

# カスケード型超並列シミュレーションに立脚した遷移経路探索法の開発と FtsZ 細胞分裂ダイナミクスへの適用

原田 隆平

筑波大学計算科学研究センター

## 1. はじめに

平成 29 年度(前期・後期)スーパーコンピューター若手・女性利用者推薦制度の採択課題「カスケード型超並列シミュレーションに立脚した遷移経路法の開発」及び「カスケード型超並列シミュレーションで探る FtsZ の細胞分裂ダイナミクス」の研究成果を報告する。報告内容は、採択期間において開発を進めた遷移経路探索法の紹介と、計算手法を適用した FtsZ 細胞分裂ダイナミクスの解析である。

## 2. イントロダクション

タンパク質の機能発現に関係する構造遷移は、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「レアイベント」である。故に、シミュレーションの時間スケールの制約から、通常の MD シミュレーションによりレアイベントを再現することは困難である場合が多い。このため、時間スケール問題の打開にむけてアルゴリズム開発が盛んに行われている。特に、タンパク質の MD シミュレーション専用の汎用計算機の開発が進んでいる。代表的な例として、米国 D. E. Shaw リサーチが保有する汎用計算機 "Anton" は、マイクロ秒からミリ秒に至る MD シミュレーションを可能にした。具体的な例として、比較的残基数の小さなタンパク質に対して、溶媒を露に含んだ全原子レベルのフォールディングシミュレーションを実現している [1-5]。しかしながら、研究室レベルの計算機を利用して機能発現の時間スケールに至る 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。また、レアイベント自体が確率過程であることから、長時間 MD シミュレーション実行しても再現出来る保証がないため、不確実性が伴う。故に、レアイベントを再現する計算手法の開発が望まれている。

研究代表者は、生体機能に重要な構造遷移であるレアイベントを再現する計算手法の開発を進めてきた。本報告で解説する「カスケード型超並列シミュレーション」は、レアイベントを再現する遷移経路探索法である。本手法では、長時間 MD シミュレーションを実行する代わりに、初期構造の異なる短時間 MD シミュレーションを超並列に実行する「構造リサンプリング」のサイクルを繰り返すことで、効率的な遷移経路探索を実現する。また、構造遷移を促進するため、サイクル毎に短時間 MD シミュレーションの初期構造を再選択し、Maxwell-Boltzmann 分布に従い初期速度を与え直すことにより、カスケード的に構造リサンプリングを繰り返す。本手法を適用し効率的な遷移経路探索を実現するためには、構造リサンプリング毎に遷移確率が高い初期

構造を選択する必要があるため、適切な反応座標の設定が重要となる。

本報告では、これまでに開発した遷移経路探索法であるカスケード型超並列シミュレーションに属する計算手法: (1) Parallel Cascade Selection MD (PaCS-MD) [6-12], (2) Fluctuation Flooding Method (FFM) [13], (3) Outlier FLOODing (OFLOOD) method [14-18], (4) TaBoo SeArch algorithm (TBSA) [19-21], (5) Structural Dissimilarity Sampling (SDS) [22,23]の中から、平成29年度(前期)に開発・改良を進めた PaCS-MD を中心に解説する。また、平成29年度(後期)の採択課題の PaCS-MD のアプリケーションとして、細菌の細胞分裂ダイナミクスに深く関係しているタンパク質である FtsZ に適用した研究成果[24]も解説する。PaCS-MD 以外の計算手法に関しては、レビュー論文[25]を参照にされたい。

