

生体分子の構造変化に関する計算物理・化学的アプローチ

藤崎 弘士

日本医科大学基礎科学物理学

Physics and Chemistry Based Computational Approach to Conformational Change of Biomolecules

Hiroshi Fujisaki

Department of Physics, Nippon Medical School

Abstract

I review several theoretical and computational approaches for calculating the conformational change of biomolecules, which is a key biomolecular event in cells. After discussing the motivation for the study of conformational change, I describe the path search algorithms in simple terms and discuss the application of the string method, a powerful path search method, to the calculation of minimum free energy pathways in an enzyme, adenylate kinase, which plays a role in converting ADP to AMP and ATP in cells. Finally I briefly mention a more advanced topic of path sampling for biomolecules.

(日本医科大学医学会雑誌 2013; 9: 202-206)

Key words: biomolecule, molecular dynamics, minimum free energy pathway, string method, parallel computing, transition path sampling

1 生体分子における構造変化の意義

タンパク質などの生体分子は特定の機能¹を発揮できるが、それは構造変化を伴うことが多い。構造変化の引き金となるのは何らかの外部からの刺激であり、光受容タンパク質であれば光刺激、ほかのタンパク質であれば、ターゲット分子との結合（広義のリガンド結合²）やリン酸化などである。その結果として、酵素反応が起こったり、細胞内でシグナルの伝達が起こる。これらは複雑な反応の連鎖であり、その各部分を詳細に理解するには原子の分解能をもつ研究手法が必要である。

本稿では、リガンド結合²などによって引き起こされる生体分子（タンパク質）の構造変化に注目する。

これはタンパク質がある安定な状態 A から、外部からの刺激によって新たに生成された安定な状態 B に変化していくという過程であり、本来ダイナミックなものである。しかし、この過程を原子レベルの分解能で追うことは非常に難しい。というのも、こういった構造変化が重要になるタンパク質は非常に巨大であり、またタンパク質の構造が時間とともに変化していくので、平衡的な統計量をとって議論する従来の実験・解析手法では扱いにくいからである。

この状況は計算機を使って構造変化を理解しようとするアプローチにとってもチャレンジングなものである。巨大分子の構造変化は、分子振動の分解能(1 fs = 10^{-15} 秒程度)で計算を進めていく従来の計算科学のアプローチからすると、非常に「遅い」過程(1 ms = 10^{-3} 秒程度)である。また、A から B への構造変化

Correspondence to Hiroshi Fujisaki, Department of Physics, Nippon Medical School, 2-297-2 Kosugi-cho, Nakaharaku, Kawasaki 211-0063, Japan

E-mail: fujisaki@nms.ac.jp

Journal Website (<http://www.nms.ac.jp/jmanms/>)

にいたる道筋は1つでなく、複数あることも予想される。いろんな条件によって、生体分子がその複数の経路を使い分けている可能性もあり、そういった経路全体を特徴づける方法が求められる。しかし、それはますます計算が難しくなるというジレンマにつながる。

そこでそういった問題をどのように定式化し、計算機に実装し、実際の実験と比較できるような結果を出すかということに関する研究が、主に化学物理・生物物理と呼ばれる分野でここ10年ほど盛んに行われるようになった。しかし、上で簡単に述べたように、実際この問題は非常に難しく、あっさり巨大生体分子(たとえば、モータータンパク質)の運動が計算できる、という段階までは至っていない。しかし、様々な新しい考えが生み出され、多様な現象に応用され始めており、筆者の最近の研究内容を含めてここで紹介したい。より技術的なまとめとしては^{3,4}、などを参照のこと。

2 経路探索のアルゴリズム

2.1 経路を求めるときの2つの基準

A地点からB地点に向かうためにはどのような経路(パス)を選べばいいのか?これが経路探索(パスサーチ)の問題である。例えば、ある点Aから別の点Bにいたる地図上の経路について考えてみよう。最も単純にはその間を直線で結ぶようなパスを考えればよい。しかし、もし現実にそのパスを通ろうとしたら、歩行者が地上を歩いている限りは(空中を浮遊したりしなければ)、いろんな障害物があって、易々とはB地点にたどり着けないだろう。通常は「道に沿って」行くことになるが、その場合、よほど近距離でなければ、経路が何通りも出てくる可能性がある。

