

令和 5 年 6 月 9 日現在

機関番号：13301

研究種目：基盤研究(C)（一般）

研究期間：2020～2022

課題番号：20K03879

研究課題名（和文）拡張アンサンブル法と液体論を結合したタンパク質の第一原理三次構造予測法の開発

研究課題名（英文）Ab initio structure prediction of proteins by combining the extended ensemble method with liquid state theory

研究代表者

三浦 伸一（Miura, Shinichi）

金沢大学・数物科学系・教授

研究者番号：10282865

交付決定額（研究期間全体）：（直接経費） 3,300,000円

研究成果の概要（和文）：本研究では拡張アンサンブル法を用いてタンパク質分子の三次構造を決定する手法の構築を試みた。まずは水溶液環境を連続体モデルにより記述した水和タンパク質のマルチカノニカル計算を実施し、対象としたタンパク質の熱力学的フォールディング転移点を決定した。この転移温度は実験から推定されているものと良い一致を示した。マルチカノニカル計算より得られた軌跡に沿ってエネルギー極小化することによりエネルギー地形の様相を明らかにし、転移温度以下では実験から知られている天然構造が支配的であること示した。また水溶液環境を3D-RISM理論により高精度化する手法の検討を行った。

研究成果の学術的意義や社会的意義

タンパク質が機能を発現する天然状態は自由エネルギー最小の状態であり、折りたたみは物理的な原理に支配されていると考えられている。これはアミノ酸残基配列情報である一次構造から機能を発現する三次構造を物理的な方法論に基づき予測できる可能性を示しており、物質科学・生命科学の最重要課題のひとつに挙げられる。本研究では水溶液環境はごく粗いモデルが用いられているが、実験で知られている天然構造を熱力学的な相転移という観点より計算のみから特定することができた。また、本研究で行われた詳細な検討により液体の積分方程式理論を用いて水溶液環境を高精度化するための方法論的な方向性を示すことができた。

研究成果の概要（英文）：In this study, we have tried to construct methods for determining tertiary structures of proteins using extended ensemble methods of molecular simulations. We have first performed multicanonical calculations of a hydrated protein using a continuum model of aqueous solutions and the associated thermodynamic folding transition was determined. The resulting transition temperature was found to be in good agreement with the experimentally estimated temperature. The energy landscape of the protein was investigated by successive energy minimizations of protein structures along the multicanonical trajectory. The experimentally known native structure was demonstrated to be dominant below the transition temperature. We have also examined methods to refine the description of aqueous solutions by the 3D-RISM integral equation theory.

研究分野：液体物性理論

キーワード：拡張アンサンブル法 液体論 タンパク質 フォールディング

## 1. 研究開始当初の背景

タンパク質は、アミノ酸がペプチド重合することにより形成される鎖状高分子である。この高分子は生理条件下で自発的に折りたたまり、その機能を発現する。Anfinsenらの実験[1]により、機能を発現する天然状態は自由エネルギー最小の状態であり、折りたたみは物理的な原理に支配されていることが明らかになった。これはアミノ酸残基配列情報である一次構造から機能を発現する三次構造を物理的な方法論に基づき予測できる可能性を示しており、物質科学・生命科学の最重要課題のひとつに挙げられる。

自然に存在するタンパク質は20種類のアミノ酸からのみなり、進化による淘汰の末に選ばれてきた精巧な分子機械である。その折りたたみに必要な時間はミリ秒程度であり、単純なランダムサーチと比較すると驚くほど速い。一般にアミノ酸がペプチド重合して形成される高分子は必ずしも十分速く自由エネルギー最小構造に折りたたまるわけではないために、タンパク質分子の自由エネルギー地形にその大きな特徴があると考えられている。この問題に対しては整合性原理[2]やファネル仮説[3]など重要な概念的なブレークスルーがあったが、未だ原子レベルで自由エネルギー地形を探索し、その天然構造を予測する方法は確立しておらず、タンパク質分子の全原子モデルをベースに天然状態の三次元立体構造を第一原理から予測する手法の開発が待たれていた。

