

生体系のシミュレーションのサンプリング手法及び解析方法の開発

光武亜代理 (慶應義塾大学理工学部物理学科, さきがけ研究者)

蛋白質は20種類のアミノ酸が共有結合であるペプチド結合でつながった1本の高分子鎖である。アミノ酸の1次配列の違いにより蛋白質の立体構造や機能が特徴づけられる。複雑な分子であるのに、生理的条件下で約100残基程度のもは自発的に1つの立体構造(天然構造)に折れ畳まれる。化学物理の視点で蛋白質の折れ畳み機構を理解することは非常に面白い。また、複数の低分子や蛋白質が非共有結合によって会合して機能を生じる現象を分子レベルで理解することも興味深い。

分子動力学法は、生体高分子の構造安定性、ダイナミクスや機能について分子レベルで調べることができる強力な方法である。蛋白質系では、以下の3つの難しさがある。1つ目は系のポテンシャル関数(分子力場)の精度である。蛋白質やまわりの溶媒を構成する原子間の相互作用を古典的に正確に取り入れることが必要である。蛋白質の分子力場は1960年代後半から開発がはじまり、現在、AMBER、CHARMMやGROMACSなどの分子力場が存在する。分子力場に結果が依存することもあるが、近年、サンプリングの向上により分子力場を評価できるようになり、精度が向上している。

2つ目は、サンプリングの問題である。研究対象の時間スケールに対応した十分なサンプリングをすることが必要である。蛋白質は多数の原子からなるヘテロな高分子で、周りの多くの水分子と相互作用をして構造変化する。つまり、水分子も含めた多自由度複雑系である。100残基程度の蛋白質の折れ畳み速度は通常ms程度で、分子シミュレーションで追うのは大変時間がかかる。折れ畳みに時間がかかるのは、自由エネルギー空間で状態がエネルギー極小状態にトラップされているからである(ms程度で自由エネルギーバリアーを超えられ

る)。効率の良いサンプリング手法を用いて自由エネルギーバリアーを乗り越えることができれば、最安定構造を得ることが計算機で実現できる(最適化問題)。最近は、専用計算機を駆使してms程度のシミュレーションが可能となっているが、会合系などの複雑な系では、依然、サンプリング問題は存在する。

3つ目は、解析手法である。生体分子の複雑な動きから巻戻りや機能に重要な要因を抜き出す必要がある。また、自由エネルギーを計算して構造安定性を調べたり、レイイベントや重要なダイナミクスを解析したりする必要がある。蛋白質の解析手法は、シミュレーション時間の増加に伴い発展してきた。静的な(時間の情報を使わない)解析手法は多く存在するが、動的解析手法は少ない。複数の極小エネルギー安定構造の特定や、どれくらいの時間スケールでどのように構造間の遷移が生じるかなどを調べる動的な解析手法が求められている。

筆者は、統計力学に基づく手法を生体高分子系の分子シミュレーションに適用・開発してきた(実験でいうと装置開発に対応する)。サンプリングの問題に関しては、効率の良いサンプリング手法である拡張アンサンブル法を蛋白質系に応用することや関連する新しい手法の開発をすることを行ってきた。(本手法は、他の複雑な分子からなる系にも応用できる。)解析手法に関しては、高分子の分野で開発された緩和モード解析という動的解析手法を蛋白質のシミュレーションに導入することを試みた。本稿では、筆者が行ってきた方法論の開発について述べる。蛋白質の構造安定性、ダイナミクスや機能の機構についてはまだ分からないことが多くある。これまで開発してきた方法は、今後さらに重要性を増すと考えられる。

—Keywords—

分子動力学法:

粒子の運動方程式(微分方程式)を数値的に解くことにより、粒子の座標 $r = \{r_1, \dots, r_N\}$ と運動量 $p = \{p_1, \dots, p_N\}$ の時間発展を追う方法。運動方程式 ($m_i \ddot{r}_i = F_i = -(\partial U / \partial r_i)$) を解くと、座標と運動量に関して全エネルギー E が一定のミクロカノニカルアンサンブルが得られる。実験系は温度と圧力が一定の系に対応するので、温度や圧力を一定にする分子動力学法が一般に使われている。

分子力場:

運動方程式のポテンシャル関数に対応するもの。蛋白質の古典系の分子シミュレーションでは一般に以下のようなエネルギー関数が使われている(AMBER14マニュアルから)。

$$U = \sum_{\text{bonds}} k (r - r_{\text{eq}})^2 + \sum_{\text{angles}} k (\theta - \theta_{\text{eq}})^2 + \sum_{\text{dihedrals}} \frac{V_n}{2} [1 + \cos(n\phi - \gamma)] + \sum_{\langle i, j \rangle} \left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^6} \right] + \sum_{\langle i, j \rangle} \left[\frac{q_i q_j}{\epsilon R_{ij}} \right]$$

詳しい変数の定義については、マニュアルを参照して頂きたいが、右辺の第3項までは共有結合項で、隣接した原子間の距離や角度や二面角に関する項である。右辺の第4項と5項は非共有結合項でレナード・ジョーンズ項と静電相互作用項である。

モンテカルロ法:

乱数を用いる方法。狭い意味では、乱数を用いてある重みに従って状態 x を生成する方法。ボルツマン因子 $\exp(-\beta E(x))$ に従って状態を生成するメトロポリスの方法はよく使われている。