そこで、いくつかの基準を設けて、その中の1つのパスを決めることになる。例えば、携帯などの電車の乗り換えアプリを使うと(電車の乗り換えも経路探索の問題である)、ある駅から別の駅まで行くときに基準となるのは、(1)速さ(かかる時間の長さ)、(2)運賃の安さ、(3)乗り換え回数などである。分子の「最適な」構造変化パスを考える上での基準もこれらと似ており、

(a) 最小自由エネルギー経路⁵

(b) 最大フラックス経路⁶

というもののみならず調べられる。誤解を恐れずに言うと、(a)は電車乗り換えの「運賃」に、(b)は「速さ」に対応している。もう少し専門的に言うと、ともに経路に沿って「ある量」を積算して、それが小さく

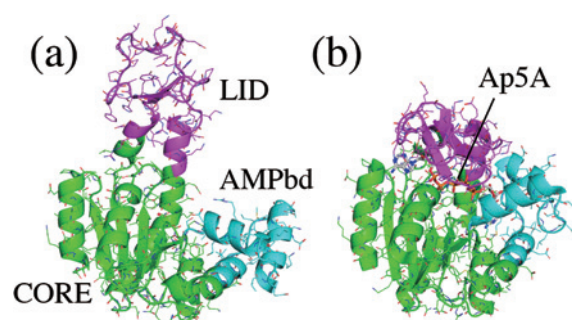


Fig. 1 Molecular structures of adenylate kinase. PDB code: 4AKE (left) and 1AKE (right). Three domains are colored green (CORE), cyan (AMP binding), and purple (LID). A quasi-substrate (Ap5A) is also shown on the right.

なるような経路を選ぶという手順(アルゴリズム)になっている。(a)に関しては、そのある量とは自由エネルギーであり、(b)に関しては、フラックス(流束)と呼ばれる単位時間あたりにモノが流れていく速度のことである。これら2つのパスは独立なものとして定義されているが、どちらのパスもよく似ている場合が多い。

さて、本稿では概念的に易しい(a)のほうを取り上げ、それを分子系に適用した具体例について以下述べていく。

2.2 最小自由エネルギー経路とストリング法

自由エネルギーは分子の安定性や反応などを熱力学的に考える上で最も重要な概念である⁷。リガンド分子とレセプター分子の結合を定量的に理解するために必要な、結合自由エネルギーに関してはすでに前稿²で解説を行っている。

生体分子などの大きな分子の構造変化を考える場合、すべての原子の情報と構造変化を関連付けることは大変であり、またあまり意味(必要)もないので、通常は反応座標(オーダーパラメータ)と呼ばれる少数の自由度で構造変化を表現することが多い。例えば、Fig. 1のタンパク質のように、大きなタンパク質はドメイン構造を持っており、構造変化は各ドメインの大きな動きとして現れてくる。そこで、例えば、動くドメインが2つであれば、最低でも12自由度を取ればよいということになり(それぞれのドメインが並進運動+回転運動で6自由度もっていると考え)、考える+計算する空間を非常に狭めることができる。

その後は、その12次元空間で自由エネルギー「地形」(自由エネルギーは反応座標の値によって、変化

するので、地図上で等高線が地形を表すように、自由エネルギーにも地形がある)を計算し、構造変化を特徴づけることになる。安定な状態 A と B をまず定義し、それらを結ぶような経路で、自由エネルギー変化ができるだけ小さくなるような経路を探すのが合理的だろう。その結果として得られる経路が最小自由エネルギー経路 (minimum free energy path = MFEP) と呼ばれるものである⁵。しかし、ここで問題が2つ出てくる: 1つは12次元でどのように自由エネルギー地形を計算するのかという問題と、それが計算できたとして、MFEPをどう計算するのかという問題である。最初の問題はかなり困難を伴う。というのも、通常は高々3次元程度の自由エネルギー地形しか計算できず、それ以上の次元の自由エネルギー地形は現在の計算資源では精度のよい計算ができないからである。

