

マルチカノニカル分子動力学法を使ったモデル Protein-G 系の研究¹

科学技術振興事業団 磯部 雅晴²

1 はじめに

タンパク質の構造や折り畳みの機構（秩序化過程：会合）は、未解決の問題で興味深い。タンパク質の3次元構造を構築するために必要なすべての情報はそのアミノ酸配列にだけ含まれており、アミノ酸間の相互作用が与えられれば、配列に対応した唯一のネイティブ構造へ構造変化する。しかし、計算機を使った研究では、与えられたアミノ酸配列からタンパク質の構造を予測することは取り得るポリペプチド鎖のコンフォメーションが、残基数に伴い急激に増大するので困難を伴う。

タンパク質の計算機を使った研究は、格子モデルで代表される極端なモデル化や、水分子を直に考え全原子間の相互作用を考慮した連続空間の現実的計算など行われてきたが、このような極端な状況下では実際にどの相互作用が折り畳みに本質的に寄与しているのかがわかりにくくなっていると思われる。そこでタンパク質を連続空間上で単純なモデル化を施し、モノマー間の各々の相互作用の相対的強さを変えることによってタンパク質の折り畳みの変化を調べ、そこからなんらかの普遍的な特徴を取り出して、包括的な理論的枠組みを構築するための基礎データを蓄積することが現段階では必要である。

そこで本研究では、連続空間上でアミノ酸残基を電荷、疎水、中性の球状モノマーに置き換えたモデル Protein-G 系 (PDB id:2gb1, 56 残基) の折り畳み過程と構造に着目しカノニカル (CMD)、マルチカノニカル分子動力学法 (MMD) を使って系統的にシミュレーションを行った。特に、溶媒の存在により生じるモノマー間の疎水相互作用がタンパク質の折り畳み過程に大きな影響を与えと考えられているため、電荷相互作用と疎水相互作用の相対的大きさによってモデルタンパク質の折り畳みの挙動がどのように変化するかに着目した。この電荷相互作用と疎水相互作用の相対的強さを変えることは、溶媒の種類を変えることに対応する。

通常の CMD においては、このような単純なモデル化により自由度の大幅な削減を施しても、平衡状態における物理量の確率分布を正確に求めるのは容易ではない。というのは実際に緩和時間がどの程度が見積もるのは難しいばかりでなく、タンパク質は準安定構造が複数存在しているので、シミュレーションが現実的に不可能なマイクロ秒のオーダーを超えてしまうからである。ところが、MMD の出現によりこのような困難は大幅に改善された。MMD においては予備的計算の後、ただ1回の RUN で広範囲な温度におけるエネルギー分布がかなりの精度で計算できる。本報告書では、実際にモデル計算に適用した場合の CMD と MMD の違いと有用性について述べる。

¹ 本研究は、科学技術振興事業団計算科学技術活用型特定研究開発事業の下で行われている。

² E-mail: isobe@cphys.s.kanazawa-u.ac.jp

2 モデル

2.1 モデルの選択（メリットとデメリット）

まず、タンパク質のモデル化について、簡単にまとめる。

1. 格子モノマーモデル（自由度小, 拘束条件大）

定性的、現象論的な議論が出来る反面、得られる物理量が限定されており、現実のタンパク質との定量的な比較は困難である。自由度が小さく、エネルギー準位が整数で表現できるため厳密な議論が出来、理論を構築しやすいと思われる。また、シミュレーションは比較的容易であるが、格子上という拘束条件が強すぎ排除体積効果の影響力が非常に大きいために、基底状態探索には特殊なアルゴリズムが必要である。

2. 連続空間上でのモノマーモデル（自由度中, 拘束条件中）

大規模計算が比較的容易で、折り畳み過程と長距離相互作用の役割との対応関係が単純にわかると期待される。適度に拘束条件を緩めて、適度に自由度を減らしたモデル。特に MD で計算した場合、並列処理が可能であり、大規模タンパク質のモデル計算に適していると思われる。

3. 溶媒、全原子を考慮した原子レベルの計算（自由度大, 拘束条件小）

すべての物理量が得られ、実験との定量的比較が可能であるが、水の自由度が大きすぎシミュレーションをする際、計算コストが膨大となる。また、原子の力場をすべて考慮するので、Native 構造は原理的に再現可能であるが、複雑すぎてポイントをしばった研究がしにくくなる。また水はタンパク質に対し、複数の役割を担っているので、水のポテンシャル関数をうまく工夫しなくてはならないと思われる。

2.2 連続空間のモノマーモデル

これまで我々は、電荷のみを持った連続空間におけるモデルポリマー系のシミュレーションを実行し、電荷相互作用が折り畳みに与える影響を研究してきた。

まず、 $N=60$ の di-block Polyampholyte が研究された。これは、正電荷と負電荷の数は同じで、全モノマーの数を N とすると、 $1 \sim N/2$ 番目のモノマーは正電荷を持ち、 $N/2 + 1 \sim N$ 番目のモノマーは負電荷を持ったものである。MMD による自由エネルギー計算の解析により低エネルギー状態において、2次元ではエネルギー障壁が見つかり、3次元系では螺旋構造が見つかった [1]。

