Pharmacophore modeling and 3D-QSAR studies of 15-hydroxyprostaglandin dehydrogenase (15-PGDH) inhibitors

Belal O Al-Najjar*\textsuperscript{a,b}, Ashok K Shakya\textsuperscript{a}, Fadi G Saqallah\textsuperscript{a} & Rana Said\textsuperscript{a}

\textsuperscript{a}Faculty of Pharmacy and Medical Sciences, Al-Ahliyya Amman University, PO BOX 263, Amman 19328, Jordan
\textsuperscript{b}Molecular Modeling and Drug Design Unit, Al-Ahliyya Amman University, PO BOX 263, Amman 19328, Jordan

E-mail: najjar.belal@gmail.com

Received 26 December 2016; accepted (revised) 11 September 2017

15-Hydroxyprostaglandin dehydrogenase (15-PGDH) plays an important role in gastric ulcer healing, bone formation and dermal wound healing, encouraging several efforts to discover and optimize new inhibitors. We explored possible pharmacophoric space of 15-PGDH using four diverse sets of inhibitors. After that, GA and MLR methods have been employed to identify the optimal pharmacophore model(s) and physicochemical descriptors able to access Quantitative structure-activity relationship equation ($r^2=0.711$, $r^2_{\text{adj}}=0.6927$, $r^2_{\text{LOO}}=0.6598$). One pharmacophore model has emerged in the Quantitative structure-activity relationship equation and has been validated by ROC curve analysis and molecular docking.

Keywords: 15-PGDH, pharmacophore modeling, quantitative structure-activity relationship, docking

The primary route of prostaglandin metabolism in the body is initiated by 15-Hydroxyprostaglandin Dehydrogenase (15-PGDH), which oxidize the hydroxyl group to keto group at position 15 (Ref 1). Such process leads to pronounced loss of biological activity, making it a key enzyme for prostaglandin biological inactivation\textsuperscript{2}. It inactivates a number of active leukotrienes and hydroxyeicosatetraenoic acids (HETEs)\textsuperscript{3}. 15-PGDH can be found in many mammalian tissues especially lung, kidney and placenta\textsuperscript{4}. Inhibition of 15-PGDH enzyme have drawn great attention in clinical management to reduce hair loss\textsuperscript{5}, gastric ulcer healing\textsuperscript{6,7}, bone formation\textsuperscript{8} and interestingly dermal wound healing\textsuperscript{9,10}. It is worth to note that there is only one X-ray 15-PGDH enzyme in the apoprotein form documented under code 2GDZ in the Protein Data Bank\textsuperscript{11}. Previously, many researchers have evaluated new compounds activities as 15-PGDH inhibitors using traditional ligand-based drug design\textsuperscript{12-16}. The continued concern in designing new 15-PGDH inhibitors prompted us to study the possibility of developing ligand-based 3D-pharmacophores integrated within QSAR models. This approach can utilize the pharmacophore models, which can be used as 3D templates to further synthesize and develop 15-PGDH inhibitors scaffolds. This approach was successfully used by our team to discover new peroxisome proliferator-activated receptor gamma (PPAR\textsubscript{\gamma}) agonists\textsuperscript{17} and estrogen receptor antagonist\textsuperscript{18}.

In this study, we utilized the CATALYST-HYPOGEN module within Discovery Studio (version 2.5)\textsuperscript{19} to prepare and analyse a variety of binding hypothesis for 15-PGDH inhibitors. Multiple linear regression (MLR) and genetic function algorithm (GFA) analysis were employed to find an optimal QSAR equation that combines pharmacophore models with other molecular descriptors. It has the capability of explaining the variation of bioactivity within a collection of diverse 15-PGDH inhibitors. Optimal pharmacophores models were validated by assessing their capabilities to successfully distinguish between active and inactive forms from a list of compounds.

Methodology

Software and hardware

Several software programs were used in the current work which include:
- ChemSketch ACD Labs release 12.01 (www.acdlabs.com).
- Autodock 4.2, Scripps Research Institute, (http://autodock.scripps.edu/).
- Discovery Studio 2.5, Accelrys Inc. (www.accelrys.com), USA.
Both Pharmacophore and QSAR modeling studies were done utilizing Discovery Studio 2.5 suite from Accelrys Inc. (San Diego, California, www.accelrys.com) installed on a Core 2 Duo Pentium PC.

