Development of the docking protocol in search for inhibitors of Niemann-Pick C1 protein as anti Ebola agents

Dawid Warszycki[1], Rafał Kafel[1], Krzysztof Rataj[1], Vincent Roy[2], Luigi A. Agrofoglio[2], Andrzej J. Bojarski[1]

[1]Institute of Pharmacology Polish Academy of Sciences, Krakow, 31-343, Poland
[2]Institut de Chimie Organique et Analytique, UMR 7311 CNRS, Université d'Orléans, 45067 Orléans, France

e-mail: warszyc@if-pan.krakow.pl

The most important aspect of the EBOV lifecycle is entry into the host cell. During the entry of EBOV into the host cell, the heavily glycosylated domain of the filovirus glycoprotein (GP) is proteolyzed by cathepsins. The proteolytically processed virus GP binds to Niemann-Pick C1 protein (NPC1) within late endosomal/lysosomal compartments and subsequent poorly characterized membrane fusion events release the viral core into the cytoplasm where replication takes place (Figure 1). Therefore, inhibition of virus entry into the host cell can effectively block spreading of the virus at an early stage of infection, thus minimizing the risk of viral evolution and drug resistance [1].

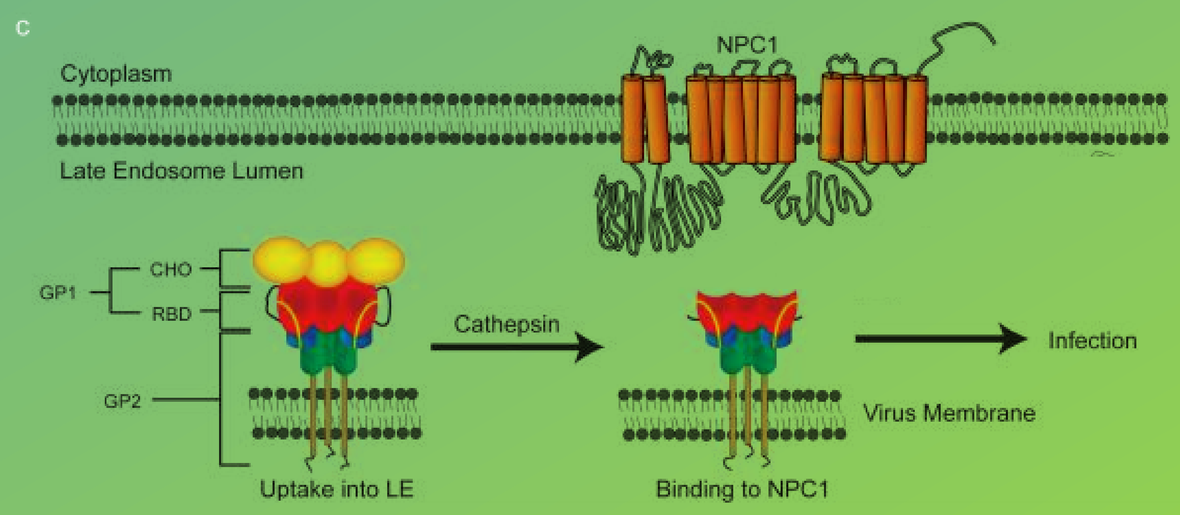

Figure 1. Proposed model of EBOV entry into the cell (according to Cote et al.[1]). Following EBOV uptake and trafficking to late endosomes, EBOV GP is cleaved by cathepsin protease to remove heavily glycosylated domains (CHO) and expose the putative receptor binding domain (RBD) of GP. The binding of cleaved GP1 to NPC1 is necessary for infection and may be blocked by EBOV inhibitors.

In this study, docking protocol consisted of ensemble of models was developed (Figure 2). Recently crystalized structure of NPC1 [2] has highly unsatisfactory efficiency in retrospective virtual screening experiments. All 1637 actives fetched from ChEMBL and PDSP database (K_i or equivalent less than 1000 nM for ChEMBL ligands or marked as an active for PDSP compounds) were clustered with the Hierarchical Clustering tool in Canvas [3]. After manual refinements compounds were splitted in 20 distinct chemical classes. Centroids from each cluster were used in induced-fit docking procedure [4] which generated one model per cluster. Screening efficiency (tested as previous, with the application of all actives and 88k decoys generated with application of DUD-E formalism [5]) was significantly improved but still not efficient enough.

However, utilization of data fusion methods combining single models into ensembles allows for generation of docking protocol with enough high screening parameters for applying in virtual screening cascade in search for new antiEBOV agents.

Summary of data fusion methods used in this study

Sum rank adds together the ranks from the different models' rank lists.
In **rank vote** each screening method votes for its 10000 highest ranked compounds. The ranking is primarily based on the number of votes each compound has received and secondarily on the compounds' sum score. **Sum score** calculates the relative score of each compound in each model by dividing all of the scores by the best score any compound acquired from that model. The calculated scores for proper models are then summed. **Pareto ranking** ranks a compound on the basis of how many other compounds are better in all analyzed models. Ties are broken using sum rank. In **parallel selection** compounds are selected from the top ranked compounds from each method in turn until the desired number of compounds is reached. If a compound that would be selected has already been selected before, the next compound from that method is chosen instead.