この問題を解決するために考え出された方法が Vanden-Eijnden らによるストリング法 (string method) である⁵。この方法では、高次元の自由エネルギー地形を計算しない。よって、1つ目の問題を避けており、また2つ目の問題も解決するようなアルゴリズムを与えている。要は、直接、MFEP (高次元空間内の1次元の曲線) だけを求めに行くアルゴリズムになっているので、高次元の自由エネルギー地形を計算しなくても済むのである。ストリング法のアルゴリズムは以下のようなものである: 例えば、Fig. 2 にあるような2次元の問題であれば、

- (1) まず、最初のパスを決めて、それを等間隔に離散化する。その各点をビーズと呼ぶ。
- (2) 各ビーズで自由エネルギーの勾配を計算する。ここが時間のかかる過程である。
- (3) その勾配にマイナスをつけた量に比例する量だけ、各ビーズを動かす。
- (4) (3) の結果、ビーズは等間隔になっていないので、等間隔になるように並べ直す。
- (5) また (2) に戻る。計算が収束するまで、(2)~(4) のサイクルを繰り返す。

となり、この繰り返しの過程が収束したときに、MFEP が得られる。ほかにもいくつか MFEP を求める方法があるが、手順が非常に簡単で、計算機に実装するのも楽であること、様々な拡張が容易であること、反応座標の数を数個から数万個まで好きに選べることなどから、現在ではよく使われるアルゴリズムとなっている。

以下で、この手法をアデニル酸キナーゼという酵素タンパク質に適用したわれわれの結果について述べる⁸。

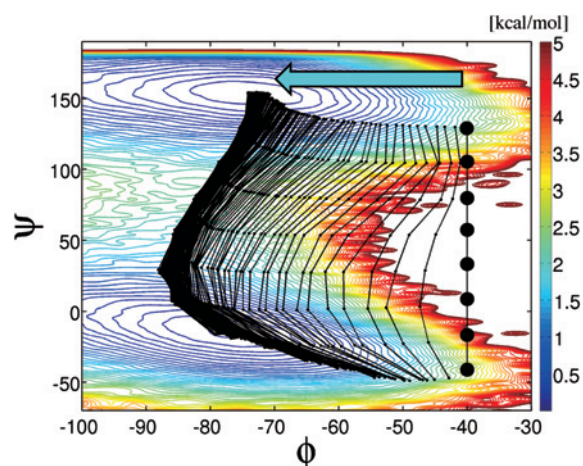


Fig. 2 String method calculation for alanine dipeptide. The free energy profile along the two dihedral angles is also depicted. The initial path is a straight line on the surface, but it converges to a minimum free energy pathway.

2.3 ストリング法を用いたアデニル酸キナーゼの構造変化シミュレーション

アデニル酸キナーゼ (adenylate kinase) は ADP から AMP, ATP への反応 ($\text{ADP} + \text{ADP} \rightleftharpoons \text{ATP} + \text{AMP}$) を生体内で触媒する酵素タンパク質である。比較的小さいタンパク質であることもあり、実験はもちろんのこと、計算でも非常によく調べられている。Fig. 1 にあるように3つのドメインを持っており、反応の際にはそのうちの2つのドメインが残りの1つのドメインの方にリガンド (基質) を抱え込むように動き、酵素反応が起こる。しかし、その構造変化の詳細についてはほとんど分かっていない。

