Improved Molecular Docking of MAO-B Inhibitors with Glide

Emilio Mateev 1,*, Maya Georgieva 1, Alexander Zlatkov 1

1 Department of Pharmaceutical Chemistry, Faculty of Pharmacy, Medical University-Sofia, Bulgaria; e.mateev@pharmfac.mu-sofia.bg (E.M.), mgeorgieva@pharmfac.mu-sofia.bg (M.G.), azlatkov@pharmfac.mu-sofia.bg (A.Z.)
* Correspondence: e.mateev@pharmfac.mu-sofia.bg (E.M.);

Abstract: Molecular docking is an essential structure-based technique that is frequently used for virtual screenings. In most cases, the precision of search algorithms of the modern docking programs is satisfactory. However, the scoring algorithms' accuracy is not sufficient to completely distinguish true positives from inactive solutions. Therefore, initial validation processes could elevate the docking enrichments. This paper proposes a validated molecular docking approach applying Glide’s HTVS mode to discover novel Monoamine Oxidase B (MAO-B) ligands. In addition, more precise docking modes (SP and XP) and the influence of free binding energy calculations (MM/GBSA) over the enrichment values were observed. The developed Glide’s docking protocol demonstrated good enrichment values, outperforming GOLD 5.3. Moreover, the XP docking mode displayed the most reliable results, while the utilization of SP and MM/GBSA simulations led to poor performances. Overall, we demonstrated that virtual screenings of large databases in 2V5Z utilizing HTVS mode followed by XP rescoring are suitable molecular docking protocols for acquiring novel MAO-B inhibitors.

Keywords: Glide; molecular docking; MAO-B; MM/GBSA; database enrichment.

© 2022 by the authors. This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Recently, molecular docking has emerged as a rapidly growing structure-based drug design (SBDD) technique, which could be utilized for hit identification and/or lead optimization [1]. In general, the docking software applies search algorithms, which generate numerous ligand poses in the receptor's active site and scoring functions that calculate the energies of the produced solutions [2]. A detailed explanation of the different docking algorithms and software types is presented elsewhere [3]. The utilization of molecular docking in the virtual screenings is a well-described process; however, initial validation of the docking software is required to observe the ability of the docking program to differentiate false positives from true positives [4].

Monoamine oxidases (MAOs) are enzymes involved in the deactivation of aliphatic and aromatic amines. The former system encompasses two similar isoforms – MAO-A and MAO-B. However, it has been found that there are significant differences in the types of active amino residues [5]. The main focus of the current study is the structure of MAO-B. After its initial crystallographic determination, it was observed that MAO-B comprises three functional domains – entrance cavity, substrate cavity, and aromatic cage (Figure 1).
Virtual screenings in the active site of MAO-B is not a heavily examined field, considering the low enrichment values, hence the inability of the docking software to distinguish actives from decoys [6]. Recently, our research group has obtained relatively poor results from the dataset enrichments with all GOLD 5.3 scoring functions [7]. In addition, unsatisfactory results were obtained when ensemble dockings with the former software were carried out [8].

Several papers had described high experimental correlations when free binding energy calculations with MM/GBSA and MM/PBSA were utilized [9,10]. Moreover, in some cases, the docking capacity of Glide has been described as higher when compared to other docking software [11,12]. Therefore, we applied Glide and the MM/GBSA recalculations to evaluate the enrichment factor of DUD-E dataset towards MAO-B crystallographic structures.

To explore the capability of Glide to place MAO-B inhibitors in top ranks correctly, detailed dataset enrichment simulations are required. Therefore, this study aims to conduct the re-docking of various co-crystallized ligands with further enrichment calculations of the most prominent PDB receptors. Moreover, we employed and evaluated the effects of Standard precision (SP), Extra precision (XP), and Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) calculations against a fragment of the dataset.

2. Materials and Methods

2.1. Molecular docking/methodology.

The docking simulations were conducted with Glide (Schrödinger, LLC, New York, NY, 2021) on the operating system – Windows 10 Pro; CPU – AMD Ryzen 5 3600 6-core 3.60 GHz, GPU – GeForce GTX 1060 3 GB and install memory (RAM) – 16 GB. Glide algorithms are based on systematic searches in the active site of the receptor [13]. The former software generates numerous conformations for each ligand. The ligands go through hierarchical filters during docking, which initially test the complementary between the ligand-receptor complex. Ligands that pass this stage undergo energy minimization and are then scored. Three different docking methodologies are embedded in Glide - High throughput virtual screening (HTVS), Standard Precision (SP), and Extra precision (XP). Standard-precision (SP) is the default option that poses moderate accuracy and speed. The high-throughput virtual screening (HTVS) is a
high-speed option, with an approximate time range of 2 seconds. It is applicable for the preliminary screening of large databases. Extra-precision (XP) docking is introduced to reduce the false-positive results with more extensive sampling.

