OPTIMIZING MOLECULAR DOCKING PROTOCOLS OF PYRROLE CONTAINING MAO-B INHIBITORS THROUGH CORRELATION COEFFICIENTS

Emilio Mateev¹, Iva Valkova², Diana Tzankova³, Maya Georgieva⁴, Alexander Zlatkov⁵

Abstract: Virtual screening is emerging as a highly applied technique for the search of hits since it significantly reduces the time required for the establishment of novel, effective compounds compared to high-throughput screening. Implementing correlation coefficients to determine if a molecular docking study is robust and reliable has been established as common practice in recent years. The aim of this work was to determine if a relevant pairwise correlation between the scoring functions (ChemPLP, GoldScore, Chemscore and ASP) of the docking software GOLD 5.2 and previously determined experimental data of pyrrole derivatives with MAO-B inhibitory activity could be achieved. In order to optimize the correlation coefficient, we calculated the Pearson's and Spearman's coefficients after each docking simulation with all four GOLD 5.2 scoring functions. Thereafter, we varied three changeable parameters – the size of the grid space, the side-chain flexibility and the presence of water molecules in the active site, to perceive if we could obtain better correlation values. The highest $R^2=0.79$ was attained with the following docking settings: scoring function ChemPLP, grid size 12 Å and no rotatable side chain residues. This work provides an applicable GOLD 5.2 docking protocol for a future virtual screening of novel MAO-B inhibitors with pyrrole moiety.

UDC Classification: 615.3, **DOI:** <https://doi.org/10.12955/pmp.v2.179>

Keywords: ensemble docking, MAO-B inhibitors, GOLD docking, correlation

Introduction

Computer-aided drug design (CADD) is a widely used term including all computational processes for the analysis and modeling of compounds (Song et al., 2009). CADD is divided into two categories – structure-based drug design and ligand-based drug design. The Structure-based drug design (SBDD) is applied when a 3D structure of the receptor is available. The most common technique utilized in SBDD is molecular docking. In general, it analyzes the interactions between ligands and macrostructures using search algorithms and scoring functions (Kumar et al., 2016). Recently the application of molecular docking is drastically increasing due to the rapid growth of resolved crystallographic receptors with co-crystallized ligands. Additionally, rising computing power is further accelerating the process of hit discovery and lead optimizations implementing molecular docking (Torres et al., 2019). However, several approximation procedures are present during molecular docking, which leads to a moderate rate of false positives (Klebe, 2006). In order to enrich the obtained top solutions, an ensemble docking simulation has been established (Ellingson et al., 2014).

The ensemble docking was first introduced in 1999 by Carlson et al. The principle of the technique is that several protein structures with different conformations are superimposed and utilized for the docking study, which somewhat addresses the problem with the flexibility of the protein during virtual screenings. The ligands are docked into the ensemble complex and the best ligand scores with the corresponding receptor are obtained (Korb et al., 2012). A study held in 2007 demonstrated that the enrichment factor with 3 proteins in an ensemble complex is higher compared to docking into one or two receptors (Rao et al., 2008).

Monoamine oxidases (MAOs) are enzymes involved in the inactivation of aliphatic and aromatic amines (Muller & Mohr, 2019). MAO-B demonstrates prominent effects in the treatment of Parkinson's disease through lowering the dopamine levels in the brain and additional neuroprotective properties (Yamada & Yasuhara, 2004; Tatton et al., 1996).

The active site of MAO-B could be divided into 3 functional gauges – entrance cavity, substrate cavity and aromatic cage. The volume of the active site is approximately 700 Å³, of which the substrate pocket poses the largest cleft with a size around 490 Å³ (Binda et al., 2003; Bonivento et al., 2010). A loop

¹ Department of Pharmaceutical chemistry, Faculty of Pharmacy, Medical University-Sofia, Bulgaria, e.mateev@pharmfac.mu-sofia.bg

² Department of Chemistry, Faculty of Pharmacy, Medical University-Sofia, Bulgaria, ivalkova@abv.bg

³ Department of Pharmaceutical chemistry, Faculty of Pharmacy, Medical University-Sofia, Bulgaria, diana_ducova@abv.bg

⁴ Department of Pharmaceutical chemistry, Faculty of Pharmacy, Medical University-Sofia, Bulgaria, maya.bg77@gmail.com

⁵ Department of Pharmaceutical chemistry, Faculty of Pharmacy, Medical University-Sofia, Bulgaria, alexbz2000@yahoo.com

consisting of several amino acids separates the entrance from the substrate cavity. Observations were made that Ile199 could be either in an open or closed conformation, which changes the activity of the inhibitor (Binda et al., 2004). The aromatic cage is situated in the substrate cavity of the active site. It consists of two amino residues – Tyr398 and Tyr435 and the cofactor of the receptor – FAD (flavin adenine dinucleotide) (Binda et al., 2001). Several papers were published addressing the importance of the aromatic cage for the stabilization of the complex. Essentially, in order to achieve greater MAO-B inhibitory activity, the ligand should be located in close vicinity to FAD (Li et al., 2006). In Figure 1, the 3D structure of MAO-B in a complex with a ligand is presented. The entrance and the substrate pockets are marked.