次に、正電荷、負電荷のみを現実のタンパク質 (Protein-g) の配列に導入したシミュレーションが行われた。これは、真空中のタンパク質に対応する。電荷モノマーが大きな影響力を持つことがわかり、電荷を持ったモノマー同士がコアを作り固体のように振舞い、中性モノマーがその周りを取り囲み液体のように振舞うことがわかった [2]。

3 溶媒中のタンパク質のモデル化

研究の次のステップとして、溶媒によりタンパク質はどのように影響をされるのかを調べることは興味深い。溶媒（水）の存在により、タンパク質は真空中と違い、相互作用に主に以下のような変化が生じると考えられる。

1. 水による遮蔽効果により、電荷モノマー同士のみかけの相互作用の強さが小さくなる。
2. 疎水モノマーの水を嫌う性質より疎水モノマーが集まり疎水コアを作る傾向にある。これは、疎水モノマー同士があたかも引力相互作用を及ぼしているかのように見える。

そこで本研究では、溶媒による影響を明らかにするため、以下の4つの2体相互作用を導入することとした。(i) 排除体積効果による斥力相互作用（ソフトコア）、(ii) 共有結合による隣接モノマー間相互作用（バネ）、(iii) 電荷モノマー間の静電相互作用（比誘電率: ϵ_c^* ）、(iv) 疎水モノマー間の有効引力相互作用（疎水相互作用の大きさ: ϵ_{hb} [kcal/mol]）。ここで疎水相互作用が、どのような有効ポテンシャル関数で表記するのが適当か（もしくはそもそも単純なポテンシャル関数で表記できるのか否かも含めて）議論の余地を残すが、今回は比較的単距離な引力（べき関数型）[3]と長距離引力（指数関数型）[4]の2つでシミュレーションを行い両者を比較することにした。シミュレーションに際し、電荷相互作用と疎水相互作用の相対的強さを制御パラメーターとし（ $\epsilon_c^*, \epsilon_{hb}$ ）で示すことにした。この制御パラメーターを変えることは、溶媒の種類を変えることとみなすことが出来、表1のように真空中と水の2つの極限系が設定できる。シミュレーションはこの2つの極限系の間で行った。

	真空中	→	水
電荷相互作用	大 ($\epsilon_c^* \sim 1$)		小 ($\epsilon_c^* \sim 80.0$)
疎水相互作用	小 ($\epsilon_{hb} \sim 0.0$)		大 ($\epsilon_{hb} \sim 3.0$ [kcal/mol])

表 1: 真空中及び水中での制御パラメーターの変化

4 マルチカノニカル分子動力学シミュレーション

マルチカノニカル分子動力学シミュレーション (MMD)[5] の CMD との唯一の違いは、内部エネルギーの確率密度分布関数があるエネルギー領域で等確率で出現させるように、各ステップにおいて力の項にスケーリングファクターをかけることである。

$$\mathcal{F}_i(t) = -\nabla_i \mathcal{E}(E) \quad (1)$$

$$= -\frac{d\mathcal{E}(E)}{dE} \nabla_i E \quad (2)$$

$$= -(1 + k_B T_0 \frac{d}{dE} \log P_{T_0}(E)) \nabla_i E. \quad (3)$$

ここで、 $\mathcal{E}(E)$ は未知なので、予備的な計算により学習させる。

$$\mathcal{E}_{new}(E) = \mathcal{E}_{old}(E) + k_B T_0 \log P_{T_0}(E) \quad (4)$$

5 オーダーパラメーター

解析に際しては、以下の2つのオーダーパラメーターを用いた。一つは慣性半径 (RG) で、これはポリマーの空間的広がりを示す量である。これにより、ポリマーが伸びた状態かコンパクトな状態かを判定できる。

$$RG = \sqrt{\frac{1}{N} \sum_{i=1}^N \mathbf{r}_i^2 - \left(\frac{1}{N} \sum_{i=1}^N \mathbf{r}_i \right)^2} \quad (5)$$

もう一つは、Distance Matrix Error(DME) と呼ばれるもので、これは構造の違いを定量的に表すことができるパラメーターである。(6) 式の $\mathbf{r}_{ij}^C$ は参照系を意味し、ここでは NMR 実験から得られた Protein-G 系の Native 構造を選んだ。これにより、慣性半径が同じ程度のコンパクトな構造同士を比較する際に、どちらがより Native 構造に近いかを定量的に判別することができる。

$$DME = \sqrt{\frac{2}{N(N-1)} \sum_{\langle i,j \rangle} (\mathbf{r}_{ij} - \mathbf{r}_{ij}^C)^2} \quad (6)$$