酵素反応のレートなどは実験から分かるが、それから直接タンパク質の分子的な情報が分かるわけではない。また、最近では単分子観測の技術が発展してきており、タンパク質内の数カ所のアミノ酸をラベルして、その間の距離の時間変化を調べることができるようになってきている。しかし、その時間分解能はそれほどよくはなく、また、原子間の距離が高々数カ所 (時間変化とともに) 分かるということであり、例えば活性部位の周りでどのような結合の組み換えが起こっているかということまで分かるものではない。原子的な詳細をフルに与えるものは X 線結晶回折法だが、それは基本的に安定な構造の付近しか分からず、構造変化をしていく途中の構造を知ることは難しい (ただし、原理的に不可能ではない)。

そこで、計算科学の出番となるが、前述のようにこ

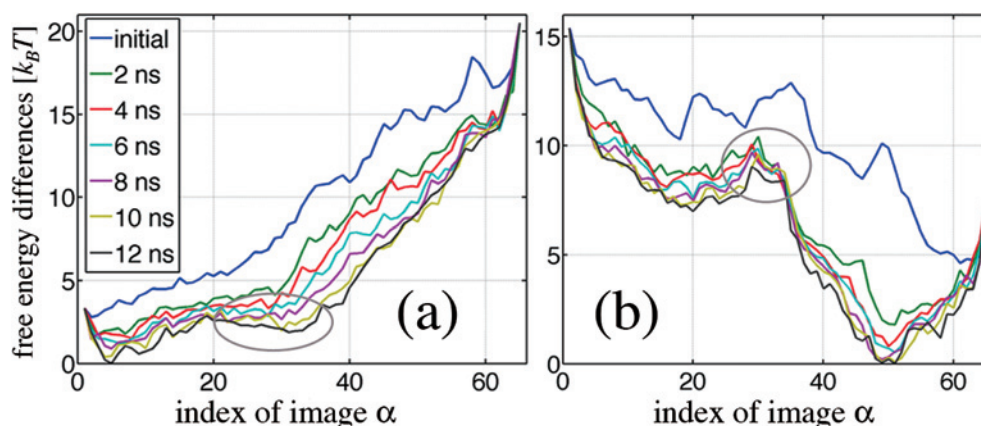


Fig. 3 Converging free energy profiles along the string paths. The black lines correspond to converged minimum free energy pathways. Left: The case without the substrate (Ap5A). Right: The case with the substrate.

のような構造変化は少なくともマイクロ秒かミリ秒（もしくはそれ以上）かかるものと考えられており、通常用いられる分子動力学法ではアクセス可能な時間領域ではない。そこでわれわれはストリング法と超並列計算を用いてこの問題に取り組んだ⁸。（ストリング法などのパスを求めるアルゴリズムは、各ビーズを計算機の各コアに割り振ると効率がよくなるので、京コンピュータのような超並列計算機との相性が非常によい。）

まず上で説明した反応座標を決める必要があるが、タンパク質の場合は、揺らぎの大きな方向というのが実は数自由度に限られているので（異方性があるという）、それを求める手法を用いてまず 20 自由度を切り出した。これは 2 つのドメインが動くということから考えてもリーズナブルな選択である（そのときは 12 自由度なので、20 自由度とっておけばまず問題はない）。そして、その 20 次元の空間で、MFEP を求めた。これは通常の計算機やアルゴリズムではできない計算であり、実際には 1 万コアほどを用いた超並列計算を実行することで結果を得ることができる。（しかし、将来的にはこういった超並列計算は GPGPU と呼ばれる並列計算アーキテクチャーの発達により、容易になるものと予想されている。）

その結果を示したものが Fig. 3 である。これは反応座標上（20 次元の MFEP 上）の自由エネルギー地形を示したものであり、左は基質がない場合、右はある場合である。この図から、基質がないときはドメインが開いた状態が安定であることが分かる。しかし、基質が結合することで自由エネルギー地形は大きく変化し、ドメインが閉じた状態が安定となる。また、その際に、開いた状態付近にも準安定な状態ができ、そ

