ABSTRACT

Objective: To find out the common structural requirement which may be responsible for biological activity of benzofurans as potent N-myristoyltransferase (Nmt) inhibitors by the generation of ligand based pharmacophore hypothesis and atom-based 3D-QSAR model.

Methods: A set of 29 compounds were taken from the literature with their in vitro Nmt inhibitory data. A training set of 24 molecules and a set of five molecules was arbitrarily set aside as the test set. To generate 3D-QSAR models, training set was used and to validate the quality of the model, test set was used. A five-point pharmacophore model has been developed with two hydrogen bond acceptors (AA), one positive ionic atom (P) and two aromatic ring residues (RR). This is denoted as AAPRR.

Results: A statistically significant 3D-QSAR model for training set of 24 compounds was obtained using pharmacophore hypothesis AAPRR, with correlation coefficient ($r^2$ = 0.916) and high Fisher ratio ($F$ = 113.9). Also, the predictive power of generated model for test set of 5 compounds was found to be significant which was confirmed by the high value of cross validated correlation coefficient ($q^2$ = 0.804) and Pearson-R (0.917).

Conclusion: The results of ligand based pharmacophore hypothesis and atom based 3D-QSAR model explore the detailed structural perceptivities and also highlight the important binding features of benzofurans with Nmt.

Keywords: Phase, Pharmacophore, Atom based 3D-QSAR, Nmt Inhibitors.

INTRODUCTION

Fungal infections have increased dramatically in recent years. Life threatening infections caused by pathogenic fungi are becoming increasingly common, especially in those individuals with immune compromised hosts, such as patients undergoing anticancer chemotherapy or organ transplants and patients with AIDS [1]. Clinically, candidosis, aspergillosis and cryptococcosis are three major fungal infections in immune compromised individuals [2, 3].

Currently available antifungal drugs for such infections essentially have three molecular targets: cytochrome P-450-dependent lanosterol 14α-demethylase, ergosterol, and 1,3-beta glucan synthase. The first is a fungistatic target vulnerable to resistance development; the second, while a fungicidal target, is not sufficiently different from the host to ensure high selectivity; and the third, a fungistatic (Aspergillus) or fungicidal (Candia) target, has limited activity spectrum and potential host toxicity that might preclude dose escalation. However, their use can suffer from limited efficiency, narrow antifungal spectrum, drug related toxicity, severer drug resistance, non-optimal pharmacokinetics and serious drug–drug interactions. Therefore, there is an emergent need to develop novel fungicidal drugs with a new mode of action.

Myristoyl-CoA: N-myristoyltransferase (Nmt; EC 2.1.3.97) is a cytosolic monomeric enzyme that catalyzes the transfer of the cellular fatty acid myristate (C14:0) from myristoyl-CoA to the N-terminal glycine amine of a variety of eukaryotic proteins [4]. Genetic studies have established that Nmt is essential for vegetative growth and survival of pathogenic fungi such as Saccharomyces cerevisiae. Nmt has also been proven to be essential for the viability of fungi, including the medically important Candida albicans [C. albicans] and Cryptococcus neoformans [5, 6].

Human Nmt has also been cloned and characterized. C. albicans Nmt has 451 amino acid residues, with a sequence identity of 45% to the human Nmt enzyme. Clear difference in the peptide–substrate specificity between fungal and human Nmts has been exploited, and Nmt has been identified as a potential chemotherapeutic target for antifungal agents [7, 8]. Hence Nmt is an attractive target for the design and development of novel antifungal agents with a new mode of action without disrupting host Nmt. Because of its novel mechanism of action, Nmt inhibitors are expected to have advantages over azole antifungals in terms of activity against azole-resistant fungal strain and lack of drug–drug interactions, which are drawbacks of azole antifungal agents [9].

We have reported molecular modeling studies on various antifungal targets like cytochrome P-450-dependent lanosterol 14α-demethylase and squalene epoxidase [10, 11] for the rational design of new antifungal agents. Also, we have generated a 3D pharmacophore model using peptidic Nmt inhibitors and successfully utilized it for the design of nonpeptidic inhibitors [12, 13]. A novel series of benzofuran analogs was reported as nonpeptidic Nmt inhibitors [14-16]. Previously, we have reported 3D-QSAR on series of benzofuran analogs as antifungals using comparative molecular field analysis (CoMFA) [17] and Genetic function approximation (GFA) [18]. Our interest continues in benzofurans because of ease of synthesis of analogs.

