

GPGPU を用いた薬物親和性評価プログラムの開発

1. 背景

医薬品の開発は大きく分けて、疾患ターゲット(原因タンパク質)の探索、そのターゲットに対して作用する化合物(医薬品となりうる化合物)の探索、臨床試験による評価の 3 段階に分けることができる。現在、疾患ターゲットに対して作用する化合物探索(医薬品となりうる化合物の探索・最適化)において、疾患ターゲットと候補化合物の親和性をコンピュータ上で仮想実験を行い評価する *in silico* 創薬を利用する動きが製薬会社やその他の研究機関で高まっている。*in silico* 創薬では実際に数百万種にも及ぶ化合物や試薬・実験器具を用意する必要がないことから、費用・環境面でのメリットや、また実在しない未知の化合物を設計し評価できるといったメリットがある。またコンピュータを用いることで 24 時間 365 日の医薬品候補化合物探索が容易になり、探索時間の短縮も大いに期待できる。

しかし *in silico* 創薬に用いられる既存の候補化合物探索ソフト(以下スクリーニングソフトウェアと呼ぶ)にはいくつかの問題点がある。それは計算時間が長いことと、計算で得られた親和性と実験で得られる結合親和性の相関が十分でないことである。既存のソフトウェアを用いると、一つの化合物を評価するのに 10~20 分程度の時間を要する。探索する化合物群の数は数万~数百万種にもなり、計算時間の短縮無しに実用的に使うことは難しい。多くの研究機関ではその問題に対して、並列クラスターを用意することで対応している。しかし並列クラスターが無い環境、または並列クラスターが用意できたとしても、並列で起動するにはライセンスがその数だけ必要になるといったことから、高速化は容易ではないと言える。

GPGPU とは General Purpose Graphic Processing Unit の略で、GPU(グラフィックボードが搭載している画像処理専用の演算装置)を一般的な演算処理にも用いることができる演算装置である。GPGPU は nVidia 社から GPGPU 専用プログラミング言語 CUDA がリリースされたことで、科学技術計算の分野でも急速に応用されてきており、今後の動向が注目されている技術の一つである。GPGPU は並列処理に特化した演算装置をもち、適したプログラム開発を行うことで CPU の何倍もの処理速度を手に入れることができる。

2. 目的

本プロジェクトでは GPGPU を用いて、薬物親和性評価の予測を高速かつ高精度に行うことを目的とする。スクリーニングソフトウェアは性質上、並列計算で高速化することが容易であるため、GPGPU を用いて高速化を図ることが出来る。また GPGPU の導入コストは並列クラスタや既存ソフトウェアのライセンス料よりはるかに安価となっている。GPGPU を用いることで個人向けの PC でもスクリーニングを行うことが可能になり、あらゆる研究グループで日常的に創薬を行うことが可能になると期待される。

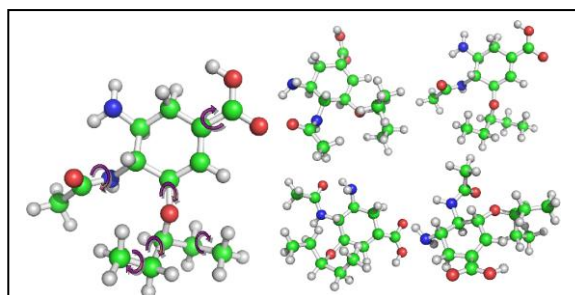


図 1: 左図は Osertamivir(タミフル)分子とその回転可能結合の例を矢印で示す。回転可能結合が回転することによって右図のような様々なポーズを取り得る。