6 結果とまとめ

様々な制御パラメーターによる構造の変化の傾向を見るため、最初に急冷 NVT アンサンブルによる CMD を行った。CMD においてエネルギー最低構造の解析をしたところ、電荷相互作用が疎水相互作用より相対的に大きいとき（すなわち、真空中に近い時）、エネルギー最低構造は C 末端が伸びた構造となり慣性半径が大きな状態をとることがわかった。ところが、疎水相互作用が電荷相互作用に比べて大ききとき（水に近いとき）は、疎水モノマー同士がコアを形成してコンパクトな構造となることがわかった。

この結果を踏まえて、次に MMD 法を行った。MMD 法により位相空間上に存在するエネルギー極小準安定状態に長時間とどまることを回避することが出来、限られた計算ステップで広範囲な位相空間の確率密度分布を正しく見積もることが可能となる。図 1,2 は、疎水相互作用をべき型で入れ、制御パラメーターを $(\epsilon_c^*, \epsilon_{hb}) = (5.0, 3.0)$ とした、温度 300[K] における慣性半径 (RG) と Distance Matrix Error(DME) の 2 次元空間における確率密度分布である（図 1: カノニカル、図 2: マルチカノニカル；サンプル数は 1 億ステップ）。マルチカノニカル RUN では、カノニカル RUN よりも広い範囲をサンプルできていることがわかる。図 3 は、MMD で得られたデータを Reweighting 法を使って求めた慣性半径の温度による変化の様子である。300K 付近では安定状態 $RG \sim 8.0, \sim 10.0[\text{\AA}]$ が 2 つ存在しており、安定構造の間には障壁が存在している。このことから、この温度領域で 1 次的な相転移をしていると考えられる。

一方、疎水相互作用を指数関数型で入れた場合はこのような転移は見られず、慣性半径は単調に増加するだけであった。

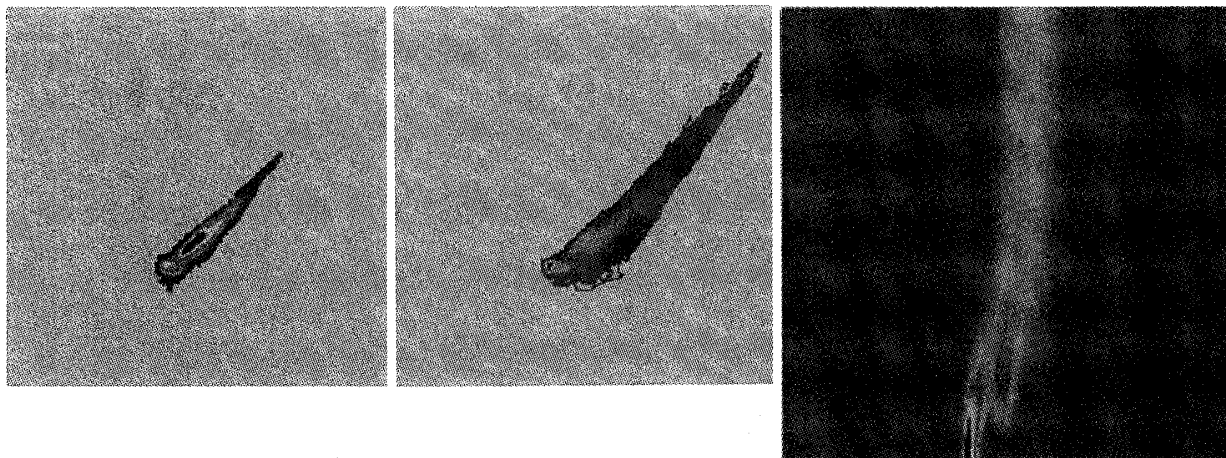


図 1: カノニカル RUN で得られた、RG-DME 空間における確率密度分布関数。横軸: RG おける確率密度分布関数。横軸: RG 0 ~ 20[Å]、縦軸: DME 0 ~ 20[Å]。図 2: マルチカノニカル RUN で得られた、RG-DME 空間における確率密度分布関数。横軸: RG 0 ~ 20[Å]、縦軸: DME 0 ~ 20[Å]。図 3: 慣性半径の分布関数の温度変化。横軸: RG 0 ~ 20[Å]、縦軸: 温度 200 ~ 600[K]

このように MMD で Reweighting 法を使って各温度における慣性半径の確率密度分布を調べた結果、疎水相互作用のポテンシャル関数のタイプにより転移の仕方が大幅に変化することが見出され、転移の傾向は溶媒の種類 ($\epsilon_c^*, \epsilon_{hb}$) 並びに関数の形にも強く影響されることがわかった。

参考文献

- [1] H. Shimizu, K. Uehara, K. Yamamoto, and Y. Hiwatari, *Mol. Sim.* **22** (1999), 285.
- [2] A. Baumketner, H. Shimizu, and Y. Hiwatari, submitted to *Eur. Phys. J. E*.
- [3] J. -E. Shea, Y. D. Nochomovitz, Z. Guo, and C. L. Brooks III, *J. Chem. Phys.* **109** (1998) 2895.
- [4] J. Israelachvili and R. Pashley, *Nature* **300** (1982) 341.
- [5] N. Nakajima, H. Nakamura, and A. Kidera, *J. Phys. Chem.* **B101** (1997), 817.