Figure 1: Crystallographic structure of MAO- with co-crystallized 1,4-diphenyl-1-butene (PDB:10J9).

Source: Binda (2003).

Our aim for this study was to see if there is an acceptable correlation between GOLD 5.2 scoring functions (ChemPLP, Chemscore, ASP and GoldScore) and the experimental data for numerous pyrroles containing MAO-B inhibitors. Several changeable docking parameters, such as waters in the active site, size of the grid space and the rotation of the side chains, were modified for further optimization of the correlation coefficient.

Materials and methods

The docking program GOLD (Genetic Optimisation for Ligand *Docking*) (Jones et al., 1997) was employed for the current docking study. GOLD 5.2 possesses four scoring functions – ChemPLP, GoldScore, ASP and Chemscore. They all have been evaluated for their ability to obtain higher Pearson's and Spearman's coefficients, and, therefore, correct ligand poses.

MAO-B inhibitors with pyrrole as a core motif and the corresponding IC_{50} were taken from several recently published papers (Regina et al., 2007; Silvestri et al., 2003). IC_{50} values are converted to pIC_{50} ($-\log_{10} IC_{50}$). The structural formulas and the corresponding pIC_{50} are given in Table 1. All compounds were initially converted to 3D structures, and pre-docking procedures such as hydrogens addition, energy minimization, and establishment of charges were carried out with the Chem3D software (PerkinElmer Informatics). Thereafter, we obtained five MAO-B structures - 3PO7, 2XFN, 3PO7, 6FW0 and 6FWC from the Protein Data Bank (PDB) (Berman et al., 2000). Utilizing the visual module Hermes, we removed monomer B from the MAO-B structures together with the ligand and the cofactor lying in the active site. Applying GOLD 5.2 wizard setup, we added hydrogen bonds, removed all water molecules and extracted the co-crystallized ligands. We continued our study by superimposing all five MAO-B receptors and defined the binding grid centering it at the N5 atom in FAD. Subsequently, we divided the prepared ligands into two groups – training and test sets. The training set was constructed out of 14 compounds. It was utilized for the initial docking procedures with GOLD 5.2, and the optimization steps were employed on that group. Linear regressions were constructed between the

experimental data and the obtained fitness scores following each optimization step. After we acquired the optimal R^2 values for the training set, we validated the docking protocol by predicting the experimental pIC_{50} of the training set.

Table 1: Structures and pIC_{50} of the training and test sets used in this study.

Structure	pIC_{50}	Structure	pIC_{50}
(1)	0.920	(5)	-0.857
(2)	0.155	(6)	0.376
(3)	0.346	(7)	-0.653
(4)	-1.699	(8)	-0.176
(9)	-0.949	(12)	-0.556
(10)	-0.845	(13)	-1.69

<p>(11)</p>	-1.845	<p>(14)</p>	-0.397
<p>(A)</p>	-0.681	<p>(B)</p>	-0.301
<p>(C)</p>	-0.903	<p>(D)</p>	0.398
<p>(E)</p>	-0.398		

Source: Authors

Results and discussion

Comparing the scoring functions

In order to assess which is the best scoring function for the acquired compounds, we employed ensemble docking procedures with all four scoring functions and three different grid sizes. After each docking simulation, the correlation coefficient was calculated using the experimental data and the obtained fitness scores from the corresponding scoring functions. In **Table 2**, the values of the pairwise correlations of GOLD 5.2 scoring functions and various grid sizes are provided. The simulations were carried out with no water in the active site and no side chain flexibility. The top solution of each scored ligand was considered, and the docking simulation was repeated 10 times.

Table 2: Pearson's correlation coefficient of the scoring functions, applying different grid sizes.

	ChemPLP	GoldScore	Chemscore	ASP
6Å	R ² =0.6325	R ² =0.7773	R ² =0.7387	R ² =0.7601
8Å	R ² =0.6348	R ² =0.7018	R ² =0.7401	R ² =0.7435
12Å	R ² =0.7973	R ² =0.7863	R ² =0.7212	R ² =0.7654

Source: Authors

As displayed in Table 2, ChemPLP demonstrated the best results when the grid size was set at 12Å. However, significantly lower values were observed when the grid space was fixed at 8Å and 6Å. Interestingly, when the binding atoms were selected within a range of 6Å, the value of the ChemPLP pairwise correlation coefficient was R²= 0.6325 (Figure 1), which was the lowest out of all scoring functions. Applying the latter grid parameter, GoldScore showed the best result – R²=0.7773. However, when we increased the size of the docking pocket to 12Å, the most optimal function was ChemPLP,

with a pairwise correlation value of 0.7973. Below we provided the scatterplot matrix of the ChemPLP and the pIC_{50} of the training set utilizing the 12Å grid space.

