

Combined ligand and structure based approaches towards developing novel renin inhibitors for the treatment of hypertension

Ambili Unni Parambath^a, Sajithalulu.Sudhakaran^{a*}, Girinath Gopinathan. Pillai^{a,b*}

^aSchool of Biosciences and Technology, VIT, Vellore, India; ^bNyro Research India, Kochi, Kerala, India

Address for correspondence: Girinath Gopinathan. Pillai, Nyro Research India,

Email: pillai@nyroindia.org Phone: +91 9944807641

Running Title: Exploration of novel chemical compounds possessing renin inhibitory activity by employing computational approaches.

Abstract

Hypertension is considered as the predominant risk factor for the onset of Cardiovascular disease (CVD) in the elder population. The chronic activation of Renin Angiotensin System (RAS) is considered as the primary causative factor for the inception of hypertension in geriatric population. Angiotensin Converting Enzyme (ACE) is a highly explored druggable target in the context of hypertension since this enzyme catalyses the conversion of angiotensin I to angiotensin II, a potent vasoconstrictor. But clinical trials conducted on ACE inhibitors reported their incompetence in the effective treatment of hypertension. Hence, recent studies are focussing on renin, which is a central component of RAS in the regulation of blood pressure. The present study focuses on the elucidation of physicochemical properties of chemical compounds essential for renin inhibition and identification of novel renin inhibitors possessing enhanced potency as well as bioavailability. We have employed Molecular Field Topology Analysis (MFTA) as well as Structure Based Drug Design (SBDD) approaches for the accomplishment of above-mentioned objectives. MFTA approach were piloted on 45 indole-3-carboxamide derivatives by elucidating the significance of charge distribution as well as molecular size of chemical species in eliciting renin inhibition. Optimal model was obtained with $N_f = 3$, $r^2 = 0.81$, $Q^2 = 0.65$. Molecular docking, atom-based binding free energy contributions and bioavailability assessments were carried out to identify most potent lead molecule among 45 compounds reported for renin inhibition. Further, new derivative molecules were predicted for the best lead molecule by employing chemical space exploration.

Keywords: Bioavailability assessment; chemical space exploration; docking ; hypertension MFTA ;renin inhibitors.

1. Introduction

Aging is an ineluctable biological phenomenon associated with the onset of chronic diseases affecting vital organ systems of the human body. Neurodegenerative, cardiovascular and immunologic diseases are highly predominant in the elder population [1]. The enhanced prevalence of Cardiovascular disease (CVD) contribute towards higher extent of mortality rate [2]. According to the estimates of World Health Organization (WHO), CVDs are considered as a highly predominant cause of enhanced global death rates [3]. Significant CVDs prevalent in the aged population include Coronary Heart Disease (CHD), Myocardial Infarction (MI), Heart Failure (HF), stroke, Peripheral Arterial Disease (PAD) and Valvular Heart Disease (VHD) [2].

Tobacco use, consumption of unhealthy diet, obesity, physical inactivity and excessive usage of alcohol are documented as the behavioural risk factors responsible for the onset of CVDs. Hypertension, diabetes and hyperlipidaemia are considered as the major physiological risk factors leading to the onset and progression of CVDs [4]. Statistical estimations reported a direct correlation between hypertension and aging [5]. Hence, hypertension act as a significant risk factor for initiating the pathogenesis of CVDs.

The pathogenesis of hypertension is associated with age-dependent chronic activation of Renin Angiotensin System (RAS), which plays a significant role in sodium homeostasis and maintenance of body fluid volume [6]. RAS comprise of angiotensinogen, renin, angiotensin I, Angiotensin Converting Enzyme (ACE), angiotensin II, Angiotensin Type I Receptor (AT1R) and Angiotensin Type 2 Receptor (AT2R). In normal conditions, decreased body fluid volume as well as hypotension triggers the expression of renin. Renin catalyses the conversion of angiotensinogen to angiotensin I (rate limiting step) and subsequent activation of angiotensin II. Angiotensin II, act as a potent vasoconstrictor and hence aids in

reconstituting body fluid volume and to maintain blood pressure to normal level [7]. But, aging is characterized by the abnormal activation of RAS which results in upregulated expression of angiotensin II, a potent vasoconstrictor [8].

