

PaCS-MD に基づく効率的リガンド結合経路探索法の開発

原田 隆平

筑波大学計算科学研究センター

1. はじめに

2020年度(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「PaCS-MD に基づく効率的リガンド結合経路探索法の開発」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標的タンパク質に対する基質(リガンド)結合経路を抽出する計算手法の開発である。

2. イントロダクション

生体内では、タンパク質をはじめとする生体分子が機能を担うことで、生命を維持している。基質(リガンド)は、タンパク質と特異的に結合する低分子であり、シグナル伝達など多彩な機能を担っている。リガンドとタンパク質の相互作用は、生命の機能解明や創薬ターゲットとして重要な役割を担う。近年、計測技術の進歩により、リガンドとタンパク質の高解像度の複合体構造が決定され、生命機能や創薬ターゲットの解明が急速に進んでいる。

しかしながら、タンパク質とリガンドの相互作用を理解するためには、複合体構造だけではなく、リガンド結合経路の探索が必要である。結合経路を探索するにあたり、分子動力学シミュレーション(MD)は強力な研究手段である。MDは、運動方程式を数値計算することで、タンパク質ダイナミクスを追跡することができ、リガンドとタンパク質との相互作用を詳細に解析することができる。しかしながら、標的タンパク質とリガンド結合は、通常のMDが追跡可能な時間スケールよりも遥かに長時間で観測される「生物学的レアイベント」であるため、抽出が困難である。故に、リガンドと標的タンパク質が結合し、複合体を形成する素過程を抽出する「レアイベントサンプリング手法」の開発が望まれている。

本研究では、リガンドと標的タンパク質の結合過程をレアイベントとして抽出する計算手法を開発する。具体的には、タンパク質の遷移経路をレアイベントとして効率的に抽出する目的で開発した「Parallel Cascade Selection Molecular Dynamics (PaCS-MD)」[1-10]をリガンド結合経路探索に拡張する。拡張にあたり、熱揺らぎの範囲において特定の反応座標方向に沿った結合経路探索を加速するため、リガンドが複数存在する環境下におけるPaCS-MDを実行可能なligand-docking-PaCS-MD(ld-PaCS-MD)を開発する。

一般的にMDを用いた標的タンパク質へのリガンド結合経路探索は、リガンドのポテンシャルを操作する手法が広く利用されている。一方において、タンパク質は溶媒環境で立体構造を変化させるため、リガンドのポテンシャル操作が結合経路探索に与える影響を考慮しなければならない。ここで、PaCS-MDはタンパク質などの構造変化(構造遷移)を効率的に探索可能なレア

イベントサンプリング手法である。図1に示すように、PaCS-MDの具体的な計算手順は、(1)複数の短時間MDから遷移確率の高い分子構造を選択し、(2)選択構造を初期構造として再び短時間MDを実行し繰り返すことで、効率的な構造探索を可能にする。非常に単純な計算スキームであるが、(1)と(2)からなるサイクルを繰り返すだけで、効率的に生体分子の構造遷移を誘起させることが可能である。常に遷移確率の高い分子構造を選択し、選択した分子構造から短時間MDを繰り返すサイクルにより、生体機能に関する重要な構造変化を効率的に引き起こすことが可能になる。PaCS-MDは様々なターゲットに適用され、生体機能に重要な構造遷移の抽出に成功してきた。[11-13] 本研究では、従来のPaCS-MDを拡張し、タンパク質1分子のみならず、リガンドと標的タンパク質が結合し、複合体を形成する素過程を抽出する。具体的には、標的タンパク質に加え、リガンドの自由度も考慮した複合系において適用可能な形式に拡張し、ld-PaCS-MDを開発する。

ld-PaCS-MDでは、標的タンパク質の周囲に複数のリガンドをランダムに配置し、短時間MDが生成したスナップショットの中から、リガンドの重心と結合部位間の距離を計算・ランキング後、最近接のスナップショットから順に上位 n 個を新たな初期構造として選択し、短時間MDをリスタートさせることで、リガンドの結合確率を上昇させる。これら一連のサイクルを繰り返す行うことで、効率的な結合経路探索を実現する。ld-PaCS-MDの独創性は、バイアスの少ない環境においてリガンド結合過程抽出の効率化が可能となる点である。また、PaCS-MDのアルゴリズムをリガンド結合経路探索に特化させ、他のリガンド結合探索手法とは異なるアプローチに基づき、リガンド結合経路の探索を効率化させる。