2.2. Protein preparation.

22 highly resolution crystallographic structures of MAO-B with the following codes: 1OJA, 1OJC, 1OJ9, 1S3B, 2BK3, 2BYB, 2C65, 2VRM, 2VRL, 2VZ2, 2V5Z, 2XFU, 3PO7, 4A7A, 4A79, 4CRT, 5MRL, 6FVZ, 6FWC, 6FW0, 6RKB, and 6RLE were downloaded from the Protein Data Bank (https://www.rcsb.org/). All the target structures were further refined for docking by applying the Protein Preparation Wizard available in Maestro [14]. Initially, complexes comprising a Heme group or covalent interactions between protein and ligand were deleted. Furthermore, hydrogen bonds and het states were generated, water molecules beyond 5Å from het groups, and all monomers except monomer A were removed. Finally, the amino acids were ionized at pH 7, and the complexes were minimized utilizing the OPLS2005 force field. The grid box was generated around each co-crystallized ligand with the receptor grid generation module of Maestro [13]. No constraints and no flexible side chains were included in the docking protocol.

2.3. Databases.

To evaluate the robustness of GOLD’s ChemPLP and Glide’s SP scoring functions in finding known inhibitors embedded in random decoys, we applied a dataset downloaded from DUD-E [15]. The former comprised 7000 decoys and 169 actives. The decoys are molecules with similar molecular weights, LogP, and rotatable bonds compared to the true inhibitors; however, the topological characteristics are dissimilar. Importantly, the decoy molecules are not experimentally tested. Therefore, some of them could display actual activities. Several limitations of the DUD dataset have been reported [16]; however, the latter still abides as a reliable benchmark database, especially after the occurrence of the improved version DUD-E [17].

2.4. Ligands preparation.

The 3D ligands were prepared utilizing the LigPrep module (Schrödinger Release 2021-2: LigPrep, Schrödinger, LLC, New York, NY, 2021.). Initially, hydrogen bonds were added, charged groups were neutralized, and tautomeric and ionization states were included. Thereafter, the energy was minimized using the OPLS2005 force field.

2.5. Re-docking simulations.

The purpose of the re-docking simulations was to identify the top MAO-B receptors which can be subsequently used for enrichment screenings. Root-mean-square deviation (RMSD) values were calculated to quantify each docking program's difference between the crystal ligand coordinates and the predicted coordinates. We utilized an RMSD threshold of 2 Å as opposed to the recently applied 1.5 Å [18] regarding the relatively high number of freely rotatable bonds in most of the co-crystallized ligands. The RMSD values of the best-ranked solutions were considered. No energy minimizations of the ligands were conducted before the re-docking simulations.
2.6. MM/GBSA.

To assess the free binding energy of the docked complexes, we applied Molecular Mechanics/Generalized Born Surface Area (MM-GBSA) recalculations with Prime [19]. MM/GBSA has been frequently utilized to evaluate docking poses, determine structural stability, and predict binding affinities [20]. The latter method incorporates the OPLS3 force field and VSGB dissolvable model to carry out the calculations [21].

2.7. Evaluation metrics.

The ideal virtual screening protocol could distinguish true-positive ligands from an inactive dataset of decoys and accurately predict the activities of the compounds by ranking the actives at the top rank positions. To validate the performance of Glide and to compare it with other previously published docking protocols, we applied two established metrics: the enrichment factor (EF) and the area under the Receiver Operating Characteristic (ROC) curve [22].

The enrichment factors consider the number of active molecules located in a specific [23]. The latter is enhanced when more actives are located in the examined fraction of the dataset. The formula of the classic EF is given below:

\[
\text{EF} = \frac{HITS_{\text{sampled}}}{HITS_{\text{total}}} \times \frac{N_{\text{sampled}}}{N_{\text{total}}}
\]

In the formula above, \(HITS_{\text{sampled}}\) stands for the active compounds which are detected in the chosen percentage of the dataset. \(HITS_{\text{total}}\) is the count of all active structures seeded in the decoys, which in our case equals 169. \(N_{\text{total}}\) is the value of the docking dataset, while \(N_{\text{sampled}}\) is the percentage of the dataset which is being observed. To examine early enrichments, we applied 1% and 5% of the benchmark set. Percentages over five are applicable when high-throughput screenings are feasible.

