

課題名 (タイトル) :

蛋白質機能解析のための自由エネルギー地形計算手法の開発

利用者氏名 : ○原田 隆平*

所属 : 情報基盤センター 計算工学応用開発ユニット

1. 本課題の研究の背景、目的、関係するプロジェクトとの関係

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(MD)で追跡可能な時間より長時間の確率過程で観測される。レアイベントを計算機上に再現し、原子レベルの解像度で解析することは、タンパク質の機能発現を理解する上で非常に有益である。本研究では、タンパク質機能解析に重要となるレアイベントを再現するための、手法論開発を目的とする。

2. 具体的な利用内容、計算方法

レアイベントを計算機上に効率的に再現するために、(1) 構造変化を誘起する可能性が高い初期構造選択と、(2) 複数の短時間 MD による初期構造のリサンプリング過程から成る方法論を提案し、様々な生体系に適用した。本手法で鍵となるのは、構造変化を誘起する可能性が高い初期構造を適切に選択するということである。本報告では、本年度に開発したいいくつかのタンパク質レアイベント探索手法から、OFLOOD 法[R. Harada, T. Nkamura, Y. Shigeta, *J. Comput. Chem.*, volume 36, pages 97-102, (2015)] について詳細に報告する。

OFLOOD 法

タンパク質の構造変化を記述する上で、反応座標を定義し、それらに射影された高次元空間における状態分布を求めることは、直接自由エネルギー地形を探索することにつながる。情報科学的に、それらの状態分布における密な領域はクラスタと呼ばれ、自由エネルギー的に出現頻度が高い準安定状態に対応する。タンパク質は多自由度複雑系であるので、高次元空間上に複数の準安定状態(クラスタ)が存在し、構造変化の際にクラスタ間を遷移する。そこで我々は、クラスタに属さない状態の一部が遷移状態に属す

る可能性があるとして仮定し、出現頻度の低い疎(スパース)な分布を集中的にリサンプリングしていくことで、構造遷移を高確率に誘起できると考えた。情報科学的にクラスタに属さないスパースな分布は、「はずれ値」あるいは "Outlier" と呼ばれる。この「はずれ値」に対応する状態を検出し、初期速度再配分による短時間 MD を用いて、構造リサンプリングを繰り返していくことにより、あたかもタンパク質の状態が洪水(Flooding)のように周囲に溢れ出していく様を模した、「**Outlier FLOODing Method (OFLOOD)**」と呼ばれるタンパク質レアイベント探索手法を提案した。OFLOOD 法の概念図を図 1 に示す。

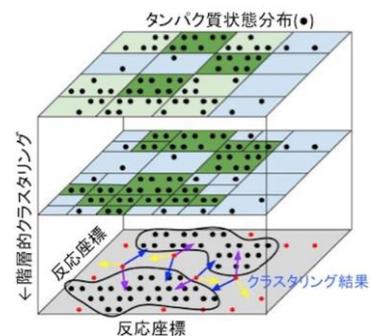


図 1: OFLOOD 法の概念図

下記に、OFLOOD 法の計算手順の詳細を記す。クラスタリングには、階層的クラスタリング手法である "FlexDice" を用いる。FlexDice は、タンパク質状態分布をユニットセルに射影した後(図 1: 最上層の黒点)、ユニットセル内の状態密度に基づき、ユニットセル間のリンク作成していくことにより、状態分布をクラスタリングする。図 1 に示すように、最下層がクラスタリング結果に対応し、黒線で囲まれたクラスタがタンパク質の準安定状態に対応する。OFLOOD 法では、クラスタに属さない Outlier(図 1: 赤点)を集中的にリサンプリングすることで、矢印に示すような準安定状態間の構造遷移を促進する。具体的には、Outlier を初期構造とする短時間 MD シミュレー

ションを初期速度再配分によりリスタートさせ、リサンプリングを繰り返す。これにより、図 1 で示す矢印の様な構造遷移が確率的に誘起され、準安定状態間の構造遷移や、新規準安定状態の探索につながっていく。構造遷移は確率的なものであるから、出来るだけ多くの **Outlier** を選択し、リサンプリングを繰り返すことが重要である。本研究では計算機資源の制約から、1 サイクルあたりおおよそ 100 個の **Outlier** を初期構造として選択し、100ps 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行った。

3. 結果

OFLOOD 法適用によるタンパク質フォールディング過程の再現

OFLOOD の適用例として、Villin(35 残基)のフォールディング経路抽出過程を示す。反応座標として、先行研究[*Proc. Natl. Acad. Sci. USA*, volume 104, pages 4925-4930, (2007)]で定義された平均二乗偏差(root mean square deviation: RMSD)を用いた。2 次元 RMSD として、Helix1-Helix2 および Helix2-Helix3 からなる部分構造(図 2 左下を参照)を segment A, segment B として、天然構造からの RMSD を計算し、状態を 2 次元空間に射影した。初期構造として完全に伸びきったアミノ酸鎖をモデリングし、溶媒を一般化ボルンモデルで取り扱った。完全に伸びきったアミノ酸鎖から出発して、階層的クラスタリングアルゴリズム”FlexDice”によりはずれ値を検出し、初期速度再配分の短時間 MD 計算によるリサンプリングを 20 サイクル繰り返した。