Pharmacophore modeling has been one of the important and successful approach for novel drug discovery since last few years. A pharmacophore model of the target binding site summarizes steric and electronic features needed for optimal interaction of a ligand with a target. Most common properties that are used to define pharmacophores are hydrogen bond acceptors, hydrogen bond donors, basic groups, acidic groups, partial charge, aliphatic hydrophobic moieties, and aromatic hydrophobic moieties. Pharmacophore features have been used extensively in drug discovery for virtual screening, de novo design, and lead optimization. A pharmacophore model of the target binding site can be used to virtually screen a compound library for putative hits. Apart from querying data base for active compounds, pharmacophore models can also be used by de novo design algorithms to guide the design of new compounds [19].

To continue our research efforts in the development of pharmacophore and 3D-QSAR of various therapeutic agents [20-22] we report here studies on pharmacophore generation and atom based...
3D-QSAR model studies of a series of benzofurans using PHASE (Pharmacophore Alignment and Scoring Engine) [23] incorporated in Schrodinger Software, USA. The objective of the present study is to find out the common features which may be responsible for biological activity of benzofurans as potent N-myristoyl transferase inhibitors. The generated ligand based pharmacophore hypothesis and the cubes generated from atom-based 3D-QSAR model highlights the important structural features required for Nmt inhibition which can be useful further for design of potent N-myristoyl transferase inhibitors.

**MATERIALS AND METHODS**

**Dataset**

In the present study a set of 29 compounds were taken from the literature with their in vitro Nmt inhibitory data. [14-16] A training set of 24 molecules (molecules 1-24, table 1) were used to generate the QSAR models. The training set molecules were selected in such a way that they contained information in terms of both their structural features and biological activity ranges. The most active molecules, moderately active, and less active molecules were included, to spread out the range of activities [24]. To assess the predictive power of the model, a set of five molecules (molecules 25-29, table 1) was arbitrarily set aside as the test set. The test molecules were selected in such way that they truly represent the training set. To generate 3D-QSAR models, training set was used and to validate the quality of the model, test set was used. Biological activity used in the present study was expressed as:

\[ \text{pIC}_{50} = -\log_{10} \text{IC}_{50} \]  \hspace{1cm} (1)

Where IC$_{50}$ is the nanomolar concentration of the inhibitor producing 50% inhibition. In all the models subsequently developed, pIC$_{50}$ values were used as the dependent variable. Structures and related inhibitory activities (IC$_{50}$ values) are reported in table 1. Actual and predicted inhibitory activities (pIC$_{50}$ values) and residual values are reported in table 1.

**Table 1: Structure and activities of the molecules in the training (1–24) and test (25–29) sets**

<table>
<thead>
<tr>
<th>Comp.</th>
<th>R</th>
<th>Observed activity</th>
<th>Predicted activity</th>
<th>Residuals</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>IC$_{50}$ (pM)</td>
<td>pIC$<em>{50}$ (Log1/IC$</em>{50}$)</td>
<td>pIC$_{50}$</td>
</tr>
<tr>
<td>1</td>
<td>O(CH$_3$)$_2$NHCH(CH$_3$)$_2$</td>
<td>50</td>
<td>-1.698</td>
<td>-1.490</td>
</tr>
<tr>
<td>2</td>
<td>OCH$_2$CH(OH)CH$_2$NHCH(CH$_3$)$_2$</td>
<td>0.98</td>
<td>0.008</td>
<td>0.08</td>
</tr>
<tr>
<td>3</td>
<td>O(CH$_3$)$_2$NHCH(CH$_3$)$_2$</td>
<td>1.7</td>
<td>-0.230</td>
<td>0.020</td>
</tr>
<tr>
<td>4</td>
<td>O(CH$_3$)$_2$NHCH(CH$_3$)$_2$</td>
<td>4.4</td>
<td>-0.643</td>
<td>-0.620</td>
</tr>
<tr>
<td>5</td>
<td>O(CH$_3$)$_2$NHCH(CH$_3$)$_2$</td>
<td>1.6</td>
<td>-0.204</td>
<td>-0.070</td>
</tr>
<tr>
<td>6</td>
<td>O(CH$_3$)$_2$NHCH(CH$_3$)$_2$</td>
<td>15</td>
<td>-1.176</td>
<td>-1.470</td>
</tr>
<tr>
<td>25*</td>
<td>O(CH$_3$)$_2$CH(OH)CH$_2$NHCH(CH$_3$)$_2$</td>
<td>4.4</td>
<td>-0.643</td>
<td>-0.030</td>
</tr>
<tr>
<td>26*</td>
<td>O(CH$_3$)$_2$CH(OH)CH$_2$NHCH(CH$_3$)$_2$</td>
<td>1.2</td>
<td>-0.079</td>
<td>0.001</td>
</tr>
</tbody>
</table>

**Ligand preparation**

All molecules were built in Maestro and were prepared using Lig Prep [25] to convert 2D structure to 3D, generate stereoisomer, determine the most probable ionization state at user defined pH, neutralize charged structures, add hydrogen and to generate the possible number of energy minimized bioactive conformers using Conf Gen by applying OPLS-2005 force field [26]. Conformational space was explored through combination of Monte-Carlo Multiple Minimum (MCMM) / Low Mode (LMOD) with maximum number of conformers...
Generation of Common Pharmacophore Hypothesis (CPH)

PHASE provides a standard set of six pharmacophoric features, hydrogen bond acceptor (A), hydrogen bond donor (D), hydrophobic group (H), negatively ionizable (N), positively ionizable (P), and aromatic ring (R) to define chemical feature of ligands. A five point common pharmacophore hypothesis (CPH) was identified from all the conformations of the active ligands having identical set of features with very similar spatial arrangement keeping minimum intersite distance of 2.0 Å.