**Dataset**

A group of 85 chemical structures of 15-PGDH inhibitors (Table A under Supplementary Data) were gathered from several published articles along with their *in vitro* IC$_{50}$s. The two-dimensional (2D) chemical structures of the ligands were drawn using ChemSketch and saved in MDL-molfile format. They were then introduced into Discovery Studio, to prepare the 3D structures and energy minimized using CHARMM force field implemented in CATALYST module within Discover Studio.

**Conformational analysis**

Each inhibitor (Table A, Supplementary Data) was explored by a proper conformational space implementing the “best conformer generation” parameter within Discovery. Such conformational poses were prepared with 250 conformers as a maximum limit and an energy threshold of 20 kcal/mol from each minimized structure.

**Pharmacophoric hypotheses generation**

Compounds from the data set with corresponding conformational models were regrouped and the activity of the molecules were assigned with uncertainty values of 2 or 3 to control the number of training compounds within the “most potent list” (see equation 1).

\[(\text{MAct x UncMAct)}= (\text{Act} / \text{UncAct}) < 0.0 \quad \text{Equation 1}\]

Four structurally diverse training subsets were selected from the dataset to perform pharmacophore modeling (Table B, Supplementary Data). Informative training sets should include at least 16 ligands as required by CATALYST, otherwise the number of compounds may produce faulty pharmacophore models. The training sets were used to perform eight modeling runs to search the pharmacophoric space of the inhibitors within their feature distance parameter which was varied between 1 and 3 Å.

**Assessment of the generated hypotheses**

During the process of hypothesis generation, CATALYST was focused on reducing the cost function, which consists of three parts: weight cost, error cost and configuration cost. Moreover, Catscramble utility within the program will further assess the quality of CATALYST_HYPOGEN pharmacophores. This validation process is based on Fischer’s randomization test, and was set to 85%. All the pharmacophores exceeding 85% threshold were considered to be acceptable.

**QSAR modeling**

The 1/IC$_{50}$ (µM) logarithm values for the ligands were used in QSAR, to correlate the data with the change of free energy. The inhibitors were imported into Discovery studio, followed by calculating different 2D and 3D descriptors. This includes electro-topological state indices, logarithm of partition coefficient, polarizability, dipole moment, molecular volume and molecular surface area, etc.). The selected pharmacophores were used to perform “ligand pharmacophore fitting” protocol for the ligands, utilizing the best fit option in CATALYST, and the fit values were uploaded along with other QSAR descriptors. The QSAR equation was created using Genetic function approximation (GFA). The Optimal GFA parameters: explore linear, quadratic and spline equations at mating and mutation probabilities of 50%; population size = 500; number of genetic iterations =30,000 and lack-of-fit (LOF) smoothness parameter = 1.0. However, to determine the optimal number of explanatory terms (QSAR descriptors), it was recommended to assess all possible QSAR models resulting from 4 to 10 descriptive terms.

**Molecular Docking**

Rigid molecular docking can be used to determine the possible intermolecular interactions between 15-PGDH protein and the best fitted conformation of compound 70 (IC$_{50}$ = 4nM, see Table A under Supplementary Data). Docking simulation was performed on the X-ray crystal structures of 15-PGDH (PDB code: 2GDZ, resolution 1.65 Å) using Autodock 4 software. The protein crystal structure was initially prepared by merging all of the non-polar hydrogens and removing water molecules. The conformation that best fit compound 70 against S3D1H9 pharmacophore model was prepared by AutoDockTools, making the compound rigid. Gasteiger and Kollman united atom charges were added to the ligand and protein, respectively, along with atomic solvation parameters.

**Receiver operating characteristic (ROC) curve analysis**

The optimal pharmacophore models emerged from QSAR equation were further validated by evaluating
their capabilities to selectively separate diverse 15-PGDH inhibitors from a testing list of actives and inactives. The testing list was downloaded from CHEMBL database (https://www.ebi.ac.uk/chembl/) which contains 130 active and inactive compounds as 15-PGDH inhibitors.

The test list was screened against each pharmacophore model utilizing the “Best flexible search” parameter, while the ligands conformations were generated using the “CAESER” option.