## 2. 研究の目的

本課題では、我々のグループで開発した拡張アンサンブル生成手法をタンパク質の全原子モデルに適用し、その天然状態での三次構造を予測する手法の開発を行う。この拡張アンサンブル生成手法は、これまでに我々のグループでモデルタンパク質の折りたたみ問題に成功裏に適用されている[4,5]。このモデルのエネルギー最小状態は図1(a)に示した $\alpha$ -バレル構造であり、このモデルの“天然構造”に対応する。我々のグループで開発した手法は、このエ

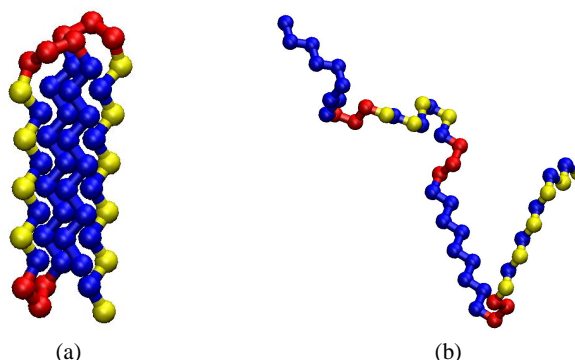


図1 タンパク質の(a)粗視化モデルの天然構造と(b)ランダムコイル構造。

ネルギー最小構造から図1(b)に示したランダムコイル構造などをカバーする広範囲な構造サンプリングを実現し、広い温度領域にわたるタンパク質の安定性を精度良く記述した。水は融解温度以下に冷却すると相転移を経て氷となるように、タンパク質は折りたたみ転移温度以下ではその天然状態に転移する。我々の方法により折りたたみ転移温度以下の天然構造は正しく同定されているが、その際にエネルギー最安定構造に関する情報は事前に用いていないことは強調したい。

タンパク質分子を全原子モデルで取り扱う場合、その周りの水溶液環境を適切に取り扱うことが重要である。水を露わに考慮しての折りたたみシミュレーションは、小さいタンパク質に対しては専用計算機を長期間占有し分子動力学計算をミリ秒程度に到るまで実行することにより可能であることが示されているが[6]、スーパーコンピュータを含む通常の計算機環境では未だ現実的ではない。また安定性に関わる自由エネルギーの評価にはほど遠い状況である。このような状況に鑑み、本課題では液体の積分方程式理論の一つである3D-RISM理論を用いて溶媒効果を取り扱い、現実的な計算時間で上述の目的を実現する手法を検討する。3D-RISM理論は溶質分子周りの溶媒の3次元密度場を記述する積分方程式理論であり、これまで生体高分子系を含む様々な溶液内分子過程に適用され、その有用性は示されている[7]。我々のグループでも、シンプルなジペプチドを用いたベンチマークを行い、水を露わに考慮する計算手法の結果と比較することにより、その信頼性を確認した[8]。本研究では、申請者のグループで開発した拡張アンサンブルの生成手法を3D-RISM理論と結合し、その天然状態での構造を分子間相互作用から決定する第一原理手法について検討する。

### 3. 研究の方法

我々は、これまで拡張アンサンブル生成手法としてレプリカ交換一般化ハイブリッドモンテカルロ (GHMC) 法およびマルチカノニカル一般化ハイブリッドモンテカルロ (GHMC) 法の開発を行い、モデル蛋白質の折りたたみ転移に対して成功裏に適用した[4,5]。本研究課題では、このシミュレーション手法をベースにタンパク質の全原子モデルを用いて折りたたみ問題にアプローチする手法を構築する。まずは水溶液環境を連続体モデルにより記述する手法によりフォールディング転移を取り扱う。

- 1) Generalized Born Surface Area (GBSA) 法などの連続体モデルに基づき溶媒和自由エネルギーをモデル化する。タンパク質の分子内相互作用にこの溶媒和自由エネルギーを加えたものを“ポテンシャル”とし、このシステムに対してレプリカ交換 GHMC 計算を行う。これにより広い温度領域での効率的な構造サンプリングが可能となる。各温度レプリカの計算結果から Weighted Histogram Analysis Method (WHAM) 法などを用いて状態密度を計算し、マルチカノニカル計算に必要な重みを求める。
- 2) ステップ 1) で求めた重みを用いてマルチカノニカル GHMC 計算を実行する。マルチカノニカル法は系の分布関数がポテンシャルエネルギーに対して一定となるような人工的な統計アンサンブルを導入することにより構成される。この分布関数のために系はポテンシャルエネルギー空間をランダムウォークし、低エネルギーから高エネルギーの構造までを広くサンプリングすることが可能になる。このステップでは分子内相互作用と連続体モデルにより溶媒和が記述された系のマルチカノニカル分布が生成されている。