Hypotheses were generated by a systematic variation of number of sites and the number of matching active compounds. CPH was considered, which indicated at least five sites common to all molecules.

Scoring pharmacophore hypothesis

Scoring procedure helps in ranking of the different hypotheses to yield the best alignment of the active ligands using an overall maximum root mean square deviation (RMSD) value of 1.2 Å with default options for distance tolerance. Thus, it helps in making rational choice regarding which hypothesis is more appropriate for further investigation. The quality of alignment was measured by a survival score, [29, 30] defined as:

$$ S = W_{vecs} + W_{site} + W_{sel} + W_{rew} - W_{AE} + W_{A} $$

where W's are weights and S's are scores; W_{site} represents alignment score, the RMSD in the site point position; W_{vecs} represents vector score, and averages the cosine of the angles formed by corresponding pairs of vector features in aligned structures; represents volume score based on overlap of van der Waals models corresponding pairs of vector features in aligned structures; score, and averages the cosine of the angles formed by

Structure activity relationships. In atom-based QSAR, a molecule is treated as a set of overlapping Vander Waals spheres. Each feature is grouped according to a simple set of rules: hydrogen attached to polar atoms are classified as hydrogen bond donors (D); carbons, halogens, and C–H hydrogen are classified as hydrophobic/non-polar (H); atoms with an explicit negative ionic charge are classified as negative ionic (N); atoms with an explicit positive ionic charge are classified as positive (P); non-ionic atoms are classified as electron-withdrawing (W); and all other types of atoms are classified as miscellaneous (X).

Validation of Pharmacophore Model

The main purpose to develop QSAR model was to predict biological activities of new compounds whereby the generated model will be statistically robust both internally as well as externally. The dataset was divided into training set and a test set. Atom-based 3D-QSAR models were generated for hypotheses using the 24 compounds in training set. The best QSAR model was externally validated by predicting activities of the 5 test set compounds.

The robustness of the developed pharmacophore hypotheses was internally validated by statistical parameters like squared correlation coefficient (r^2), q^2 (r^2 for test set), Standard deviation of regression (SD), Pearson correlation coefficient (Pearson-R), Statistical significance (P) and variance ratio (F). The predicted pIC50 at 2*m PLS factor is given in table 3. Further increase in the number of PLS factors did not improve the statistics or predictive ability of the model[31, 32].

RESULTS AND DISCUSSION

Pharmacophore modeling and QSAR studies were performed using PHASE module of Schrodinger suite. The study is performed to generate 3D pharmacophoric features of benzofurans responsible for binding to target and to predict biological activity as N-myristoyl transferase inhibitors using atom based 3D-QSAR model.

Pharmacophore generation and 3D-QSAR model analysis

Common pharmacophore hypothesis (CPH) was developed using a set of 7 active ligands in "pharmset" as they contain important structural features crucial for binding to the receptor binding site.
We have used four minimum sites in find pharmacophore step and five maximum sites to generate optimum combination of features common to the most active compounds. The three and four featured pharmacophore hypothesis was rejected due to low value of survival score, as they were unable to define the complete binding space of the selected molecules. Five featured pharmacophore hypotheses were selected and subjected to stringent scoring function analysis.

CPH were generated with different combination of variants: AAPRR, AAHRR, AAARR, AARRR and 2455 CPH were generated with the variant AAPRR which was further considered for QSAR model generation and the results are given in table 2. Among the various generated pharmacophores, best hypothesis was identified by aligning the active compounds on hypothesis and calculating the survival score which helps in ranking all the hypotheses. The selected hypothesis distinguishes between the active and inactive molecules. Further, to confirm that the pharmacophore hypothesis maps with more active than inactive features, they were aligned to inactive compounds and scored. Hypothesis showing poor score were rejected. Therefore, adjusted survival score was calculated by subtracting the inactive score from survival score of this hypothesis (Table 2).

