Improved Virtual Screening Performance through Docking Scoring Fusion in the Discovery of Dual Target Ligands for Parkinson’s Disease

Yunierkis Perez-Castillo*,1,2, Aliuska Morales Helguera2, M. Natália D. S. Cordeiro3, Eduardo Tejera4, Cesar Paz-y-Miño4, Aminael Sánchez-Rodriguez5, Fernanda Borges*,6, Maykel Cruz-Monteagudo4,6

1 Sección Físico Química y Matemáticas, Departamento de Química, Universidad Técnica Particular de Loja, San Cayetano Alto S/N, EC1101608 Loja, Ecuador;
2 Molecular Simulation and Drug Design Group, Centro de Bioactivos Químicos (CBQ), Central University of Las Villas, Santa Clara, 54830, Cuba;
3 REQUIMTE, Department of Chemistry and Biochemistry, Faculty of Sciences, University of Porto, 4169-007 Porto, Portugal;
4 Instituto de Investigaciones Biomédicas (IIB), Universidad de Las Américas, 170513 Quito, Ecuador;
5 Departamento de Ciencias Naturales, Universidad Técnica Particular de Loja, Calle París S/N, EC1101608 Loja, Ecuador;
6 CIQUP/Departamento de Química e Bioquímica, Faculdade de Ciências, Universidade do Porto, Porto 4169-007, Portugal.

* Author to whom correspondence should be addressed; E-Mail: yperez@utpl.edu.ec (YPC);
  fborges@fc.up.pt (FB)
  Tel.: +351220402502; Fax: +351220402659.

Published: 4 December 2015

Abstract: Virtual methodologies have become essential components of the drug discovery pipeline. Specifically, structure-based drug design methodologies exploit the 3D structure of molecular targets to discover new drug candidates through molecular docking. Recently, dual target ligands of the Adenosine A2A Receptor and Monoamine Oxidase B enzyme have been proposed as effective therapies for the treatment of Parkinson’s disease. To the best of our knowledge, no theoretical study has been devoted to developing structure-based virtual screening methodologies for the discovery of dual A2AAR antagonists and MAO-B inhibitors. In this communication we propose a structure-based methodology for the discovery this type of molecules.
Keywords: Parkinson’s disease, Dual target ligands, Molecular docking, Scoring fusion.

1. Introduction

During the last decades, Virtual Screening (VS) methodologies have emerged as efficient alternatives to the expensive, in terms of time and money, High Throughput Screening (HTS) approaches for the discovery of new drug candidates [1]. In terms of efficiency, the hit rates obtained when VS tools are employed to filter large databases of chemical compounds are considerably higher than those obtained with HTS techniques [2]. Literature reports where VS experiments conducted to the identification of hit molecules in a wide range of application can be found elsewhere [3,4].

VS techniques can be divided into two main categories: Structure-Based VS (SBVS) and Ligand-Based VS (LBVS) [5]. The first one includes all the modeling approaches such as Molecular Docking and Molecular Dynamics that depend on the structure of a molecular receptor. This kind of approach uses the three-dimensional structure of the receptor to study its interactions with a set of putative ligands. Depending on the amount of ligands to study and the available computational resources a more or less detailed representation of the receptor-ligands interactions is chosen to estimate the stability of the receptor-ligand complexes. The computed scores serve then to order the investigated compounds from higher to lower probabilities of binding to the receptor [6].

One of the factors that negatively affect the performance of Molecular Docking SBVS studies is the accuracy of the scoring functions. Given that no scoring function can capture all the information relevant for the receptor-ligand binding process, the fusion of different scoring functions has been proposed as an alternative to improve the performance of SBVS methods [7]. These applications range from general ones intended for obtaining the best consensus strategy for any SBVS problem [8,9] to others proposed for specific researches [10,11]. In all these reports the proposed ensemble (fusion) methods outperform the VS performance obtained with a single scoring fusion.

In addition, usually SBVS methodologies are evaluated employing only a small set of decoy molecules. In the case of the standard DUD-E database only 50 decoy molecules can be selected per ligand [12]. This ligands/decoys proportion is far from what is observed in a real screening scenario where the ratio of active molecules is ranges from 0.01 to 0.14% [13]. To address this situation we have previously proposed a home-made algorithm for the generation of larger decoys sets resembling the ligands/decoys ratio of a real screening campaign [14].

