Implementation of a Webserver for Ensemble Docking with SARS-CoV-2 Proteins

Mauricio M Rigo1, Sarah Hall-Swan1, Dinler A. Antunes1, Didier Devaurs2, Lydia E. Kavraki1, and Geancarlo Zanatta3

1 Department of Computer Science, Rice University, Houston, TX; 2 Université Grenoble Alpes - Inria, Grenoble, France; 3 Department of Physics, Federal University of Ceará, Brazil

Background

• The consequences of COVID-19 pandemic have been disastrous, and efforts are currently being made to develop a vaccine and identify active medications.

• Most of these efforts rely on in silico studies targeting SARS-CoV-2 proteins. One of these proteins is the Main protease (Mpro).

The problem

Need for computational methods that take into account the inherent flexibility of Mpro towards drug design.
**Generation of Ensembles:** Mpro ensembles account for the inherent protein flexibility

- Retrieval of 158 Mpro structures from Protein Data Bank (PDB)
- Molecular dynamics with CHARMM36 force field
- Molecular dynamics with GROMOS53a6 force field
- 100,000 conformations extracted using a python library called Mdtraj\(^1\)
- Dimensionality reduction using PCA and clustering (K-means)

**PDB structures**

- CHARMM36 force field simulation
- GROMOS53a6 force field simulation

Representative conformations (red circles) of Mpro ensembles along PC1 and PC2. Free energy was estimated with PyEMMA\(^2\) and is represented on the right.

Crystal ensemble presents low flexibility; CHARMM36 ensemble presents intermediate flexibility (0.89-2.69 Å); GROMOS53a6 ensemble shows the largest variation (1.38-4.08 Å).
**Molecular Docking**: sampling performed on parallelized threads and output scoring with AutoDock Vina, AutoDock4, and Vinardo

The algorithm was validated in a subset of 36 Mpro ligand conformations.

No. | PDB ID | PubChem CID | RMSD\(^a\) (Å) |
---|---|---|---|
1 | 5R7Y | 118569 | 1.91 |
2 | 5R7Z | 405042899 | 1.49 |
3 | 5R80 | 89847 | 0.58 |

... |
35 | 6M2N | 5281605 | 2.71 |
36 | 6W63 | 145998279 | 0.54 |

**Mean** | | | 1.4 |

\(^a\)RMSD, Root Mean Square Deviation

Our approach explores a range of receptor conformations within the top scoring binding modes for each of the scoring functions used for rescoring.
**DINC-COVID: a Webserver for Ensemble Docking with SARS-CoV-2 Proteins**

**Conclusion and Resources**

- **DINC-COVID** is an ensemble docking solution that accounts for the receptor’s conformational flexibility without the burden of running molecular dynamics simulations for docking calculations.

- The server offers a ready-to-use solution for researchers to test their own compounds against the SARS-CoV-2 main protease.

- **DINC-COVID** is available at dinc-covid.kavrakilab.org.

**Acknowledgements**

We thank the Center for Research Computing (CRC) at Rice University. This work was funded in part by the National Science Foundation IIBR:Informatics:RAPID program (2033262), the Cancer Prevention & Research Institute of Texas (RP170593 and T15LM007083-29), by the National Council for Scientific and Technological Development (Brazil, 437373/2018-5), and by Rice University funds.

**References**


3. Devaurs, D. et al. **Using parallelized incremental meta-docking can solve the conformational sampling issue when docking large ligands to proteins.** BMC Molecular and Cell Biology, 2019.