

科学研究費助成事業 研究成果報告書

平成 28 年 5 月 19 日現在

機関番号：12102

研究種目：若手研究(B)

研究期間：2014～2015

課題番号：26840056

研究課題名(和文) タンパク質構造揺らぎに基づく構造変化予測法の構築

研究課題名(英文) Developments of methods for predicting transitions of proteins based on structural fluctuations

研究代表者

原田 隆平 (Harada, Ryuhei)

筑波大学・数理物質系・特別研究員(PD)

研究者番号：60612174

交付決定額(研究期間全体)：(直接経費) 3,200,000円

研究成果の概要(和文)：タンパク質の構造揺らぎに基づく構造変化予測法を開発した。計算手順として、下記2過程を繰り返す。構造変化を引き起こす可能性が高い分子構造選択、複数の短時間MDシミュレーションによる選択分子構造の構造探索。遷移確率が高い分子構造から短時間MDシミュレーションを繰り返すことで、タンパク質の構造遷移を促進し、効率的に構造変化を予測する。

上記の計算手順に基づき、様々な構造変化予測法を開発し、実際の生体系に適用した。具体的には、ミニタンパク質シニョリンの天然構造予測に適用し、高精度での天然構造予測に成功した。また、加水分解酵素T4リゾチームの開閉運動予測に成功した。

研究成果の概要(英文)：Methods for predicting conformational transitions have been developed based on structural fluctuations of proteins. As computational schemes, the following two processes are repeated: (I) Selections of structures of proteins that have high potential to transit to other meta-stable states, where the candidates are selected based on a set of appropriate reaction coordinates. (II) Short-time molecular dynamics simulations restarting from the structures as conformational search. The above processes are powerful for predicting conformational transitions of proteins efficiently.

According to the strategy, we have developed several sampling methods for predicting conformational transitions, and applied them to biological systems. For instance, our methods have successfully predicted the native state of a mini-protein, chignolin. Furthermore, large-amplitude domain motions of T4 lysozyme in explicit water have been reproduced with our methods.

研究分野：生物物理学

キーワード：分子動力学シミュレーション タンパク質の天然構造予測 タンパク質の構造変化予測

1. 研究開始当初の背景

タンパク質の機能発現に重要な構造変化は「レアイベント」であり、通常の分子動力学(Molecular Dynamics: MD)シミュレーションが到達可能な時間スケールと比較して、長時間の確率過程において観測される。従って、従来のMDシミュレーションに基づきレアイベントを再現することは、時間スケールの制約から困難な場合が多い。時間スケールの制約を打破するために、「ANTON」に代表される生体分子の分子動力学シミュレーション専用の汎用計算機の開発が急速に進んでおり、マイクロ秒からミリ秒に至るタンパク質フォールディングシミュレーションも実現可能になってきている。しかしながら、これらの高速計算機は誰もが利用出来る訳ではない。現状として、研究室レベルの計算機を用いてタンパク質フォールディングに至る時間スケールまでMDシミュレーションを実行することは、依然として困難である。故に、高速計算機に代わる戦略として、タンパク質のレアイベントを再現するための計算手法開発が必要である。

本研究では、汎用計算機を用いた長時間MDシミュレーションを実行する代わりに、初期構造の異なる超並列な短時間MDシミュレーションをカスケード的に実行する「カスケード型超並列シミュレーション」の概念を提案する。本戦略では、タンパク質の長時間ダイナミクスを直接追跡することは困難であるが、時間的制約によるタンパク質構造探索を打開することが可能である。何故ならば、複数の初期構造の異なる短時間MDシミュレーションをリスタートさせることで、遷移確率を上昇させることが出来るからである。本研究では、カスケード型超並列シミュレーションを駆使し、タンパク質機能に重要なレアイベントを効率的に再現する計算手法を開発することを目指す。