Angiotensin II is also responsible for the production of superoxide ions via the activation of intracellular NADPH oxidase. Superoxide ions belong to the category of Reactive Oxygen Species (ROS), which is mainly responsible for triggering the onset of age-associated diseases. Hence, angiotensin II is an effector molecule of RAS pathway, which causes a detrimental effect in the aged population by promoting the upregulated production of ROS and the onset of hypertension [9]. Hence, inhibition of angiotensin II expression could prevent the pathogenesis of hypertension and elevated production of superoxide ions. The production of angiotensin II is mediated by ACE. Hence, inhibition of ACE is a widely accepted therapeutic approach for controlling angiotensin II expression. Clinically approved drugs such as captopril, enalapril, lisinopril, and fosinopril act as ACE inhibitors. But clinical trials conducted on ACE inhibitors have reported incomplete RAS blockade which further leads to the rise in plasma renin activity and activation of angiotensin II [10]. The reduced feedback inhibition of renin release by act as a causative factor for the incomplete RAS blockade by ACE inhibitors [11].

But, direct renin inhibitors bind to the catalytic site of renin and inhibits the conversion of angiotensinogen to angiotensin I, which is the rate limiting step in RAS pathway leading to the formation of angiotensin II [12]. Therapeutic strategies reported on direct renin inhibition leads to the complete blockade of RAS pathway and decreased angiotensin I and angiotensin II activity [13]. Hence, this study focus on elucidating physicochemical features of chemical compounds responsible for renin inhibition and to predict novel renin inhibitors with improved potency and bioavailability by employing ligand and structure based approaches respectively.

2. Materials and methods

2.1 Dataset

45 indole-3-carboxamide compounds reported for renin inhibitory activity were considered for this study [14]. The potency of indole-3-carboxamide derivatives for eliciting renin inhibition was expressed as IC₅₀ (half maximal inhibitory concentration) values. Logarithmic values of IC₅₀ were used as response variables for activity modelling studies.

2.2 Molecular Field Topology Analysis (MFTA) for renin inhibitors

MFTA was employed on establishing Quantitative Structure Activity Relationships (QSAR) of structurally related chemical compounds [15]. In MFTA, molecular structures are represented as molecular graphs, which are subsequently superimposed into a Molecular Super Graph (MSG). MSG takes accounts for the common scaffolds shared by chemical compounds and distribution of local descriptors along the molecular structure [15].

2.3 MFTA- Training Set Preparation

Two Dimensional (2D) co-ordinates of 45 compounds present in the dataset were drawn using Marvin Sketch (<https://chemaxon.com/products/marvin>) and saved as .sdf format. The activity file depicting potency values for renin inhibition (LogIC₅₀ values – dependent variables) were prepared in Comma Separated Value (CSV) file format. SDF file of chemical structures were merged with the activity file using StarDrop (<https://www.optibrium.com/stardrop/>). The merged file consisting of chemical structures and activity values which were loaded as input to generate MFTA dataset (MDS file).

2.4 MFTA – Descriptor Calculation

MFTA is based on the distribution of local molecular descriptors on homogenous chemical compounds. The local molecular descriptors (independent variables) calculated by considering the properties of atoms and bonds in chemical structures include Charge.Q (Q), VdW.Re (Re), HBond.Hd (H_d), HBond.Ha (H_a) and Lipo.Lg (L_g). Q depicts effective atomic charge calculated by Electronegativity Equalization Method (EEM), which determines atomic charges by considering electronegativity of atoms present in the molecule [16]. ‘Re’ depicts effective van der Waals radius of atom by considering steric requirements of the central non-hydrogen atom and other attached atoms. H_d and H_a represent the ability of atom to act as hydrogen bond donor and hydrogen bond acceptor respectively. Lipo.Lg accounts for group lipophilicity considering the contributions of central non-hydrogen atom and the attached hydrogen atoms. Hence Q accounts for electrostatic interactions, VdW.Re for steric interactions, Lipo.Lg for hydrophobic interactions, Hd & Ha for hydrogen bond interactions of chemical compounds [17]. Execution of MFTA aided in elucidating the distribution of above-mentioned local descriptors on chemical compounds and its influence on targeted biological activity. Descriptor combinations recommended for running MFTA calculations are provided in (Table 1).

