Consensus Scoring for Challenging GPCR Virtual Ligand Screenings

Koriazhkina, A.[1], Marin E.[1], Luginina A.[1], Gusach A.[1], Lyapina E.[1], Mishin A.[1], Borshchevskiy V.[1]

[1]Research Center for Molecular Mechanisms of Aging and Age-Related Diseases, Moscow Institute of Physics and Technology (National Research University), Dolgoprudny, Russia

G protein-coupled receptors (GPCRs) represent the largest superfamily of more than 800 receptors. GPCRs provide numerous physiological and pathophysiological effects, which makes these receptors very attractive targets for drug design. GPCRs interact with trimeric G proteins (Gα, Gβ, and Gγ subunits), which bind to the intracellular side of the receptor. The binding of endogenous ligands to GPCR triggers conformational changes of the receptor, which, in turn, promotes the exchange of Gα-bound GDP to GTP, and subsequent dissociation of Gα and G$\beta\gamma$. As a result, numerous intracellular downstream signaling cascades are activated in response to the binding of the endogenous ligand. GPCR ligands have a very diverse nature and include ions, amines, nucleotides, peptides, lipids, etc.[1]

Modern drug discovery involves comprehensive screening of many candidate chemical compounds and normally starts with virtual ligand screening (VLS) campaigns. VLS campaigns enable a time- and cost-effective initial check of large chemical libraries with a few million compounds to select a small subset of molecules for further experimental evaluation of their activity towards the selected target. VLS campaigns have already shown relatively high hit rates ($> 25\%$) for some GPCRs. [2] However, the situation is quite different in the case of lipid GPCRs. They bind large hydrophobic molecules and have a large and open ligand-binding pocket with distinct hydrophobic and hydrophilic regions. For the above reasons, standard VLS campaigns are not likely to achieve high hit rates. Indeed, a recent VLS campaign for cysteinyl leukotriene GPCRs CysLT1R and CysLT2R yielded 155 candidate compounds, 139 of which were synthesized, but only 5 of them have proven their potency in experimental trials. Nevertheless, the design of drugs, targeting CysLT1R and CysLT2R, is of great interest, as these GPCRs play a key role in allergic disorders, cardiovascular diseases, and cancer.[3]

Lipid GPCRs remain a very interesting class of targets for drug design. There are already 26 approved drugs that target lipid GPCRs, including compounds for the treatment of neurological, neuropsychiatric, metabolic, cardiovascular, allergic, immune, and many other diseases. Moreover, many drugs targeting lipid GPCRs are currently under evaluation in more than 100 clinical trials.[1]

Therefore, optimization of scoring algorithms for VLS campaigns for lipid GPCRs is of high interest. Many algorithms for ligand pose prediction and docking score assigning, which estimates the likelihood of the ligand binding to the target, have been developed to date. Combining the results from different docking programs is found to be a very successful strategy to obtain a docking score, which frequently outperforms individual docking algorithms.[4] The goal of this study is to develop a novel consensus scoring method, which will facilitate docking and improve VLS hit rates for lipid GPCRs.

Before the model building and training, an appropriate benchmark should be chosen. For our work, we have chosen GPCR Ligand Library with its accompanying GPCR Decoy Database (GLL/GDD). This database is built for docking of GPCRs on the same principles as A Database

of Useful Decoys: Enhanced (DUD-E), which is one of the most commonly used benchmarks for testing the performance of docking and VLS techniques in the discrimination of ligands and decoys. Moreover, the database has distinct sets of agonists and antagonists for GPCRs, as active and inactive forms of receptors significantly differ.[5] However, numerous experimental screenings of compounds activity against GPCR targets have been conducted since the GLL/GDD development, and a significant extension of the benchmark can be done using this data. Therefore, in our work, we attempt to extend the existing datasets for model evaluation. Moreover, it will enlarge the number of accessible lipid GPCR targets for model evaluation, as only 8 datasets from GLL/GDD with available high-resolution structures of receptors in the required activity state were found.

Primarily, we hypothesized that the predictive ability of docking can be improved with the re-weighting of energy terms, which constitute the docking score of the ICM program. As ligands of lipid GPCRs have quite special physicochemical properties, this strategy can be quite beneficial. Although re-weighted coefficients did not give any interpretable insights into the peculiarities of ligand binding of lipid GPCR, distinct models have proven its benefits in some cases. For some receptors, the models with re-weighted coefficients were more successful, than discrimination of ligands versus decoys using standard ICM docking score. These results can be further utilized for the preliminary sifting of candidate compounds, which would be further screened using the consensus scoring method, which combines ICM docking results with estimations of other algorithms. The optimal consensus scoring scheme is under development now. Combined with extensive datasets, it will help to facilitate an effective method for VLS campaigns with improved hit rates for lipid GPCRs. This work was supported by the Russian Science Foundation project no. 19-14-00261.

References

[1] Yasuyuki Kihara. *Introduction: Druggable Lipid Signaling Pathways*, volume 1274. 2020.

[2] Miles Congreve, Chris de Graaf, Nigel A. Swain, and Christopher G. Tate. Impact of GPCR Structures on Drug Discovery, apr 2020.

[3] Arman A. Sadybekov, Rebecca L Brouillette, Egor Marin, Anastasiia V Sadybekov, Aleksandra Luginina, Anastasiia Gusach, Alexey Mishin, Élie Besserer-Offroy, Jean Michel Longpré, Valentin Borshchevskiy, Vadim Cherezov, Philippe Sarret, and Vsevolod Katritch. Structure-based virtual screening of ultra-large library yields potent antagonists for a lipid gpcr. *Biomolecules*, 10(12):1–15, 2020.

[4] Karen Palacio-Rodríguez, Isaias Lans, Claudio N. Cavasotto, and Pilar Cossio. Exponential consensus ranking improves the outcome in docking and receptor ensemble docking. *Scientific Reports*, 9(1):1–14, 2019.

[5] Edgar A. Gatica and Claudio N. Cavasotto. Ligand and decoy sets for docking to G protein-coupled receptors. *Journal of Chemical Information and Modeling*, 52(1):1–6, 2012.