次に、水溶液環境の記述の高精度化手法を検討する。本研究課題では液体の積分方程式理論のひとつである 3D-RISM 理論をとりあげ[7]、以下の方法により取り扱う。

上述のステップ 2) で生成された各々の構造に対して、液体の積分方程式理論を用いて特定の温度での溶媒和自由エネルギーを求める。連続体モデルから生じるバイアスを取り除くために、積分方程式から求めた溶媒和自由エネルギーを用いて Reweighting を実行し、この温度での 3D-RISM 計算と整合するカノニカル平均を求める。ここでステップ 2) においてタンパク質の空間的にコンパクトな状態からランダムコイルの状態までをカバーする広範囲にわたるサンプリングが実現されていることが前提となっている。対象とする温度を変えて積分方程式の計算を実行することにより、広い温度領域でのタンパク質の安定性を求めることができ、フォールディング転移温度以下での天然状態を同定することができる。

### 4. 研究成果

タンパク質のシミュレーションを始めるにあたって、まずは一般化ハイブリッドモンテカルロ (GHMC) 法の計算効率の検討を行った[9]。GBSA 法により水溶液環境が記述された水和アラニンジペプチドを例にとり、ベンチマークを行った。GHMC 法では試行配位は運動方程式を用いて生成される。多重時間刻み (r-RESPA) 法と hydrogen mass repartitioning (HMR) 法を併用することにより効率を損なうことなく運動方程式を数値的に解く際の時間刻みを 4 倍程度長くとれることを示した。r-RESPA 法および HMR 法の利用なし (GHMC-1 法)、利用あり (GHMC-3 法) の場合で、それぞれ比較的長い GHMC 計算を行った。ここで GHMC ステップ数は同一に設定されており、それぞれの計算時間は実質的に同じであることに注意しておこう。ジペプチドの二つの二面角の関数として計算した密度分布を図 2 に示す。GHMC-3 法では GHMC-1 法ではサンプルされていない領域にも分布がみられ、格段の計算の効率化を確認することができる。

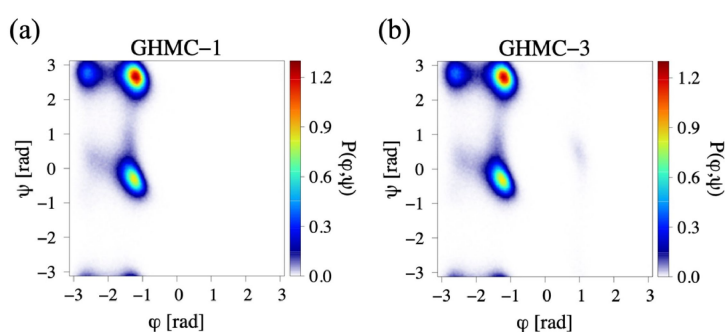


図 2 二つの二面角に関する分布。(a) GHMC-1 法、(b) GHMC-3 法による計算結果。文献[9]より転載。



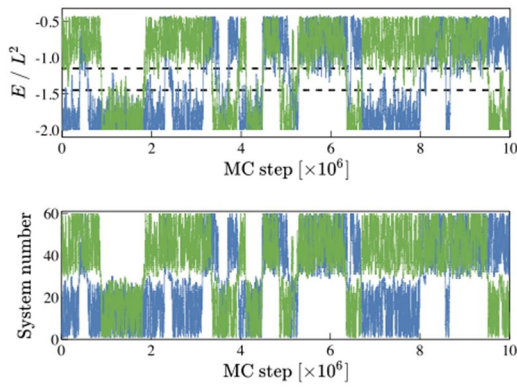


図3 ガウシアン-カノニカルハイブリッドレプリカ交換法による特定のレプリカのエネルギーをモンテカルロ軌跡に沿って図示(上図). 青は低温の秩序相から, 緑は高温の無秩序相から出発したレプリカのエネルギー. 下図は同様にレプリカが訪問したシステム番号. 文献[10]より転載.

