

TSUBAME 共同利用 令和7年度 産業利用 成果報告書

利用課題名 IT 創薬向け並列計算化学シミュレーション手法の開発

英文: Development of parallel computational chemistry simulation methods for IT-based drug discover

利用課題責任者

秋山泰

Yutaka Akiyama

所属

アヘッド・バイオコンピューティング株式会社

Ahead Biocomputing, Co. Ltd.

<https://ahead-biocomputing.co.jp>

本課題では、低分子創薬における大規模 Virtual Screening (VS) の高速化を目的として、疑似量子アニーラを用いた Fragment-based VS (FBVS) ワークフローの実行可能性を検証した。商用化合物ライブラリを使用し、フラグメントドッキング、QUBO 行列の作成、疑似量子アニーラによる組合せ最適化、化合物構造の復元、ドッキングの評価までを一連の処理として確認した。

This project verified the feasibility of a fragment-based virtual screening (FBVS) workflow using a pseudo-quantum annealer for large-scale screening in IT-based drug discovery. The workflow covers fragment docking, QUBO matrix construction, combinatorial optimization, reconstruction of compound structures from annealing results, and subsequent docking evaluation. The study demonstrated that commercial compound libraries can be processed by the FBVS protocol.

Keywords: Virtual screening / Protein-ligand docking / Fragment-based / QUBO / Quasi-quantum annealer

背景と目的

創薬における化合物探索では、膨大な化学空間の中から実験候補を効率良く選抜することが重要である。計算機を用いた Virtual Screening (VS) は実験前の候補化合物の絞り込みに広く用いられており、標的タンパク質構造を利用する Structure-based VS (SBVS) ではドッキング計算が中心的な手法として利用される。一方で、商用ライブラリ全体のような大規模データを対象に高精度ドッキングを網羅的に実施するには、計算時間と計算資源が大きな制約となる。

このような制約から、ドッキング計算の前段階で候補を減らす事前スクリーニングが重要となる。本課題では、化合物をフラグメントに分割し、フラグメント単位のドッキング結果を組み合わせで候補構造を探索する Fragment-based VS (FBVS) に着目した。特に、タンパク質リガンドドッキングを Quadratic Unconstrained Binary Optimization (QUBO) 問題として表現し、疑似量子アニーラにより組合せ最適化を行う手法^{1, 2}を、大規模化合物ライブラリに適用する

ための計算ワークフローとして検証した。本課題では、疑似量子アニーラを用いた FBVS を実行し、必要リソースを見積もることを目的として研究を実施した。

概要

本課題では、疑似量子アニーラに入力する QUBO 問題を作成する前処理、疑似量子アニーラへのジョブ投入、得られた計算結果から化合物構造を復元する後処理、および復元構造に対するドッキング評価を確認した。各処理は Apptainer コンテナ上で実行し、実行環境の再現性を確保した。

各工程ではデータ規模に応じて CPU 数、メモリ量、入出力ファイルサイズが大きく異なるため、TSUBAME 4.0 で利用可能な資源タイプ(表 1)を適切に選択し、TSUBAME 4.0 の通常のジョブの実行時間上限である 24 時間以内に完了するジョブとなるように設計した。実計算では百万件オーダーの商用化合物ライブラリを対象とし、フラグメントドッキングから化合物復元、ドッキングまでの接続性を確認した。

表 1. 東京科学大学 TSUBAME 4.0 における CPU キューの CPU コア数及びメモリ容量の一覧

Resource type	Physical CPU cores	Memory(GB)
cpu_4	4	9.2
cpu_8	8	18.4
cpu_16	16	36.8
cpu_40	40	92
cpu_80	80	184
cpu_160	160	368

結果および考察

はじめに、計算に必要となるリソースを確認するため、各処理に対して速度測定及び消費リソースの測定を実施した。各工程の速度および消費リソースを測定し、工程ごとに適切な資源タイプと並列数を決定した。TSUBAME 4.0 では資源タイプごとに使用可能な物理 CPU コア数、メモリ容量、ローカルスラッチ領域が定義されているため、単純並列性の高い工程では CPU 数を増やす一方、メモリ使用量が増えやすい工程では 1 ジョブあたりのデータ件数を抑える設計とした。特に、化合物構造を復元する処理の前処理でメモリ使用量が大きくなりやすく、ジョブ分割と資源タイプ選定が計算効率に大きく影響することが分かった。

実際に行った計算は、(1) REstretto を用いたフラグメントドッキングによる事前計算、(2) フラグメント配置および相互作用に基づく QUBO 行列の作成、(3) SQBM+を用いた組合せ最適化、(4) 最適化結果からの化合物構造復元、(5) 復元構造に対する AutoDock Vina を用いたドッキング計算である。図 1 に、本課題の計算プロトコルを示す。

本プロトコルでは、全化合物を高精度に直接ドッキングするのではなく、フラグメント単位の事前計算と QUBO 最適化により候補組合せを選抜する。そのため、後段のドッキング計算では、疑似量子アニーラにより選抜・復元された化合物構造を対象として評価を行うことができる。今回の実計算により、商用化合物ライブラリに対してワークフローを適用し、復元した候補構造を通常のドッキング評価へ接続できることを確認した。

一方で、構造復元前処理や QUBO 行列作成では、入力データの増加に伴ってメモリ使用量およびファイル

I/O が増大する。今後、大規模データに対して安定的に運用するには、ジョブ分割単位、入出力ファイル形式、失敗時の再実行単位をさらに整理する必要がある。また、SQBM+のような外部サービスを利用する場合は API 接続や投入制限も全体スループットに影響するため、ローカル実行可能な最適化手法や QUBO 以外の表現も比較検討する必要がある。

疑似量子アニーラを用いた
Fragment-based Virtual Screening 計算プロトコル



図 1. 疑似量子アニーラを用いた Fragment-based Virtual Screening 計算プロトコル

まとめ、今後の課題

本課題では、疑似量子アニーラを用いた FBVS ワークフローを TSUBAME4.0 上で構築し、商用化合物ライブラリに対する大規模スクリーニング処理が実行可能であることを確認した。特に、フラグメントドッキング、QUBO 行列作成、SQBM+による組合せ最適化、構造復元、AutoDock Vina による評価を接続したことで、FBVS を実用的なスクリーニングプロトコルとして運用するための計算上の課題を整理できた。

1. 外部 API に依存しない、または依存度を下げた組合せ最適化実行方法の検討
2. 選抜化合物のドッキングスコア、既知活性化合物、実験データを用いた性能評価
3. 各処理のメモリ使用量削減と I/O 最適化
4. QUBO 以外の表現方法の検討
5. ワークフロー全体の自動化、再実行性、ログ管理の強化

謝辞

本研究は、NEDO 量子・古典ハイブリッド技術のサイバー・フィジカル開発事業の支援を受けて実施された。

参考文献

1. Yanagisawa, K.; Fujie, T.; Takabatake, K.; Akiyama, Y. QUBO Problem Formulation of Fragment-Based Protein–Ligand Flexible Docking. *Entropy* 2024, 26 (5), 397.
2. 柳澤 溪甫；藤江 拓哉；高畠 和輝；秋山 泰；Frasco-VS: フラグメントに基づく薬剤候補化合物選抜の量子アニーリングによる実現. 研究報告 量子ソフトウェア (QS) 2026, 2026-QS-17 (5), 1–8.