On the other side, Parkinson Disease (PD) causes chronic disability and it is the second commonest degenerative condition of the nervous system. The standard treatment for PD is levodopa, which helps to increase the dopamine levels in the brain[15]. However there is a need of finding alternative therapies since levodopa has many side effects and can become ineffective over time. To this end, multicomponent therapies (combination of different drugs) have been used. However the discovery of new multi-target drugs (a single molecule that acts on multiple targets) is attracting more and more attention[16]. Multi-target drugs, compared with the use of combinations of different drugs, have more predictable pharmacokinetic and pharmacodynamic relationships as a
consequence of the administration of a single drug [17].

Antagonists of A2A adenosine receptors (A2AAR) with monoamine oxidase B (MAO-B) inhibitory activity are a class of promising dual-target drugs for PD [18]. Thus, there is a need for developing novel and diverse drugs, antagonists of A2AAR with MAO-B inhibitory activity for PD.

In this report we propose a structure-based methodology, which is extensively validated, for the discovery this type of molecules. The proposed methodology involves the molecular docking to both A2AAR and MAO-B of a set of 25744 molecules containing 16 known dual target ligands and decoy molecules. The obtained docking poses are rescored using six different scoring functions for the two molecular targets. Then we investigate several aggregation schemes with the objective of maximizing the enrichment of known ligands at the beginning of the ranked list they produce. Finally, we show that the developed methodology provides high values of enrichment of known ligands, which outperform that of the individual scoring functions. At the same time, the obtained ensemble can be translated in a sequence of steps that should be followed to maximize the enrichment of dual target dual A2AAR antagonists and MAO-B inhibitors.

2. Results and Discussion

The receptors, ligands and decoys were prepared for molecular docking calculations as described in the Materials and Methods section. The validation dataset consisting of the combination of the 16 known dual MAO-B inhibitors– A2AAR Antagonists and decoy molecules was docked to both receptor structures following the protocol described in the Materials and Methods.

In all cases analyzed from here on, the best molecular docking protocol was selected as the scoring scheme providing the highest value of BEDROC among those achieving the maximum EF at three different selection sizes (1, 5 and 10 percent of screened data). We separately analyzed the results obtained for the raw and weighted by number of heavy atoms scores. Scoring schemes were produced by fusing the ranks derived from the scoring functions using either arithmetic or geometrical mean as described in the Materials and Methods section.

Different Fusion Schemes (FS) were assayed in this investigation and they can be classified into two groups. The first group consisted in fusing the scoring functions maximizing the enrichment of dual ligands for the A2AAR and MAO-B enzyme separately. The application of the optimal scoring scheme of each target yields one fused ranking of compounds for each one. Then these two fused ranks were aggregated in one final rank. By employing this first fusion scheme we ensure that the final ranking will be based upon information derived from both the A2AAR and the MAO-B enzyme. This fusion scheme will be referred as Fusion Scheme 1 (FS1) from here on.

The second group consisted in evaluating the performance of all possible ensembles resulting from all possible combinations of the individual scoring functions ranks obtained for both targets at the same time. Since the number of scoring functions employed in this study is small, it was possible to evaluate all their possible combinations of size 1 to 2N, being N the number of computed scoring functions per target. For this second approach no constrain is imposed during the modeling process regarding the need of information from both targets in the final
ensemble. Therefore, there is the possibility that, in opposition to the expected behavior, the best performing ensemble would contain information from only one of the two molecular targets. This fusion scheme will be referred as Fusion Scheme 2 (FS2) from here on.

As mentioned before, we tested the arithmetic and geometric means as fusion operators. FS1 contains three aggregation steps: the aggregation of $A_{2A\text{AR}}$ scoring functions, the aggregation of MAO-B scoring functions and the aggregation of the rankings obtained for both targets. In this case all possible combinations of both fusion operators were tested. That is, scoring functions were first aggregated using the same fusion operator, either arithmetic or geometric mean, for each target separately. Then in the second step the aggregated ranking for each target was fused using both aggregation operators. Considering that the aggregation experiments are conducted with the raw scores and with the scores weighted by number of heavy atoms, the proposed setup provides eight different variants of FS1. These variants are summarized in Table 1.

For FS2, since the scores derived for both targets are considered together, there is only one rankings fusion step. Thus, considering that we studied the raw scores and the scores weighted by number of heavy atoms, four different setups were assayed. The different FS assayed in this scenario are summarized in Table 2.