2.5 Assessment of model quality and selection of optimal model

MFTA employs Partial Least Squares (PLS) method for generating predictive models. PLS method is based on the linear transition from a large number of original descriptors to a new variable space based on small number of orthogonal factors (latent variables). In the PLS method, latent variables are chosen in such a way as to provide maximum correlation with the dependent variable [18]. The selection of optimal model was performed by considering

model predictivity and interpretability. Model predictivity was analysed by considering statistical parameters such as Q^2 and Root Mean Square Error (RMSE) [19]. Models possessing higher Q^2 and low Root Mean Square Error (RMSE) values were chosen. Interpretability of MFTA models is associated with the distribution of local descriptors and their influence on the biological activity.

2.6 Identification of active site residues and molecular docking studies

Computational tools such as DoGSiteScorer and ScanProsite were employed for the active site prediction. DoGSiteScorer server aids in the prediction of binding sites by analysing the geometric and physico-chemical properties of pockets by utilizing Support Vector Machine (SVM) algorithm [20]. ScanProsite detects functional domains of proteins [21]. Active site residues identified from these computational tools were compared with existing literature references on renin inhibitors.

Molecular docking studies were executed by FlexX, which exploits incremental construction algorithm (IC). IC algorithm samples the conformation space of flexible ligand and utilize tree search technique for the accurate placement of ligand fragments incrementally into the protein active site [22].

2.7 Assessment of torsional alerts and atom-based free energy contributions

SeeSAR (SeeSAR version 10.0; BioSolveIT GmbH, Sankt Augustin, Germany, 2020, www.biosolveit.de/SeeSAR) was employed for calculating the torsional alerts and binding free energy contributions for all the compounds in dataset. SeeSAR is a computational tool which aids in the visual compound prioritization as well as compound evolution. SeeSAR is an interactive tool which helps in the visual inspection of binding free energy contributions of each atoms of the ligand. The binding free energy was calculated by HYDE scoring function,

which accounts for the hydrogen bond as and dehydration energies of protein-ligand complexes [23].

2.8 Bioavailability assessment

The ranking of highly potent lead like molecules was carried out by analysing physicochemical features influencing bioavailability. Bioavailability assessment was carried out by utilizing StarDrop's ADME module (<https://www.optibrium.com/stardrop>), which provides outstanding predictive ADME models for diverse class of chemical compounds. StarDrop's ADME module predicts key physicochemical properties including logP (Octanol/Water), Intrinsic Aqueous Solubility (logS), Human Intestinal Absorption (HIA) Classification, Cytochrome P450 Affinities and hERG pIC50 . The toxicity assessment was predicted by utilizing vNN web server (<https://vnnadmet.bhsai.org/>) which employs k-nearest neighbor (k-NN) method for the accurate prediction of toxicity alerts including cytotoxicity and mutagenicity [24]. The structural features of chemical compounds responsible for enhancement of physicochemical properties encoding bioavailability was predicted by StarDrop's glowing molecule module.

2.9 Identification of derivatives

New derivative compounds were predicted from the most effective lead molecule among 45 compounds in the dataset. Effective lead molecule was identified by considering multiple parameters such as higher dock score, appropriate binding interactions with target protein (renin) and good bioavailability profiles. Identification of derivative compounds were accomplished by exploiting infiniSee tool (<https://www.biosolveit.de/infiniSee/>). InfiniSee explores chemical spaces by utilizing FTrees similarity search engine, which employs a fuzzy pharmacophore descriptor for identifying structurally distinct molecules in pharmacophore

space [25]. Identification of most promising derivative compounds was performed by hit to lead optimization strategies including molecular docking, binding affinity, torsional alerts prediction and bioavailability assessment studies.

3. Results and Discussion

3.1 MFTA – Renin inhibitors

MFTA is an approach based on local atomic descriptors from homogenous dataset. MFTA calculations elucidated molecular supergraphs, which depicted the influence of local descriptors on chemical structures. The training set comprising of 45 compounds reported for renin inhibitory activity were compiled from the literature. Renin inhibitory activities of 45 compounds were expressed in IC_{50} (half maximal inhibitory concentration) values. Logarithmic values of IC_{50} were employed for further analysis. MFTA QSAR models were generated for all descriptor combinations listed in Table 1. QSAR models of all descriptor combinations resulted in similar statistical quality as well as coherent picture of local descriptor influence on the activity. The prioritization of optimal model was prepared by considering the biological as well as statistical significance of generated QSAR models. Optimal model was obtained for the descriptor combination (Charge.Q, VdW.Re). The parameters illustrating statistical significance of QSAR model is provided in (**Table 2**). Details regarding experimental and predicted values of compounds present in the dataset is provided in Supplementary Table S1.

