International Journal of Pharmacy and Pharmaceutical Sciences

Print ISSN: 2656-0097 | Online ISSN: 0975-1491 Vol 13, Issue 8, 2021

Original Article

DATABASE ENRICHMENTS OF MAO-B THROUGH ENSEMBLE DOCKING

EMILIO MATEEV*, IVA VALKOVA, MAYA GEORGIEVA, ALEXANDER ZLATKOV

Department of Pharmaceutical Chemistry, Faculty of Pharmacy, Medical University–Sofia
Email: e.mateev@pharmfac.mu-sofia.bg

Received: 01 May 2021, Revised and Accepted: 02 Jun 2021

ABSTRACT

Objective: The recent growth of highly resoluted crystallographic structures, together with the continuous improvements of the computing power, has established molecular docking as a leading drug design technique. However, the problems concerning the receptor flexibility and the lowered ability of docking software to correctly score the occurred interactions in some receptors are still relevant.

Methods: Recently, several research groups have reported an enhancement in enrichment values when ensemble docking has been applied. Therefore, we utilized the latest technique for a dataset of Monoamine Oxidase–B (MAO-B) inhibitors. The docking program GOLD 5.3 was used in our study. Several docking parameters (grid space, scoring functions and ligand flexibility) were altered in order to achieve the optimal docking protocol.

Results: The results of 200 000+docking simulations are represented in a modest table. The ensembled simulations demonstrated low ability of the docking software to correctly score the actives seeded in the dataset. However, the superimposed complex-**1S3B-1OJA-1OJC**, achieved a moderate enrichment value equaled to 9. No significant improvements were noted when five complexed receptors were employed.

Conclusion: As a conclusion, it should be noted that in some cases the ensemble docking enhanced the database enrichments, however overall the value is not suitable for future virtual screening. Further investigations in that area should be considered.

Keywords: Molecular docking, MAO-B, Ensemble docking, Enrichment factor

DOI: https://dx.doi.org/10.22159/ijpps.2021v13i8.41956. Journal homepage: https://innovareacademics.in/journals/index.php/ijpps.

INTRODUCTION

Molecular docking is the most common structure-based technique utilized for the virtual design and development of novel molecules. In general, it observes the interactions between a ligand and a receptor applying searching and scoring algorithms [1]. The recent growth of highly resoluted crystallographic structures, together with the enhanced computer power, established molecular docking as a leading drug design technique [2]. However, few downsides such as receptor flexibility [3] and the inability of the docking softwares to correctly score the occurring bonds [4] are still relevant.

The problem with the receptor flexibility has been presented as hard to deal with, considering the exponential growth of computational time [5]. The ensemble docking (ED) somewhat charges that issue, since it applies several different conformations of the receptor during the docking simulations [6]. The technique has demonstrated promising results viewed from the papers reporting high enrichments when ED has been applied [7-9]. However, the inducement of ensembles cannot always enhance the enrichment values as reported by Rao *et al.* [10].

Binda *et al.* were the first to report the crystallographic structure of MAO-B (fig. 1) [11]. The paper has identified three functional domains–Entrance cavity, substrate cavity and aromatic cage, of which the substrate cavity comprises the biggest volume. The authors also described a loop of four amino acids, which separates the entrance from the substrate cavity. Moreover, the amino residue Ile199 has been found to operate in two different states–opened and closed [12, 13]. The aromatic cage was built of (Flavin Adenine Dinucleotide) FAD and two tyrosine amino residues–Tyr398 and Tyr435 [14]. Numerous papers have described the effects of the aromatic cage towards the stabilization of the ligand-receptor complex [15-17].

The lack of ensemble docking screenings in the active site of the crystallographic structure of MAO-B, together with the prominent reports about improved enrichments after ensemble docking simulations, have encouraged our research group to investigate the role of the latest technique in the virtual enrichments of MAO-B. The aim of this work was to observe if ED could increase the enrichment values in MAO-B crystallographic structures. For that purpose, most of the resolved 3D monoamine oxidases were taken from the Protein Data Bank (PDB) and ensembled for further docking simulations.

