SUCCESSFUL APPLICATION OF COMPUTING METHODS TO DEVELOPMENT SARS-COV-2 INHIBITORS

Vladimir Sulimov

DIM ψ NTA

The work is financially supported by the Russian Science Foundation, Agreement no. 21-71-20031

Since the COVID-19 outbreak, an avalanche of publications on anti-SARS-CoV-2 inhibitors has begun

- Our team has been involved in this work since 2020:
  - A.V. Sulimov, et al. Supercomput. Front. Innov. **7** (2020), pp. 41–56
  - A. Sulimov, et al. Molecules **27** (2022), pp. 5732
- By the end of 2024, the total number of publications will be about a thousand
- We have prepared a review: A.V. Sulimov et al. Docking and other computing tools in drug design against SARS-CoV-2, SAR and QSAR in Environmental Research, **35** (2024) 2, 91-136
- Here I will present the main findings of this review.

Structure-Based Computer-Aided drug design is widely used in inhibitors development

- Docking is the main modelling tool of hit discovery
- The accuracy of docking programs should be increased to raise efficiency of the early stage drug development
- Supercomputer resources have become increasingly used for docking.
- To increase the efficiency of docking, it should be supplemented by post-processing of the top scored compounds

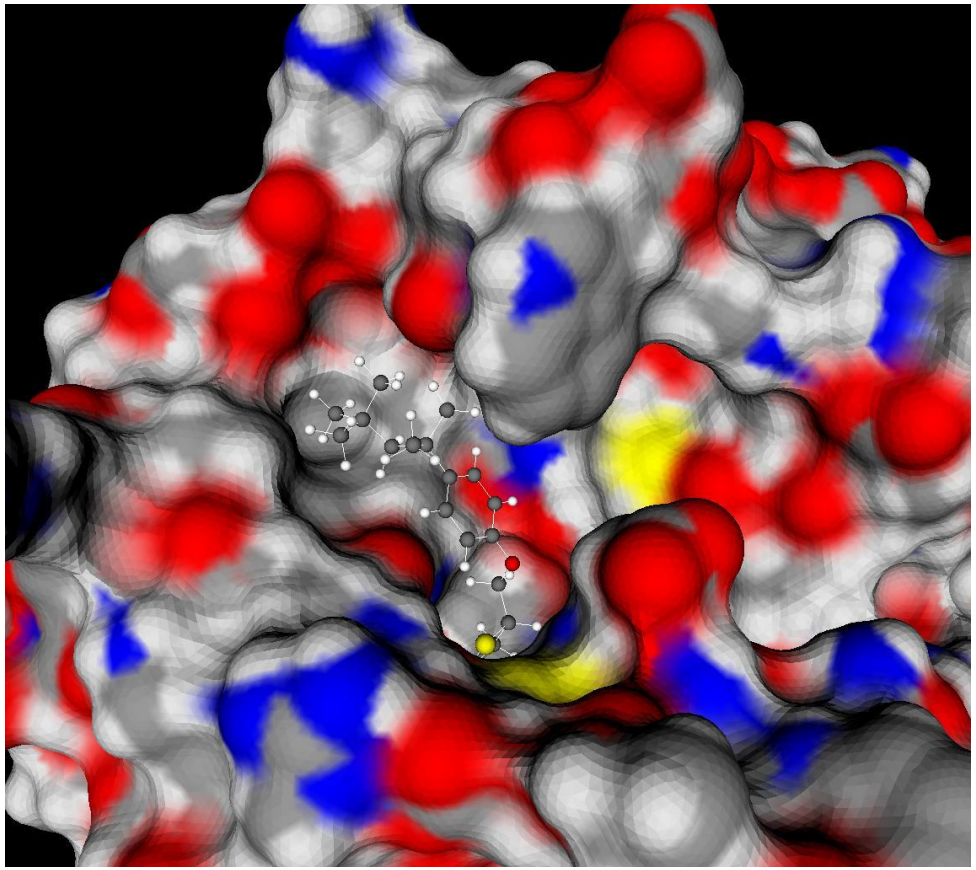
Docking is a software used for the drug development at the initial stage

Docking:

- Ligand positioning in the target protein
- Estimation of the protein-ligand binding energy ΔG_{bind} – Scoring function, Soring, Score
- Positioning accuracy – ***satisfactory***
- Accuracy of the calculations of the protein-ligand binding energy ΔG_{bind} – ***bad***