2. 研究の目的

レアイベントを効率的に再現するため、カスケード型超並列シミュレーションでは、下記の計算手順を繰り返す。(1) 遷移確率が高い初期構造を反応座標に基づき選択。(2) 選択した初期構造から短時間MDシミュレーションをリスタートすることによる構造リサンプリング。上記の計算手順は単純であるが、遷移確率が高い候補構造を初期条件として選択し、選択した初期構造から短時間MDシミュレーションをリスタートさせることで、効率的にレアイベントを再現することが可能である。また、遷移確率を上昇させるために、初期構造を適切に選択しなければならない。具体的には、適切な反応座標を指定することで、構造遷移に重要な候補構造を選択する。

3. 研究の方法

研究期間内にタンパク質レアイベント探

索するタンパク質構造変化予測法として、下記3つの構造サンプリング法を開発した。

- (1) Fluctuation Flooding Method (FFM).
- (2) Outlier FLOODing (OFLOOD) method.
- (3) TaBoo SeArch (TBSA) algorithm.

(1) Fluctuation Flooding Method (FFM)

FFM (*J. Chem. Phys.*, 140, 125103(1-7) (2014)) では、タンパク質の大振幅構造揺らぎに基づき、遷移確率が高い初期構造を特徴付け、選択した初期構造から短時間MDシミュレーションをリスタートすることで、構造リサンプリングを繰り返し、レアイベントを再現する。具体的には、タンパク質の異方的な構造揺らぎが機能発現に重要であるとして主成分分析により抽出し、振動モードに状態を射影した上で、構造的に大きく揺らいでいるスナップショットを構造リサンプリングにおける初期構造として選択する。上記の計算手順により、初期構造を選択した後、初期速度再生成による短時間MDシミュレーションをリスタートする。このサイクルを繰り返すことで、主成分分析で抽出した振動モード方向に構造変化が促進され、効率的にレアイベントを再現することが出来る。自由エネルギー地形的には、構造的に大きく揺らいでいる構造は地形の縁に位置しているため、これらのタンパク質構造を初期条件として短時間MDシミュレーションを繰り返していくことで、隣接する準安定状態への構造遷移が促進される。

(2) Outlier FLOODing (OFLOOD) method

OFLOOD法 (*J. Comput. Chem.*, 36, 97-102 (2015), *Chem. Phys. Lett.*, 639, 269-274 (2015)) では、タンパク質の状態分布に注目し、確率密度が疎な状態分布に着目する。一般的に、タンパク質の状態分布は反応座標により射影された構造空間における高次元構造分布として記述される。この高次元構造分布において、密な分布は出現確率が高いクラスタとして認識され、タンパク質の安定状態に対応する。これに対して、クラスタに属さない密度が疎な分布は、はずれ値(Outlier)と呼ばれ、遷移状態近傍に存在している可能性が高い。OFLOOD法では、これら出現確率が低いOutlierに対応するタンパク質構造を構造リサンプリングの初期構造として選択し、短時間MDシミュレーションに構造探索していくことで、構造遷移を促進する。Outlierを検出するために、タンパク質の高次元構造分布を階層的クラスタリング手法である「FlexDice」によりクラスタリングし、クラスタに属さない構造分布を求め、Outlierとした。Outlierは元々準安定状態間の遷移状態近傍に位置しているため、初期速度の再分配により運動エネルギーを与え直すことにより、容易に隣接する準安定状態に構造変化する。計算手順としては、Outlierの

構造リサンプリングにより得られたトラジェクトリを用いてタンパク質構造分布をアップデートし、Outlier を逐次検出していく。最終的に、考えている高次元構造空間におけるタンパク質状態分布が収束するまで Outlier の検出とその構造リサンプリングを繰り返す。

(3) TaBoo SeArch (TBSA) algorithm

TBSA (*J. Comput. Chem.*, 36, 763-772 (2015) *Chem. Phys. Lett.*, 630, 68-75 (2015))では、タンパク質状態分布の逆分布を用いることにより、出現確率の小さいタンパク質状態を検出し、構造リサンプリングの初期構造に選択する。具体的には、エネルギー空間に射影された状態分布を考え、その逆分布を定義し、この逆分布を重みとして、構造リサンプリングのための初期構造を選択する。計算手順としては、構造リサンプリングにより得られるスナップショットを用いて逆分布関数をアップデートしながら初期構造選択を繰り返し、短時間 MD シミュレーションを実行することで構造リサンプリングを繰り返す。収束条件として、タンパク質状態分布関数が増加しなくなるまで初期構造選択と構造リサンプリングを繰り返す。TBSA の拡張系として、1 次元逆分布関数を多次元版に拡張することで更なる構造リサンプリング効率の上昇に成功した。

