Article

3D Pharmacophore-based Virtual Screening, Docking Approaches toward the Discovery of Novel HPPD Inhibitors

Ying Fu¹, Yi-Na Sun¹, Ke-Han Yi¹, Ming-Qiang Li¹, Hai-Feng Cao¹, Jia-Zhong Li², Fei Ye*,¹

¹ Department of Applied Chemistry, College of Science, Northeast Agricultural University, Harbin 150030, China; fuying@neau.edu.cn
² School of Pharmacy, Lanzhou University, 199 West Donggang Rd., 730000 Lanzhou, China; lijiazhong@lzu.edu.cn

* Corresponding Author: Fei Ye, Email address: yefei@neau.edu.cn; Tel.: +86-451-55191507

Abstract: \(\text{p-Hydroxyphenylpyruvate dioxygenase (HPPD) is not only the useful molecular target in treating life-threatening tyrosinemia type I, but also an important target for chemical herbicides. A combined in silico structure-based pharmacophore and molecular docking based virtual screening were performed to identify novel potential HPPD inhibitors. The complex based pharmacophore model (CBP) with 0.721 of ROC used for screening compound showed remarkable ability to retrieve known active ligands from decoy molecule. The ChemDiv database was screened using CBP-Hypo2 as a 3D query, and the best-fit hits subjected to molecular docking with two methods of LibDock and CDOCKER in Accelrys Discovery Studio 2.5 (DS 2.5) to discern interactions with key residues at the active site of HPPD. 4 Compounds with top rank in HipHop model and well-known binding model were finally chosen as identification of lead compounds with potentially inhibitory effects on active site of target. The results provided powerful insight to the development of novel HPPD inhibitors herbicides using computational techniques.}

Keywords: HPPD inhibitors; pharmacophore model; molecule docking; HipHop model; virtual screening; ChemDiv

1 Introduction

Weeds compete with crops for sunshine, water, nutrients, and space, which influence the growth of crops and undermine both crop quality and yield. In agrochemical research, the discovery of novel high-activity and low-toxicity herbicide lead compounds is still remained as a challenge. 4-Hydroxyphenylpyruvate dioxygenase belongs to the non-heme Fe(II)-dependent dioxygenase family [1,2]. As an important enzyme correlated to the pigment synthesis and tyrosine catabolism in most organisms, HPPD is important in drug discovery in both agricultural and therapeutic areas [3-5]. HPPD catalyzes the conversion of 4-hydroxyphenylpyruvate (HPPA) to homogentisate (HGA), and this transformation involving decarboxylation, aromatic hydroxylation, and substituent migration in a single catalytic cycle is unique in nature [6]. In plants, HGA can be further transformed into tocopherol and plastoquinone, both of them are crucial for the normal growth of plants [7]. Inhibition of HPPD will block photosynthesis, which leads to the deficiency in isoprenoid redox cofactors such as plastoquinone and tocopherol, and finally causes growth inhibition, necrosis and death of treated plants [8-11].

Herbicides, which inhibit HPPD, represent one of the newest classes of herbicides available for use in crop production [12]. HPPD-inhibiting herbicides show many advantages, such as low application rate, low toxicity, broad-spectrum weed control (including herbicide-resistant weed biotypes), excellent crop selectivity and benign environmental effects [13,14]. Several of them are currently used as selective broad leaf herbicides including triketones, pyrazoles, isoxazoles, diketone nitriles and benzophenones over the last two decades. [15,16]. The triketone herbicides have spurred a variety of commercialized HPPD inhibitors through chemical modification, such as
sulcotrione, mesotrione and benzobicylon [17-19]. But the main problem associated with the use of herbicides is the occurrence of herbicide resistant weeds. Therefore, it is necessary to develop efficient herbicides with novel structures against HPPD.