For each known ligand, 1607 decoys were selected following the procedure described in our previous publication [14]. This amount of decoys provides a ratio of active to decoy compounds of 0.06%, which resembles a real screening scenario [13]. In this case the maximum EF that any of the individual scoring functions can achieve is when the selection size is set to 1% of the ranked list is 6.23. The best performing FS is presented in Table 3.

Our results show that the best scoring schemes are obtained when the docking scores to MAO-B and $A_{2A\text{AR}}$ are considered together during the scoring fusion procedure and scores fusion using arithmetic mean provides better results than fusion using geometrical mean. Also, in all the examined cases the best scoring scheme obtained with the raw docking scores outperforms that obtained fusing the scores weighted by the number of heavy atoms. It should also be noted that in almost every case, the best enrichment is derived from more than one scoring function through their fusion.

For the current validation setup the maximum values that the EF can reach are 100, 20 and 10 when 1%, 5% and 8% of screened data are selected respectively. Taking this into consideration it can be seen that if 1% of the screened data is selected for further analyses the resulting virtual screening protocol is able of achieving 31.17% of the theoretical maximum enrichment. Following the same reasoning, when the 5% and 8% of screened data are selected the corresponding virtual screening tools achieve 75% and 87.5% of the theoretical maximum of the EF respectively. Last but not least, the BEDROC values obtained for these virtual screening protocols are away from random (BEDROC=0). The accumulative curves corresponding to the three optimal virtual screening protocols are presented in Figure 1.

The obtained results provide a set of approaches from which we can select the optimal one for the virtual screening of databases of chemical compounds in the search of dual MAO-B inhibitors- $A_{2A\text{AR}}$ Antagonists. For example, if we were to select the appropriate virtual screening protocol for screening a database of chemicals and select 1% of data for further analysis, we should follow this procedure:
1. Dock the database to both MAO-B and A2AAR.
2. Select the best pose of each compound in each target according to the grid-based scoring function.
3. Rescore the best poses in A2AAR using the GB/SA Score, Continuous Score; Amber Score (everything rigid) and Amber Score (flexible ligand) scoring functions.
4. Rescore the best poses in MAO-B using the GB/SA Score and SA_Descriptor Score scoring functions.
5. Generate the individual ranking produced by the scoring functions GB/SA Score, Continuous Score; Amber Score (everything rigid) and Amber Score (flexible ligand) for A2AAR and GB/SA Score and SA_Descriptor Score for MAO-B.
6. Fuse the obtained individual rankings using arithmetic mean.

### Table 1. Variants of FS1 assayed

<table>
<thead>
<tr>
<th>Fusion Scheme</th>
<th>Scores Type</th>
<th>Target Scores Fusion</th>
<th>Final fusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>FS1.1</td>
<td>Raw</td>
<td>Arithmetic Mean</td>
<td>Arithmetic Mean</td>
</tr>
<tr>
<td>FS1.2</td>
<td>Raw</td>
<td>Arithmetic Mean</td>
<td>Geometric Mean</td>
</tr>
<tr>
<td>FS1.3</td>
<td>Raw</td>
<td>Geometric Mean</td>
<td>Arithmetic Mean</td>
</tr>
<tr>
<td>FS1.4</td>
<td>Raw</td>
<td>Geometric Mean</td>
<td>Geometric Mean</td>
</tr>
<tr>
<td>FS1.5</td>
<td>Weighted</td>
<td>Arithmetic Mean</td>
<td>Arithmetic Mean</td>
</tr>
<tr>
<td>FS1.6</td>
<td>Weighted</td>
<td>Arithmetic Mean</td>
<td>Geometric Mean</td>
</tr>
<tr>
<td>FS1.7</td>
<td>Weighted</td>
<td>Geometric Mean</td>
<td>Arithmetic Mean</td>
</tr>
<tr>
<td>FS1.8</td>
<td>Weighted</td>
<td>Geometric Mean</td>
<td>Geometric Mean</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Fusion Scheme</th>
<th>Scores Type</th>
<th>Target Scores Fusion</th>
<th>Final fusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>FS1.1</td>
<td>Raw</td>
<td>Arithmetic Mean</td>
<td>Arithmetic Mean</td>
</tr>
<tr>
<td>FS1.2</td>
<td>Raw</td>
<td>Arithmetic Mean</td>
<td>Geometric Mean</td>
</tr>
<tr>
<td>FS1.3</td>
<td>Raw</td>
<td>Geometric Mean</td>
<td>Arithmetic Mean</td>
</tr>
<tr>
<td>FS1.4</td>
<td>Raw</td>
<td>Geometric Mean</td>
<td>Geometric Mean</td>
</tr>
<tr>
<td>FS1.5</td>
<td>Weighted</td>
<td>Arithmetic Mean</td>
<td>Arithmetic Mean</td>
</tr>
<tr>
<td>FS1.6</td>
<td>Weighted</td>
<td>Arithmetic Mean</td>
<td>Geometric Mean</td>
</tr>
<tr>
<td>FS1.7</td>
<td>Weighted</td>
<td>Geometric Mean</td>
<td>Arithmetic Mean</td>
</tr>
<tr>
<td>FS1.8</td>
<td>Weighted</td>
<td>Geometric Mean</td>
<td>Geometric Mean</td>
</tr>
</tbody>
</table>