4. 研究成果

(1) FFM の加水分解酵素 T4 リゾチームへの適用

FFM の構造リサンプリング効率の検証として、164 残基の加水分解酵素 T4 リゾチームの全原子系に適用した。T4 リゾチームは、準安定状態として Open 構造と Closed 構造が存在し、2 構造間を構造遷移することが知られている。ただし、構造遷移の時間スケールが長時間であるため、通常の MD シミュレーションにより Open/Closed 構造変化を再現することは困難である。本研究では FFM を適用することで、野生型構造の Open 構造からスタートし、Closed 構造へ至る遷移経路を再現することを試みた。具体的には、Open 構造周辺の大規模構造揺らぎを主成分分析で抽出し、構造的に大きく由来している 10 構造を初期構造として選択した後、初期構造あたり 100 ps の短時間 MD シミュレーションをリスタートさせることで構造リサンプリングを繰り返した。計算コストとしては、1 サイクルあたり 1 ns (10 初期構造 X100 ps) である。FFM を適用することで、10 サイクル (10 ns) 程度の計算コストで Open 構造から Closed 構造へ至る遷移経路を再現することが出来た。従来の MD シミュレーションとの計算効率の比較を行うため、Open 構造スタートの MD シミュレーションを 1 マイクロ秒まで継続したところ、Open 構造から Closed 構造へ至る遷移経路を再現することは出来なかった。上記の計算

結果を比較すると、FFM はナノ秒オーダーの計算コストで T4 リゾチームのレアイベントを再現出来たことになり、通常の MD シミュレーションと比較して、飛躍的な計算効率の向上を示すことが出来た。

(2) OFLOOD 法のタンパク質フォールディングへの適用

OFLOOD 法の構造リサンプリング効率の検証として、小タンパク質のフォールディング経路の再現を試みた。選択した小タンパク質は、汎用計算機 ANTON で用いられた中から選択した。その 1 例として、35 残基の Villin の計算結果について記述する。フォールディングシミュレーションの初期構造としては、完全に伸びきった変性構造をアミノ酸配列からモデリングし、タンパク質周りの溶媒効果は一般化ボルンモデルにより取り込んだ。反応座標として、天然構造の 2 次構造を参照構造とした部分的平均自乗距離 (Root Mean Square Deviation: RMSD) を採用し、構造空間を定義した。この構造空間に構造リサンプリングにより得られるトラジェクトリを射影し、構造分布を作成した後、FlexDice により Outlier を検出した。1 サイクルあたり 100 個の Outlier を検出し、初期構造あたり 100 ps の短時間 MD シミュレーションをリスタートさせた (サイクル当たりの計算コストは、10 ns: 100 初期構造 X100 ps)。計算結果として、50 サイクル (500 ns) 以内で、Villin のフォールディング経路を再現することが出来た。特に、主フォールディング経路に加えて、先行研究では報告されていない副フォールディング経路の再現に成功した。計算コストとしては、ANTON を用いた長時間 MD シミュレーションではフォールディング経路の再現にマイクロ秒オーダーを要していたが、OFLOOD 法はナノ秒のオーダーでレアイベントの再現に成功した。

(3) TBSA のタンパク質フォールディングへの適用

TBSA の構造リサンプリング効率の検証として、OFLOOD 法と同様に小タンパク質のフォールディング経路の再現を試みた。その 1 例として、10 残基の Chignolin の計算結果について記述する。フォールディングシミュレーションの初期構造としては、完全に伸びきった変性構造をアミノ酸配列からモデリングし、タンパク質周りに水分子を露に配置した。反応座標として、慣性半径と初期構造から測定した RMSD を採用し、構造空間を定義した。この構造空間において逆分布を定義した後、初期構造選択の重みとして用いた。1 サイクルあたり 100 個の初期構造を検出し、初期構造あたり 100 ps の短時間 MD シミュレーションをリスタートさせた (サイクル当たりの計算コストは、10 ns: 100 初期構造 X100 ps)。計算結果として、10 サイクル (100 ns) 以内で、Chignolin のフォールデ

