. 1: Model of PSA/GAc

PSA/GAc は、GA のオペレータを用いた SA であるため、SA の探索点を個体 (Individual)、探索点の総数 (SA の並列数) を個体数 (Population size) と呼ぶこととする。本手法をいくつかのテスト関数に適用したところ、高い解探索能力を示すことを確認した。

3 タンパク質立体構造予測

本研究では、最小エネルギー値既知のタンパク質である Met-enkephalin と、最小エネルギー値が未知のタンパク質であるヒト副甲状腺ホルモン (PTH) のフラグメントを立体構造予測の対象とした。

3.1 Met-enkephalin の立体構造予測

Met-enkephalin は、5 個のアミノ酸残基からなるものである。ECEPP/2 エネルギー関数に基づいた気相中において、 $E \leq -11kcal/mol$ の領域で最小エネルギー構造をしている。本実験では、Met-enkephalin の主鎖における 10 個の二面角と、側鎖における 9 個の二面角をそれぞれ設計変数とした。二面角のとり得る値は $[-180^\circ, 180^\circ]$ の範囲で表現した。つまりこのタンパク質は 19 個の設計変数を持っており、1MCsweep によって 19 回の Metropolis 判定が課されるものとした。

岡本らは SA を用いて Met-enkephalin の立体構造予測を行い、10 試行中 5 回の確率で $E \leq -11kcal/mol$ 以下の最小エネルギー構造を得た。PSA/GAc による Met-enkephalin の立体構造予測では、10 試行中 9 回の確率で最小エネルギー構造が得られた。この結果から、最適構造を得る確率は SA を用いる場合よりも PSA/GAc を用いた場合の方が高いことが明らかとなった。

3.2 PTH(1-34) の立体構造予測

PTH のフラグメント (1-34) は、NMR(核磁気共鳴) 実験によって 2 つの α -helix の存在が示唆されている。エネルギー関数 ECEPP/2 に基づいた気相中において行われた岡本らの SA を用いた実験では、得られた最小エネルギー構造に 2 つの α -helix が出現し、そのときのエネルギー値は $-210kcal/mol$ であることが示されている。PTH(1-34) は主鎖に 68 個、側鎖に 110 個の二面角を持ち、これらを設計変数としたため、1MCsweep によって 178 回の Metropolis 判定が課される。

本実験によって得られた最も低いエネルギー値は $-240.0kcal/mol$ であった。構造は Fig. 2 のようになり、2 つの α -helix と 1 つの β -sheet を持っている。

岡本らの実験で得られた最小エネルギー構造や、NMR 実験によっても 2 つの α -helix が確認されており、PSA/GAc による立体構造予測でも 2 つの α -helix が出現した。このため、PTH(1-34) は 2 つの α -helix を持つと考えられる。また、岡本らの実験で得られた構造には存在しなかったが、PSA/GAc によって得られた構造では β -sheet が出現した。 α -helix や β -sheet は安定構造であるが、誤った部分に出現

するとエネルギー値が高くなる。PSA/GAc によって得られた構造のエネルギー値は、岡本らの結果よりも低い値であるため、PTH(1-34) の立体構造は 2 つの α -helix と 1 つの β -sheet を持つ可能性が高いと考えられる。この結果から、PTH(1-34) の立体構造予測において PSA/GAc は SA より解探索能力が高いことが明らかとなった。

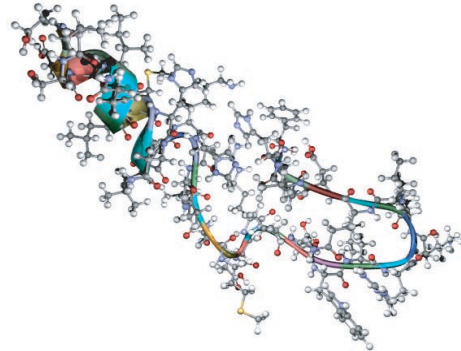


Fig. 2: Lowest-energy conformation of PTH(1-34) gained by PSA/GAc

4 結論

本研究では、遺伝的交叉を用いた並列シミュレーテッドアニーリング (PSA/GAc) によってタンパク質立体構造予測を行った。

5 個のアミノ酸残基からなる Met-enkephalin を対象に立体構造予測を行った数値実験では、従来用いられていた SA よりも PSA/GAc の解探索能力が高いことが明らかとなった。またアミノ酸残基 34 個のヒト副甲状腺ホルモンについて立体構造予測を行った場合には、従来用いられていた SA よりもよりエネルギー値の低い立体構造を予測することに成功した。

タンパク質のエネルギー関数は連続空間に存在し、局所的に無数の極小値、大域的にもいつかの極小値を持つ。並列 SA は局所的な探索が得意かつ計算の高速化を図ることができ、また GA は大域的な探索が得意かつ部分解の組み合わせで最適解が得られる問題に有効である。そのため、並列 SA に GA オペレータを取り入れた本アルゴリズムは、SA よりもタンパク質立体構造予測において有効であったと考えられる。

これらの結果から、PSA/GAc はタンパク質立体構造予測に有効な手法であることが示された。

謝辞

本研究は文部省からの補助を受けた同志社大学の学術フロンティア研究プロジェクト「知能情報科学とその応用」における研究の一環として行った。

出典：

日本機械学会 第 15 回計算力学講演会 講演論文集,
No.02-02,
pp.53-54,
(2002 年 11 月)

問い合わせ先：

同志社大学工学部/ 同志社大学大学院工学研究科
知的システムデザイン研究室
(<http://mikilab.doshisha.ac.jp>)