

MM・MD・QM/MM法による生体分子シミュレーション

東 雅大

分子シミュレーションでは、計算機の中で物理法則に従って分子を動かして「実験（シミュレーション）」を行う。近年の計算機技術の著しい発展により、タンパク質や核酸、糖質などの生体高分子のさまざまな物性や反応性を分子シミュレーションにより解明することが可能になりつつある。さらに、創薬の分野でも、無数に存在する化合物から標的タンパク質と結合する薬の候補化合物を探す（スクリーニングする）際に分子シミュレーションがよく用いられている。分子シミュレーションは実際の実験で試す候補化合物の数を大幅に減らすことで、開発コストの削減に大きく貢献している（図1）。

本稿では、これらの分子シミュレーションに多く用いられている分子力学法、分子動力学法、quantum mechanics / molecular mechanics (QM/MM) 法の概略と近年の発展について紹介する。

分子力学法

分子シミュレーションを行うためには、まず、分子のさまざまな構造（原子核配置）に対するポテンシャルエネルギーを評価する必要がある。分子のポテンシャルエネルギーは原理的には分子軌道法や密度汎関数法などの量子力学計算によって評価できるが、タンパク質などの巨大な分子の量子力学計算は現在の計算機環境でも非常に時間がかかる。そこで、タンパク質などの巨大な分子のポテンシャルエネルギーの評価には、主に分子力学（molecular mechanics, MM）法が用いられる。

MM法では、分子の結合の形からポテンシャルエネ

ルギー V を以下のような単純な式で表現する。

$$\begin{aligned}
 V = & \sum_{\text{bond}} K_r (r - r_0)^2 + \sum_{\text{angle}} K_\theta (\theta - \theta_0)^2 \\
 & + \sum_{\text{dihedral}} K_\phi [1 + \cos(n\phi - \gamma)] \\
 & + \sum_{\text{nonbond}} \left[\frac{q_i q_j}{R_{ij}} + \frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^6} \right]
 \end{aligned} \quad (1)$$

右辺の第1～4項は、それぞれ結合距離 r 、結合角 θ 、二面角 ϕ 、結合のない2原子間の距離 R_{ij} に対するポテンシャルエネルギーを表す。それ以外の式の中の変数はパラメータであるが、主に高精度量子化学計算や実験データを再現するように決められる。タンパク質などの生体分子に適用できる汎用パラメータとして、AMBER (assisted model building with energy refinement)¹⁾ と CHARMM (Chemistry at HARvard Macromolecular Mechanics)²⁾ が有名である。（なお、CHARMMの開発者のKarplusは、2013年にノーベル化学賞を受賞した。）このようなMM法のパラメータの開発は、古くは1970年代から行われているが、現在でも生体分子をより高精度に記述可能なパラメータの開発が行われている。