The use of computational techniques in drug discovery and development has become the most effective methods. Among them, virtual screening is a conventional method used in drug discovery, which screen large collections of compounds to identify molecular structures that are most likely to bind into a particular biological target [20]. It has been reported that molecular docking, pharmacophore modeling, and structure-based virtual screening in drug discovery have been successful applied. Structure-based virtual screening has emerged as an efficient strategy in identifying potential a natural product-like STAT3 dimerization inhibitor from a database of natural product and natural product-like compounds, and molecular docking analysis suggested that compound 1 might putatively function as an inhibitor of STAT3 dimerization by binding to the SH2 domain [21]. The novel TLR1–TLR2 inhibitors were obtained through molecular docking from a database of natural product and natural product-like compounds and the results of activity experiments show that compound 1 was the most effective in inhibiting TNF-a and IL-6 secretion induced by Pam3CSK4 in RAW 264.7 cells [22]. High-throughput, ligand-docking based virtual screening methods were applied to identify small agents targeting menin–MLL binding from a natural product/natural product-like chemical database. From the activity assay, compound 1 which was tested in a bimolecular fluorescence complementation (BiFC) assay emerged as the top candidate for inhibiting menin–MLL interaction. Moreover, a high degree of shape complementarity is observed between compound 1 and the binding pocket of menin, suggesting that this protein–ligand interaction could also be stabilized by significant hydrophobic interactions [23]. Rutin, as a promising lead compound, would be further developed into an antidyslipidemic molecule as a good alternative to statins using a docking-based strategy and MD stimulation [24]. Metadynamics-based protocol was developed to investigate the unbinding mechanism of an inhibitor of the pharmacologically relevant target p38 MAP kinase. The result of calculations showed that the salvation of the ligand and of the active site played crucial roles in the unbinding process and demonstrating that metadynamics could be a powerful tool in designing new drugs with engineered binding/unbinding kinetics [25].

The virtual filtered strategy graph is shown in Figure 1. The goal of this study is to identify the novel and potential structure of HPPD inhibition through 3D pharmacophore models based on the known crystal complex of HPPD (PDB ID: 1TFZ). CBP-Hypo2 with quality=0.721 (Fair) was selected as the best hypothesis, which included one hydrogen bond donor (HBD), one ring aromatic (RA) and two hydrophobic features (HY). Subsequently, the reliable pharmacophore hypotheses were used in virtual screening ChemDiv databases to identify potential HPPD inhibitors. The virtual screened hit compounds were then docked into active pocket of HPPD in DS2.5. Further, the selected screened hits were performed binding energy calculation and precision docking. 9 Compounds were obtained with good affinity. The 9 hits obtained were matched to HipHop model. Finally, 4 compounds displayed good match to ligand-based pharmacophore HipHop-hypo2.
2 Results and discussion

2.1 Pharmacophore model generation and validation

CBP that was generated based on protein 1TFZ and inhibitor DSA869 was used as a virtual screening model to discovery novel HPPD inhibitors. 22 active compounds and 38 inactive compounds were used as testing set to validate the receptor-ligand pharmacophore automatically. Six hypotheses were generated. As shown in Figure 2, CBP-Hypo2 desired Quality (0.712, Fair) was selected as the best hypothesis. One hydrogen bond donor (HBD), one ring aromatic (RA) and two hydrophobic features (HY) were regarded as the critical features of the model. During HipHop pharmacophore generation, six highly active inhibitors were selected from the literature to serve as training set (Table 1). 10 hypotheses were generated and ordered by ranking score. All 10 hypotheses ranked scores ranging from 81.54 to 72.16 (Table 1). The rank values and feature pharmacophore of HipHop-Hypo1 were same as HipHop-Hypo2.

ROC (Receiver Operating Characteristic) curve was used to evaluate the degree of false positivity of the model screening compound. The curve was obtained by plotting false positive rate for x-axis against true positive rate on y-axis in Figure 2. The accuracy of the test was shown by measuring the area under the curve (AUC). The result of the model represented with excellent AUC score of 0.721. For HipHop model, test composed of active compounds and inactive compounds was used evaluate the selective model. HipHop-Hypo2 was considered as the best chemical hypothesis due to the model better distinguishing active and inactive compounds. As can be seen from the Figure 3, 12 active compounds Fit value was above 2.0, rather than the inactive molecules distributed in blue area.
2.2 CBP Pharmacophore model-based virtual screening