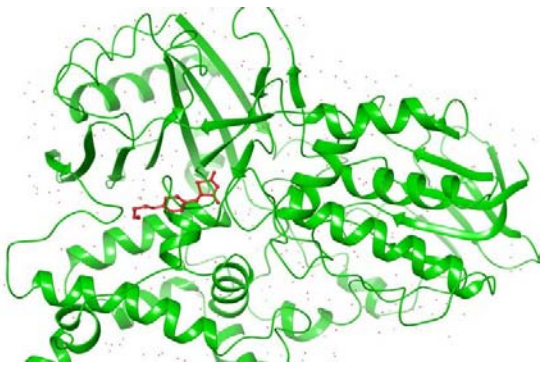

Fig. 1: Crystallographic structure of MAO-B with co-crystallized ligand (red) [18]

MATERIALS AND METHODS

Docking program

For the current study, we utilized (Genetic Optimization for Ligand Docking) GOLD 5.3 as a molecular docking program [19]. The software operates with four scoring functions–ChemPLP, GoldScore, Chemscore and ASP. Gold Score is the first GOLD function and it considers Van der Waals clashes as well as calculations of hydrogen bonds between the ligands and the receptor. Chemscore is an empirical algorithm calibrated from numerous complexes with known binding affinities. The knowledge-based function is ASP. The

most recent and successful scoring algorithm employed in GOLD is ChemPLP, which uses piecewise linear potential to score the contacts in the ligand-receptor complex [20]. The program includes a wizard with pre-defined docking parameters; however, the binding site is user-defined.

Hardware

Operating system–Windows 10 Pro; CPU-AMD Ryzen 5 3600 6-core 3.60 GHz; GPU–GeForce GTX 1060 3 GB; Install memory (RAM)–16 GB

Pre-docking setup

All the crystallographic structures of MAO-B with the co-crystallized ligands (1OJA, 1OJC, 1S3B, 2BK3, 2C65, 3PO7, 4A7A, 6FVZ, 6FW0, 6FWC) were downloaded from the protein data bank (PDB) [21]. The collected receptors were resolved with resolutions under 2 Å. Firstly, from all the MAO-B receptors, monomer B was deleted together with the ligands and the co-factor. If a covalent bond between FAD N5 and the co-crystallized ligand was present, it was removed. Utilizing the GOLD setup wizard, we added hydrogen bonds, removed all water molecules and extracted the co-crystallized ligands. Thereafter, we superimposed the MAO-B protein structures. The grid space was firstly altered between 6, 8 and 10 Å and it was centered on FAD's N5 atom. The search efficiency was switched to "virtual screening" (30%) due to the time-demanding simulations of the "very flexible (200%)" GOLD docking option. The ligand flexibility was also modified between fixed and free rotatable bonds.

Decoys

The benchmarking set of active compounds and decoys was taken from the (Directory of Useful Decoys-Enhanced) DUD-E [22]. It contains more than 100 crystallographic proteins with the corresponding active and inactive (decoy) molecules. The decoys are classified as similar to the active compounds, but with no activity towards the observed receptor. No preliminary energy minimizations of the receptors and the ligands were performed. It should be taken into account that the decoys are not experimentally tested. They are compounds with the same physical properties however, they could contain some kind of activity towards the receptors and should not be considered as completely not active.

Docking accuracy

Initial re-docking procedures were carried out applying the default GOLD 5.3 settings. The (root mean square diameter) RMSD values of the original conformation of the co-crystallized ligand and the solutions obtained with the software were examined with Hermes.

The enrichment factor (EF) represents the quantification of the reliability of a docking program. We have calculated both the modified enrichment factor (EF`) [7] and the classical EF [23]. The latter enrichment value does not take into consideration the rankings of the seeded actives, while EF` provides higher values when the active compounds are located in top-ranked positions. The formulas of EF and EF`(N) are defined as:

$$EF = (HITS_{sampled}/HITS_{total})/(N_{sampled}/N_{total}),$$

Where $HITS_{sampled}$ are the active ligands located in the chosen percentage of the dataset. $HITS_{total}$ is the sum of all actives seeded in the decoys. N_{total} is the number of the whole dataset (7031), while $N_{sampled}$ is the chosen part of the dataset (in our case 100).

$$EF`(N) = (50\%/APR_{sampled}) \times (Hits_{sampeld}/Hits_{total}),$$

Where N is the percent of the active compounds set; ARP stands for "average percentile rank'' of $Hits_{total}$. In this study we calculated the EF` value of 6% of the seeded active compounds, therefore the $Hits_{sampeld}$ corresponds to 10; $Hits_{total}$ equals 6931. The maximum number of EF`(6) would be 37.6, considering the fact the top 10 rankings were occupied from 10 seeded actives.