### Table 2. Variants of FS2 assayed

<table>
<thead>
<tr>
<th>Fusion Scheme</th>
<th>Scores Type</th>
<th>Target Scores Fusion</th>
<th>Final fusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>FS1.1</td>
<td>Raw</td>
<td>Arithmetic Mean</td>
<td>Arithmetic Mean</td>
</tr>
<tr>
<td>FS1.2</td>
<td>Raw</td>
<td>Arithmetic Mean</td>
<td>Geometric Mean</td>
</tr>
<tr>
<td>FS1.3</td>
<td>Raw</td>
<td>Geometric Mean</td>
<td>Arithmetic Mean</td>
</tr>
<tr>
<td>FS1.4</td>
<td>Raw</td>
<td>Geometric Mean</td>
<td>Geometric Mean</td>
</tr>
<tr>
<td>FS1.5</td>
<td>Weighted</td>
<td>Arithmetic Mean</td>
<td>Arithmetic Mean</td>
</tr>
<tr>
<td>FS1.6</td>
<td>Weighted</td>
<td>Arithmetic Mean</td>
<td>Geometric Mean</td>
</tr>
<tr>
<td>FS1.7</td>
<td>Weighted</td>
<td>Geometric Mean</td>
<td>Arithmetic Mean</td>
</tr>
<tr>
<td>FS1.8</td>
<td>Weighted</td>
<td>Geometric Mean</td>
<td>Geometric Mean</td>
</tr>
</tbody>
</table>

*a Fusion scheme identifier

*b Type of score the rankings are derived from, either the raw scores or the scores weighted by the number of heavy atoms

*c Fusion operator employed to fuse the rankings derived of each scoring function in each target

*d Fusion operator employed to aggregate the fused rankings obtained for each target
Table 3. Enrichment metrics for the best performing FS

<table>
<thead>
<tr>
<th>FS Method a</th>
<th>EF b</th>
<th>BEDROC c</th>
<th>AUAC d</th>
<th>Fused Scoring Functions e</th>
</tr>
</thead>
<tbody>
<tr>
<td>FS2.1</td>
<td>31.17</td>
<td>0.11</td>
<td>0.87</td>
<td>A2AAR: 3, 5, 6, 7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>MAO-B: 3, 4</td>
</tr>
</tbody>
</table>

a Employed fusion method. See Tables 2 and 3 for the detailed setup of each method  
b Enrichment Factor for the best scoring scheme.  
c BEDROC for the best scoring scheme. Alpha value is set to 160.9.  
d Area Under the Accumulative Curve for the best scoring scheme.  
e Scoring functions fused in the best scoring scheme. The following numbering is employed for scoring functions: 1) Grid Score; 2) PB/SA Score; 3) GB/SA Score; 4) SA_Descriptor Score; 5) Continuous Score; 6) Amber Score, everything rigid and 7) Amber Score, flexible ligand.

Figure 1. Accumulative curves obtained for the best virtual screening protocol when 1%, 5% and 8% of screened data are selected for further analysis. A) Complete curves. B) Curves for the first 10% of screened data.

3. Materials and Methods

Receptor preparation

The crystallographic structures of the A2AAR in complex with the antagonist ZM241385, PDB code 3PWH and of the MAO-B in complex with a coumarin inhibitor, PDB code 2V61, were obtained from the Protein Data Bank (www.wwpdb.org) database [19]. Receptor preparation was carried out with UCSF Chimera software.[20] During receptor preparation all water molecules and ligands were removed and
hydrogen atoms and charges were added. For both receptors the ligand binding pocket was defined as any residue lying at a distance below 5Å from the crystallographic ligand structure.