The ROC curve analysis defines the sensitivity (Se or true positive rate, Equation 2)) for any probable change in the number of selected ligands (n) as a function of (1-Sp). Sp is defined as specificity or true negative rate (Equation 3).

\[ Se = \frac{\text{Number of selected actives}}{\text{Total number of actives}} = \frac{TP}{TP + FN} \]  
Equation 2

\[ Sp = \frac{\text{Number of Discarded Actives}}{\text{Total number of Inactives}} = \frac{TN}{TN + FP} \]  
Equation 3

where, TP is the number of active compounds selected by the screening method (true positives), FN is the number of active compounds rejected by the screening method, TN is the number of rejected inactives, while FP is the number of selected inactives.

**Results and Discussion**

HYPOGEN protocol models ligand enzyme interaction via information derived from only the ligand chemical structure. It detects a 3D presentation from a maximum of five pharmacophoric features that are shared within active training ligands. Hydrogen bond donors (HBD), hydrogen bond acceptor (HBA), aliphatic and aromatic hydrophobes (Hbic) and aromatic rings (RingArom) are the chemical features considered in this study. The flexibility of ligand inhibitors are prepared by generating multiple conformers to ensure representative models over specified energy parameters. Pharmacophore models have been used as a three dimensional queries to filter-out structural databases for novel inhibitors.

In this study, we have created a variety of hypothesis for several 15-PGDH inhibitors compounds. 85 compounds in total were used in this work (Table A, Supplementary Data). Four training subsets were carefully chosen from the dataset, with each subset containing potent, moderate activity and inactive compounds.

**Exploration of 15-PGDH pharmacophoric space**

Each compound conformation was sampled using the poling algorithm embedded within Discovery studio 2.5 using “Best” module. The pharmacophoric space can be evaluated by the configuration (config.) cost calculated for each protocol run. It is generally suggested that the configuration cost of any CTALYST-HYPOGEN run should not exceed 17 (Ref 25, 26). Each pharmacophoric run will be automatically ranked according to their related “total cost” value, the sum of error cost and configuration cost. Error cost offers the main influence to total cost and it is directly related to the ability of the specific pharmacophore as 3D-QSAR model, for example in correlating the chemical structures to the corresponding biological responses. HYPOGEN also calculates null hypothesis cost with the understanding that there is no correlation in the data and the biological activities are normally distributed around their mean. The larger the difference from the null hypothesis cost (Table I) the more likely it is that the correlation between the fit values and actual activities are not a random occurrence. The validation

<table>
<thead>
<tr>
<th>Training set</th>
<th>Interfacial Distance</th>
<th>Hypothesis</th>
<th>Total Cost</th>
<th>Cost of null hypothesis</th>
<th>Residual cost</th>
<th>Cat-scramble (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>109.121</td>
<td>137.311</td>
<td>28.19</td>
<td>85</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>100.908</td>
<td>137.311</td>
<td>36.403</td>
<td>85</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>109.121</td>
<td>137.311</td>
<td>28.19</td>
<td>85</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>109.121</td>
<td>137.311</td>
<td>28.19</td>
<td>85</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>108.32</td>
<td>137.311</td>
<td>28.991</td>
<td>85</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>109.121</td>
<td>137.311</td>
<td>28.19</td>
<td>85</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>102.58</td>
<td>137.311</td>
<td>34.731</td>
<td>85</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>109.121</td>
<td>137.311</td>
<td>28.19</td>
<td>85</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td>109.121</td>
<td>137.311</td>
<td>28.19</td>
<td>85</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>105.313</td>
<td>137.311</td>
<td>31.998</td>
<td>85</td>
<td></td>
</tr>
</tbody>
</table>

*Bolded pharmacophore hypothesis appeared in the best QSAR equation (see section 3.2)
technique based on Fisher’s randomization test introduced into CATALYST-Cat. Scramble was performed\textsuperscript{21}. In this test the bioactivities and the related chemical structures are scrambled numerous times and the program is challenged to create pharmacophoric models from the randomized data\textsuperscript{25-27}.

At the end of modeling runs 75 pharmacophore models were identified from nine automatic HYPOGEN runs. These models were uploaded in QSAR modeling (see section 3.2.). Table I shows a sample of pharmacophore hypotheses performances generated for 15-PGDH. From Table I, the models shared similar features and acceptable statistical success criteria. Many pharmacophore models were optimal and statistically similar which shows that ligands are capable of adopting multiple pharmacophoric binding modes within the binding site. Therefore, it is relatively challenging to select any particular pharmacophore model as an exclusive representative of the binding mode.