### 3. カスケード型超並列シミュレーション

#### 3.1 遷移経路探索法の設計指針

カスケード型超並列シミュレーションは、「初期構造の選択」と「選択した初期構造からリスタートする短時間 MD シミュレーション」による「構造リサンプリング」を繰り返すことにより遷移経路を探索する。本手法の特徴的な点として、バイアスをかけることなく構造間の遷移経路を探索可能である。従来法では、構造遷移を促進するため系の温度を高温にして MD シミュレーションを実行するか、系の特定の部分に外力や外部ポテンシャルといった「バイアス」を付加することがほとんどであった [26,27]。バイアスは、付加する位置や付加する大きさ、向き、時間などパラメーターを最適化した上で適用しなければならないので、パラメーター最適化の準備計算が必要となる。また、系のサイズが大きくなるにつれて、パラメーターの最適化自体が困難となる場合もあるため、構造探索の計算時間よりもパラメーターの最適化にかかる計算時間の方が長くなる可能性もある。本手法は、反応座標を設定してしまえば、遷移確率が高い初期構造を選択し、短時間 MD シミュレーションにより構造リサンプリングを繰り返していだけなので、バイアスは使用しない。更に、パラメーターを最適化するための準備計算も必要としない。本手法により効率的に構造探索が可能であるのは、構造リサンプリングを通して「稀にしか発生しない分子構造」の出現確率を選択的に上昇させているからである。

図1に本手法の概念図を示す。図1に示す様に、始構造からスタートして終構造へ遷移する可能性が高い候補構造を反応座標に基づき選択し、短時間 MD シミュレーションをリスタートさせ

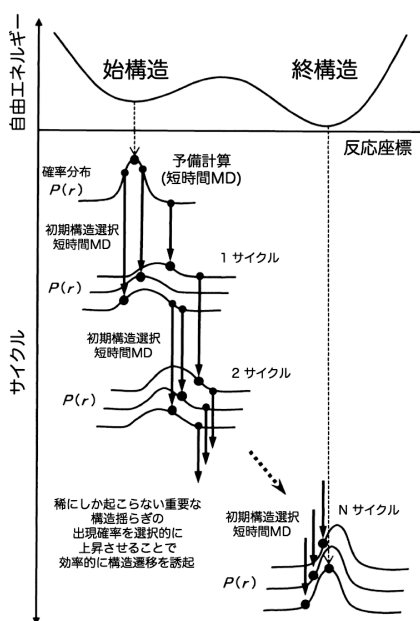


図1: カスケード型超並列シミュレーション

るサイクルを繰り返すことで、徐々に構造分布が終構造近傍に近づいていく。手法適用にあたり様々な反応座標が考えられるが、最も単純な例として、終構造が分かっている場合には、終構造に対する平均自乗距離(Root Mean Square Deviation: RMSD)が考えられる。構造リサンプリングにより得られるトラジェクトリをRMSDの値でランキングし、値の小さい順にスナップショットを複数個選択し、短時間MDシミュレーションの初期構造として構造探索していくことにより、徐々に終構造へ近づいていく。図2に本手法のフローチャートを示す。

図2では、反応座標に射影した値(measures)に基づき、N個の初期構造(seed-1, seed-2, ..., seed-N)を選択し、短時間MDシミュレーションを並列に実行することで遷移経路を探索する。

サイクル毎に、収束条件(終構造に対するRMSDや構造分布)を確認し、条件を満たしていればサイクルを終了する。そうでない場合は、収束条件を満たすまでカスケード的に構造リサンプリングを繰り返す。遷移経路探索後は、必要に応じてアンブレラサンプリング(Umbrella Sampling: US) [26,27]と多重ヒストグラム法(Weighted Histogram Analysis Method: WHAM) [28-30]を併用することで、自由エネルギーを計算する。

その他の自由エネルギー計算法として、マルコフ状態モデル(Markov State Model: MSM)を利用することが可能である。具体的には、MDシミュレーションの各スナップショットを離散化された確率過程の1状態と見なすことで、トラジェクトリがマルコフ連鎖で記述できると仮定する。この仮定に基づけば、生体分子の状態空間はマイクロな状態に離散化され、状態間の遷移確率をMD計算のトラジェクトリから容易に見積もることが出来る。状態空間の離散化は、トラジェクトリをクラスタリングし、マイクロな状態を定義することで行う。具体的には、適切な反応座標を決定後、MDシミュレーションから得られるタンパク質構造をマイクロな状態に割り当てる。ここで、MDシミュレーションから系の時間発展が得られているので、マイクロな状態間の遷移を数え上げることが出来る。また、得られた遷移回数を元に遷移行列(T)を最尤法により推定する。系がマルコフ過程で記述出来るとすると、平衡状態に達したとき、 $\pi = T\pi$  ( $\pi$ : 平衡状態の確率分布)で表される。最終的に、上記の固有値問題を解くことで固有値及び固有ベクトルを求めれば平衡状態(定常状態)の確率分布が分かるため、自由エネルギーを見積もることが出来る。計算の詳細は、文献[7,8,31]を参照されたい。US/WHAMを用いる方法と比較して、MSMを併用する方法は追加計算なしに自由エネルギー計算が可能であるため、計算コストを抑えることが出来る。