A notable drawback of the classical EF is that it does not consider the rankings of the active ligands. Therefore, we included calculations of a modified enrichment factor, which contemplates the fitness scores of the obtained active molecules. \(\text{EF}^*(x)\) is defined below:

\[
\text{EF}^*(x) = \frac{50\%}{\text{APR}_{\text{sampled}}} \times \frac{\text{Hits}_{\text{sampled}}}{\text{Hits}_{\text{total}}}
\]

Here \(x\) is the percent of the explored active compounds; \(\text{APR}_{\text{sampled}}\) is the “average percentile rank” of \(\text{Hits}_{\text{total}}\). In this study, we calculated the \(\text{EF}^*\) value of 5% of the dataset.

The receiver operating curve (ROC) represents the relationship between the selectivity (Se) and specificity (Sp) for a range of continuous values. Overall, the ratio of true positives in the function of the false positives is observed. The area-under-curve (AUC) of ROC was utilized as a quantification measure. Values under 0.50 indicate a random selection, whereas AUC value of 1.0 represents the perfect distribution of the active molecules in the simulated dataset.

3. Results and Discussion

3.1. Workflow.

In the current paper, high throughput virtual screening (HTVS) simulations were applied, followed by the more precise Glide’s docking modes – SP and XP, and free binding energy recalculations with MM/GBSA. The enrichments obtained after utilizing all scoring methods were evaluated and discussed.
3.2. Redocking simulations.

Initial validation procedures for each docking software should be conducted, considering the latter programs' inconsistent searching and scoring algorithms [24]. The re-docking procedure is one of the most commonly applied internal validation procedures, which test the accuracy of the search and score algorithms implemented in the docking software. During these simulations, and experimentally resolved conformation of a co-crystallized ligand is docked back into the active site of a receptor, and RMSD values between the initial conformation and the docked poses are determined [18].

![Figure 2. Glide re-docking simulations of 22 MAO-B proteins. The native co-crystallized conformation is denoted in green, while the conformation obtained with Glide is in red.](https://biointerfaceresearch.com/)

To test the accuracy of Glide against MAO-B crystallographic structures, each co-crystallized ligand of 1OJA, 1OJC, 1OJ9, 1S3B, 2BK3, 2BYB, 2C65, 2VRM, 2VRL, 2VZ2, 2V5Z, 2XFU, 3PO7, 4A7A, 4A79, 4CRT, 5MRL, 6FVZ, 6FWC, 6FW0, 6RKB, and 6RLE PDB structures was docked back into its binding cavity. The top-scored docking solutions were
compared with the experimentally resolved active conformations of the respective ligands. For this study, a threshold of 2 Å was set; therefore, complexes comprising values over that number were considered incorrect. In Figure 2 we provided the native co-crystallized poses superimposed with the conformations retried from Glide. In addition, the RMSD values of each solution are denoted.

Of the 22 MAO-B crystallographic structures, 68% displayed good re-docking RMSD values under 2Å. In three cases (10JA, 2V5Z, and 3PO7), the superimpositions after the re-docking simulations displayed values under 1Å. The top eight MAO-B crystallographic structures were considered in the following validation stages.

An in-depth re-docking and enrichment study of MAO-B receptors with Glide's docking program has not been reported so far. However, a previous study of our research group has demonstrated that GOLD 5.3 provides good re-docking RMSD values for the following PDB structures: 1OJ9, 1S3B, 2BK3, 2V5Z, 3PO7, 4A7A, 4A79, 6FVZ, 6FW0 and 6FWC [7]. When compared to the results acquired in the current study, we noticed a significant difference between GOLD 5.3 and Glide re-docking simulations in one case. In particular, the latter could not correctly place the co-crystallized ligand of the PDB structure 2BK3. Another study conducted in 2007 demonstrated that DOCK 6.0 could correctly place the co-crystallized ligands of 1OJ9, 1OJC, 1OJD, and 2BK3 [25]. The first two PDB codes demonstrated good redocking results in the current study.

Overall, the re-docking simulations should not be considered fully reliable, especially when the further screened and examined structures are chemically dissimilar from the co-crystallized core moieties [26]. In the latter case, implementing the decoy dataset could provide good preliminary information about the robustness of the docking software.