Docking accuracy → **Drug discovery efficiency**

Docking paradigm: the ligand binds in the active site of the target protein in close proximity of the global energy minimum of the protein-ligand complex



Docking is the search for the global minimum of the energy of the protein-ligand complex

Reviews:

- **Sulimov V.B., et al. // Curr. Top. Med. Chem., 2021**
- **Sulimov V.B., et al. // Curr. Med. Chem., 2019**
- **Sulimov A.V. , et al. // Supercomput. Front. Innov., 2019**

Therapeutic target proteins of SARS-CoV-2

1. Main protease: M^{pro} or $3CL^{pro}$
2. Papain-like protease: PL^{pro}
3. Viral non-structural protein 14: nsp14
4. Nonstructural uridylate-specific endoribonuclease: NendoU or nsp15
5. 2’-O-methyltransferase: nsp16
6. Viral helicase: nsp13
7. RNA-dependent RNA polymerase: RdRp
8. Spike (S) protein

After docking – post-processing

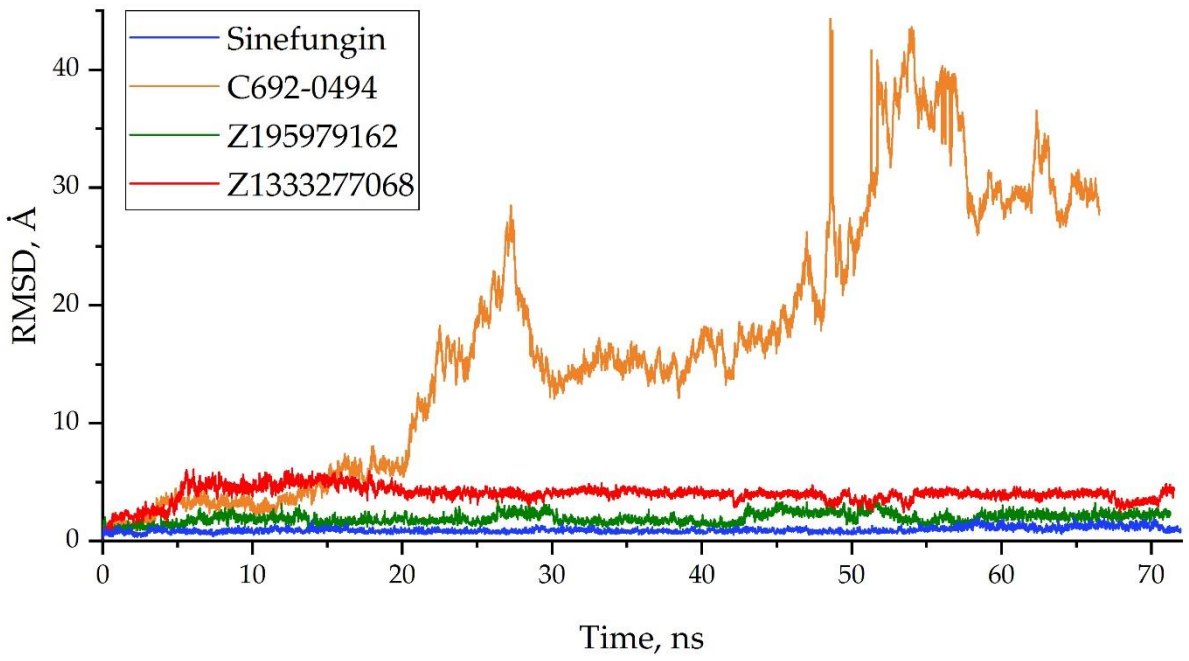
- To increase the efficiency of docking, it is supplemented by post-processing of the top scored compounds
- Post-processing:
  - Quantum-chemical calculations of the protein-ligand binding enthalpy – **using docked posed of ligands**
  - Molecular Dynamics: the study of kinetics of protein-ligand system from **docked ligand positions** along long MD trajectories $\approx 100\ ns$
  - MM/GBSA (MM/PBSA): Molecular Mechanics (a force field)+Solvent
  - Free Energy perturbation(**FEP**) or Thermodynamics Integration (**TI**) MD methods **were rarely used in post-processing**
  - Using programs for clustering chemically similar compounds to increase the chemical diversity of candidates for experimental testing
  - Counting the number of hydrogen bonds between the ligand and the protein, analyzing other interactions, pharmacophore models, *etc*.