こからバリアを登って閉じた状態に移っていく様子を想像できる。この山の頂点のことを遷移状態（transition state）と呼ぶが、これは（酵素）反応を司る関所のような場所である。よって、遷移状態を同定し、その性質を調べることは非常に重要である。われわれの計算によって、初めてこの酵素の遷移状態を決定し、その分子レベルの性質を（計算機上で）調べることが可能となった。この手法は様々な酵素や巨大タンパク質に適用可能であり、現在そのような応用計算が行われている。これによって、様々な酵素反応の遷移状態の詳細や、もっと大きなタンパク質の生体機能を理解することが可能になるものと思われる。

3 その他の手法：パスサンプリング

上で紹介したストリング法は構造変化を計算で求めることのできる非常に優れた方法であり、将来的にも広く使われていくことが予想されるが、いくつかの弱点がある。そのうちの最も重要なものは、構造変化に関わる時間スケールや、ダイナミックな分子の動きなどをとらえることができないということである。そこでそのような動的なパスを求めるためのアルゴリズム（パスサンプリングと総称される）もここ 10 年ほどで化学物理・生物物理の分野で開発されてきている。

その中で最も一般的でポピュラーなものは transition path sampling (TPS) と呼ばれる手法である⁹。これはパスに重みをつけて、それに基づくサンプリングを行っていくという手法であり、やや込み入っているが非平衡のダイナミクスを統計的に理解する上で出発点となる考え方である。われわれはパスサンプリングについても研究を行っており、TPS やそ

れに類する手法の生体分子への応用を目指した開発を行っている¹⁰⁻¹²。その結果として、生体分子が機能するときの、よりダイナミックで分子的な描像が得られるものと期待している。

謝辞：松永康佑博士（理化学研究所）には Fig. 1~3 を提供して頂き、また本稿に有益なコメントをして頂いた。ここに感謝します。

文 献

1. Petsko GA, Ringe D: タンパク質の構造と機能, 2005; メディカル・サイエンス・インターナショナル.
2. 藤崎弘士: タンパク質へのリガンド結合に関する計算物理・化学的アプローチ. 日本医科大学医学会雑誌 2013; 9: 135-139.
3. Fuchigami S, Fujisaki H, Matsunaga Y, Kidera A: Protein Functional Motion: Basic Concepts and Computational Methodologies. Adv Chem Phys 2011; 145: 35-82.
4. 藤崎弘士: 生体分子におけるパスサーチおよびパスサンプリングについて. 日本医科大学基礎科学紀要 2011; 40: 83-98.
5. Maragliano L, Vanden-Eijnden E: On-the-fly string method for minimum free energy paths calculation. Chem Phys Lett 2007; 446: 182-190.
6. Huo S, Straub JE: The MaxFlux algorithm for calculating variationally optimized reaction paths for conformational transitions in many body systems at finite temperature. J Chem Phys 1997; 107: 5000-5006.
7. Zuckerman DM: Statistical Physics of Biomolecules: An Introduction, 2010; CRC Press.
8. Matsunaga Y, Fujisaki H, Terada T, Furuta T, Moritsugu K, Kidera A: Minimum Free Energy Path of Ligand-Induced Transition in Adenylate Kinase. PLoS Comput Biol 2012; 8: e1002555.
9. Dellago C, Bolhuis PG, Geissler PL: Transition Path Sampling. Adv Chem Phys 2002; 123: 1-84.
10. Fujisaki H, Shiga M, Kidera A: Onsager-Machlup action-based path sampling and its combination with replica exchange for diffusive and multiple pathways. J Chem Phys 2010; 132: 134101.
11. Fujisaki H, Shiga M, Moritsugu K, Kidera A: Multiscale enhanced path sampling based on the Onsager-Machlup action: Application to a model polymer. J Chem Phys 2013; 139: 054117.
12. Shiga M, Fujisaki H: A quantum generalization of intrinsic reaction coordinate using path integral centroid coordinate. J Chem Phys 2012; 136: 184103.

(受付：2013年7月4日)

(受理：2013年8月1日)