**QSAR modeling**

The pharmacophoric hypothesis elucidates the ligand-receptor recognition, and can be used as search queries for new biologically active scaffolds. In this study, we performed classical QSAR analysis to help us find the best combination of pharmacophores and other 2D and 3D descriptors capable of explaining the biological activity variation within the whole list of collected ligands (Table A under Supplementary Data). Genetic function approximation and multiple linear regression QSAR analysis were performed to look for an optimal QSAR equation. The fit values obtained by pharmacophoric mapping of the 85 ligands were uploaded together with nearly 100 other physicochemical descriptors\textsuperscript{17,28,29}. All QSAR models were cross-validated automatically using the leave-one-out cross-validation in Discovery Studio 2.5 (Ref 19).

Equation (4) shows the details of the optimal QSAR model.

\[
\log(1/IC_{50}) = -0.5043 + 0.75621 \times \text{Count<ECFP_6:-2091421151>} + 0.54296 \times \text{Count<FCFC_6:-910120391>} - 0.47941 \times \text{Count<LCFC_6:2234325>} + 1.8879 \times \text{Count<LCFC_6:-117657316>} + 0.3136 \times S3D1H9
\]

Equation 4

\[r^2 = 0.711, r^2_{\text{(adj)}} = 0.6927, r^2_{\text{(LOO)}} = 0.6598\]

where, \(r^2\) is the correlation coefficient against 84 training compounds, \(r^2_{\text{LOO}}\) is the leave-one-out correlation coefficient, \(r^2_{\text{adj}}\) is \(r^2\) adjusted for the number of terms in the model\textsuperscript{19,30} \(S3D1H9\) represent the fit values of the training ligands against the ninth pharmacophoric hypotheses, using the interfacial distance of 1 Å from the third training subset.

Table I shows the pharmacophore modeling runs and the statistical criteria of output models. Bolded runs in Table I correspond to the QSAR selected pharmacophore S3D1H9. The fit values were calculated based on Equation (5).

\[
\text{Fit} = \Sigma \text{mapped hypothesis features} \times W \left[1 - \Sigma (\text{disp/tol})^2\right]
\]

Equation 5

Figure 2 shows S3D1H9 and how it maps two potent training compounds, namely, 70 (IC\textsubscript{50} = 4 nM),
2 (IC$_{50}$ = 22 nM), while Table II shows the X, Y, Z coordinates of the pharmacophore. ECFP$_6$ is the atom type extended connectivity fingerprints count up to diameter 6, FCFC$_6$ is the functional class extended connectivity fingerprints count up to diameter 6 and LCFC$_6$ is the ALogPtypes extended connectivity fingerprint count up to diameter 6 (Ref 19,31). Emergence of the orthogonal pharmacophoric model S3D1H9 in equation 3 suggests that it represents the optimal binding mode of the ligands within the binding pocket of 15-PGDH enzyme. This implies that the pharmacophore can optimally explain the bioactivities of the compounds. Similar conclusions were made for several targets about the binding pockets utilizing QSAR analysis$_{17,28,29}$. The model is composed of hydrogen bond acceptor (HBA), aromatic ring (RingArom), hydrophobes (Hbic) and excluded volume spheres (Figure 2). These features suggest the importance of such interaction between the ligand and the amino acids in 15-PGDH binding site.

For further insight on the intermolecular interaction, a rigid molecular docking was performed on the 15-PGDH binding site and the fitted conformation of compound 70 (Table I). Figure 3 shows that compound 70 performed hydrogen bond interaction with SER193, ASN95 and GLN148. Aromatic interaction was observed between