Quantum-chemical post-processing

- Binding enthalpy calculations by the quantum-chemical **PM7** method with the **COSMO** continual solvent model using the docked ligand pose
- $\Delta H = H_{complex} - H_{protein} - H_{ligand}$, $\mathrm{H_X}$ – the heat of formation in the **MOPAC** program
- Number of top scored ligands being processed $\sim 10^2 - 10^3$
- PM7 is a semiempirical method describing well dispersion and hydrogen bonds interactions

Ligand stability in the docked pose

- Several top scored ligands
- Molecular dynamics simulation with explicit water model along long trajectories ~ 100 ns RMSD ≤ 5 Å along the trajectory



Z…62 K_d = 2 μM

Z…68 K_d = 5 μM

Sinefungin IC_{50} = 0.86 μM

Kuojun Zhang, Alexey V. Sulimov, Ivan S. Ilin, et al. Development of Novel Nsp16 Inhibitors …, Supercomputing Frontiers and Innovations, 2024, **11**, 51 - 66

C692-0494 is inactive compound

Virtual screening of databases of ligands

- Virtual screening of databases of existing compounds (on-the-shelf compounds) by **DOCKING** helps to discovery hits:
  - Selecting best candidates, ordering compounds, experimental testing the inhibitory activity, discovery hits
  - Avoid synthesis of new compounds at the initial stage of drug development; to save time needed for synthesis
- Virtual screening of different databases containing from several dozen to billions of compounds: our review contains references to 16 databases used for screening
- The protein structure-based methodology was preferred over ligand-based methods: the former produced a greater number of experimentally confirmed active molecules
- Many studies combined strengths of both approaches: the protein structure-based and ligand-based

Virtual screening

Creating a protein model

Creating 3D structures of chemical compounds

SOL program docking $\sim 10^4$ - 10^6 ligands

Postprocessing: Quantum Chemistry, Molecular Dynamics

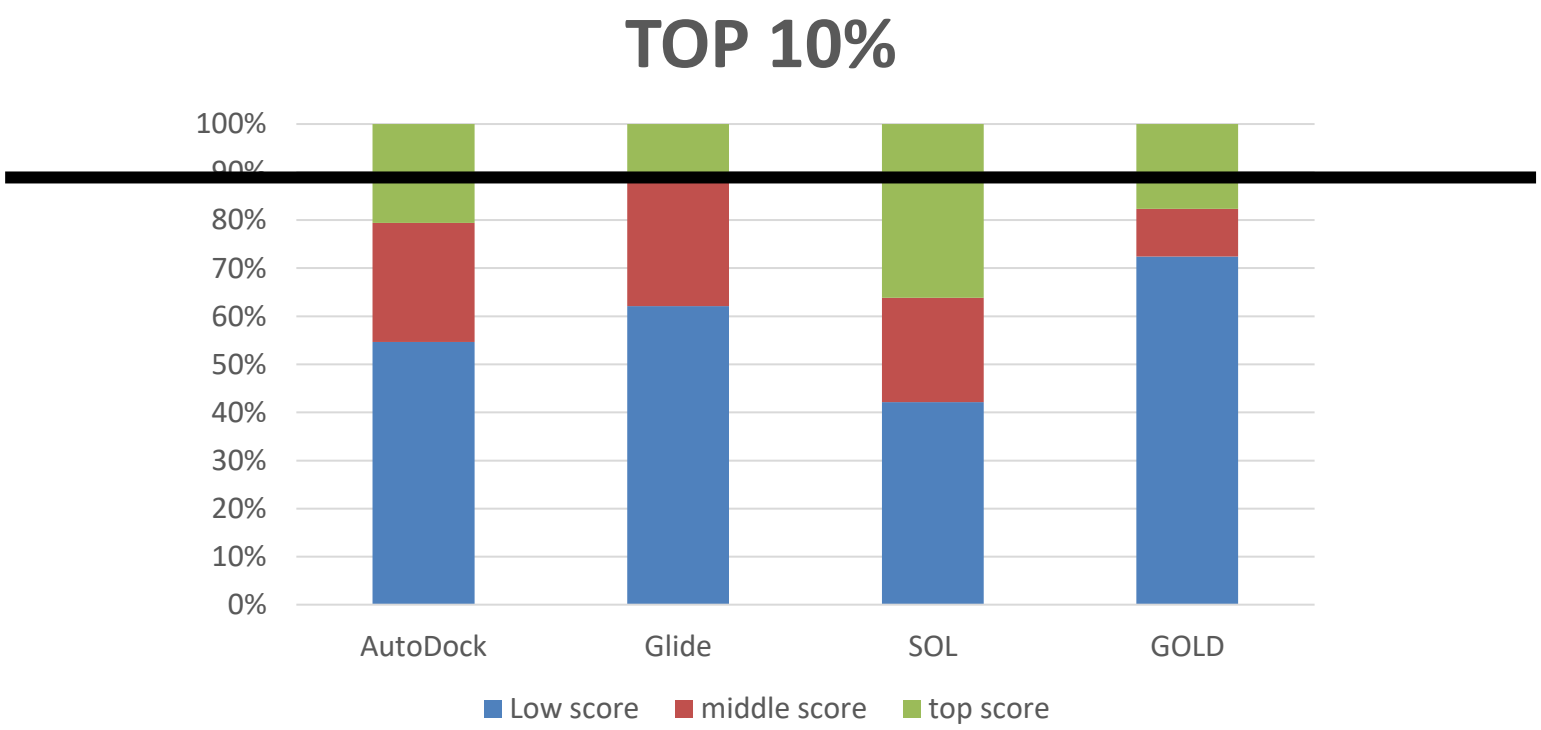
Binding enthalpy calculations by the quantum-chemical PM7 semi-empirical method with the COSMO solvent model