RESULTS

Self-docking

We started our study with the pre-docking operations, which are described in the materials and methods section of this paper. We proceed with self-docking simulations in order to evaluate the ability of GOLD 5.3 to correctly position and score the removed co-crystallized ligands. In all examined cases the RMSD values were under 2 Å, which clearly demonstrates the effectiveness of the program to recreate the original pose of the ligands. The downloaded receptors were with a chemical diverse set of co-crystallized ligands so that different conformers of the active site residues are present.

Ensemble docking

Ensembles of three and five receptors were generated and applied for the docking of 6931 decoys and 169 active MAO-B inhibitors. We utilized the classical enrichment factor, as well as a modified version of it, which considers the rankings of the ligands. Exhaustive details of both were provided in the "materials and methods" section.

All the PDB codes of the receptors, together with the optimal scoring functions, size of the binding gauge, EF and EF`(6) are given in table 1**.** From the observed cases, ChemPLP showed the best results. Moreover, it was the fastest scoring function. For the first ensemble complex, we altered 4 different sizes of the binding gorge in order to find the optimal one. Increasing the cavity volume achieved slightly better enrichments, however, the running time was significantly increased. Considering the latter observation, we employed 6 Å for the rest of the complexes.

Table 1: Calculated enrichments of 9 receptor complexes through ensemble docking. The highest enrichment value was achieved by the 1S3B-1OJA-1OJC comp superimposed lex

Receptor/Ensemble	Scoring algorithm	Size of the binding pocket	EF		EF`(6)	
			Flexible ligands	Rigid ligands	Flexible ligands	Rigid ligands
3PO7-6FW0-6FWC	ChemPLP	6 Å	4.17	5.8	4.19	8.13
		8 Å	4.17	5.8	4.11	8.04
		10 Å	4.58	6.25	4.25	8.17
2BK3-2C65-4A7A	ChemPLP	6 Å	5	5	5.9	5.72
4A7A-6FW0-6FVZ	ChemPLP	6 Å	4.58	4.58	4.65	4.72
1S3B-1OJA-1OJC	ChemPLP	6 Å	7.25	6.25	**9**	8.5
1S3B-3PO7-4A7A-6FVZ-6FW0	ChemPLP	6 Å	5.42	6.25	5.5	7.9
10JA-2BK3-2C65-4A7A-6FWC	ChemPLP	6 Å	5.42	5.42	3.99	6.4
1S3B-2XFN-3PO7-6FW0-6FWC	ChemPLP	6 Å	4.17	5.42	3.8	5.2
1S3B-1OJA-1OJC-4A7A-6FWO	ChemPLP	6 Å	5.42	4.50	5.06	4.32
1S3B-1OJA-1OJC-2C65-6FWC	ChemPLP	6 Å	5	5.8	5	7.43

Initially, we docked the seeded active molecules and the decoys into the ensemble complex of 3PO7, 6FW0 and 6FWC receptors. As shown above, when set to 6 and 8 Å, the docking software was able to correctly predict 10 active molecules situated in the top 100 solutions. Interestingly, when we fixed the rotatable bonds of the ligands, four additional active ligands were located in the rankings,

thus the enrichment factors were enhanced. Furthermore, the seeded compounds were with higher ranks demonstrated by the increased value of EF`(6).

The next ensemble complex-2BK3-2C65-4A7A, revealed a slightly higher enrichment value. Here 12 of the active ligands were located in the top 100 solutions. The rigid docking did not demonstrate better enrichment in that case. We followed up with the docking simulations of a 4A7A-6FW0-6FVZ ensemble. Compared to the previous case, the enrichment factor was lowered. Moreover, the rigid docking in this simulation showed no improvements in EF value. Higher enrichment was observed when we superimposed 1S3B with the PDB structures 1OJA and 1OJC. Here the fully flexible ligands demonstrated better enrichment in comparison to fixed rotatable bonds. As shown in table 1, the classical enrichment factor was calculated to be 7.25. Moreover, taking the rankings of the actives into consideration has led to an enhanced enrichment value of 9.