3.3. Database enrichment.

3.3.1. HTVS docking.

Performance assessment in virtual screening has been the subject of vigorous debate, particularly over the past several years; thus, we measured both overall and early screening enrichments and the area under the ROC curve. We took the top eight MAO-B crystallographic structures from the re-docking analysis and conducted database enrichment in this validation step. A group of 169 active MAO-B inhibitors seeded in 7000 decoys (inactive ligands) taken from the DUD-E website were selected to investigate the ability of Glide to distinguish true-positive from inactive ligands. Initially, the fastest docking mode in Glide (HTVS) was employed. After that, the influence of the more accurate SP, XP docking options was observed together with the free binding energy recalculation (MM/GBSA).

<table>
<thead>
<tr>
<th>Receptor</th>
<th>Num. of actives in top 1% of results</th>
<th>EF 1%</th>
<th>EF 5%</th>
<th>EF(5)</th>
<th>AUC-ROC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1OJC</td>
<td>18</td>
<td>15</td>
<td>6.4</td>
<td>6.1</td>
<td>0.49</td>
</tr>
<tr>
<td>1S3B</td>
<td>9</td>
<td>7.4</td>
<td>3.9</td>
<td>5.4</td>
<td>0.59</td>
</tr>
<tr>
<td>2C65</td>
<td>8</td>
<td>6.6</td>
<td>3.8</td>
<td>5.7</td>
<td>0.66</td>
</tr>
<tr>
<td>2V5Z</td>
<td>15</td>
<td>12</td>
<td>4.2</td>
<td>9.2</td>
<td>0.74</td>
</tr>
<tr>
<td>3PO7</td>
<td>5</td>
<td>4.1</td>
<td>3.1</td>
<td>3.5</td>
<td>0.47</td>
</tr>
<tr>
<td>4A7A</td>
<td>7</td>
<td>5.8</td>
<td>4.1</td>
<td>5.8</td>
<td>0.65</td>
</tr>
<tr>
<td>4A79</td>
<td>11</td>
<td>9</td>
<td>4.9</td>
<td>7.5</td>
<td>0.70</td>
</tr>
<tr>
<td>6RLE</td>
<td>6</td>
<td>4.9</td>
<td>4.3</td>
<td>6.2</td>
<td>0.73</td>
</tr>
</tbody>
</table>
Table 1 provides the number of active ligands located in 1% of the database, the enrichment values at 1% and 5% of the dataset, the modified enrichment factor, and the AUC-ROC value.

We examined the modified enrichments and the ROC values to compare the PDB structures. 2V5Z, 4A79, and 6RLE demonstrated the highest capability of distinguishing the actives from the decoys. An interesting case was noted after the virtual screening of 1OJC. Glide successfully obtained 18 actives in the top 1% of the database; however, the modified enrichment showed poor performance when the rank positions of the active ligands were considered. Therefore, 1OJC was not further evaluated. We observed that the most prominent MAO-B crystallographic structure was with PDB code 2V5Z. The latter was accomplished to accommodate 15 actives in the 1% of the dataset. Moreover, the modified enrichment equaled 9.2, while the AUC-ROC value was 0.74.

In a recent study, we assessed the ability of GOLD 5.3 to correctly score active MAO-B inhibitors in a dataset of decoys utilizing most of those above crystallographic MAO-B structures [7]. After re-docking, cross-docking, and dataset enrichments, it was found that 1S3B was the most prominent crystallographic MAO-B structure, which could, to some extent, successfully accommodate a diverse set of active inhibitors. The results above demonstrated that 1S3B is not a suitable choice when Glide's virtual screening is carried out. Undesirable modified enrichment of 5.4 and ROC value of 0.59 were examined, placing the latter protein in sixth place. In addition, ensemble docking simulations with various MAO-B proteins applying GOLD 5.3 as a docking program were also analyzed by our research group [8]. The best ensemble of 1S3B-1OJA-1OJC MAO-B structures demonstrated EF 1% values of 7.25, which is drastically lower than the Glide above results. In the current paper, we explored that the PDB structure 2V5Z could achieve enrichment of 12 when Glide’s HTVS mode is utilized.

3.3.2. SP docking.

In the next stage of our study, we evaluated the influence of Glide's more precise SP docking mode on the dataset enrichment. Three PDB structures that demonstrated the highest enrichments from the previous simulations were applied in this stage (Table 2). The entire dataset of ligands (7000+ molecules) was examined.