3. 計算手法の構築

3.1 Parallel Cascade Selection MD (PaCS-MD)

PaCS-MD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」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。これまでに我々が開発してきた計算手法に関しては、レビュー論文 [14,15] を参照にされたい。本手法は、バイアスをかけることなく構造間の遷移経路を探索する。従来法では、構造遷移を促進するため系の温度を高

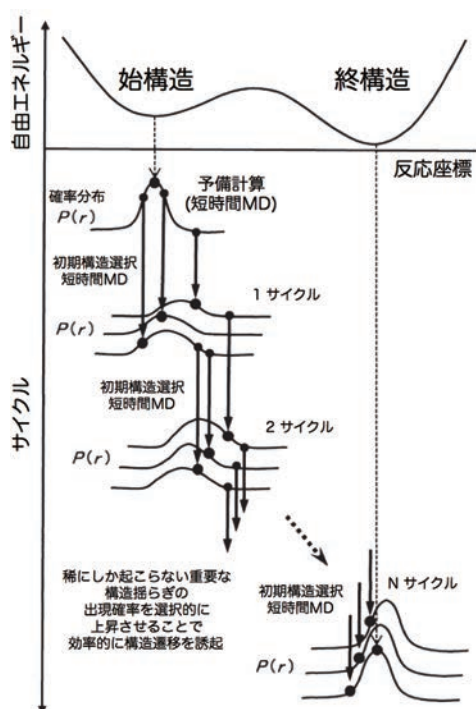


図1: PaCS-MDの概念図

温にして MD 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る [16,17]. 従来法に対して本手法は、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 により構造リサンプリングを繰り返していただくだけで、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 をリスタートさせ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、終構造が分かっている場合に関して、終構造に対する平均自乗距離 (root-mean-square deviation: RMSD) が考えられる. 例えば、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 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き、レアイベントを抽出することができる. ここで、反応座標は任意に指定可能であり、RMSD に限らず指定できる. 実際の適用では、抽出したいレアイベントを的確に記述する合理的な反応座標を指定することが重要になることに注意されたい.

3.2 ld-PaCS-MD によるリガンド結合過程の抽出

ld-PaCS-MD は、リガンドが複数個存在する混雑環境において PaCS-MD を実行し、リガンド結合に伴う複合体形成プロセスの抽出する. 図 2 に示すように、ld-PaCS-MD は活性部位とリガンドの重心の距離を反応座標として、重心間距離が近い MD スナップショットを意図的に選択し、短時間 MD を繰り返す. 具体的には、サイクル毎に活性部位-リガンド重心間距離で MD スナップショットをランキングし、上位構造を複数個選択し、短時間 MD の初期構造とする. 常に重心間距離が小さい MD スナップ

ショットから構造リサンプリングを繰り返す. ここで、複数のリガンドを配置することで、より効率的にリガンド結合プロセスを抽出することができる. また、リガンド濃度があまりにも高くなりすぎると、リガンド同士が凝集してしまう可能性があるため、ある程度低濃度の環境下において ld-PaCS-MD を実行する必要がある.

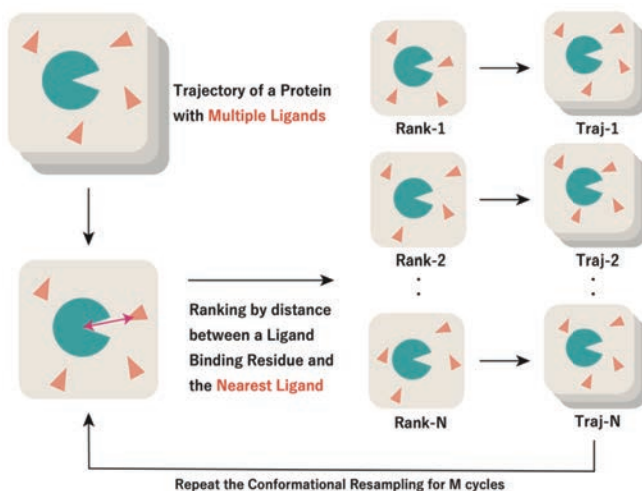


図 2: ld-PaCS-MD によるリガンド結合過程抽出

3.3 Id-PaCS-MD によるリガンド結合過程抽出

Id-PaCS-MD のデモンストレーションとして、加水分解酵素 (T4 リゾチーム) にリガンド (ベンゼン) が結合する過程を抽出した. T4 リゾチームがベンゼンに結合した複合体構造として, X 線結晶構造を参照し, 複合体構造に到達するように Id-PaCS-MD のサイクルを進めた.