The maximum calculated pairwise correlation was 0.7973, as shown in Figure 3. The Spearman's coefficient was calculated to be 0.7437. We noticed that the docking software was unable to correctly score ligand 21, which is colored in blue. Further molecular dynamics simulations could be done to calculate the root mean square fluctuations (RMSF) and to observe the stability of the complex.

For further optimization of the fitness/experimental correlation data, the flexibility of the side chain residues was altered. The 5.2 version of GOLD provides a possibility to modify the rotation of up to 10 side-chain residues. We set the flexibility of PHE343, TYR326, TYR60, PHE168, LEU171, CYS172, ILE199, TYR188, TYR 398, TYR435 amino residues to "library", which is defined as a compilation of the most commonly observed side-chain conformations (Lovell et al., 2000). However, the docking simulations with semi-flexible receptors showed no significant change in the correlation coefficients, and no further examinations in that area were done. Furthermore, including water molecules in the active site demonstrated no positive effect on the correlation coefficient.

Figure 3: Correlation results when the grid space was set at 12Å.

Overall, the settings for the final optimized docking protocol are ChemPLP, grid space 12Å and no rotatable side chains. To validate the explored parameters, we used a test set of 5 ligands with similar structures. After the docking study, we calculated the correlation between the obtained ChemPLP fitness score and the experimental *in vitro* data. The results are shown in Figure 4. A good pairwise correlation of 0.6683 was achieved.

Conclusion

Ensemble docking is emerging as a prominent technique attempting to resolve the major problem of molecular docking, that is to say, the flexibility of the proteins. Nevertheless, validation work is necessary, considering the wide variety of changeable software parameters and the numerous available

receptors in the Protein Data Banks. Achieving high correlation coefficients between the experimental *in vitro* data and the *in silico* docking fitness scores is considered a prominent verification method of the reliability of the virtual screening.

In this work, good correlation coefficients were achieved between the pIC50 values of 14 pyrrole derivatives with MAO-B activity and their fitness scores applying the docking software GOLD. After optimization of the docking protocol, in terms of scoring functions, grid spaces and rotatable side residues, we obtained a pairwise correlation of 0.79. Moreover, we noticed that only compound **9** demonstrated a false-positive result when docked into MAO-B. For further validation of the final docking settings, we examined the correlation coefficient of a test set containing 5 compounds with similar structures.

Figure 4: Calculated correlations for the test set.

Source: Authors

Acknowledgements

This work was supported by the Bulgarian Ministry of Education and Science under the National Program for Research "Young Scientists and Postdoctoral Students".

References

- Altintop, M. D., Sever, B., Osmaniye, D., Sağlık, B. N., & Özdemir, A. (2018). Design, synthesis, in vitro and in silico evaluation of new pyrrole derivatives as monoamine oxidase inhibitors. *Archiv der Pharmazie*, 351(7), 1800082. doi:10.1002/ardp.201800082
- Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., . . . Bourne, P. E. (2000). The Protein Data Bank. *Nucleic Acids Res*, 28(1), 235-242. doi:10.1093/nar/28.1.235
- Binda, C., Hubálek, F., Li, M., Edmondson, D. E., & Mattevi, A. (2004). Crystal structure of human monoamine oxidase B, a drug target enzyme monotonically inserted into the mitochondrial outer membrane. *FEBS Letters*, 564(3), 225-228. doi:10.1016/s0014-5793(04)00209-1
- Binda, C., Hubálek, F., Li, M., Herzig, Y., Sterling, J., Edmondson, D. E., & Mattevi, A. (2004). Crystal Structures of Monoamine Oxidase B in Complex with Four Inhibitors of the N-Propargylaminoindan Class. *Journal of Medicinal Chemistry*, 47(7), 1767-1774. doi:10.1021/jm031087c