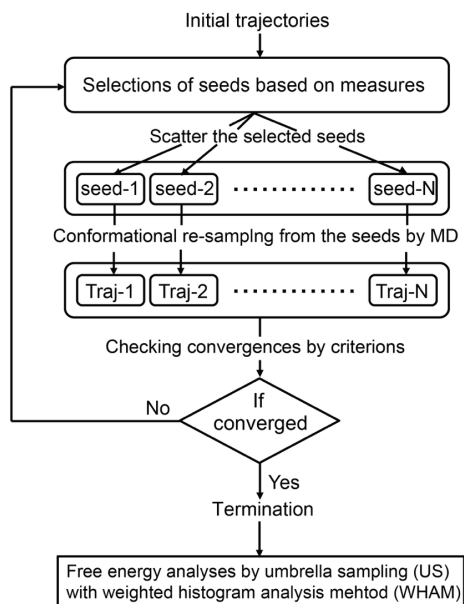


図2: フローチャート

### 3.2 Parallel Cascade Selection MD (PaCS-MD)

PaCS-MD は、始構造と終構造が既知の場合に構造類似性を表す反応座標を設定し、トラジェクトリが終構造へ収束していくように短時間 MD シミュレーションを繰り返す。前述の様に、最も単純な反応座標として構造間の構造類似性を表す RMSD が考えられる。もちろん、様々な反応座標を設定することも可能である。PaCS-MD では、遷移確率の高い初期構造から短時間 MD シミュレーションを用いた構造サンプリングを実行することで外部摂動に頼ることなく遷移経路を探索可能である。PaCS-MD の手順は下記の通りである。(1) 終構造に対する RMSD の値により MD のスナップショットをランキング。(2) 構造類似性が高い(RMSD の値が小さい)上位の複数のスナップショットを初期構造とし、短時間 MD シミュレーションを独立に実行。実際の計算では、サイクルあたり 10 初期構造を選択し、初期構造あたり 100 ps の短時間 MD シミュレーションを実行した。(1), (2)のサイクルを RMSD の値が十分小さくなるまで繰り返す。また、 $RMSD < 1.0 \text{ \AA}$  を収束条件とした。

全原子モデルを用いた PaCS-MD のアプリケーションとして、人工タンパク質 Chignolin(10 残基)のフォールディング過程と加水分解酵素 T4 lysozyme(164 残基)の open-closed 構造変化を再現した計算結果を報告する。Chignolin の始構造として完全に伸びきったアミノ酸構造をモデリングし、天然構造を終構造とした。T4 lysozyme は野生型の open 構造と closed 構造を始・終構造とし、両方向の構造遷移を再現した。終構造に対する RMSD を反応座標と

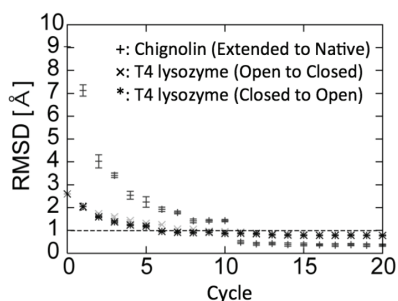


図3: 終構造対に対する RMSD の変化

して、サイクルあたり RMSD の値が小さい上位 10 個のスナップショットを初期構造に選択し、短時間 MD シミュレーションをリスタートさせることにより遷移経路を探索した。図3に終構造に対する RMSD の変化を示す。図3から分かるように、どちらのシステムも 10 サイクル程度(1 ns 秒程度)で RMSD が  $1.0 \text{ \AA}$  以内に収束しており、始構造から終構造へ至る遷移経路を再現することが出来ている。