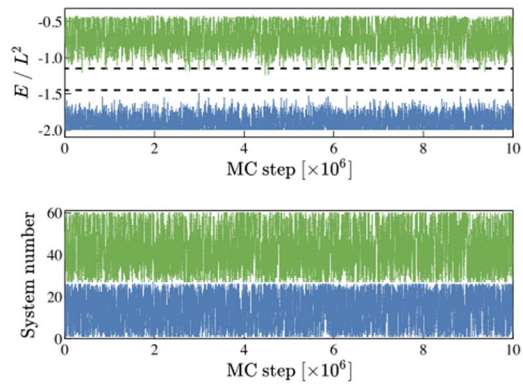


図4 カノニカルレプリカ交換法による特定のレプリカのエネルギーをモンテカルロ軌跡に沿って図示(上図). 青は低温の秩序相から, 緑は高温の無秩序相から出発したレプリカのエネルギー. 下図は同様にレプリカが訪問したシステム番号. 文献[10]より転載.

一方、後述のマルチカノニカル計算を実施する際にマルチカノニカル重み因子の精度のよい決定が鍵であることを痛感した。タンパク質のフォールディング転移は一次相転移様の現象であると考えられているために、一次相転移を効率的に取り扱う手法の構築が望まれる。カノニカルアンサンブルでは一次相転移温度では、エネルギー分布に秩序相と無秩序相に対応するピークがあり、その間に潜熱に由来する大きなギャップが開いている。このギャップ領域にアクセスする手法の開発が肝要である。標準的なカノニカルアンサンブルに基づくレプリカ交換法はこのギャップ領域を挟んでのレプリカの交換が進まず、一次相転移点付近を十分サンプルすることができない。本研究ではガウシアンアンサンブルと呼ばれる人工的な統計集合を用いて、一次相転移現象に対しても有効に働くレプリカ交換法の開発を行った[10]。ガウシアンアンサンブルは熱浴のサイズを制御することにより、ミクロカノニカル分布とカノニカル分布を補間するような統計集合を実現することができる。我々の新規手法は、相転移点付近はガウシアンアンサンブルで取り扱い、それ以外の状態はカノニカルアンサンブルで取り扱うハイブリッドレプリカ交換法である。二次元強磁性ポッツ模型をとりあげてベンチマークを実施した。図3にハイブリッドレプリカ交換法の結果を示す。この手法では系は図の破線で挟まれたギャップ領域をまたいで広くサンプリングされていることが見てとれる。比較のため、標準的なカノニカルレプリカ交換法の結果も図4に示す。この手法では低温から出発したレプリカは計算のあいだずっと低温相に局在しており、高温相から出発したレプリカもギャップ領域を超えられないことが確認できる。

これらの結果をもとに3節で説明した方法で小さなタンパク質分子のシミュレーションを実施した。対象はシニョリンと呼ばれるタンパク質である。まずはGBSA法により水溶液環境を記述したシニョリンのフォールディング転移温度の評価とそのエネルギー地形の探索を行った。計算より得られたフォールディング転移温度は実験から推定されているものと良い一致を示し、転移温度以下の構造は天然構造が支配的であることを示した。またいくつかの典型的な構造に対して3D-RISM計算を行い、溶媒和モデルにより安定性の順序等が変わることはないことを確認した。この結果は近くまとめて公表する予定である[11]。この計算で得られた構造をもとに3D-RISM法を用いたReweightingを行った。この手法はアラニンジペプチドのようなシンプルなものに対して有効に働くことは確認していたが、シニョリンの場合は残念ながら十分な精度が得られないことがわかった。これはマルチカノニカル計算から得られたシニョリンの構造に適用すると、特定の構造の重みが大きくなりすぎ統計的に意味のある結果が簡単には得られないことによる。この問題を克服するために上述の一般化ハイブリッドモンテカルロ(GHMC)法を用いて、3D-RISM法の溶媒和自由エネルギーと整合する分布を直接生成する手法の開発を行っている。この手法では水溶液環境の記述を二段階に分けて行う。ひとつはGBSA法による粗い記述、もうひとつは3D-RISM法による高精度な記述である。GHMC法では運動方程式を用いて試行配位が生成されるが、その部分はGBSA法による粗い記述に基づき行い、生成された試行配位の採択は3D-RISM法による高精度な記述により行う。ここでGBSA法によるバイアスは採択を決定する段階で取り除かれるために、この手法は3D-RISM法による溶媒和自由エネルギーと整合するカノニカル分布を生成することに注意しておく。シンプルなジペプチドに対するベンチマークは2022年日本物理学会秋季大会で報告したが[12]、今後はタンパク質のフォールディング問題に