Initially, the CBP model was used as a query to search the ChemDiv databases with 151047 compounds. Fit Value was a measure of the overlap between the features in the pharmacophore and chemical features in the molecule, which assisted in understanding the chemical meaning of the pharmacophore hypothesis [26]. According to Fit Value greater than 2.5, 1196 hit compounds mapping on to the pharmacophore model CBP-Hypo2 were retrieved, which included some compounds structurally similar to that of the existing HPPD inhibitors and some novel scaffolds. As shown in Figure 4, the obtained compounds were well matched to the CBP model and formed $\pi-\pi$ interaction with the Phe360 and Phe403 residues. Simultaneously, residues Pro259 and Met314 generated hydrophobic interaction with aromatic ring or methyl.
2.3 Molecular docking

In order to reduce the number of false positive screened virtual hits, docking analysis was performed at the active site of AtHPPD using DS2.5. The ligand in the protein 1TFZ was extracted and the hydrogen atom was added. The docking method was carried out applying two docking methods which were LibDock and CDOCKER. Subsequently, binding poses of the docking compound were compared with the ligand in the crystallographic complex and calculating the RMSD values were 0.74 and 0.55, respectively. As can be seen from the Figure 5, the ligands docked by two docking methods could be well aligned with the ligands in the crystallographic complex, so the two methods demonstrated the accuracy and reliability of the docking. These virtual 1196 molecules retrieved after pharmacophore-based screenings were subjected to receptor-based virtual screening by using LibDock methods. Docking experiments was applied to compare the binding affinities of known inhibitors with that of the screened hits and to rank the screened hits on the basis of interactions with amino acid residues of the active site. 287 Hit compounds were chosen that showed LibDockScore values above 129. Further, the selected screened hits were subsequently submitted for their binding energy value calculation and precision docking in the ‘Calculate Binding Energies module’ and ‘CDOCKER module’ of DS2.5, respectively. Finally, according to binding modes, binding affinity, 9 hits compounds with the highest docking score and lowest binding energy were selected as the target hits. The different significant chemical interactions, viz., Pi-alkyl, Pi-Pi, hydrogen bonds etc of the best hits have been presented in the following figures.

Figure 4. Mapping of each of the best hits to CBP-Hypo2. The color of pharmacophore feature, namely, HBD, RA and HY were pink, orange and cyan, respectively. (A) compound L503-0533 (B) compound G622-0791 (C) compound G883-0326 (D) compound G883-0470.
Figure 5. aligning the docked ligands with the ligands in the crystallographic complex. (A) The ligand by LibDock docking method (B) The ligand by CDOCKER docking method. docked ligands were green, the ligand in the crystallographic complex was yellow.

As shown in Figure 6, for compound L503-0533 molecule, Phe360 generated $\pi$-$\pi$ with benzene ring and Arg269 interacted with fluorine via hydrogen bond. Compound G622-0791 was found to fully embed into the active pocket in the Figure 7, and interact with amino acids Gln272, Phe398 and Lys400 via H-bonds, meanwhile, the two benzene rings formed two pairs of sandwiches interacting with Phe360 and Phe403 at the binding site. Compound G883-0470 formed $\pi$-$\pi$ stacking interaction with Phe398, Phe403 and Phe406 and generated hydrogen bond interaction with His287 and Phe398 as depicted in Figure 8. Compound G883-0326 formed $\pi$-$\pi$ stacking with benzyl ring of Phe398, Phe403 and Phe360. His287 interacted with carbonyl via hydrogen bond was shown in Figure 9. The compound G622-0791 was finally selected as the most potent HPPD inhibitor based on its least binding energy (-167.41 kcal/mol), the –CDOCKER score of this compound was -39.18 with a Fit Value (pharmacophore-based on CBP-Hypo2) of 2.97.

Figure 6. The receptor-ligand interaction of screening compounds L503-0533 with HPPD active site.
2.4 HipHop pharmacophore model-based virtual screening

9 Compounds obtained were matched to HipHop model in the figure 9, and the results were found that 4 compounds were well matched to the ligand-based pharmacophore HipHop-Hypo2 and all the colors of other 5 compounds with low fit values in the heat map were light blue.

Figure 7. The receptor-ligand interaction of screening compounds G622-0791 with HPPD active site.

