

# 3次元最短経路探索法を取り入れた Supervised Molecular Dynamics による タンパク質ーリガンド結合シミュレーション

2016 年 3 月提出

東京大学新領域創生科学研究科 メディカル情報生命専攻

47-146903 小河原惇

指導教員 広川貴次 客員准教授

キーワード：Supervised Molecular Dynamics、最短経路探索法、  
タンパク質ーリガンド相互作用、創薬支援

## 1 はじめに

近年、創薬支援技術の一つとして、分子動力学シミュレーションによってタンパク質とリガンドの結合シミュレーションに関する研究が盛んに行われている。これは、予めリガンド結合ポケットを決定し、その空間内で配置・配向を計算するドッキング計算とは異なり、溶液中にランダムに配置したリガンドから結合位置を物理化学的手法で探索することで、最終的にタンパク質-リガンド複合体を予測するのみならず、タンパク質-リガンド結合パスウェイを可視化できることが特徴である。

D.E.Shaw 研究所の Shan らは、分子動力学計算専用のスーパーコンピュータを利用して分子動力学シミュレーションを行うことにより、タンパク質 (Src kinase) とリガンド (PP1、Dasatinib) の結合シミュレーションに成功した<sup>[1]</sup>。ただし Shan らの MD (Molecular Dynamics) は、数〜数十  $\mu$ s 分の long-time simulation が必要であり、また、MD 専用機を使って計算をしているということで汎用性に欠ける。

一方最近では、Sabbadin らによって数十 ns でタンパク質とリガンドの結合シミュレーションに成功している例も報告されている<sup>[2]</sup>。Sabbadin らの論文では教師付分子動力学計算 (Supervised Molecular Dynamics、以下 SuMD) という手法を用いている。これは予めリガンド結合位置を与え、その部位に近接するトラジェクトリだけを採択しながらタンパク質ーリガンド結合シミュレーションをする方法であり、標的タンパク質としたリガンドの結合シミュレーションでは、長くとも 100 ns 程度で X 線の結合構造が再現できることがわかっている。しかし、SuMD でのシミュレーションについては、リガンドが近づいてくる方向が想定されている GPCR のような膜タンパク質への適用結果しか報告されておらず、多様な結合パスウェイが想定される水溶性タンパク質に適用可能か検証が必要である。

以上のことから本研究では、水溶性タンパク質にも適用できる拡張型 SuMD によるタンパク質ーリガンド結合シミュレーションを開発することを目的とする。

## 2 方法

本研究で提案する拡張型 SuMD では、リガンドと結合ポケットの距離を、従来の Sabbadin らの方法のように直線的に測りながら SuMD するのではなく、三次元の立体構造を考慮して最短経路を探索し、その経路上を通るように SuMD を実行することで、局所構造に陥る問題を回避する。本研究における拡張型 SuMD の主な手順を図 1 に示す。

本システムでは、大きく 3 つの手続きを行った：(1) MD 環境の準備、(2) 三次元最短経路探索<sup>[3]</sup>、(3) SuMD。また、分子動力学シミュレーションプログラムには GROMACS (ver. 5.0.4)、力場には ff99SB-ILDN (タンパク質) と gaff (リガンド)、溶媒環境には TIP3PBOX を使い、計算は温度 300K、1bar の NPT アンサンブルで行った。

また、ベンチマークセットとして、タンパク質は Checkpoint kinase 1 (CHK1、PDBID: 4FTA)、リガンド



表 1

リガンド	経路ID	辺の本数	リガンドとポケット 重心の最終距離 (Å)	R <sub>edge</sub>	R <sub>pocket</sub>	R <sub>total</sub>
中 活 性 (H5K)	1	10	-	-	-	-
	2	3	-	-	-	-
	3	6	4.91	24 / 64 (0.39)	49 / 235 (0.21)	74 / 299 (0.25)
	4	1	11.13	12 / 24 (0.50)	40 / 381 (0.11)	52 / 405 (0.13)
	5	8	13.46	17 / 34 (0.50)	9 / 82 (0.11)	26 / 116 (0.22)
	6	1	16.20	5 / 14 (0.36)	45 / 274 (0.16)	50 / 288 (0.17)
	7	6	-	-	-	-
	8	6	-	-	-	-
	9	4	13.52	22 / 122 (0.18)	13 / 125 (0.10)	35 / 247 (0.14)
	10	7	-	-	-	-
高 活 性 (HK3)	1	10	-	-	-	-
	2	3	18.68	9 / 23 (0.39)	12 / 82 (0.15)	21 / 105 (0.20)
	3	6	7.73	33 / 134 (0.25)	31 / 165 (0.15)	64 / 299 (0.21)
	4	1	10.93	9 / 24 (0.38)	22 / 184 (0.12)	31 / 208 (0.15)
	5	8	7.08	21 / 55 (0.38)	31 / 159 (0.20)	52 / 214 (0.24)
	6	1	7.09	13 / 44 (0.30)	36 / 242 (0.15)	49 / 286 (0.17)
	7	6	-	-	-	-
	8	6	4.97	14 / 28 (0.50)	58 / 383 (0.15)	72 / 411 (0.18)
	9	4	4.99	13 / 47 (0.28)	27 / 91 (0.30)	40 / 138 (0.29)
	10	7	10.80	24 / 57 (0.42)	15 / 180 (0.08)	39 / 237 (0.17)

※R<sub>i</sub> は「採択ステップ数 / i 段階での総ステップ数 (採択率)」の形式で表示

※**ハイライト**は辺に対する SuMD を完了したシミュレーション、**赤枠**はポケット重心までリガンドが5 Å以内に近接したシミュレーション

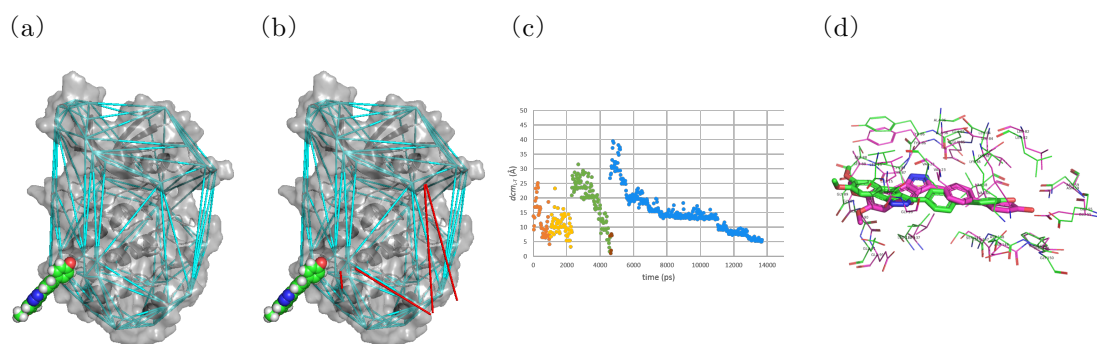


図2 SuMDの実行例。(a) ドロネー三角形分割法によるCα原子を頂点としたタンパク質3次元構造の多面体近似(cyan stick)とリガンド初期位置(sphere)。(b) 最短経路探索結果。red stickは最短経路上の辺をタンパク質表面の立体障害を考慮して、近似多面体の辺よりマージン10Å確保して配置した辺。(c) 目標地点ーリガンド距離( $d_{cm,L-T}$ )とtimeのグラフ。目標地点ごとに色を分けている。(d) 計算によって得られた最終結合構造(megenda stick)とX線構造解析によるリガンド結合構造(green stick)。