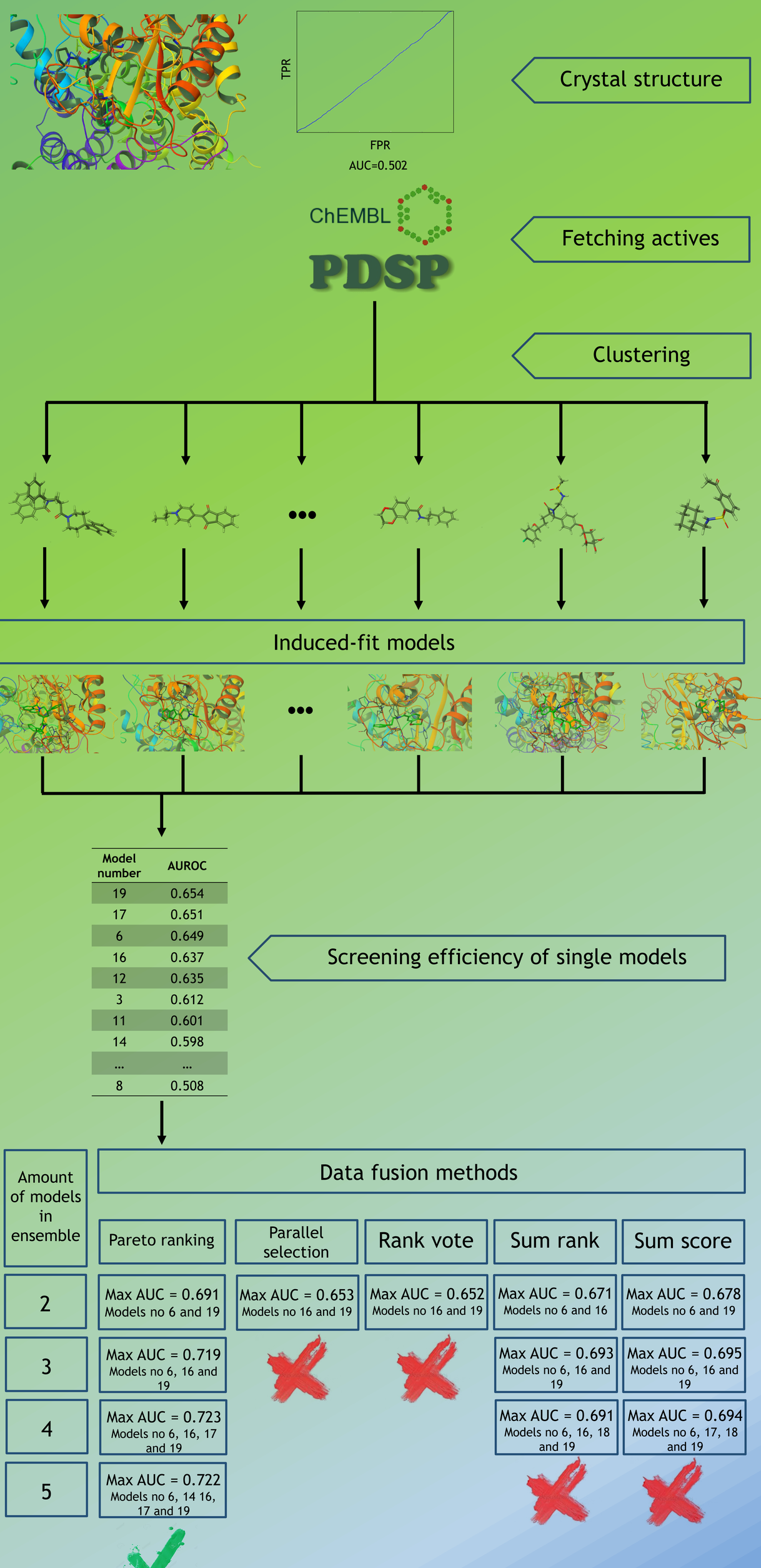

Figure 2. Workflow of NPC1 docking protocol development with application of induce-fit docking and data fusion methods.

References

[1] Cote, M. et al., Nature, 2011, 477, 344-348.
[2] Li, X. et al., Proc. Natl. Acad. Sci. 2017, 114, 9116-9121.
[3] Canvas, Schrödinger, LLC, New York, NY, 2017.
[4] Sherman, W. et al., J. Med. Chem., 2006, 49, 534-553.
[5] Mysinger, M. J. Med. Chem., 2012, 55, 6582-6594.

Acknowledgements

The work was supported by the National Science Centre (Poland) grant no. 2016/21/N/NZ2/01725.

NATIONAL SCIENCE CENTRE POLAND