それぞれのサイクルでどのように初期構造が選択されていくかを追跡し、トラジェクトリをつなぎ合わせるにより、始構造から終構造へ至る擬似的な遷移経路 (リアクティブトラジェクトリ) を生成することが出来る。リアクティブトラジェクトリの物理的解釈として、現実的に可能な構造変化経路候補の 1 つと考えることができ、様々な用途に利用可能である。例えば、構造遷移経路のリファインメントとして Transition Path Sampling [32,33]や String 法[34,35]の初期経路パスとして利用することが可能である。また、アンブレラサンプリングを用いてリアクティブトラジェクトリ周辺の自由エネルギー地形を計算することも可能である。図4に PaCS-MD で生成した Chignolin のリアクティブトラジェクトリの射影を示す。この射影から、変性構造(Extended)からスタートし、天然構造構造(Native)へ至る Chignolin のフォールディング経路が生成されている。

ることが分かる。10本のリアクティブトラジェクトリの中には、フォールディングの途中でミスフォールド(Mis-fold)構造を経由して天然構造へ折り畳まれる遷移経路も存在していることが分かる。

尚, PaCS-MD を拡張することで、終構造が未知の場合でも適用可能である non-targeted PaCS-MD も開発している。紙面の都合上、その詳細は割愛させて頂くが、興味のある方は文献[36]を参照されたい。また、冒頭で記載した OFLOOD 法, TBSA, SDS などの構造探索法も終構造が未知でも適用可能であり, PaCS-MD を拡張した計算手法である。

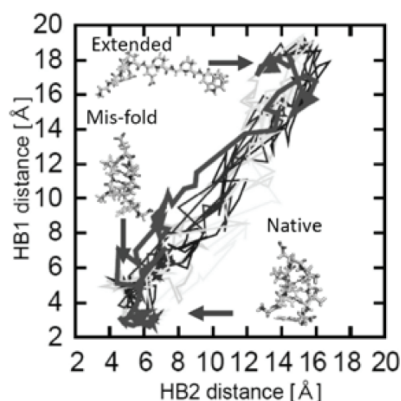


図4: 10本のChignolinのフォールディング経路の反応座標空間(主鎖水素結合距離:HB1, HB2)への射影。矢印は、遷移経路における代表構造を表す。

### 3. PaCS-MD 適用に基づく FtsZ 細胞分裂ダイナミクスの解析

PaCS-MD のアプリケーションとして、細菌の細胞分裂に深く関係しているタンパク質 FtsZ に適用した研究成果[24]を紹介する。FtsZ は、細胞膜の内側にリング状のフィラメントである Z リングを形成し、ダイナミックに離合集散を繰り返すことで細胞膜に陥入を生じさせる(図5左上)。この細胞膜の陥入はZリングの収縮により起きると考えられるが、その分子メカニズムには未解明な部分が多い。平成平成29年度(後期)は、PaCS-MD のアプリケーションとして立命館大学・松村教授の研究グループが X 線構造解析により決定した黄色ブドウ球菌 FtsZ の結晶構造をもとに、細菌の細胞分裂タンパク質の動的秩序解析を行った。

FtsZ に関して興味深いことに、X 線結晶構造解析から同一結晶中に立体構造が大きく異なる 2 つの状態(T-state, R-state)が存在していることが判明した(図5左下)。同一種で状態の異なる 2 構造が得られた研究報告はなく大変貴重な実験結果であるため、PaCS-MD を適用することで FtsZ モノマーの T-R 状態間の構造遷移経路を探索し、細胞分裂タンパク質の動的秩序過程の解析を行った。

PaCS-MD で探索した T-R 状態間の遷移経路を解析したところ、状態遷移における重要なアミノ酸残基のメカニズムを突き止めることができた。具体的には、29番目のアルギニン残基(Arg29)に