---

**Table II — The weights, tolerances and X, Y, Z coordinates of S3D1H9 pharmacophoric features**

<table>
<thead>
<tr>
<th>Definitions</th>
<th>HBA</th>
<th>Ring Arom</th>
<th>Hbic</th>
<th>Hbic</th>
<th>Hbic</th>
<th>Excluded volume</th>
<th>Excluded volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weights</td>
<td>1.0287</td>
<td>1.0287</td>
<td>1.0287</td>
<td>1.0287</td>
<td>1.0287</td>
<td>1.0287</td>
<td>1.0287</td>
</tr>
<tr>
<td>Tolerances</td>
<td>1.60</td>
<td>2.20</td>
<td>1.60</td>
<td>1.60</td>
<td>1.60</td>
<td>1.60</td>
<td>1.60</td>
</tr>
<tr>
<td>Coordinates X</td>
<td>-6.08</td>
<td>-6.49</td>
<td>-1.50</td>
<td>-1.31</td>
<td>2.08</td>
<td>2.68</td>
<td>0.20</td>
</tr>
<tr>
<td>Y</td>
<td>2.67</td>
<td>4.22</td>
<td>0.22</td>
<td>1.25</td>
<td>-1.72</td>
<td>1.48</td>
<td>2.48</td>
</tr>
<tr>
<td>Z</td>
<td>-3.87</td>
<td>-6.45</td>
<td>-0.09</td>
<td>2.72</td>
<td>0.68</td>
<td>3.16</td>
<td>-1.44</td>
</tr>
</tbody>
</table>

---

![Figure 2 — Pharmacophoric features of (A) S3D1H9 against (B) 70 (IC$_{50}$ = 4 nM) and (C) 2 (IC$_{50}$ = 22 nM). HBA shown as green vectored spheres, Hbic as light blue spheres, RingArom as vectored orange spheres and excluded volume spheres are shown as grey spheres.](image-url)
compound 70 and TYR151, whereas ALA140 and VLA145 participated in hydrophobic interaction. The importance of these amino acids in the catalytic reaction of prostaglandins has been discussed in previous literatures\textsuperscript{21,32}.

**Receiver operating characteristic (ROC) curve analysis**

For additional validation of selected pharmacophoric model by QSAR, we subjected S3D1H9 model to receiver operating curve (ROC) analysis. In such analysis, the capability of the pharmacophoric model to selectively classify a group of ligands as actives and inactives is determined by the area under the curve (AUC) of the corresponding ROC curve, including, specificity, accuracy, true positive rate and false negative rate\textsuperscript{23,24}. Table III and Figure 4 show the results of ROC analysis of the selected pharmacophore S3D1H9 by QSAR. S3D1H9 illustrated very good overall performances with ROC-AUC value 74%.
Figure 3 — Stick representation of the interacting residues of 15-PGDH active site with compound 70

Table III — List of ROC criteria for S3D1H9 pharmacophore model

<table>
<thead>
<tr>
<th>ROC criteria</th>
<th>ROC^a-AUC^b</th>
<th>ACC^c</th>
<th>SPC^d</th>
<th>TPR^e</th>
<th>FNR^f</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.744</td>
<td>0.708</td>
<td>0.676</td>
<td>0.785</td>
<td>0.323</td>
</tr>
</tbody>
</table>

^a ROC: receiver operating characteristic curve.
^b AUC: area under the curve.
^c ACC: overall accuracy.
^d SPC: overall specificity.
^e TPR: overall true positive rate.
^f FNR: overall false negative rate.

Figure 4 — ROC curves of S3D1H9 pharmacophore model

**Conclusion**

15-PGDH inhibitors currently play an important role in gastric ulcer healing, bone formation and dermal wound healing. One orthogonal pharmacophore model appeared after performing pharmacophore modeling followed by QSAR analysis. The pharmacophore model suggested molecular features that are important for binding interaction, such as...
hydrogen bond acceptor, hydrophobes and aromatic rings. These features were further analyzed by identifying the intermolecular interaction between 15-PGDH crystal structure (PDB code: 2GDZ) and the best fitted conformation of compound 70 against S3D1H9 model. The docking procedure showed that the ligand formed interaction with critical amino acids in the binding site which supports the pharmacophore model quality. This pharmacophore model and QSAR equation can be used as a search query for ligand databases to identify and predict the activity of unknown ligands, thus facilitating the drug discovery process toward HPGD inhibitors.

Supplementary Information

Supplementary information is available in the website http://nopr.niscair.res.in/handle/123456789/60.

Acknowledgements

This study was funded by Al-Ahliyya Amman University (AAU), research grant number: 42/19/2015-2016. Special thanks also go to Prof. Mutaseom O. Taha, University of Jordan, for all the help in this work.

References

9 Seo S Y, Han S-I, Bae C S, Cho H&Lim S C, Prostaglandins, Leukotrienes and Essential Fatty Acids, 97, 35.