インク経路を再現することが出来た。

現在、更なる構造探索効率向上のため、よりの確に出現確率が小さなタンパク質状態を検出出来るように逆分布の改良を行っている。この改良は、よりサイズの大きなタンパク質への適用のために必要であり、今後は巨大系へ適用し 生体機能解析へ繋げていく予定である。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計 13 件)

R. Harada, A. Kitao, Non-targeted Parallel Cascade Selection Molecular Dynamics for Enhancing the Conformational Sampling of Proteins. *J. Chem. Theory Comput.*, 11, 5493-5502 (2015). 査読有, DOI: 10.1021/acs.jctc.5b00723

R. Harada, T. Nakamura, Y. Shigeta, Automatic Detection of Hidden Dimension in OFLOOD Method. *Chem. Phys. Lett.*, 639, 269-274 (2015). 査読有, DOI: 10.1016/j.cplett.2015.09.031.

R. Harada, Y. Takano, Y. Shigeta, Efficient Conformational Sampling of Proteins based on a Multi-Dimensional Inverse Histogram: An Application to Folding of Chignolin in Explicit Solvent. *Chem. Phys. Lett.*, 630, 68-75 (2015). 査読有, DOI: 10.1016/j.cplett.2015.04.039.

Y. Shigeta, R. Harada, M. Kayanuma, M. Shoji, Quantum Cumulant Dynamics for Real-Time Simulations of Quantum Many-Body Systems. *Int. J. Quant. Chem.*, 115, 300-308 (2015). 査読有, DOI: 10.1002.qua.24820.

M. Feig, R. Harada, T. Mori, I. Yu, K. Takahashi, Y. Sugita, Complete Atomistic Model of a Bacterial Cytoplasm for Integrating Physics, Biochemistry, and System Biology. *J. Mol. Graph. Modell.*, 56, 1-9 (2015). 査読有, DOI: 10.1016/j.jmgm.2015.02.004.

R. Harada, Y. Takano, T. Baba, Y. Shigeta, Simple, yet Powerful for Conformational Sampling of Proteins. *Phys. Chem. Chem. Phys.*, 17, 6155-6173 (2015). 査読有, 招待執筆,

DOI: 10.1039/C4CP05262E

R. Harada, Y. Takano, Y. Shigeta, Enhanced Conformational Sampling Method for Proteins on the TaBoo SeArch algorithm: Application to the Folding of a Mini-Protein Chignolin. *J. Comput. Chem.*, 36, 763-772 (2015). 査読有, DOI: 10.102/jcc.23854

R. Harada, T. Nakamura, Y. Takano, Y. Shigeta: Protein Folding Pathways Extracted by OFLOOD Method. *J. Comput. Chem.*, 36, 97-102 (2015). 査読有, DOI: 10.1002/jcc.23773

R. Harada, Y. Nishihara, N. Wakai, A. Kitao: Conformational Transition Pathway and Free Energy Analyses of Proteins by Parallel Cascade Selection Molecular Dynamics (PaCS-MD). *International Conference of Computational Methods in Sciences and Engineering 2014 (ICCMSE 2014) AIP Conf. Proc.*, 1618, 86-89 (2014). 査読有, DOI: 10.1063/1.4897682

西原泰孝, 原田隆平, 北尾彰朗: カスケード型超並列シミュレーションによるタンパク質構造遷移のパスウェイ探索, 「統計数理」, 62, 273-284 (2014). 査読有, 招待執筆.