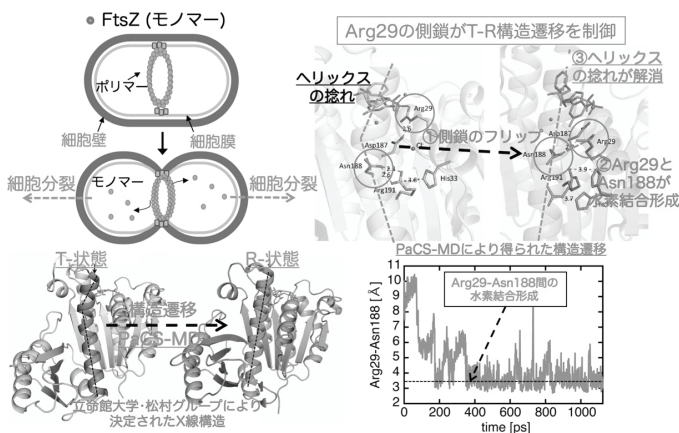


図5: 黄色ブドウ球菌由来の FtsZ モノマーの状態遷移メカニズム



注目したところ、側鎖のフリップがスイッチとなり状態遷移を制御していることが明らかになった。具体的には、図 5 右上に示す様に Arg29 の側鎖がフリップすることで Asn188 の側鎖と水素結合を形成し、FtsZ 中央に存在しているヘリックスが捩れた構造から捩れが解消された直線的な形状に構造遷移するメカニズムを解明した。一般的に、タンパク質側鎖のフリップは長時間スケールスケールで観測されるレアイベントであり、通常の MD シミュレーションで抽出することは困難を伴う。実際に、PaCS-MD で抽出した Arg29-Asn188 の水素結合距離の時系列データを解析したところ、T-R 状態遷移で Arg29-Asn188 間に水素結合が形成されていることが分かる(図 5・右下)。更に遷移経路を解析したところ、FtsZ は 2 状態状態の構造遷移において基質である GDP を段階的に認識・解除しており、中間体構造を経て多段階に状態遷移していることを突き止めた。

本研究において、実験と計算化学が密に連携することで FtsZ の構造揺らぎと T-R 状態間構造遷移におけるメカニズムを解明し、FtsZ ポリマーの離合集散の関係を解明するための足がかりを築くことが出来た。本報告は FtsZ モノマーを中心とした解説であったが、現在 FtsZ ポリマーの計算を継続し計算結果を解析している段階にあり、今後はより生体環境に近いモデルを用いて詳細な細胞分裂過程の動的秩序解明を進めていく。

#### 参考文献

- [1] D.E. Shaw, M.M. Deneroff, R.O. Dror, J.S. Kuskin, R.H. Larson, J.K. Salmon, C. Young, B. Batson, K.J. Bowers, J.C. Chao, M.P. Eastwood, J. Gagliardo, J.P. Grossman, C.R. Ho, D.J. Ierardi, I. Kolossvary, J.L. Klepeis, T. Layman, C. Mcleavey, M.A. Moraes, R. Mueller, E.C. Priest, Y.B. Shan, J. Spengler, M. Theobald, B. Towles, S.C. Wang, *Commun. Acn.* **51** (2008) 91-97.
- [2] J.L. Klepeis, K. Lindorff-Larsen, R.O. Dror, D.E. Shaw, *Curr. Opin. Struct. Biol.* **19** (2009) 120-127.
- [3] D.E. Shaw, P. Maragakis, K. Lindorff-Larsen, S. Piana, R.O. Dror, M.P. Eastwood, J.A. Bank, J.M. Jumper, J.K. Salmon, Y.B. Shan, W. Wriggers, *Science* **330** (2010) 341-346.
- [4] K. Lindorff-Larsen, S. Piana, R.O. Dror, D.E. Shaw, *Science* **334** (2011) 517-520.
- [5] S. Piana, K. Lindorff-Larsen, D.E. Shaw, *Proc. Natl. Acad. Sci. U. S. A.* **110** (2013) 5915-5920.
- [6] R. Harada, A. Kitao, *J. Chem. Phys.* **139** (2013) 035103.
- [7] R. Harada, Y. Nishihara, N. Wakai, A. Kitao, *AIP Conf. Proc.* **1618** (2014) 86-89.
- [8] A. Kitao, R. Harada, Y. Nishihara, D.P. Tran, *AIP Conf. Proc.* **1790** (2016) 020013.
- [9] R. Harada, Y. Shigeta, *J. Comput. Chem.* **38** (2017) 2671-2674.
- [10] R. Harada, Y. Shigeta, *Bull. Chem. Soc. Jpn.* **90** (2017) 1236-1243.
- [11] R. Harada, Y. Shigeta, *Chem. Lett.* **46** (2017) 862-865.
- [12] R. Harada, Y. Shigeta, *Mol. Simul.* **44** (2018) 206-212.