Figure 8. The receptor-ligand interaction of screening compounds G883-0326 with HPPD active site.

Figure 9. The receptor-ligand interaction of screening compounds G883-0470 with HPPD active site.
Compound L503-0533 exhibited the highest matching value of 3.8. Finally, 4 new compounds with diverse scaffolds were as possible candidates for the designing of potent HPPD inhibitors (Table 1). The values of the four compounds were higher than those of the reference compound with Binding Energy, LibDock ENERGY, -CDOCKER ENERGY, Fit Value. Further investigations on these 4 compounds about testing in vitro and in vivo against HPPD are currently underway in our laboratories.

Figure 9. Heat map of the ten hypotheses from docked compounds and ligand of HPPD.

<table>
<thead>
<tr>
<th>Name</th>
<th>Structure</th>
<th>Binging Energy</th>
<th>LibDock Score</th>
<th>-CDOCKER ENERGY</th>
<th>Fit Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>crystallographic ligand</td>
<td><img src="image" alt="Structure" /></td>
<td>-68.857</td>
<td>130.542</td>
<td>21.08</td>
<td>2.35</td>
</tr>
<tr>
<td>L503-0533</td>
<td><img src="image" alt="Structure" /></td>
<td>-130.39</td>
<td>151.48</td>
<td>31.75</td>
<td>2.75</td>
</tr>
<tr>
<td>G622-0791</td>
<td><img src="image" alt="Structure" /></td>
<td>-167.41</td>
<td>138.71</td>
<td>39.18</td>
<td>2.97</td>
</tr>
<tr>
<td>G883-0326</td>
<td><img src="image" alt="Structure" /></td>
<td>-125.71</td>
<td>141.43</td>
<td>21.93</td>
<td>2.56</td>
</tr>
<tr>
<td>G883-0326</td>
<td><img src="image" alt="Structure" /></td>
<td>-133.97</td>
<td>138.34</td>
<td>22.73</td>
<td>3.02</td>
</tr>
</tbody>
</table>
3. Materials and methods

3.1 Data collection and preparation

The X-ray crystal structure of complex HPPD with an inhibitor (PDB code: 1TFZ) was downloaded from the RCSB Protein Data Bank (www.rcsb.org). The inhibitor was removed from the complex as an active ligand in building the CBP models.

Virtual screening using the ChemDiv database was performed to identify novel potential HPPD inhibitors. Based on the published literatures [27-29], 22 HPPD inhibitors with IC₅₀ less than 0.1nM were collected as the active ligands of CBP for verifying the model. Meanwhile, to validate the selectivity of the obtained pharmacophore models, 38 compounds randomly selected from the HPPD-decoys set were chosen as the inactive ligands.

Based on the known HPPD inhibitor [29-31] six typical compounds (Table 2) with bioactivity value IC₅₀ lower 0.1µM and molecular properties were identified as a training set to generate common feature based pharmacophore models of these inhibitors. 27 Compounds were selected as a test set, among, 12 compounds was HPPD inhibitors with .ligand file, 17 compounds randomly obtained from ZINC database were used as inactive molecules with .zinc file. All the molecules were prepared and optimized using SYBYL-X 2.0.

Table 2. Chemical structures and molecular properties of the training set compounds in HipHop model and screening compounds