$\Delta H = H_{complex} - H_{protein} - H_{ligand}$: $\sim 10^2 - 10^3$

Experimental *in vitro* verification (≈20 compounds)

Consensus docking

- Docking a given library of ligands into a target-protein using several different docking programs
- Selection of only those ligands which are in the top 10% of the results of most docking programs



Virtual screening of databases of ligands

- The database content greatly influences the success of inhibitor discovery
- Screening of very large databases has not yet led to high hit detection efficiency
- Grug Repurposing: Virtual screening of the libraries of existing and approved drugs, drug candidates being at Phase III clinical trials. Being confirmed in experiments the approval is faster.
- For screening of very large databases (up to tens of billions of molecules), neural networks and other machine learning tools helps to select compounds with supposedly good docking scores and avoid docking of ligands with obviously bad scores; pharmacophore-based screening before docking.

Virtual screening of databases of ligands

- The VirtualFlow platform enables docking without any AI over a billion of ligands: massive parallelization, leveraging hundreds of thousands of computing cores via Google Clouds.
- Fast docking of large ligand databases: AutoDock Vina, AutoDock GPU and Quick Vina
- For more accurate docking: LibDock, FlexX, Dock 3.7, SOL, ICM, MOE, FRED CDOCKER, and AutoDock Flexible Receptory; Glide and AutoDock 4.2 were the most popular
- The larger the size of the database, the more important it is to use postprocessing to reduce the number of top-scored molecules to a level that can be processed experimentally

Docking as an auxiliary tool

- Docking plays a secondary, although important, role in the hit-to-lead optimization.
- Even in the experimental search for inhibitors, docking is still useful – it allows one to determine the position of inhibitors and study the pattern of their interactions with the protein, and create analogues with higher activity.
- Many of the discovered inhibitors of recombinant target proteins **do not act against live coronavirus in cell culture**.
- In some publications, the candidate inhibitors of a given coronavirus target proteins discovered in virtual screening were tested immediately on live virus, without experiments on viral recombinant proteins

Conclusions

- Docking is a main computing tool at the initial stage of drug development
- Central to the success is the coordinated and focused work of:
  - a good quality target protein structure
  - molecular modellers: construct target models, docking etc.
  - biochemists and biophysicists: experimental test systems that determine the binding of selected candidates to the target protein
  - synthetic chemists creating new compounds
  - virologists working with live viruses in cell cultures and with animal models.

Thank you for attention!

Our team – authors of the review:
Docking and other computing tools in drug design against SARS-CoV-2, SAR and QSAR in Environmental Research, **35** (2024) 2, 91-136

Alexey Sulimov **Ivan Ilin** **Anna Taschilova** **Olga Kondakova** **Danil Kutov**

The work is financially supported by the Russian Science Foundation, Agreement no. 21-71-20031