An employment of 5 superimposed MAO-B receptors into the docking simulations did not conduct the needed improvements in the enrichments coefficients. Primarily, we examined the enrichment with flexible ligands in the complexes 1S3B-3PO7-4A7A-6FVZ-6FW0 and 10JA-2BK3-2C65-4A7A-6FWC. The simulations led toa coefficient of 5.25. That corresponds to 13 active molecules located in the first 100 solutions. However, the rankings of the actives between the two ensemble sets did significantly differ. The affirmation of that finding came from the calculation of EF`(6), which considers the fitness score scores of the active molecules. The value of 1S3B-3PO7-4A7A-6FVZ-6FW0 was calculated to be 5.5 with comparison to EF`(6)=3.99 in the 10JA-2BK3-2C65-4A7A-6FWC ensemble set. Considering the fact that the highest enrichment was obtained when the 1S3B-1OJA-1OJC ensemble was superimposed, we obtained two complexes with 5 ensembled receptors, including the above sequence. In both cases, unsatisfying results were analyzed and no further simulations were conducted.

DISCUSSION

Numerous papers have reported improved database enrichments when ED was utilized in the docking simulations [7-9, 24] as it somewhat deals with the flexibility of the side residues of the crystallographic structure. Most of the simulations in our work did confirm these reports. An example is the superimposed complex 1S3B-1OJA-1OJC which achieved a moderate enrichment value. Furthermore, the actives in the latter case were situated at the top positions, which drastically increased the modified enrichment factor (EF`) [7] to 9.

Considering the recent work of Sheng-You Huang [25], it is evident that in some cases the flexible docking does not provide higher enrichments compared to rigid simulations. Several of the above-analyzed ensemble complexes confirmed that statement and better MAO-B enrichments were calculated when rigid docking was carried out. Interestingly, in some instances, the active ligands were scored with drastically higher fitness scores compared to the decoys. This confirms that the more flexible docking can still score false positives and it should not be considered as a prior to the rigid docking technique in all cases. Furthermore, the exponential growth of the computational time when high numbers of amino residues are flexible [26] put into consideration the applicability of fully flexible docking.

Overall, the ensembled simulations in the MAO-B crystallographic receptors demonstrated low ability of the docking software GOLD 5.3 to correctly score the actives seeded in the dataset. The enrichment values are not suitable for a future virtual screening and further work in that direction should be conducted.

CONCLUSION

In this study, more than 210 000 dockings (7060 ligands x 37 ensembled receptors) were carried out in order to analyze the ability of the ensemble docking to correctly score the active seeds in a dataset of decoys. We calculated both the standard enrichment (EF) and a modified enrichment factor (EF`) which takes into account the placement of the active ligands between the decoys. As conclusion, it should be noted that in some cases the ensemble docking increased the calculated enrichments, however, as a whole the value is not suitable for future virtual screening. Further investigations in that area should be considered.

ACKNOWLEDGEMENT

This work was supported by the Medical Science Council of Medical University of Sofia, Grant No.:D-7901/2020.

FUNDING

Nil

AUTHORS CONTRIBUTIONS

Emilio Mateev conducted the pre-docking procedures and the docking simulations. Iva Valkova provided the commercial docking program GOLD 5.3 and she took part in the refinements during the docking simulations. Maya Georgieva and Alexander Zlatkov were the major contributors to the idea of the experiment and participated in the analysis of the obtained data. All authors were engaged in the writing and editing of the manuscript.