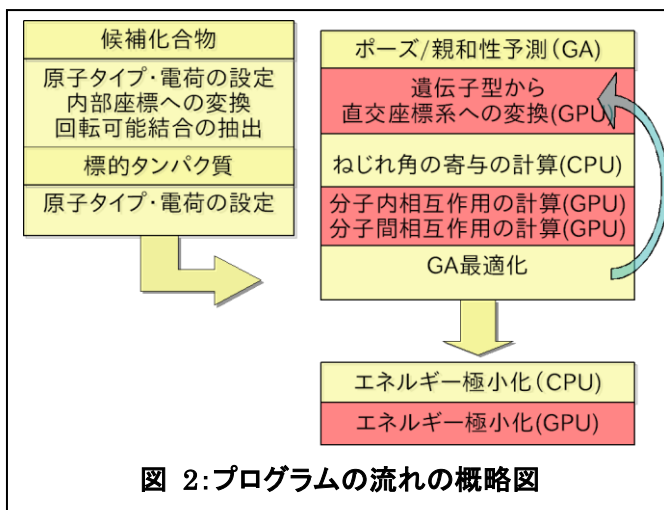


図 2: プログラムの流れの概略図

3. 開発の内容

nVidia から提供されている GPGPU 専用のプログラミング言語 CUDA を用いて、薬物親和性評価プログラムを開発する。薬物の親和性評価とは化合物とタンパク質の三次元データを用いて「どのようなポーズで」「どれほどの強さで」結合するかを予測することである。通常、化合物は回転可能結合を持ち、1 つの化合物でも様々なポーズをとることが出来る(図 1)。プログラムはまず化合物がどのような自由度を持つのか判定する。次に疾患ターゲットに対して様々なポーズを生成し、それらのポーズに対して親和性を評価する。親和性は分子力学(Molecular Mechanics: MM)を用いてエネルギー変化として計算する。MM では力場(Force Field)と呼ばれる原子の種類(原子タイプ)に依るパラメーターセットを用いてエネルギーの算出を行う。本プロジェクトでは精度の向上を目指して、全原子の MM 計算を行い親和性の予測を行う。全体的なプログラムの流れを図 2 に示す。

結合ポーズは、分子の位置や回転可能結合などを一定の刻み幅で変化させることで生成する。しかしその組み合わせは膨大となり網羅的に探索するのは現実的ではない。本プロジェクトでは遺伝的アルゴリズムを用いて最適なポーズを探索する。

またポーズは一定の刻み幅で探索されるため、最終的には生成された構造を、力(エネルギーの微分)の方向に動かし微調整する必要がある。これをエネルギー極小化という。エネルギー極小化計算も GPGPU を用いて高速化を図る。

CPU と GPGPU を用いた場合のプログラム全体の実行速度の比較を表 3 に示す。CPU は Intel Core2 Quad 3.0GHz のうち、1コアを用いた。GPGPU は nVidia Geforce GTX 295 を用いた。また同グラフィックカードは1枚のボード上に2つの GPGPU を搭載しているので、CPU2 コアと GPGPU2 コアを用いたときの実行速度も示した。表 3 より GPGPU を追加することで 80 倍程度の高速化を達成することができた。GPGPU1コアと GPGPU2 コアの速度比較も約 2 倍となり、これは GPGPU のコア数を増やしていくことで更に高速化が見込めることを示している。

表 1 実行速度比較

	CPU	GPGPU	GPGPU*2
実行速度	27.6 min	20.7 sec (80.0 倍)	10.8sec (153.3 倍)

次に本プログラムによって導き出された結合ポーズと実験で得られた化合物との比較を表 2 に示す。タンパク質と化合物のセットは一般的に知られている 6 セットを用いた。再現性は計算と実験で求めたポーズのずれを RMSD (平均二乗偏差) で示した。RMSD は 2 以下のとき、構造がかなり再現できていると言え、4 以下であればそれなりに良い結果であると言える。それ以上のときはあまり再現できているとは言えない。表 2 より、6 セット中 5 セットで良い再現性が得られた。

表 2 ポーズの再現性

タンパク質名	Acetylcholine Esterase	CDK2	Cyclo-oxygenase-1	Cycro-oxygenase-2	HIV-1 Protease	HIV-1 Reverse Transcriptase
RMSD	2.07	3.48	1.72	1.33	7.84	1.39

4. 従来の技術(または機能)との相違

本プログラムは GPGPU をスクリーニングソフトウェアに応用した史上初のプログラムで、従来のソフトウェアよりも短時間で化合物の評価を行うことができる。また GPGPU は低コストで導入ができるため、今まで PC クラスタがない環境では行えなかった計算が、家庭用 PC でも実用レベルで行えるようになった。

現時点では、既存のソフトウェアと同様にタンパク質側の自由度を無視した計算を行っている。タンパク質側の自由度を考慮することは計算時間が数倍に増えることを意味するが、GPGPU を用いることで追加の計算時間を最小限に抑えることが出来ると予想され、高精度の予測ができるようになる。

5. 期待される効果

低コスト・短時間で高精度な薬物親和性評価を行うことができるため、あらゆる研究グループで in silico 創薬を行うことができるようになった。in silico 創薬が身近になることで創薬研究は効率化され、一つの医薬品に十数年かかると言われている開発期間が短縮する方向に働くことは間違いない。また創薬だけに関わらず、結合ポーズを提示することで化合物の合成戦略に有用な知見を与えることができる。

6. 普及(または活用)の見通し

今後、本研究室の Web サーバーにソフトウェアを無償公開する予定である(URL:
<http://bukka.p.chiba-u.jp/>)。またこのソフトウェアを用いた創薬研究を行い、論文の執筆や学会発表を繰り返していくことで、ソフトウェアの有用性を検証し、普及させることに努める。

7. 開発者名(所属)

山岸 純也(千葉大学 医学薬学府 総合薬品科学専攻 薬品物理化学研究室)

参考:千葉大学 医学薬学研究院 薬品物理化学研究室 <http://bukka.p.chiba-u.jp/>