In Silico Screening for Potential Inhibitors of SARS-CoV-2 Nsp13 Helicase

Arnošt Polák, Jan Faflík, Radoslav Jochman, Sofya Safiulina and Adam Kuniak

Charles University, Prague

Introduction

This project addresses the unprecedented challenge posed by the global COVID-19 pandemic caused by SARS-CoV-2 focusing on the identification of potential inhibitors for the virus's replication process through *in-silico* high throughput screening. The primary focus of the projects is to target Nsp13 helicase, an enzyme playing a key role in the replication process, using three distinct virtual screening methods—AutoDock Vina, DiffDock, and NRGDock—utilizing a diverse set of ligands. The screened dataset comprised both 11,586 FDA-approved drugs from the DrugBank database and 48.2 million compounds from Enamine REAL Diversity set. By repurposing already existing drugs, known for their clinical safety, the objective is to enhance the efficiency of the drug discovery process. The investigation includes a detailed analysis of four Nsp13 conformations and their respective binding sites. The findings provide insights into potential drug candidates for inhibiting the virus.

Method

Our approach focuses on drug repurposing using the DrugBank database, which consists of 11,586 FDA-approved drugs. Drug repurposing involves investigating compounds previously tested for specific clinical indications as potential candidates for treating other medical conditions. To identify optimal ligands for specific targets, we employed an *in silico* approach, conducting virtual screening through three distinct molecular docking methods—AutoDock Vina, DiffDock, and NRGDock.

When selecting protein structures for screening, we prioritized the newest and highest-resolution structures based on literature discussing Nsp13 helicase conformations, The engaged conformation, capable of binding downstream t-RNA, was our primary focus due to its functional significance. Using PrankWeb, we identified binding sites with substantial similarities among all conformations.

The research pipeline unfolds in two pathways. In the first, we pre-screened the whole DrugBank database using AutoDock Vina. The best results were then redocked through DiffDock on the 7RDY structure. The second pathway employs NRGDock on two binding sites of 7RDY, followed by re-screening top molecules with AutoDock Vina to facilitate a comprehensive comparison between the two methods.

Results

Our examination produced multiple result sets and findings, revealing insights into potential inhibitors targeting the SARS-CoV-2 Nsp13 helicase. Firstly, the top 100 hits from the AutoDock Vina screening on the 7RDY(engaged) conformation revealed molecules with notable binding affinities. The top 10 hits, comprising drugs from various classes like antibiotics, anti-inflammatory, and anticancer agents. They share promising characteristics, including high molecular weights, neutral physiological charges, and multiple rings in their structures.

In order to enhance our understanding of the binding features, we reexamined the top 100 molecules identified in the AutoDock Vina screen for 7RDY by subjecting them to redocking on three additional conformations of Nsp13 - 7RDX (1B-open), 7RDZ (apo) and 7RE0 (swiveled). Although a slight trend appeared, suggesting that the 7RDX had stronger binding affinities, the overall ranking of ligands displayed not-insignificant variability across various conformations.

Additionally, DiffDock, an AI-based molecular docking method, was employed to redock the top 100 molecules on the 7RDY structure. Despite substantial challenges in generating, DiffDock provided confidence scores for 55 ligands. Notably, no correlation was observed between AutoDock Vina and DiffDock scoring, possibly caused by the distinct scoring methodologies employed by these docking methods.

Lastly, we incorporated NRGDock, a high-throughput screening method, into our study to analyze the Enamine REAL Diversity set on the 7RDY(engaged) structure. The comparative analysis with AutoDock Vina results revealed minimal correlation, possibly caused by the different scoring approach of NRGDock. Utilizing MACCS fingerprints, we identified 8 ligands from the Autodock Vina screen and 11 ligands from the NRGDock screen that displayed significant structural similarity. Additionally, CHEESE Embeddings were employed to forecast the physicochemical and ADMET properties of the 11 ligands from the NRGDock screen, indicating potential challenges for these molecules to function as drugs.

Discussion

During our research, we encountered some challenges that are worth mentioning and should be considered in future studies. Notably, Diffdock and NRGDock stand out. DiffDock showed us great potential in the future due its unique approach and features. However, challenges may arise, particularly with larger molecules, the need for increased computing power and of course the dissimilarities in results compared to AutoDock Vina.

By utilizing NRGDock in conjunction with AutoDock Vina, we were able to successfully screen a large amount of diverse compounds that could potentially yield valuable results. The differences between results of NRGDock and AutoDock Vina come from different scoring algorithms and approaches, however, it’s very important to note that neither method is inherently incorrect.

We implemented a 2-minute cutoff on AutoDock Vina runs, resulting in skipped molecules, potentially due to issues such as large size or conversion errors. About 3% of the DrugBank database was affected in the initial docking phase, with an additional 30% during redocking. Another 5% molecule loss was caused by AutoDock Vina lacking configurations for specific atoms, emphasizing the importance of fitting configurations and adopting a more detailed strategy in setting up binding pockets to enhance the precision and efficiency of the docking process.

Interestingly by analyzing both the DrugBank database and the Enamine REAL Diversity Set we discovered that the majority of potentially inhibiting molecules did not abide by Lipinski's rule of five. This observation led us to the conclusion that the rule may not be as reliable as an indicator of molecule suitability for drug use as commonly thought.

Altogether we can say that our research was successful in finding new molecules that can be used in drug development against Covid-19. Moreover, our demonstration of the lack of correlation between the three aforementioned docking methods underscores the importance of employing a docking method with a strong theoretical background, meticulous method selection, and statistically robust results, recognizing that each method serves a distinct purpose and research strategies should align accordingly.