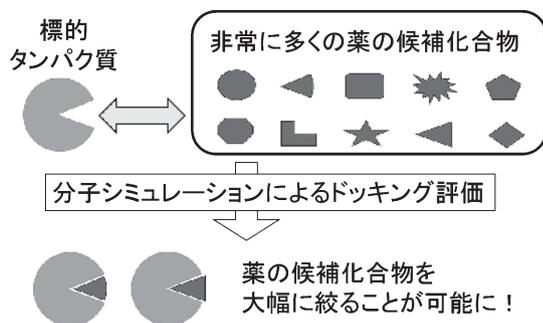


図1. 分子シミュレーションを用いた創薬の概念図

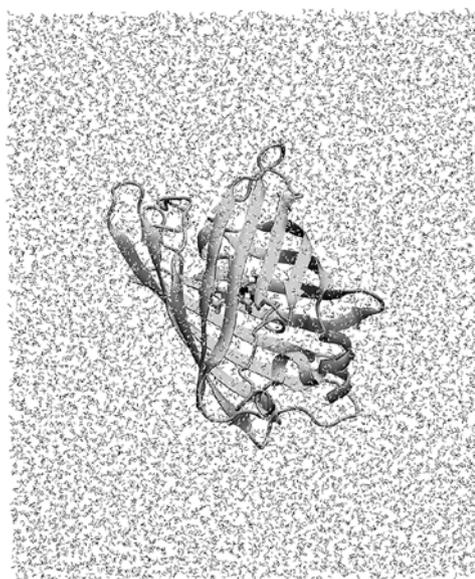


図2. 緑色蛍光タンパク質（GFP）の分子シミュレーション

MM法により、タンパク質などの巨大な分子でも高速にポテンシャルエネルギーを評価することが可能である。たとえば図2のような緑色蛍光タンパク質 (GFP) と水分子からなる約5万原子の巨大な系でも現在の計算機環境なら1つの構造につき、わずか数秒で計算できる。前述の創薬におけるスクリーニングの場合、単純なものでは標的タンパク質の活性部位にさまざまな向きで候補化合物をはめこみ、式(1)の最後の項でタンパク質と化合物の相互作用エネルギーを評価している。また、ポテンシャルエネルギーの構造に対する微分から各原子に働く力が求められ、構造の最適化も可能である。(ただし、最適化で得られる構造は初期構造に大きく依存する。また、熱揺らぎを考慮した上で安定性を議論する場合は、後述する分子動力学法を用いた方がよい。)

近年のMM法を用いた新しい試みとして、Foldit³⁾と呼ばれるタンパク質の立体構造を予測するパズルゲームがある。Folditでは、プレイヤーはタンパク質の分子モデルを動かし、より安定な構造を作るほど高得点が得られるが、その得点の評価にMM法が用いられている。2011年、研究者が10年以上解けなかったタンパク質の立体構造がFolditプレイヤーによりわずか3週間で解明されたと話題になった⁴⁾。

MM法の欠点は、分子の結合の形からポテンシャル関数を作成するために、結合の生成や解離が起きる化学反応は取り扱えないことである。さらに、光を吸収して電子励起状態になった色素分子なども、従来の基底状態を想定したパラメータは適用できないために取扱いは困難である。これからの解決策として、後述するQM/MM法がある。

分子動力学法

MM法を用いて、熱揺らぎを考慮した上で構造の安定性を議論したい、また動的な物性を議論したい場合は分子動力学 (molecular dynamics, MD) 法を用いることが多い。MD法では、ニュートンの運動方程式

$$m_i \frac{d^2 \mathbf{R}_i}{dt^2} = \mathbf{F}_i = - \frac{dV}{d\mathbf{R}_i} \quad (2)$$

を数値的に解くことで、分子を時間に沿って動かしていく。具体的には、ある時刻 t での構造、速度、各原子に働く力から、わずかな時間 Δt だけ進んだ時刻 $t + \Delta t$ での構造と速度が求められる。さらに、時刻 $t + \Delta t$ での構造で各原子に働く力を計算することにより、時刻 $t + 2\Delta t$ での構造と速度が分かる。この繰り返しで分子を動かしていく。このようにして得られた分子の軌跡 (トラジェ

クトリー) から熱揺らぎを考慮した構造の安定性やタンパク質周囲の水の分布などさまざまな物性を解析可能である。

また、MD法では実際に時間に沿って分子を動かすため、分子の動きや構造変化など動的な物性も計算可能である。しかし、シミュレーション可能な時間スケールは著しく限られる。運動方程式を正しく解く (系の全エネルギーが保存する) ためには、時間の刻み幅 Δt を十分短くする必要がある。分子の運動を考える場合、 Δt は1フェムト秒か2フェムト秒程度が限界である。すなわち、1ミリ秒のトラジェクトリーを得たい場合、数億回の力の計算が要求されることを意味する。MM法により高速で分子のポテンシャルエネルギーや力を計算できるとはいえ、これは現在の計算機環境をもってしても容易ではない。さらに統計平均を充分に取るためには見たい時間オーダーの数千から数万倍のトラジェクトリーが必要なことも考えると、現在、容易に取り扱える動的な現象は、アミノ酸残基の側鎖の構造変化などピコ秒やナノ秒オーダーの動きが限度であろう。しかし、現象がそれより遅くとも、タンパク質とリガンドのドッキングのように始状態と終状態が分かっており、その間の「反応座標」を適切に定義することが可能な場合は、反応座標方向に人為的に力を加えたシミュレーションを行い、そのトラジェクトリーから自由エネルギーの変化量を計算することは可能である。一方、タンパク質のフォールディングのように分子の動きがナノ秒よりもはるかに遅く、かつ反応座標を定義することが困難な場合は、取扱いが困難である。そのため、このような遅い現象を取り扱える手法の開発が行われている。