Scatter plot illustrating the correlation between experimental and predicted endpoints is provided in (**Figure 1**). Other relevant plots obtained by performing modelling is provided in Supplementary Figure S1-S4.

Since, primary objective of the present work was to identify the influence of physicochemical features of chemical compounds in eliciting renin inhibition, descriptor combination representing the significance of electrostatic and steric interactions (Charge.Q, VdW.Re) were chosen for performing MFTA calculations. Electrostatic and steric interactions play significant role in establishing stable protein-ligand interactions [26]. Electrostatic interactions belong to the category of non-covalent interactions which accounts for the attractive or repulsive interactions between chemical species possessing electric charges [27].

Electrostatic interactions are highly predominant in proteins. The amino acids, building blocks of proteins contribute for the electrostatic interactions in proteins. The strength of electrostatic interactions in and around proteins are explained by coulomb's law, the Poisson equation, the Boltzmann distribution and Born model [28]. The small molecule drugs interacting with proteins also possess high charge density as well as they undergo frequent protonation changes upon binding to their target protein. Hence, stabilization of electrostatic interactions between protein and small molecule drugs (ligands) confers stability for protein-ligand complexes [29]. Steric effects are classified as nonbonding interactions, which play vital role in determining the conformation and reactivity of organic compounds [30]. The presence of bulkier moieties on organic compounds results in slowing down of chemical reactions due to the higher magnitude of repulsive forces between the overlapping electronic clouds. Hence, electrostatic as well as steric effects play key role in determining the stability and reactivity of organic compounds. The present work focused on identifying the contribution of electrostatic as well as steric interactions in establishing stable interactions between renin and indole-3-carboxamide compounds.

Charge Q descriptor, illustrating the significance of electrostatic interactions was calculated by Electronegativity Equalization Method (EEM), which is a fast-empirical method for atomic charge calculation. The calculation of atomic charges depends on electronegativity,

charge distribution as well as distance between pair of atoms [31]. VdW.Re descriptor accounts for effective van der Waals radius of atom environment taking into account the steric requirements of central non-hydrogen atom and other attached atoms.

The bioactivity model constructed after performing MFTA calculations explained the contribution of local descriptors towards eliciting a specific biological activity. The positions marked with blue coloured circles in activity maps illustrated that an increase in descriptor tend to decrease the activity, while positions marked with red coloured circles indicated that increase in descriptor value tend to increase the activity. We have employed molecular docking studies to comprehend the biological significance of activity model obtained after performing MFTA calculations, since docking studies provide details regarding significant interactions made by chemical compounds with renin. The bioactivity model constructed for renin inhibitors explained the significance of nitrogen atoms in establishing specific interactions with aspartate residues present in the catalytic site of renin [32]. The presence of blue coloured circles (for charge Q descriptor) in the regions of nitrogen atoms explained the specificity of these groups in their respective positions. Blue coloured circles for charge Q descriptor in the position of nitrogen atoms signified the fact that replacement of nitrogen atoms with chemical species possessing higher electronegativity or charge distribution tend to decrease renin inhibitory activity by failing to inhibit key aspartate residues present in renin catalytic site. Molecular docking studies revealed the prevalence of hydrogen bond and hydrophobic interactions of target protein (renin) with carboxamide derivatives. This observation was correlated with the MFTA results by analysing the contribution of local descriptors. Hydrogen bond interactions were found between electronegative oxygen atoms and NH groups of carboxamide derivatives. The presence of blue coloured circles around these atoms signified that the replacement of chemical species having higher charge distribution than these atoms tend to decrease the potential to form hydrogen bond

interactions with the target protein. From the molecular docking studies, it was evident that the stability of protein-ligand complex is attributed to the hydrophobic interactions of carboxamide derivatives with renin protein. Hydrophobic interactions arise due to the presence of non-polar molecules, where the charge distribution has a null role. So, the regions of carboxamide derivatives making hydrophobic interactions with renin were also marked by the presence of blue coloured circles. Hence, observations from molecular docking studies and MFTA bioactivity model were found to be highly coherent. The Descriptor, VdW.Re accounts for the molecular size and steric effects of organic compounds. The presence of red coloured circles in the bioactivity model indicated that increase in the molecular size of rings positively contribute for the renin inhibitory activity. The compound S8000024 was chosen as reference for molecular supergraph and influence the local descriptor depiction, as this compound exhibited good binding energy as well as significant interactions with renin (obtained after conducting molecular docking studies).