適用できるようアルゴリズムの改良を進める予定である。

<引用文献>

- [1] C. B. Anfinsen, *Science* **181**, 223 (1973).
- [2] N. Go, *Annu. Rev. Biophys. Bioeng.* **12**, 183 (1983).
- [3] J. N. Onuchic, Z. Luthey-Schulten and P. G. Wolynes, *Annu. Rev. Phys. Chem.* **48**, 545 (1997).
- [4] N. Mukuta and S. Miura, *J. Chem. Phys.* **149**, 072322 (2018).
- [5] N. Mukuta and S. Miura, *Biophys. Physicobiol.* **17**, 14 (2020).
- [6] K. Lindorff-Larsen, S. Piana, R. O. Dror and D. E. Shaw, *Science* **334**, 517 (2011).
- [7] F. Hirata (ed.), *Molecular Theory of Solvation* (Kluwer Academic Publishers, Dordrecht, 2003).
- [8] H. Iwasaki, S. Yamaguchi and S. Miura, *Mol. Simul.* **43**, 13 (2017).
- [9] D. Suzuki, T. Hori, and S. Miura *J. Phys.: Conf. Ser.* **2207**, 012018 (2022).
- [10] D. Suzuki, D. Suzuki, and S. Miura, *J. Phys. Soc. Jpn.* **91**, 044006 (2022).
- [11] Y. Yamada, D. Suzuki and S. Miura (in preparation).
- [12] 石井完樹, 三浦伸一, “多重レベル一般化ハイブリッドモンテカルロ法の開発と生体分子への適用”, 日本物理学会 2022 年秋季大会.

5. 主な発表論文等

〔雑誌論文〕 計2件（うち査読付論文 2件/うち国際共著 0件/うちオープンアクセス 1件）

1. 著者名 Suzuki D, Hori T, Miura S	4. 巻 2207
2. 論文標題 A simple method to accelerate configurational sampling for a generalized hybrid Monte Carlo method	5. 発行年 2022年
3. 雑誌名 Journal of Physics: Conference Series	6. 最初と最後の頁 012018 ~ 012018
掲載論文のDOI (デジタルオブジェクト識別子) 10.1088/1742-6596/2207/1/012018	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Suzuki Daisuke, Suzuki Daisuke, Miura Shinichi	4. 巻 91
2. 論文標題 An Efficient Replica Exchange Monte Carlo Method Using the Gaussian Ensemble for First-Order Transitions	5. 発行年 2022年
3. 雑誌名 Journal of the Physical Society of Japan	6. 最初と最後の頁 044006-1 ~ 5
掲載論文のDOI (デジタルオブジェクト識別子) 10.7566/JPSJ.91.044006	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計19件（うち招待講演 1件/うち国際学会 1件）

1. 発表者名 石井完樹, 三浦伸一
2. 発表標題 多重レベル一般化ハイブリッドモンテカルロ法の開発と生体分子への適用
3. 学会等名 日本物理学会2022年秋季大会
4. 発表年 2022年

1. 発表者名 山田悠友, 鈴木大輔, 三浦伸一
2. 発表標題 シニョリンの熱力学的フォールディング転移とエネルギー地形 III
3. 学会等名 第16回分子科学討論会
4. 発表年 2022年