近年のMD法に関連する大きな出来事として、Antonと呼ばれるMD法専用のスーパーコンピュータがShawらにより開発された⁵⁾。Antonは従来のスーパーコンピュータと異なり、MD法しか実行できないが、その分MD法の実行速度は非常に速く、1日で10万原子程度の系の5マイクロ秒ものMDシミュレーションが実行できる⁶⁾。Antonによりすでに小さなタンパク質のフォールディングを再現することに成功しており⁷⁾、イオンチャネルにおけるイオン透過⁸⁾やタンパク質と薬剤分子とのドッキング⁹⁾など、今まで人為的な力を加えるなど工夫しないと取り扱えなかった時間スケールの現象も、ある意味「力技」で取り扱えるようになりつつある。また、スーパーコンピュータを用いなくても、パソコンのグラフィック描画に用いられるgraphics processing unit (GPU) をMD法で利用することで、比較的安価に長時間のシミュレーションを行うことも可能となってきた。

しかし、さまざまな現象を人為的な力を加えずに起こすことができても、なぜそれが起きるのかを理解することが困難なことも多い。今後は長時間のトラジェクトリーをどのように解析するかが重要になっていくのかもしれない。

QM/MM法

前述したようにMM法の大きな欠点は、結合の生成や解離が起きる酵素反応や電子励起状態を伴う光反応などを取り扱えないことである。そのような反応の取扱いには量子力学計算が必要となるが、タンパク質のような巨大な系のすべてを量子力学で計算することは非常に難しい。また、反応の多くは酵素の活性中心やタンパク質中の色素など局所的に起こっており、系すべてを量子力学で扱う必要もない。そこで、反応が起きる重要な部分のみを量子力学 (quantum mechanics, QM) で扱い、残りの大部分をMM法で扱うQM/MM法が有効である。

QM/MM法では系をQM領域とMM領域の2つに分割し、ポテンシャルエネルギーを以下のように表す。

$$\begin{aligned} V(\mathbf{R}^{\text{QM}}, \mathbf{R}^{\text{MM}}) = & V^{\text{QM}}(\mathbf{R}^{\text{QM}}) \\ & + V^{\text{QM-MM}}(\mathbf{R}^{\text{QM}}, \mathbf{R}^{\text{MM}}) \\ & + V^{\text{MM}}(\mathbf{R}^{\text{MM}}) \end{aligned} \quad (3)$$

右辺の第1～3項は、それぞれQM領域のエネルギー、QM-MM間の相互作用エネルギー、MM領域のエネルギーである。さらに、QM-MM間の相互作用エネルギーは、静電相互作用、ファンデルワールス相互作用、共有結合相互作用 (QMとMM領域の境界に共有結合がある場合) の3つに分けられる。このうち、量子力学で取り扱うのは反応に大きな影響を与える静電相互作用だけのことが多い。また、特にQMとMM領域の境界に共有結合がある場合、境界周辺の取扱いは困難なことが多く、現在でもさまざまな手法の開発が行われている。

この「系をQMとMMに分割する」というアイデアを最初に生み出したのがWarshelとKarplusであり¹⁰⁾、現在のQM/MM法とほぼ同じ計算を最初に行ったのがWarshelとLevittである¹¹⁾。現在では凝縮相中の化学反応の解析にQM/MM法はなくてはならない手法となっており、2013年、Karplus, Levitt, Warshelの3名はノーベル化学賞を受賞した。

QM/MM法により、酵素反応の遷移状態や光受容タンパク質の吸収エネルギーなどが解析可能である。しかし、酵素反応の自由エネルギー計算や光受容タンパク質における光励起ダイナミクスなど「揺らぎ」を露わに考

慮する必要がある現象の取扱いは現在でも非常に困難である。QM/MM法の欠点は、MM法と比較して計算コストが遥かに高いことである。QM計算は、最低でもMM計算の数倍以上の計算コストを要求する。したがって、QM領域で扱う原子数がごくわずかでも、残りの大部分のMM領域の計算よりもQM領域の計算のコストは遥かに高い。そのため、MD法などの統計力学的な手法をQM/MM法で用いることは困難であり、QM/MMでMDシミュレーションを行うとすると、現在の計算機環境でも、QM計算の精度を落としてもピコ秒のオーダーがやっとである。さらに、電子励起状態のQM計算には基底状態の計算よりも遥かに計算コストが要求され、状況は一層深刻である。揺らぎを露わに考慮するためには、更なる手法の開発が必要である。

QM/MM法における揺らぎ

QM/MM法を用いて揺らぎが重要な現象を理解するために、さまざまな手法の開発が行われてきた。それらを大きく分けると2つに分けられる。1つはQM領域の熱揺らぎを露わに考慮することを諦め、MM領域のみ熱揺らぎを考慮する手法である。この場合、たとえばMDシミュレーションでは、QM領域の構造や電荷分布は固定したままMM領域のみ分子を動かすため、QM計算を行う回数を大幅に削減可能である。また、もう1つは従来の統計力学手法が使えるように、QM領域の計算コストを大幅に減らす手法である。一般的にはQM計算のコストを減らすと精度も落ちてしまうため、結果の信頼性も落ちてしまう。しかし、適切なポテンシャル関数などでQM計算を上手く置き換えれば精度を落とさずにコスト削減が可能である。我々も、この両方の立場から新たな手法の開発を行ってきた。以下に概要を紹介する。