図 2 にはずれ値と短時間 MD によるリサンプリングから得られたトラジェクトリの射影を示す。サイクルを繰り返すごとに、はずれ値が分布の端を広げていき、10 サイクル程で Villin の天然構造をサンプル出来ていることが分かる。最終的な X 線構造からの最小 C_{α} RMSD は 0.60 \AA であり、非常に精度良く天然構造をサンプル出来た。更に、先行研究[*Proc. Natl. Acad. Sci. USA*, volume 104, pages 4925-4930 (2007)]では報告されていない副フォールディング経路(図 2:

Minor folding 経路)も再現できた。計算効率として、天然構造を再現するまで(C_{α} RMSD $< 1.0 \text{ \AA}$)の合計の計算時間が 135.6 ns であり、先行研究[*Proc. Natl. Acad. Sci. USA*, volume 104, pages 4925-4930 (2007)]のレプリカ交換 MD(8 μ s)と比較して、非常に効率的にフォールディング経路を探索できた。その他の適用例として、Chignolin(10 残基), Trp-cage(20 残基), FBP28 WWdomain(28 残基), BBA(28 残基), 1BBD(46 残基)に関しても、完全に伸びきった構造を出発し、数百 ns 程度の計算時間で天然構造を再現することが出来た。

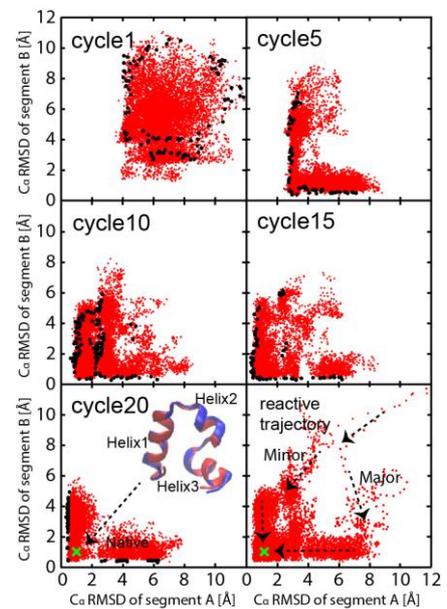


図 2. OFLOOD による探索した Villin のフォールディング経路。

4. まとめ

本研究で開発したレアイベント探索手法の 1 つである OFLOOD 法を用いて、タンパク質フォールディング経路を再現し、そのフォールディング機構を解析した。

5. 今後の計画・展望

今後の展望としては、より残基数の大きい通常のタンパク質への手法適用、タンパク質の大規模構造変化への適用を目指し、手法の改良を継続していく予定である。

平成 27 年度 利用研究成果リスト

【論文、学会報告・雑誌などの論文発表】

[査読付き論文 6 件]

1. **Ryuhei Harada** and Akio Kitao, “Non-targeted parallel cascade selection molecular dynamics (nt-PaCS-MD) for enhancement of the conformational sampling of proteins”, *J. Chem. Theory Comput.*, volume 11, pages 5493-5502, (2015), 査読有.
2. **Ryuhei Harada**, Tomotake Nakamura, and Yasuteru Shigeta, “Automatic detection of hidden dimension in OFLOOD method”, *Chem. Phys. Lett.*, volume 639, pages 269-274, (2015), 査読有.
3. **Ryuhei Harada**, Yu Takano, and Yasuteru 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.*, volume 630, pages 68-75, (2015), 査読有.
4. **Ryuhei Harada**, Yu Takano, and Yasuteru Shigeta, “Enhanced conformational sampling method for proteins based on the TaBoo SeArch algorithm: Application to the folding of a mini-protein, chignolin”, *J. Comput. Chem.*, volume 36, pages 763-772, (2015), 査読有.
5. **Ryuhei Harada**, Tomotake Nakamura, Yu Takano, and Yasuteru Shigeta, “Protein folding pathways extracted by OFLOOD method”, *J. Comput. Chem.*, volume 36, pages 97-102, (2015), 査読有.
6. **Ryuhei Harada**, Tomotake, and Yasuteru Shigeta, “Sparsity-weighted Outlier FLOODing (OFLOOD) method: Efficient rare event sampling method using sparsity of distribution”, *J. Comput. Chem.*, DOI: 10.1002/jcc.24255, (2015), 査読有.

【国際会議などの予稿集、proceeding】

該当無し

【国際会議、学会などでの口頭発表】

[国際学会発表 1 件]

1. **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 2015, 2015 年 12 月 19 日, 米国ハワイコンベンションセンター

【その他（プレスリリース、学術会議以外の一般向けの講演など）】

[特許出願 3 件]

1. 中村朋健, 重田育照, **原田隆平**, 情報処理シミュレーションプログラムおよびシミュレーション方法, 出願番号:2015-0332 (出願年月日:2015 年 2 月 20 日)
2. 中村朋健, 重田育照, **原田隆平**, 情報処理装置, シミュレーション手法, およびシミュレーションプログラム, 出願番号:2015-15602 (出願年月日:2015 年 8 月 7 日)
3. 中村朋健, 重田育照, **原田隆平**, 情報処理装置, 指標次元抽出プログラム, 出願番号:2015-15603 (出願年月日:2015 年 8 月 7 日)