The molecular supergraph and distribution of local descriptors of the compound S8000024 is provided in the (Figure 2).

3.2 Molecular docking studies

Molecular docking studies were executed to identify significant binding interactions of chemical compounds with target protein (renin). Active site residues of renin identified from computational tools such as DoGSiteScorer and ScanProsite were compared with the existing literatures on renin inhibitors. Thr12, Gln13, Asp32, Gly34, Ser35, Trp39, Tyr75, Ser76, Thr77, Pro111, Phe112, Phe117, Val120, Asp215, Ser219, Tyr220, His287 and Met289 were identified as active site residues. Asp32 and Asp215 residues were required for the catalytic activation of renin.

Molecular docking studies were performed by utilizing FlexX, which employs Incremental Construction (IC) algorithm for the accurate placement of ligand fragments into the protein active site. Active site definition was followed by execution of IC algorithm by providing optimized 3D coordinates of ligands. Top 20% docked poses of indole-3-carboxamide derivatives were generated by considering three different stereo modes (racemization) such as E/Z, R/S and pseudo R/S. Binding of ligand to protein was driven by enthalpy-entropy hybrid approach with 20,000 solutions per fragmentation. Binding affinity and ligand efficiency were calculated by performing HYDE assessment. Molecular docking studies were executed to explicate binding affinity and significant binding interactions between target protein (renin) and ligands. Compounds S8000024, S8000022 and S8000005 were identified as best binders of renin protein by considering binding affinity as well as interactions with key amino acid residues. Compound S8000024 exhibited higher binding energy. The binding affinity of this compound was found in nanomolar range. But all the three compounds showed interactions with Asp32 and Asp215 catalytic residues..

Dock score, binding energy, ligand efficiency and binding affinity range for best binders of target protein are provided in **(Table 3)**.

Dock score, binding energy, ligand efficiency and binding affinity range for all the compounds in the dataset is provided in Supplementary Table S2.

2D and 3D depictions of ligand-protein interactions of S8000024, S8000005 and S8000022 generated from PoseView are provided in **(Figure 3)**, **(Figure 4)** and **(Figure 5)** respectively. 2D depictions of binding interactions for all compounds in the dataset is provided in Supplementary Figure S5. Prioritization of effective lead molecules were done by considering few check points such as higher binding energy, good binding affinity (nanomolar range) and interactions with catalytic residues Asp32 and Asp215. The binding affinity of co-crystal

ligand was calculated in mM- μ M range with a dock score of -50.05 . The binding interactions of co-crystal ligand is depicted in the (**Figure 6**). Molecular docking studies were followed by the assessment of torsion alerts and atom-based free energy contributions.

3.3 Assessment of torsions and atom-based free energy contributions

The torsional alerts and atom-based free energy calculations were performed for all the compounds in the dataset using SeeSAR software. In addition to torsional alerts and atom-based free energy calculations, SeeSAR aided in the estimation of inter and intra clashes of ligands. Inter-clashes occur within ligands while intra-clashes occur with the target protein, with which ligand is interacting. The occurrence of intra-clashes results in the inappropriate placement of ligands, the ligand cannot even be placed into the protein's active site. The presence of inter-clashes is related with the stability of ligands. Torsion is considered as one of the significant structural parameters of ligands, which explicates the significance of ligand flexibility. Good/moderate torsions and clashes are favourable for making stable protein-ligand interactions. The values of torsion angle are considered to be positive values if torsion angle depicts a clockwise rotation is performed with the molecule and it will be negative when an anti-clockwise rotation is performed with the molecule in its plane. SeeSAR provides visual inspection of favourable/unfavourable torsion angles by the means of colouring scheme. Torsion angles falling within the first tolerance value around a local minimum are coloured green. Those between the first and second tolerance values are orange, and those beyond the second tolerance are coloured red [33].