まず、前者の手法として、タンパク質のような不均一な環境でも線形応答理論を用いて効率的に自由エネルギーを計算可能なQM/MM linear response free energy (LRFE) 法を開発した¹²⁾。溶液中など均一な環境下での分子の自由エネルギーを計算する手法は、液体の積分方程式理論とQM計算を組み合わせたreference interaction site model self-consistent field (RISM-SCF) 法¹³⁾などいくつかあるが、タンパク質など不均一な環境に適用できるものはほとんど存在しなかった。そこで我々は、QM-MM間の相互作用に線形応答理論を用いることで、効率的にQM領域の構造を最適化し、自由エネルギーを計算可能な手法を開発した。また、この手法を酵素chorismate mutase内のchorismateのクライゼン転位反応に適用し、酵素は反応が起こりやすいように

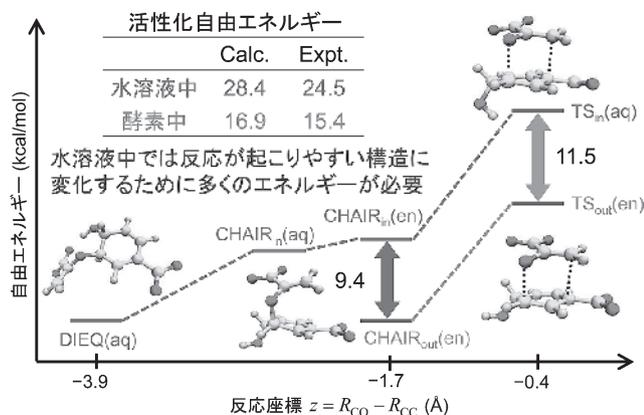


図3. 水溶液中と酵素中における chorismate のクライゼン転位反応の自由エネルギープロファイル。

chorismate の構造を制限していることを明らかにした¹⁴⁾ (図3)。さらに近年、この手法は線形応答理論を用いないより厳密な自由エネルギー計算手法に改良され、酵素反応に伴うタンパク質の大きな動きの解析¹⁵⁾や溶液中のクロロフィル色素の吸収エネルギーの解析¹⁶⁾に適用されている。

また、後者の手法として、QM領域の計算コストを大幅に減らすために、凝縮相の化学反応のポテンシャル面を高精度・高効率に生成する electrostatically embedded multi-configuration molecular mechanics (EE-MCMM) 法を開発した¹⁷⁾。EE-MCMM法では、MM法を用いて反応物と生成物のポテンシャル面を大まかに描き、QM計算との差を Shepard 内挿法で補正することで、わずかな QM/MM法の計算結果から化学反応の大域的なポテンシャル面を高精度・高効率に生成することが可能である。この EE-MCMM法を酵素 haloalkane dehalogenase 内の脱ハロゲン化反応に適用し、精度をほとんど落とさずに QM 計算を約 1000 倍高速化し、反応の活性化自由エネルギーを定量的に計算することに成功している¹⁸⁾。

EE-MCMM法のような QM 領域の計算コストを大幅に削減する手法のメリットは、自由エネルギーを計算可能なことだけでなく、光受容タンパク質の光励起ダイナミクスなど凝縮相の励起状態ダイナミクスを露わに取り扱い可能なことである。基底状態の反応と比較して遥かに高コストの QM 計算を必要とする凝縮相中の励起状態反応ダイナミクスの解析は、従来の手法では不可能であり、ほぼまったく未開拓の分野であった。たとえば、我々が行った溶液中の励起状態分子内プロトン移動の解析¹⁹⁾は、従来法なら 25000 年もの時間がかかる。それを EE-MCMM法によりわずか 1 か月に短縮することにより、解析を可能にしている。