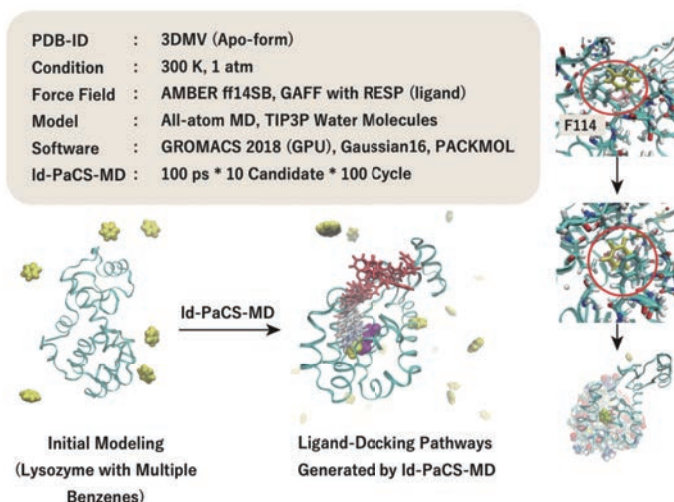


図 3: Id-PaCS-MD による抽出した T4 リゾチームとベンゼンの結合過程

図 3 に示すように, T4 リゾチームの周りに 6 個のベンゼンをランダムに配置し, Id-PaCS-MD を適用し, リガンド結合過程を抽出した. 計算効率の比較として, 通常の MD, 通常の PaCS-MD のリガンド結合過程の抽出プロファイル (活性部位とリガンド重心間距離 vs. サイクル数) を図 4 に示す. 図 4 (右) に示すように, 通常の MD (100 ns) はサイクルに対して活性部位とリガンドの重心距離が収束しておらず, T4 リゾチームに結合する結合過程を抽出できていない. 図 4 (中) に示すように, 通常の PaCS-MD (リガンド 1 個を配置) は 3 回のトライアルのうち 2 回のみプロファイルが収束しており, 1/3 の確率でリガンド結合過程を抽出できている. これらに対し, Id-PaCS-MD (リガンドを 6 個配置) は, トライアル全て (3/3 回) においてプロファイルが収束し, リガンド結合過程を抽出できており, 計算効率の高さを示している.

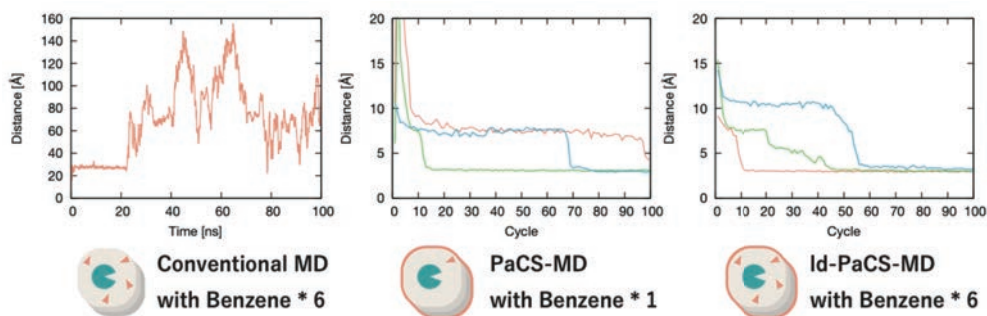


図 4: (左) 通常の MD. (中) PaCS-MD. (右) Id-PaCS-MD のリガンド結合プロファイル

次に、アプリケーションとして SARS-CoV2 メインプロテアーゼ 図 5 (左) のリガンド結合プロセスを Id-PaCS-MD により抽出した。図 5(右) に示すように、4 種類のリガンド (A, B, C, D) について、Id-PaCS-MD のサイクルに対して活性部位とリガンドの重心距離が減少・収束しており、4 つのリガンドに対して 1 つの試行 (リガンド C) を除き、全ての試行でリガンド結合プロセスの抽出に成功していることが分かる。これらのアプリケーションは、Id-PaCS-MD のサンプリング効率の高さ、および実用性を示している。