- Binda, C., Li, M., Hubalek, F., Restelli, N., Edmondson, D. E., & Mattevi, A. (2003). Insights into the mode of inhibition of human mitochondrial monoamine oxidase B from high-resolution crystal structures. *Proceedings of the National Academy of Sciences*, 100(17), 9750-9755. doi:10.1073/pnas.1633804100
- Bonivento, D., Milczek, E. M., McDonald, G. R., Binda, C., Holt, A., Edmondson, D. E., & Mattevi, A. (2010). Potentiation of Ligand Binding through Cooperative Effects in Monoamine Oxidase B. *Journal of Biological Chemistry*, 285(47), 36849-36856. doi:10.1074/jbc.M110.169482
- C. Basak, S. (2013). Editorial (Recent Developments and Future Directions at Current Computer Aided Drug Design). *Current Computer - Aided Drug Design*, 9(1), 1-1. doi:10.2174/157340913804998766
- Carlson, H. A., Masukawa, K. M., & McCammon, J. A. (1999). Method for Including the Dynamic Fluctuations of a Protein in Computer-Aided Drug Design. *The Journal of Physical Chemistry A*, 103(49), 10213-10219. doi:10.1021/jp991997z
- Ellingson, S. R., Miao, Y., Baudry, J., & Smith, J. C. (2014). Multi-Conformer Ensemble Docking to Difficult Protein Targets. *The Journal of Physical Chemistry B*, 119(3), 1026-1034. doi:10.1021/jp506511p
- Klebe, G. (2006). Virtual ligand screening: strategies, perspectives and limitations. *Drug Discovery Today*, 11(13-14), 580-594. doi:10.1016/j.drudis.2006.05.012
- Korb, O., Olsson, T. S. G., Bowden, S. J., Hall, R. J., Verdonk, M. L., Liebeschuetz, J. W., & Cole, J. C. (2012). Potential and Limitations of Ensemble Docking. *Journal of Chemical Information and Modeling*, 52(5), 1262-1274. doi:10.1021/ci2005934
- Kumar, R., Långström, B., & Darreh-Shori, T. (2016). Novel ligands of Choline Acetyltransferase designed by in silico molecular docking, hologram QSAR and lead optimization. *Scientific Reports*, 6(1). doi:10.1038/srep31247
- La Regina, G., Silvestri, R., Artico, M., Lavecchia, A., Novellino, E., Befani, O., . . . Agostinelli, E. (2007). New Pyrrole Inhibitors of Monoamine Oxidase: Synthesis, Biological Evaluation, and Structural Determinants of MAO-A and MAO-B Selectivity. *Journal of Medicinal Chemistry*, 50(5), 922-931. doi:10.1021/jm060882y
- Li, M., Binda, C., Mattevi, A., & Edmondson, D. E. (2006). Functional Role of the Aromatic Cage in Human Monoamine Oxidase B: Structures and Catalytic Properties of Tyr435 Mutant Proteins. In: Worldwide Protein Data Bank.
- Lovell, S. C., Word, J. M., Richardson, J. S., & Richardson, D. C. (2000). The penultimate rotamer library. *Proteins: Structure, Function, and Genetics*, 40(3), 389-408. doi:10.1002/1097-0134(20000815)40:3<389::aid-prot50>3.0.co;2-2
- Rao, S., Sanschagrin, P. C., Greenwood, J. R., Repasky, M. P., Sherman, W., & Farid, R. (2008). Improving database enrichment through ensemble docking. *Journal of Computer-Aided Molecular Design*, 22(9), 621-627. doi:10.1007/s10822-008-9182-y
- Rondon, P., Arguello, H., & Torres, R. (2011). *Development of a zoned genetic algorithm for semi-flexible protein-ligand docking in drug design*. Paper presented at the 2011 6th Colombian Computing Congress (CCC). <http://dx.doi.org/10.1109/colomcc.2011.5936321>
- Silvestri, R., La Regina, G., De Martino, G., Artico, M., Befani, O., Palumbo, M., . . . Turini, P. (2003). Simple, Potent, and Selective Pyrrole Inhibitors of Monoamine Oxidase Types A and B. *Journal of Medicinal Chemistry*, 46(6), 917-920. doi:10.1021/jm0256124
- Tatton, W. G. (1996). Modulation of gene expression rather than monoamine oxidase inhibition: (-)-Deprenyl-related compounds in controlling neurodegeneration. *Neurology*, 47(Issue 6, Supplement 3), 171S-183S. doi:10.1212/wnl.47.6_suppl_3.171s
- Torres, P. H. M., Sodero, A. C. R., Jofily, P., & Silva-Jr, F. P. (2019). Key Topics in Molecular Docking for Drug Design. 20(18), 4574.
- Vlachakis, D. (2018). Introductory Chapter: Molecular Docking - Overview, Background, Application and What the Future Holds. In *Molecular Docking: InTech*.
- Yamada, M. (2004). Clinical Pharmacology of MAO Inhibitors: Safety and Future. *NeuroToxicology*, 25(1-2), 215-221. doi:10.1016/s0161-813x(03)00097-4