Finally the hypothesis with maximum adjusted survival score and lowest relative conformational energy was selected for generating atom-based 3D-QSAR model. The best model with good predictive power was found to be associated with the five points which consist of two hydrogen bond acceptor (AA), one positive ionic atom (P) and two ring feature (RR). This is denoted as AAPRR. The pharmacophore hypothesis showing distance between pharmacophoric sites is depicted in Fig. 1. As observed in Fig. 2, among the two hydrogen bond acceptor groups (AA) one feature is mapped to ‘O’ of benzofuran ring, second one is observed on C=O group of side chain and two aromatic rings (RR) are present on benzene ring attached to side chain and in benzofuran ring.

The alignment generated by best pharmacophore hypothesis AAPRR is used for generation of 3D-QSAR model. From Fig. 2 and Fig. 3 it can be observed that active ligands have better alignment than inactive ligands.

Atom-based 3D-QSAR model was generated with 2nd PLS factor having good statistical significance and predictive power. PLS factor was increased up to two since up to second factor there is incremental increase in predictive power and statistical value of model.

A statistically significant 3D-QSAR model was obtained using this pharmacophore hypothesis with correlation coefficient value ($r^2 = 0.916$) and high Fisher ratio ($F = 113.9$) for the training set of 24 compounds. Also, the predictive power of generated model was found to be significant which was confirmed by the high value of cross validated correlation coefficient ($q^2 = 0.804$) and Pearson-R (0.917) for the test set of 5 compounds.

The high value of $F$ (113.9) indicates a statistically significant regression model, which is also supported by the low value of the variance ratio (P) indicating a high degree of confidence. The summary of atom based 3D-QSAR results is shown in table 3. Graphs of observed versus predicted biological activity of training and test set molecules are shown in fig. 4a and 4b respectively.

In order to visualize the generated 3D-QSAR model and to study its correlation with inhibitory activity one or more ligands from the series having diverse inhibitory activity were taken into consideration. The cubes were generated which represents important structural features required for interaction of ligand with active site of receptor which is shown in fig. 5 and 6 respectively for the most active and least active compounds in the series. In these
representations, the blue cubes indicate favorable regions while red cubes indicate unfavorable regions for biological activity. By using 3D-QSAR model the comparison was done to study favorable and unfavorable interaction using most active (compound 23) and least active (compound 22) as shown in Fig. 5 and 6 respectively.

On comparing fig. 5 and 6 it is observed that blue cubes are seen on C=O group and imidazole ring attached to benzo[4]furan ring of compound 23 which suggest that highly polar groups are favored in this region. While, in compound 22 the imidazole ring is replaced by less polar ethoxy group and red cubes were seen in this region, which leads to tremendous decrease in the activity of compound 22.

Hence, this suggests that more electronegative and polar groups are needed in this region and replacement with bulky groups will lead to decrease in activity. This is confirmed from the activities of compounds 13-20, 24, 27 and 28 which are having more polar and electronegative groups in this region hence they are active. While, the compounds 1-12, 21-22, 25-26, 29 have less polar and less electronegative groups, hence they are less active compounds. Also, blue cubes were observed around nitrogen and benzene ring which suggest that presence of nitrogen and aromatic residue is essential for activity which is revealed from the activities of compounds 13-20, 23-24 and 27-28.

On comparing fig. 5 and 6 it is observed that blue cubes are seen on C=O group and imidazole ring attached to benzo[4]furan ring of compound 23 which suggest that highly polar groups are favored in this region. While, in compound 22 the imidazole ring is replaced by less polar ethoxy group and red cubes were seen in this region, which leads to tremendous decrease in the activity of compound 22.

Hence, this suggests that more electronegative and polar groups are needed in this region and replacement with bulky groups will lead to decrease in activity. This is confirmed from the activities of compounds 13-20, 24, 27 and 28 which are having more polar and electronegative groups in this region hence they are active. While, the compounds 1-12, 21-22, 25-26, 29 have less polar and less electronegative groups, hence they are less active compounds. Also, blue cubes were observed around nitrogen and benzene ring which suggest that presence of nitrogen and aromatic residue is essential for activity which is revealed from the activities of compounds 13-20, 23-24 and 27-28.

Hence, this suggests that more electronegative and polar groups are needed in this region and replacement with bulky groups will lead to decrease in activity. This is confirmed from the activities of compounds 13-20, 24, 27 and 28 which are having more polar and electronegative groups in this region hence they are active. While, the compounds 1-12, 21-22, 25-26, 29 have less polar and less electronegative groups, hence they are less active compounds. Also, blue cubes were observed around nitrogen and benzene ring which suggest that presence of nitrogen and aromatic residue is essential for activity which is revealed from the activities of compounds 13-20, 23-24 and 27-28.

The authors are grateful to Dr. Shivajirao S. Kadam, Vice Chancellor, Bharati Vidyapeeth University, Pune, Maharashtra, India for providing the computational facility.

ACKNOWLEDGEMENT
The authors report no conflicts of interest in this work.

REFERENCES