<table>
<thead>
<tr>
<th>Receptors</th>
<th>Num. of actives in top 1% of results</th>
<th>EF 1%</th>
<th>EF 5%</th>
<th>EF (5)</th>
<th>AUC-ROC</th>
</tr>
</thead>
<tbody>
<tr>
<td>2V5Z</td>
<td>8</td>
<td>6.6</td>
<td>3.9</td>
<td>5.9</td>
<td>0.67</td>
</tr>
<tr>
<td>4A79</td>
<td>4</td>
<td>3.3</td>
<td>4.3</td>
<td>4.1</td>
<td>0.53</td>
</tr>
<tr>
<td>6RLE</td>
<td>5</td>
<td>4.1</td>
<td>3.1</td>
<td>4.1</td>
<td>0.62</td>
</tr>
</tbody>
</table>

Analyzing the obtained results, we noted that the faster HTVS function in Glide demonstrated significantly better enrichments when compared to the more precise SP function. Similar findings have been discussed in a recent paper by Pandey et al. [27]. The latter work has reported that the less time demanding mode in Glide - SP, exerted better results when compared to XP docking in the screening of antileishmanial drugs.

3.3.3. MM/GBSA and XP simulations.

As a final stage of our study, we examined if free binding energy recalculation with MM/GBSA and docking with the most precise mode of Glide - XP could positively affect the
ranks of the active molecules. Several free binding energy calculation methods, such as molecular mechanics generalized Born surface area [28], have been implemented for free binding energy calculations since they are more accurate and higher enrichments could be achieved [29]. In many cases, the utilization of both molecular docking and MM/(PB)GBSA rescoring has proven to be a promising strategy in identifying correct binding poses and reliable rankings of ligands [20,30]. However, a significant increase in computer power is the major drawback of conducting the latter calculations. Considering the significant increase in the hardware power, we considered only the top 1000 ranks from the previously obtained HTVS results. The number of actives found in the top 20 positions with HTVS, XP, and MM/GBSA against the receptors 2V5Z, 4A79, and 6RLE, are provided in Table 3.

Table 3. The number of active ligands detected in the top 20 ranks after utilizing HTVS, XP, and MM/GBSA simulations.

<table>
<thead>
<tr>
<th>Receptors</th>
<th>HTVS docking</th>
<th>XP docking</th>
<th>MM/GBSA</th>
</tr>
</thead>
<tbody>
<tr>
<td>2V5Z</td>
<td>10</td>
<td>15</td>
<td>6</td>
</tr>
<tr>
<td>4A79</td>
<td>5</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>6RLE</td>
<td>3</td>
<td>6</td>
<td>3</td>
</tr>
</tbody>
</table>

The results demonstrated that the XP mode significantly increased the number of active ligands in the top 20 ranks when 2V5Z and 6RLE were employed. Importantly, when 2V5Z was utilized, the XP docking mode was able to place five additional active ligands compared to the HTVS docking. We noted the weak performance of the free binding energy recalculations when 2V5Z, 4A79, and 6RLE were observed. In all cases, the results obtained with HTVS were equal or better when compared to the MM/GBSA recalculations.

4. Conclusions

Molecular docking is emerging as a frequently applied technique in discovering novel and effective hits. However, in most cases, a major drawback is the inability of the docking software to distinguish true positives from inactive molecules. This work demonstrated that Glide can identify most of the active MAO-B inhibitors in a dataset of decoys. In addition, the most prominent MAO-B crystallographic structure was with PDB code 2V5Z. Interestingly, when the more precise and hardware demanding SP mode of Glide was applied, the enrichment values were drastically reduced. The recalculations of the binding free energy with MM/GBSA demonstrated poor results. Overall, the best docking enrichment was obtained when the XP mode was utilized towards the PDB structure 2V5Z. Further consensus docking of MAO-B proteins could be employed to search for even better results.

Funding

This work was supported by the Medical Science Council of the Medical University of Sofia, Contract No 100/04.06.2021, Project No 7901/19.11.2020.

Acknowledgments

This research has no acknowledgment.

Conflicts of Interest

The authors declare no conflict of interest.
References


28. Chen, J.; Zhang, S.; Wang, W.; Sun, H.; Zhang, Q.; Liu, X. Binding of Inhibitors to BACE1 Affected by pH-Dependent Protonation: An Exploration from Multiple Replica Gaussian Accelerated Molecular Dynamics and MM-GBSA Calculations. *ACS Chemical Neuroscience* 2021, 12, 2591-2607, [https://doi.org/10.1021/acschemneuro.0c00813](https://doi.org/10.1021/acschemneuro.0c00813).