原田隆平, 北尾彰朗: 理論/実権技術 “タンパク質の柔らかな運動を誘起するシミュレーション”, 日本生物物理学会誌「生物物理」, 54, 167-171 (2014). 査読有, 招待執筆, DOI: 10.2142/biophys.54.167

R. Harada, Y. Takano, Y. Shigeta: Fluctuation Flooding Method (FFM) for Accelerating Conformational Transitions of Proteins. *J. Chem. Phys.*, 140, 125103(1-7) (2014). 査読有, DOI: 10.1063/1.4869594

T. Baba, R. Harada, M. Nakano, Y. Shigeta: On the Induced-Fit Mechanism of Substrate-enzyme Binding Structures of Nylon-Oligomer Hydrolase. *J. Comput. Chem.*, 35, 1240-1247 (2014). 査読有, DOI: 10.1002/jcc.23614

[学会発表](計 11 件)

原田隆平: カスケード型超並列シミュレ

ーションを用いた生物学的レアイベント探索手法の開発, 第2回凝縮系の理論化学, 2016年3月10日~11日, 琉球大学(沖縄県那覇市), 招待講演.

原田隆平, 鷹野優, 重田育照: Simple, yet Powerful Conformational Sampling Methodologies for Proteins, 環太平洋化学会議 Pacificchem2015, 2015年12月15日~20日, ハワイコンベンションセンター(米国ハワイ州).

原田隆平, 鷹野優, 重田育照: 逆状態分布に基づく効率的タンパク質構造サンプリング手法の開発, 第29回分子シミュレーション討論会, 2015年11月30日~12月2日, 朱鷺メッセ(新潟県新潟市).

原田隆平: Simple yet Powerful Conformational Sampling Methods for Reproducing Biologically Rare Events, 第53回日本生物物理学会年会, 2015年9月13日~15日, 金沢大学(石川県金沢市), 招待講演.

原田隆平, 鷹野優, 重田育照: レアイベント探索手法"OFLOOD"で解明するタンパク質フォールディング機構, 第15回日本蛋白質科学会年会, 2015年6月24日~26日, あわぎんホール(徳島県徳島市).

原田隆平, 鷹野優, 重田育照: タンパク質レアイベント抽出のための効率的構造サンプリング手法, 第28回分子シミュレーション討論会, 2014年11月11日~14日, 仙台市民会館(宮城県仙台市).

原田隆平, 鷹野優, 重田育照: タンパク質レアイベントを抽出する効率的構造サンプリング法, 第52回日本生物物理学会年会, 2014年9月25日~27日, 札幌コンベンションセンター(北海道札幌市).

馬場剛史, 原田隆平, 中野雅由, 重田育照: 酵素と基質結合と大規模構造変化に関する理論的研究, 第8回分子科学討論会 2014, 2014年9月21日~23日, 広島大学(広島県広島市).

馬場剛史, 原田隆平, 中野雅由, 重田育照: ナイロンリゴマー分解酵素の大規模構造変化に対するアミノ酸変異導入の効果, 第14回日本蛋白質科学会年会, 2014年6月25日~27日, ワークピア横浜(神奈川県横浜市).

原田隆平, 鷹野優, 重田育照: 蛋白質構造揺らぎを用いた構造変化加速法:FFM, 第14回日本蛋白質科学会年会, 2014年6

月25日~27日, ワークピア横浜(神奈川県横浜市).

馬場剛史, 原田隆平, 中野雅由, 重田育照: 大規模構造変化を誘起するMD手法を利用したナイロンオリゴマー分解酵素の誘導適合過程の解明, 第17回理論化学討論会, 2014年5月22日~24日, 名古屋大学(愛知県名古屋).

〔図書〕(計 0 件)

〔産業財産権〕
出願状況(計 2 件)

名称: 情報処理装置, シミュレーション方法, およびシミュレーションプログラム
発明者: 中村朋健, 原田隆平, 重田育照
権利者: 富士通, 筑波大学
種類: 特許
番号: 2015-156702
出願年月日: 2015年8月7日
国内外の別: 国内外

名称: 情報処理装置, 指標次元抽出方法, および指数次元抽出プログラム
発明者: 中村朋健, 原田隆平, 重田育照
権利者: 富士通, 筑波大学
種類: 特許
番号: 2015-156703
出願年月日: 2015年8月7日
国内外の別: 国内外

取得状況(計 0 件)

〔その他〕
ホームページ等
該当無し

6. 研究組織

(1) 研究代表者

原田 隆平 (HARADA Ryuhei)
筑波大学・数理物質系・特別研究員(PD)
研究者番号: 60612174