創薬支援ソフトウエアmyPrestoを用いた 分子シミュレーション演習

株式会社バイオモデリングリサーチ 中村寛則 nakamura@biomodeling.co.jp





2020/1/28 19:40-21:10



Docking -1: インフルエンザウィルスのノイラミニダーゼとタミフル代謝物活性体
Docking -2: SARS コロナウィルスの3CLプロテアーゼと阻害剤
分子動力学計算(MD): Docking-2のドッキング結果を用いてMD計算を行う

本日は、MolDesk Screeningを使った計算実習を行います。 MolDesk Screeningの内部では、myPrestoを使っています。



myPresto/MolDeskでのインシリコ・スクリーニングは、ドッキング計算を多数のリガンドに対し て大規模に行います。そのため、インシリコ・スクリーニングの計算手順(計算準備)はドッキングと ほとんど同じです。

myPrestoとMolDeskについて



トップ Top	Basic 詳細 MolDesk Basic	Screening 詳細 MolDesk Screening	機能比較 Compare	71	ライセンス License	
solay Color Ontion Event	Simile Screening Pressontion Window Linds Helin				- 0	
	ange anterne repeated whose check hep	MTS010: 4 MTS DOCK001 : 16 Docking	untited : 9 Add 22 😁 🗖 📴 Console 🗖 Doc	king Info 💠 🗖 Screenin	g info	
n: A	Inset from File	\$. ()		A deltaG scor	e RMSD 195 1.04	
3	Find Pocket					

myPrestoは、経済産業省、NEDO及び AMEDからの委託プロジェクトの中で開 発された医薬品開発支援の分子シミュ レーションシステム。無料で利用可能。 現在は次世代天然物化学技術組合が開発。 <u>https://www.mypresto5.jp</u> コマンドで実行するプログラム群。 機能制限付きのGUIソフトmyPresto Portalも利用可能

MolDeskは株式会社情報数理バイオ が開発した商用のGUIソフトウェア。 内部で**myPresto**を実行。 https://www.moldesk.com

myPresto5に含まれるプログラム/ データ



赤字:主要プログラム、青字:準備用プログラム、緑字:化合物ライブラリ

myPrestoの特長と成功例

特長:

- ドッキングシミュレーションと分子動力学シミュレーションの連携が容易
- •200万化合物程度のインシリコ・スクリーニング計算が 容易
- •計算ツール群や化合物データベース等が充実

成功例:

 合計11標的蛋白質に対して200以上の新規活性化合物を 3~50%のヒット率で発見

myPresto5の計算例: ドッキング計算



myPresto5の計算例: MD計算



実際のスライドは動画

myPresto実行方法の選択肢

コマンド実行(無料) オリジナルのmyPresto myPresto Portal (GUI, 無料)

MolDesk Basic (GUI, 有料) 情報数理バイオ社製品 MolDesk Screening (GUI, 有料) 情報数理バイオ社製品

MF myPresto (GUI, 有料) フィアラックス社製品

myPresto PortalとMolDesk Basic/Screeningは、インターフェースはほぼ同じです。 利用可能な機能の種類に違いがあります。

高性能な計算機が無くても、 有料のソフトウェアが無くても、 myPrestoがあれば、 インシリコ・スクリーニングは可能です。

MolDeskがあれば、myPrestoをより効率的に活用することができます。

計算時間は長くなりますが、ノートPCでも計算可能です。 クラウドPCを使えば、一時的に高性能の計算機を利用可能。 myPrestoのプログラムは、ノートPC、計算サーバー、スパコン、クラウドPCで 実行可能です。

ローカルPCと計算サーバーの組み合わせは 作業効率が良い





計算サーバー

計算の準備はGUIが便利。 時間がかかる大規模な計算は 計算サーバーを使った方が 短い時間で計算を完了できます。

分子シミュレーション実行には、

タンパク質分子の準備、 リガンドの準備、

が必須であることが多い。

タンパク質分子の準備

水素原子を含む全原子の座標情報と部分電荷情報が必要

X線結晶構造のタンパク質立体構造情報には、水素原子の座標が含まれていない。 また、実験データにはタンパク質の一部の原子情報が欠損している場合がある。

PDBに記録されていない情報で、シミュレーションに必要な情報は、 別のファイル(トポロジーファイル)に記録されている。 myPrestoも含め、多くの分子動力学計算プログラムでは、トポロジーファイルを使用する。 他の多くのドッキングプログラムはトポロジーファイルを使用しないが、 myPrestoのドッキングプログラムはトポロジーファイルを使用する。