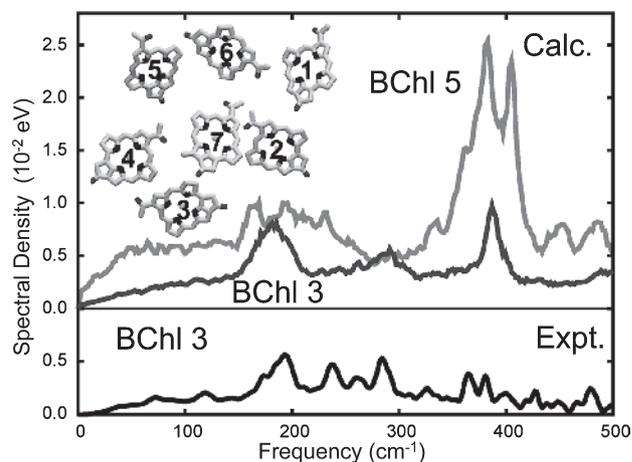


図4. FMOタンパク中の色素バクテリオクロロフィルaの励起エネルギーの揺らぎの大きさを表す Spectral Density。

また現在、我々は光捕集アンテナの1つである Fenna-Matthews-Olson (FMO) タンパクにおける励起エネルギー移動の解析を行っている。励起エネルギー移動において重要なファクターは、タンパク質中の異なる環境に置かれた各色素の励起エネルギーと色素間の励起子相互作用、そしてそれぞれの揺らぎである。しかし、各色素の励起状態が近接し、また強く相互作用しているため、それらすべてを実験結果から得ることは難しい。我々は EE-MCMM法を応用した新手法を開発し、周囲の環境によって色素の励起エネルギーの揺らぎが異なることを明らかにしている (図4)。現在、励起エネルギー移動の分子論的機構を明らかにするために、さらなる解析に取り組んでいる (投稿準備中)。

おわりに

本稿では、生体分子のシミュレーションに多く用いられている分子力学法、分子動力学法、QM/MM法の概略と近年の発展について紹介した。このような分子シミュレーション手法の多くは、表1に示すようなパッケージソフトウェアに組み込まれており、気軽に使うこ

表1. 主な分子シミュレーションソフトウェア

ソフトウェア名	URL
AMBER	http://ambermd.org/
CHARMM	http://www.charmm.org/
NAMD	http://www.ks.uiuc.edu/Research/namd/
GROMACS	http://www.gromacs.org/
TINKER	http://dasher.wustl.edu/tinker/

とができる。

ここ10年の計算機の進歩と方法論の発達により、10年前までは不可能と思われていたことが分子シミュレーションにより解析可能になりつつある。しかし、QM/MM法における揺らぎの問題は、QM計算のコストが高過ぎるためにまだ気軽に扱える問題ではない。だが、我々も含めて現在も世界で活発に研究が進められており、今後10年で気軽に取り扱えるようになるかもしれない。

文 献

- 1) Cornell, S. J. *et al.*: *J. Am. Chem. Soc.*, **117**, 5179 (1995).
- 2) MacKerell Jr., A. D. *et al.*: *J. Phys. Chem. B*, **102**, 3586 (1998).
- 3) Foldit: <https://fold.it/portal/>
- 4) Khatib, F. *et al.*: *Nat. Struct. Mol. Biol.*, **18**, 1175 (2011).
- 5) Shaw, D. *et al.*: *Commun. ACM*, **51**, 91 (2008).
- 6) Anton RFP: <https://www.psc.edu/index.php/anton-rfp>
- 7) Lindorff-Larsen, K. *et al.*: *Science*, **334**, 517 (2011).
- 8) Jensen, M. Ø. *et al.*: *Science*, **336**, 229 (2012).
- 9) Dror, R. O. *et al.*: *Nature*, **503**, 295 (2013).
- 10) Warshel, A. and Karplus, M.: *J. Am. Chem. Soc.*, **94**, 5612 (1972).
- 11) Warshel, A. and Levitt, M.: *J. Mol. Biol.*, **103**, 227 (1976).
- 12) Higashi, M. *et al.*: *J. Chem. Phys.*, **126**, 144503 (2007).
- 13) Ten-no, S. *et al.*: *Chem. Phys. Lett.*, **214**, 391 (1993).
- 14) Higashi, M. *et al.*: *Chem. Phys. Lett.*, **437**, 293 (2007).
- 15) Kosugi, T. and Hayashi, S.: *J. Am. Chem. Soc.*, **134**, 7045 (2012).
- 16) Higashi, M. *et al.*: *J. Phys. Chem. B*, **117**, 10906 (2014).
- 17) Higashi, M. and Truhlar, D. G.: *J. Chem. Theory Comput.*, **4**, 790 (2008).
- 18) Higashi, M. and Truhlar, D. G.: *J. Chem. Theory Comput.*, **5**, 2925 (2009).
- 19) Higashi, M. and Saito, S.: *J. Phys. Chem. Lett.*, **2**, 2366 (2011).