- [13] R. Harada, Y. Takano, Y. Shigeta, *J. Chem. Phys.* **140** (2014) 125103.
- [14] R. Harada, T. Nakamura, Y. Shigeta, *Chem. Phys. Lett.* **639** (2015) 269-274.
- [15] R. Harada, T. Nakamura, Y. Takano, Y. Shigeta, *J. Comput. Chem.* **36** (2015) 97-102.
- [16] R. Harada, T. Nakamura, Y. Shigeta, *J. Comput. Chem.* **37** (2016) 724-738.
- [17] R. Harada, T. Nakamura, Y. Shigeta, *Bull. Chem. Soc. Jpn.* **89** (2016) 1361-1367.
- [18] R. Harada, Y. Takano, Y. Shigeta, *J. Comput. Chem.* **38** (2017) 790-797.
- [19] R. Harada, Y. Takano, Y. Shigeta, *Chem. Phys. Lett.* **630** (2015) 68-75.
- [20] R. Harada, Y. Takano, Y. Shigeta, *J. Comput. Chem.* **36** (2015) 763-772.
- [21] R. Harada, Y. Takano, Y. Shigeta, *J. Chem. Theory Comput.* **12** (2016) 2436-2445.
- [22] R. Harada, Y. Shigeta, *J. Chem. Theory Comput.* **13** (2017) 1411-1423.
- [23] R. Harada, Y. Shigeta, *J. Comput. Chem.* **38** (2017) 1921-1929.
- [24] J. Fujita, R. Harada, Y. Maeda, Y. Saito, E. Mizohata, T. Inoue, Y. Shigeta, H. Matsumura, *J. Struct. Biol.* **198** (2017) 65-73.
- [25] R. Harada, Y. Takano, T. Baba, Y. Shigeta, *Phys. Chem. Chem. Phys.* **17** (2015) 6155-6173.
- [26] G.M. Torrie, J.P. Valleau, *J. Comput. Phys.* **23** (1977) 187-199.
- [27] G.M. Torrie, J.P. Valleau, *J. Chem. Phys.* **66** (1977) 1402-1408.
- [28] S. Kumar, D. Bouzida, R.H. Swendsen, P.A. Kollman, J.M. Rosenberg, *J. Comput. Chem.* **13** (1992) 1011-1021.
- [29] M. Souaille, B. Roux, *Comput. Phys. Commun.* **135** (2001) 40-57.
- [30] J.S. Hub, B.L. de Groot, D. van der Spoel, *J. Chem. Theory Comput.* **6** (2010) 3713-3720.
- [31] D.P. Tran, K. Takemura, K. Kuwata, A. Kitao, *J. Chem. Theory Comput.* **14** (2018) 404-417.
- [32] C. Dellago, P.G. Bolhuis, P.L. Geissler, *Advan. Chem. Phys.*, **123** (2002) 1-78.
- [33] P.G. Bolhuis, D. Chandler, C. Dellago, P.L. Geissler, *Annu. Rev. Phys. Chem.* **53** (2002) 291-318.
- [34] L. Maragliano, A. Fischer, E. Vanden-Eijnden, G. Ciccotti, *J. Chem. Phys.* **125** (2006) 024106.
- [35] L. Maragliano, E. Vanden-Eijnden, *Chem. Phys. Lett.* **446** (2007) 182-190.
- [36] R. Harada, A. Kitao, *J. Chem. Theory Comput.* **11** (2015) 5493-5502.