CONFLICT OF INTERESTS

Declared none

REFERENCES

1. Pinzi L, Rastelli G. Molecular docking: shifting paradigms in drug discovery. Int J Mol Sci 2019;20:4331.
2. Sudha R, Kanakam C, Nithya G. Molecular docking studies and bioactivity of various benzilic acids and its analgogs. Asian J Pharm Clin Res 2018;11:463-6.
3. Awuni Y, Mu Y. Reduction of false positives in structure-based virtual screening when receptor plasticity is considered. Molecules 2015;20:5152-64.
4. Li J, Fu A, Zhang L. An overview of scoring functions used for protein-ligand interactions in molecular docking. Interdiscip Sci 2019;11:320-8.
5. Alhossary A, Handoko SD, Mu Y, Kwoh CK. Fast, accurate, and reliable molecular docking with QuickVina 2. Bioinformatics 2015;31:2214-6.
6. Jeyam M, Karthika G, Poornima V, Sharanya M. Molecular understanding and Insilico validation of traditional medicines for Parkinson's disease. Asian J Pharm Clin Res 2012;5:125-8.
7. Lovell T, Chen H, Lyne PD, Giordanetto F, Li J. On evaluating molecular-docking methods for pose prediction and enrichment factors. J Chem Inf Model 2006;46:401−15.
8. Korb O, Olsson TSG, Bowden SJ, Hall RJ, Verdonk ML, Liebeschuetz JW, *et al.* Potential and limitations of ensemble docking. J Chem Inf Model 2012;52:1262-74.
9. Craig IR, Essex JW, Spiegel K. Ensemble docking into multiple crystallographically derived protein structures: an evaluation based on the statistical analysis of enrichments. J Chem Inf Model 2010;50:511-24.
10. Rao S, Sanschagrin PC, Greenwood JR, Repasky MP, Sherman W, Farid R. Improving database enrichment through ensemble docking. J Comput Aided Mol Des 2008;22:621-7.
11. Binda C, Newton Vinson P, Hubalek F, Edmondson DE, Mattevi A. Structure of human monoamine oxidase B, a drug target for the treatment of neurological disorders. Nat Struct Biol 2002;9:22-6.
12. Hubalek F, Binda C, Khalil A, Li M, Mattevi A, Castagnoli N, *et al.* Demonstration of isoleucine 199 as a structural determinant for the selective inhibition of human monoamine oxidase b by specific reversible inhibitors. J Biol Chem 2005;280:15761-6.
13. Milczek EM, Binda C, Rovida S, Mattevi A, Edmondson DE. The 'gating' residues Ile199 and Tyr326 in human monoamine oxidase B function in substrate and inhibitor recognition. FEBS J 2011;278:4860-9.
14. Li M, Binda C, Mattevi A, Edmondson DE. Functional role of the aromatic cage in human monoamine oxidase b: structures and catalytic properties of tyr435 mutant proteins. Biochemistry 2006;45:4775-84.
15. Akyüz MA, Erdem SS, Edmondson DE. The aromatic cage in the active site of monoamine oxidase B: effect on the structural and electronic properties of bound benzylamine and p-nitrobenzylamine. J Neural Transm 2007;114:693-8.

16. Binda C, Mattevi A, Edmondson DE. Structure-function relationships in flavoenzyme-dependent amine oxidations. J Biol Chem 2002;277:23973-6.
17. Smith TE, Weissbach H, Udenfriend S. Studies on monoamine oxidase: the mechanism of inhibition of monoamine oxidase by iproniazid. Biochemistry 1963;2:746-51.
18. Ogidigo J, Anosike CON, Omotuyi O, Oyekanmi N, Metibemu D. *In silico* molecular docking and pharmacokinetic studies of some selected Phyto-constituents of bryophyllumpinnatum as a potential selective inhibitor of monoamine oxidase–B (MAO-B). Pharmacologyonline 2018;3:272-85.
19. Liebeschuetz JW, Cole JC, Korb O. Pose prediction and virtual screening performance of GOLD scoring functions in a standardized test. J Comput Aided Des 2012;26:737-48.
20. Nurisso A, Bravo J, Carrupt PA, Daina A. Molecular docking using the molecular lipophilicity potential as a hydrophobic descriptor: impact on GOLD docking performance. J Chem Inf Model 2012;52:1319-27.
21. Helen B, Kim H, Namura H. Announcing the worldwide protein data bank. Nat Struct Biol 2003;10:980.
22. Mysinger MM, Carchia M, Irwin JJ, Shoichet BK. Directory of useful decoys, enhanced (DUD-E): better ligands and decoys for better benchmarking. J Med Chem 2012;55:6582-94.
23. Zhou Z, Felts AK, Friesner RA, Levy RM. Comparative performance of several flexible docking programs and scoring functions: enrichment studies for a diverse set of pharmaceutically relevant targets. J Chem Inf Model 2007;47:1599-608.
24. Huang SY, Zou X. Ensemble docking of multiple protein structures: considering protein structural variations in molecular docking. Proteins 2007;66:399-421.
25. Huang SY. Comprehensive assessment of flexible-ligand docking algorithms: current effectiveness and challenges. Brief Bioinform 2018;19:982-94.
26. Huang SY, Zou X. Advances and challenges in protein-ligand docking. Int J Mol Sci 2010;11:3016-34.