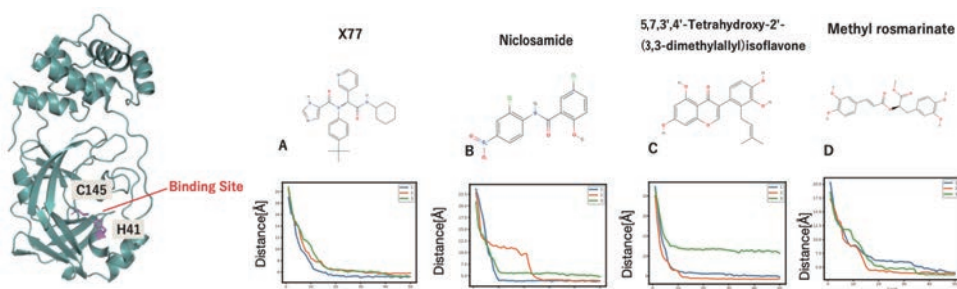


図 5: Id-PaCS-MD により抽出した SARS-CoV2 メインプロテアーゼのリガンド結合過程 (リガンド 4 種類)

4. まとめと今後の展望

Id-PaCS-MD により、標的タンパク質とリガンドの詳細な結合過程が解明されれば、実際に薬剤がタンパク質と結合する素過程を観察することが可能になる。このような結合過程は、実験で観察することが困難であるため、貴重な情報となる。また、結合過程においてどのような相互作用が働き、リガンドと標的タンパク質が結合しているかに関して、時々刻々相互作用の移り変わりが解析できるので、実際の薬剤設計における設計指針を立てやすくなる。このように、Id-PaCS-MD の開発により、創薬研究への応用が期待される。

5. 謝辞

本研究は、東京大学情報基盤センター「若手・女性利用者推薦」2020 年度 (後期) の採択課題として、Reedbush-H を用いて行われた。

参考文献

- [1] R. Harada, A. Kitao, J. Chem. Phys. 139 (2013) 035103-035101-035110.
- [2] R. Harada, Y. Nishihara, N. Wakai, A. Kitao, ICCMSE 2014 1618 (2014) 86-89.
- [3] R. Harada, A. Kitao, J. Chem. Theory Comput. 11 (2015) 5493-5502.
- [4] R. Harada, Y. Shigeta, J. Comput. Chem. 38 (2017) 2671-2674.

- [5] R. Harada, Y. Shigeta, *Bull. Chem. Soc. Jpn.* 90 (2017) 1236-1243.
- [6] R. Harada, A. Shigeta, *J. Chem. Theory Comput.* 14 (2018) 3332-3341.
- [7] R. Harada, V. Sladek, Y. Shigeta, *J. Chem. Theory Comput.* 15 (2019) 5144-5153.
- [8] R. Harada, V. Sladek, Y. Shigeta, *J. Chem. Inf. Model.* 59 (2019) 5198-5206.
- [9] R. Harada, R. Yoshino, H. Nishizawa, Y. Shigeta, *J. Mol. Graph. Model.* 92 (2019) 94-99.
- [10] R. Harada, K. Yamaguchi, Y. Shigeta, *J. Chem. Theory Comput.* 16 (2020) 6716-6725.
- [11] A. Kitao, R. Harada, Y. Nishihara, D.P. Tran, *Aip Conf Proc* 1790 (2016).
- [12] R. Sato, R. Harada, Y. Shigeta, *Biophys. Physicobiol.* 13 (2016) 311-319.
- [13] J. Fujita, R. Harada, Y. Maeda, Y. Saito, E. Mizohata, T. Inoue, Y. Shigeta, H. Matsumura, *J. Struct. Biol.* 198 (2017) 65-73.
- [14] R. Harada, Y. Takano, T. Baba, Y. Shigeta, *Phys. Chem. Chem. Phys.* 17 (2015) 6155-6173.
- [15] R. Harada, *Bull. Chem. Soc. Jpn.* 91 (2018) 1436-1450.
- [16] G.M. Torrie, J.P. Valleau, *J Comput Phys* 23 (1977) 187-199.
- [17] G.M. Torrie, J.P. Valleau, *J. Chem. Phys.* 66 (1977) 1402-1408.