**Ligand preparation**

Sixteen known dual MAO-B inhibitors-A₂AR Antagonists were compiled from the literature [21,22]. Three dimensional conformers for the compounds were generated using the OMEGA software [23]. A maximum of 500000 conformations per molecule were generated using an energy window of 100 kcal/mol. All rotatable bonds were considered during the torsion search using the Merck Molecular Force Field (MMFF) and duplicate conformers were discarded based on a RMS value of 0.5 Å. A maximum number of 200 conformers were saved for each compound. Afterwards, AM1-BCC charges were added to each conformer using the MOLCHARGE programs that is part of the QUACpac package.[24].

**Decoy molecules selection**

Decoys selection was a based on a desirability-based home-developed algorithm that has been previously employed in the selection of tailored decoy sets for the validation of virtual screening strategies [14]. Decoys were prepared for molecular docking following the same protocol described above for ligands.

**Molecular Docking**

Molecular docking was performed with the DOCK v6.6 software.[25] A maximum of 2000 orientations per ligand was explored allowing a maximum of two bumps between the ligand and the receptor. Bumps were defined as any pair of atoms closer than the 75% of the sum of their Van der Waals radii. The energy grid-based scoring function was selected for poses quality evaluation. The pose with the lowest score for each ligand conformer was saved, allowing for a maximum of 200 saved poses.

For interaction energies calculation, a grid was pre-computed for the receptor binding pocket region. The grid spacing was set to 0.3 Å and the attractive and repulsive Van der Waals coefficients were set to 6 and 12 respectively. Calculations were performed considering an all-atoms model.

**Molecular docking post-processing**

The molecular docking protocol described above results in 200 docked conformations of each compound being saved. For every compound the best scored conformation was selected for further rescoring using six scoring functions implemented in DOCK. The scoring functions used for poses rescoring were: PB/SA Score, AMBER Score considering the whole complex as rigid, AMBER Score considering the ligand as flexible, Hawkins GB/SA Score and Solvent Accessible Surface Area (SASA) Score. These rescoring calculations plus the previous grid-based scoring employed for poses evaluation and selection provide seven different ways of evaluating the ligand-receptor interaction energies. In addition to the raw docking scores, the scoring value of each compound was weighted by the number of heavy atoms on it.

The seven computed scoring functions were used for the implementation of a consensus ranking scheme. Instead of combining the raw scoring values coming from different scoring functions, the ranks produced by these scoring functions were combined following the procedure described next. Firstly, the rank derived from each scoring function was produced. Then, for a specific combination of scoring functions, a fused rank was computed as either the arithmetic or geometric mean of the compound’s rank in the individual models.

To evaluate the performance of the developed models in a virtual screening scenario the
following metrics were computed: Area Under the Accumulation Curve (AUAC); Area under the Receiver Operating Characteristic Curve (ROC); Enrichment factor (EF) and Boltzmann-enhanced discrimination of ROC (BEDROC).[26] Here the same definitions proposed by Truchon et al. are used.[26]

4. Conclusions

We investigated different variants of docking scores fusion for maximizing the enrichment of dual target ligands of the Adenosine A2A Receptor and the Monoamine Oxidase B enzyme in virtual screening experiments. Our results show that for achieving high values of dual ligands enrichment, information relative to docking scores to both targets have to be combined. In addition, no single scoring function can be employed for achieving good virtual screening performance. Instead, combining the rankings derived from different scoring functions proved to be a valuable strategy for improving the enrichment relative to single scoring function in virtual screening experiments.

Acknowledgments

MC-M acknowledges the postdoctoral grant [SFRH/BPD/90673/2012] financed by the FCT – Fundação para a Ciência e a Tecnologia, Portugal, co-financed by the European Social Fund.

Conflicts of Interest

The authors declare no conflict of interests.

References and Notes


OpenEye Scientific Software, I. Omega, 2.3.2; Santa Fe, NM, USA, 2008.


© 2015 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions defined by MDPI AG, the publisher of the Sciforum.net platform. Sciforum papers authors the copyright to their scholarly works. Hence, by submitting a paper to this conference, you retain the copyright, but you grant MDPI AG the non-exclusive and un-revocable license right to publish this paper online on the Sciforum.net platform. This means you can easily submit your paper to any scientific journal at a later stage and transfer the copyright to its publisher (if required by that publisher). (http://sciforum.net/about ).