<table>
<thead>
<tr>
<th>Structure</th>
<th>AlogP</th>
<th>Weight</th>
<th>Num-H Acceptors</th>
<th>Num-H Donors</th>
<th>Num-H Rotatable Bonds</th>
<th>Molecular Polar SurfaceArea</th>
<th>IC₅₀</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Structure 1" /></td>
<td>3.429</td>
<td>444.54</td>
<td>7</td>
<td>2</td>
<td>6</td>
<td>104.32</td>
<td>-</td>
</tr>
<tr>
<td><img src="image2.png" alt="Structure 2" /></td>
<td>2.53</td>
<td>329.23</td>
<td>5</td>
<td>0</td>
<td>4</td>
<td>97.03</td>
<td>0.28</td>
</tr>
<tr>
<td><img src="image3.png" alt="Structure 3" /></td>
<td>2.216</td>
<td>372.82</td>
<td>6</td>
<td>0</td>
<td>5</td>
<td>102.95</td>
<td>0.01</td>
</tr>
<tr>
<td><img src="image4.png" alt="Structure 4" /></td>
<td>2.827</td>
<td>369.39</td>
<td>7</td>
<td>2</td>
<td>4</td>
<td>137.48</td>
<td>0.01</td>
</tr>
<tr>
<td><img src="image5.png" alt="Structure 5" /></td>
<td>1.883</td>
<td>328.77</td>
<td>5</td>
<td>1</td>
<td>3</td>
<td>93.72</td>
<td>0.01</td>
</tr>
<tr>
<td><img src="image6.png" alt="Structure 6" /></td>
<td>2.336</td>
<td>442.91</td>
<td>7</td>
<td>1</td>
<td>7</td>
<td>115.35</td>
<td>0.01</td>
</tr>
<tr>
<td><img src="image7.png" alt="Structure 7" /></td>
<td>1.965</td>
<td>458.91</td>
<td>8</td>
<td>1</td>
<td>6</td>
<td>124.57</td>
<td>0.04</td>
</tr>
</tbody>
</table>
3.2 Pharmacophore model generation

CBP model was developed within CBP module in Catalyst using known crystal complex of HPPD (PDB ID: 1TFZ) and ligand (DSA869), and both were chosen as the receptor and ligand, respectively. Validation option was turned into True. 22 Reported active HPPD inhibitors and 38 inactive compounds were used as the active ligands and inactive ligands, respectively. The rest parameters were set to default values. Six hypotheses were generated, and CBP-Hypo2 with desired Quality (0.721, Fair) was used to screening the library, subsequently.

HipHop pharmacophore models were generated from a set of known molecules with promising activity towards HPPD. Based on the atom-types presented in the molecule, HipHop selected the key common chemical features for creating 3D-pharmacophore models. The principal value 1 for all the ligands and maximum-omit feature as 2. The common features pharmacophore generation module “Feature Mapping” was used to identify the important chemical features of the training set compounds before building HipHop pharmacophore model. Hydrogen bond acceptor (HBA), hydrogen bond donor (HBD), hydrophobic features (HY) and ring aromatic (AR) were considered for generation of the pharmacophore model. The diverse conformation option was applied and 250 conformations within 20 kcal mol⁻¹ cutoff were generated using the “BEST”. The final common feature 3D-pharmacophore models were ranked based on pharmacophore fit value. The fit value of the ten chemical hypothesis generated along with the key 3D-pharmacophoric chemical features are presented in Table 3.

Table 3. Chemical feature compositions for the ten hypotheses generated from six known HPPD inhibitors.

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Features</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>HipHop-Hypo1</td>
<td>RHAAAAA</td>
<td>81.854</td>
</tr>
<tr>
<td>HipHop-Hypo2</td>
<td>RHAAAAA</td>
<td>81.854</td>
</tr>
</tbody>
</table>
Validating the hypothesis is one of the significant methods in pharmacophore generation. Test set including active compounds and inactive compounds was prepared using the same protocol as training set prepared and used to determine whether the hypothesis was able to discern active compounds. Fit value was used as an important evaluation criterion [32].

3.3 Pharmacophore-based virtual screening

About 151047 small molecules were obtained from ChemDiv database (www.chemdiv.com) and subjected to virtual screening. All the compounds which were optimized in DS2.5 were used as virtual screening library. Subsequently, the first screening of CBP-Hypo2 was used to retrieve the database in ‘Search, Screen and Profile module’ of DS2.5. The number of conformations was set to 200, while the conformation method was set to BEST, which provided a complete and improved coverage of the conformational space by performing a rigorous energy minimization and optimizing the conformations in both torsional and Cartesian space by the poling algorithm [33]. Minimum Interfeature Distance was set to 2. Limit Hits was set to First N, and Maximum Hits was set to 500. The rest parameters were set to default values.