PDB ファイル トポロジー ファイル

GUIでは、トポロジーファイルの存在を意識しなくても計算が できるように設計されている。 GUIで計算準備をして、他の計算機で長い計算をする場合には、 トポロジーファイルも保存して移動させる。

水素原子付加前







myPrestoでは、水素原子付加と同時に、部分電荷情報も付加されます。 側鎖の原子座標が欠損していても、補完されます。 主鎖の原子座標が欠損している場合は、補完されません。 欠損がある場合には、別のプログラムで、モデリングします。

リガンドの準備

水素原子を含む全原子の座標情報と部分電荷情報の必要。 シミュレーション以外の用途で用意された化合物ファイルでは、非極性水素は省略、部分電荷が与えら れていない場合が多い。

基本的には、mol2ファイル形式の化合物ファイルを用いる。 SDファイル (拡張子が.sdfのもの)、MOLファイル(.mol)のものは、mol2ファイルに変換して用いる。



JChemPaintで描画し、 test.molという名前で保存

https://jchempaint.github.io



Test.molをMolDeskで 読み込んだ図



MolDeskで 3D構造にしたもの. 部分電荷構造も付加

MOLフォーマット





myPreseto/MolDeskでは、 MOLフォーマットのファイルは、 そのままではシミュレーションに使えない。

MOL2フォーマット

各原子の座標

@ <tripos>MOLECUL inp0.mol2 19 19 0 0 0 SMALL GASTEIGER</tripos>	E						
@ <tripos>ATOM 1 C1 2 H1 3 H2 4 C2 5 H3 6 C3 7 H4 8 H5 9 C4 10 H6 11 H7 12 C5 13 H8 14 H9 15 C6 16 H10 17 H11 18 01 19 H12</tripos>	-2.5 -2.4 -2.5 -1.3 -1.3 -1.3 -0.4 -1.1 -2.6 -2.7 -2.6 -3.8 -4.8 -3.8 -3.8 -3.8 -4.0 -4.7 -0.1	$\begin{array}{cccccccccccccccccccccccccccccccccccc$	0.114 1.198 -0.350 -0.403 -1.498 0.150 -0.288 1.236 -0.151 -1.234 0.315 0.356 0.057 1.453 -0.1866 -1.271 0.2566 -0.0344 -0.499	C.3 H C.3 H C.3 H C.3 H C.3 H C.3 H C.3 H C.3 H C.3 H C.3 H	1 LGD 1 LGD	-0.0269 0.0290 0.0553 0.0593 -0.0269 0.0290 0.0290 0.0290 -0.0507 0.0266 0.0266 -0.0529 0.0265 -0.0529 0.0265 -0.0507 0.0266 0.0266 -0.3918 0.2098	
@ <tripos>BOND 1 1 2 4 3 6 4 9 5 12 6 1 7 4 8 1 9 1 10 4 11 6 12 6 13 9 14 9 15 12 16 12 17 15 18 15 19 18</tripos>	4 1 6 1 9 1 12 1 15 1 15 1 15 1 18 1 2 1 3 1 5 1 7 1 8 1 10 1 11 1 13 1 14 1 16 1 17 1 19 1				各原子の 電荷情報	部分	
× T	法合情	雪 報					

SARSについて

重症急性呼吸器症候群(Severe acute respiratory syndrome; SARS) は、SARSコロナウイルス (SARS coronavirus; SARS-CoV) によって引き起 こされるウイルス性の呼吸器感染症。

2002年11月から2003年7月の期間に、 中華人民共和国南部を中心に流行 8,096人が感染、774人が死亡。

SARSコロナウィルスの3CLプロテアーゼ阻害剤の報告例

2002年11月から2003年7月 流行期間

2003年11月に発表された論文 最初の複合体構造、リガンドは基質アナログの阻害剤(5残基のペプチドを修飾したもの) PDB ID: 1UK4(登録日2003-08-14, 公開日2003-11-18) PNAS 2003, 100, 23,13190–13195 ASN-SER-THR-LEU-GLN

2006年に発表された論文 Virtual Screening with GOLD IC50: 0.3 µ M, 3µM (PDB ID: 2GZ7,2GZ8) J. Med. Chem. 2006, 49, 5154-5161





2007年に発表された論文 IC50: 50nM J. Med. Chem. 2007, 50, 1850-1864



IC50 = 50 nM

3CLプロテアーゼをターゲットとしたインシリコ・スクリーニングの例



J. Med. Chem. 2006, 49, 5154-5161

SRAS Coronavirus 3CL protease をターゲットとした活性化合物の情報は ChEMBLに登録されています。

https://www.ebi.ac.uk/chembl/



IC50のデータは133件が登録されています(2020/1/28現在)。

MolDeskの 画面の 説明