1. 発表者名 鈴木大輔 , 三浦伸一
2. 発表標題 シニョリンの熱力学的フォールディング転移とエネルギー地形 II
3. 学会等名 日本物理学会第77回年次大会
4. 発表年 2022年

1. 発表者名 吉森匠 , 三浦伸一
2. 発表標題 ベイスンホッピング法を用いたアラニンジペプチドのエネルギー地形探索
3. 学会等名 第35回分子シミュレーション討論会
4. 発表年 2021年

1. 発表者名 鈴木大輔 , 三浦伸一
2. 発表標題 シニョリンの熱力学的フォールディング転移とエネルギー地形
3. 学会等名 第35回分子シミュレーション討論会
4. 発表年 2021年

1. 発表者名 吉森匠 , 三浦伸一
2. 発表標題 ベイスンホッピング法を用いた生体分子のエネルギー地形探索
3. 学会等名 日本物理学会秋季大会 (物性)
4. 発表年 2021年

1. 発表者名 鈴木大輔, 三浦伸一
2. 発表標題 拡張アンサンブル法を用いた小さなタンパク質のフォールディングシミュレーション
3. 学会等名 日本物理学会秋季大会 (物性)
4. 発表年 2021年

1. 発表者名 鈴木 大輔 , 三浦 伸一
2. 発表標題 微小タンパク質のエネルギー地形とフォールディング転移
3. 学会等名 第15回分子科学討論会
4. 発表年 2021年

1. 発表者名 Daisuke Suzuki , Shinichi Miura
2. 発表標題 Thermodynamic folding transition of a small protein molecule
3. 学会等名 Conference on Computational Physics 2021 (国際学会)
4. 発表年 2021年

1. 発表者名 鈴木 大輔 , 鈴木 大介 , 三浦 伸一
2. 発表標題 ガウス型統計集合に基づくレプリカ交換法の開発と一次相転移系への適用
3. 学会等名 第23回理論化学討論会
4. 発表年 2021年



1. 発表者名 鈴木大介, 三浦伸一
2. 発表標題 ガウス型統計集合とレプリカ交換法を結合した新規シミュレーション手法の開発
3. 学会等名 日本物理学会 2020年秋季大会 (物性)
4. 発表年 2020年

1. 発表者名 吉森匠, 三浦伸一
2. 発表標題 ベイスンホッピング法を用いた水和生体分子のポテンシャルエネルギー地形探索
3. 学会等名 日本物理学会 2020年秋季大会 (物性)
4. 発表年 2020年

1. 発表者名 鈴木大輔, 三浦伸一
2. 発表標題 マルチカノニカル一般化ハイブリッドモンテカルロ法を用いたシニョリン分子のフォールディング解析
3. 学会等名 日本物理学会 2020年秋季大会 (物性)
4. 発表年 2020年

1. 発表者名 鈴木大輔, 三浦伸一
2. 発表標題 拡張アンサンブル法と液体の積分方程式理論を結合した水和生体高分子の効率的な構造サンプリング手法の開発
3. 学会等名 日本物理学会 2020年秋季大会 (物性)
4. 発表年 2020年

1. 発表者名 鈴木大介 , 三浦伸一
2. 発表標題 ガウス型統計集合とレプリカ交換法を結合した新規シミュレーションの開発と高効率化
3. 学会等名 第34回分子シミュレーション討論会
4. 発表年 2020年

1. 発表者名 吉森匠 , 三浦伸一
2. 発表標題 ベイスンホッピング法を用いた水和生体分子のエネルギー地形探索
3. 学会等名 第34回分子シミュレーション討論会
4. 発表年 2020年

1. 発表者名 鈴木大輔 , 三浦伸一
2. 発表標題 水和シニョリン分子のエネルギー地形とフォールディング転移
3. 学会等名 第34回分子シミュレーション討論会
4. 発表年 2020年

1. 発表者名 鈴木大介 , 三浦伸一
2. 発表標題 ガウス統計集合とレプリカ交換法を結合した新しいシミュレーション手法の開発と一次相転移系への適用
3. 学会等名 日本物理学会第76回年次大会
4. 発表年 2021年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
--	---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------