3.4 Molecular docking

The AtHPPD crystal structure (PDB ID: 1TFZ) with resolution of 1.8Å was used for molecular docking studies. The protein was prepared by removing the water and some other co-crystallized small molecules, and potentials were assigned using CHARMm force field, the missing atoms / residues were building using the ‘Build and Edit Protein’ module, and cleaning protein were prepared in ‘Prepare Protein module’. The protein structure was energy minimized for 5000 steps (with the heavy atoms constrained) using the conjugate gradient algorithm with the ‘Minimize and Refine Protein’ module in DS2.5. After the protein preparation, the binding site of the protein was defined based on volume occupied by the known ligand pose already in an active site. The obtained

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Phosphorylation</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>HipHop-Hypo3</td>
<td>RHA AAA</td>
<td>81.788</td>
</tr>
<tr>
<td>HipHop-Hypo4</td>
<td>RA AAA</td>
<td>72.750</td>
</tr>
<tr>
<td>HipHop-Hypo5</td>
<td>RA AAA</td>
<td>72.347</td>
</tr>
<tr>
<td>HipHop-Hypo6</td>
<td>RA AAA</td>
<td>72.293</td>
</tr>
<tr>
<td>HipHop-Hypo7</td>
<td>RA AAA</td>
<td>72.234</td>
</tr>
<tr>
<td>HipHop-Hypo8</td>
<td>RA AAA</td>
<td>72.188</td>
</tr>
<tr>
<td>HipHop-Hypo9</td>
<td>RA AAA</td>
<td>72.188</td>
</tr>
<tr>
<td>HipHop-Hypo10</td>
<td>RA AAA</td>
<td>72.167</td>
</tr>
</tbody>
</table>

R (ring aromatic), H (hydrophobic features), A (Acceptor)
receptor was used as the “Input Receptor Molecule” parameter. All hit compounds subjected to fist filtering processes were saved as .sd file. The saved structures were chosen as “Input Ligand” and docked into the active site of HPPD. Docking was performed to ensure the proper binding orientation and placement of each ligand and to confirm the geometric fit of each ligand inside the active site. During the docking process top 10 conformations were saved for each ligand based on dock score value after the energy minimization using the smart minimize method through LibDock and CDOCKER methods.

4. Conclusions

In this study, the strategy for the selection of new chemical compounds with HPPD inhibition by virtual screening was performed. Virtual screening was divided into receptor based virtual screening and ligand based virtual screening. Receptor based virtual was more effective in detecting novel chemical scaffolds and more commonly used in academic labs.

The CBP model was generated based on receptor HPPD enzyme and active ligand that dig out the essential structural features required for inhibition, which was helpful for screening of novel molecules having inhibitory activity against HPPD based on receptor, CBP-Hypo2 was used to screen the ChemDiv library to find potential HPPD inhibitors and Fit Value was believed to an important criterion. The hydrophobic groups (benzene and methyl) of the four compounds obtained from ChemDiv database formed hydrophobic interactions with Met314 and Pro205, and the intermediate aromatic ring generated \( \pi-\pi \) interaction with Phe403 and Phe360. Further, molecular docking was performed to provide insights into molecular recognition via protein–ligand interactions. The result was analyzed based on the docking score, binding modes, and molecular interactions with active site residues. Subsequently, binding free energy of selected compounds relevant to ligand and receptor was calculated, 9 hits of novel scaffolds with good docking score and low binding energy were chosen. Screening compounds could be completely embedded into the HPPD active pocket and interact with the active site of Phe360, Phe403, Arg269, Phe398 and Asn402 and so on. Finally, compounds obtained through docking were matched HipHop model, 4 hits with high Fit value further optimization could be used as potential leads in designing new HPPD inhibitors herbicides. This study provided a set of guidelines that will greatly help in designing the novel and more potent HPPD inhibitors herbicides.

Acknowledgments: This work was supported by the National Nature Science Foundation of China (31572042) and the Research Science Foundation in Technology Innovation of Harbin (2015RAYXJ010).

Author contributions: Y. Fu and F. Ye developed the concept of the work. Y.N. Sun and K.H. Yi carried out the pharmacophore screening work. M.Q. Li and H.F. Cao conducted the molecule docking assay. Y.N. Sun and J.Z Li discussed and analysed the results. Y. Fu wrote the paper.

Conflict of interest: The authors have no conflicts of interest to declare.
References