The atom-based free energies were calculated by considering the change in entropy, change in enthalpy and split of the free energy. SeeSAR software aids in the visual inspection of

atom-based free energy contributions by providing coloured corona representations around atoms of ligand. The presence of green coloured corona on an atom depicts favourable contribution and the presence of red coloured corona on atom represents unfavourable contribution. The depiction of favourable and unfavourable contributions is based on the desolvation terms and free energy of binding of each atom present in the ligand. Negative values of free energy of binding represents favourable atom contributions (green coloured corona) while positive values of free energy of binding represents unfavourable atom contributions (red coloured corona).

Since, free energy of binding is an additive thermodynamic property, we have calculated total unfavourable contributions made by the atoms of S8000024, S8000005 and S8000022. The sum of unfavourable contributions made by atoms of S8000024 is 9.6 kJ/mol (6.4 kJ/mol + 1.2 kJ/mol+2.00 kJ/mol), 17.7 kJ/mol for S8000005 and 19.2 kJ/mol for S8000022. The comprehensive analysis of free energy contributions of atoms of the top ranked compounds (S8000024, S8000005 and S8000022) revealed least unfavourable atom contributions for the compound S8000024. The atom-based free energy contributions of co-crystal ligand present in renin was also calculated. A total of 22.9 kJ/mol of unfavourable atom contribution was estimated for the crystal ligand. Hence, unfavourable atom contributions of top binders (S8000024, S8000005 and S8000022) were found to be significantly lesser than the co-crystal ligand. The histogram peaks depicted in the figures represent most frequently observed torsion values. Compounds S8000024 and S8000022 exhibited favourable torsion angles (green colour). But S8000005 exhibited unfavourable torsion angle of -56° (red colour). Hence, all the top ranked compounds obtained from docking studies possessed appreciable structural features for making protein-ligand interactions. Even though compound S8000022 showed very good structural parameters, atoms present in this compound exhibited unfavourable contributions with renin (19.2 kJ/mol free energy of binding). The torsion,

interclash and intraclash of co-crystal ligand were calculated as moderate, good and good respectively.

(**Figure 7**), (**Figure 8**) and (**Figure 9**) represent atom-based free energy contributions and torsion angles for top ranked compounds obtained after conducting molecular docking studies.

The torsions, interclash and intraclash for S8000024, S8000005 and S8000022 are provided in (**Table 4**). The torsion alerts and clashes for all the compounds present in the dataset is provided in Supplementary Table S3.

3.4 Bioavailability assessment

Bioavailability assessment was carried out by StarDrop's (Absorption, Distribution, Metabolism, Elimination) ADME module software. The physicochemical features such as aqueous solubility, hERG (human Ether-a-go-go-Related Gene) liability, Human Intestinal Absorption (HIA) and affinity towards cytochrome p450 enzyme isoform were predicted. In addition, the toxicity alerts associated with the chemical compounds were predicted using vNN server.

Solubility is an essential physicochemical feature responsible for proper distribution of lead molecules in the biological compartments. LogP descriptor accounts for lipophilicity of chemical compounds. Lipophilicity is accountable for crossing lipoproteinaceous biological membranes. Hence logS and logP descriptors accounts for the penetration of lead like molecules into biological compartment [34]. 2C9 pKi descriptor accounts for the affinity of organic compounds towards cytochrome isoform. The metabolic stability of lead molecules was assessed by analysing affinity towards cytochromep450 isoforms [35]. Lead molecules should possess low/medium affinity towards cytochromep450 isoforms. Higher affinity

towards cytochrome p450 leads to faster rate of metabolic degradation of lead molecules before eliciting its therapeutic action. hERG (human Ether-a-go-go-Related Gene) liability assessment associated with the cardiotoxicity alert of compounds were analysed. hERG encodes alpha subunit of potassium ion channel, which contributes to the electrical activity of heart. Drugs inhibiting the biological functioning of this gene contribute to the clinical condition called ventricular tachyarrhythmia [36].

The bioavailability assessment was carried out using StarDrop's ADME module. Among the top 3 lead compounds, aqueous solubility was found to be higher for S8000024. The HIA profile, logP and affinity towards cytochrome isoform 2C9 were highly acceptable for S8000024. But, hERG liability was found to be higher for all the 3 lead molecules. All the 3 lead compounds were free from the toxicity endpoints such as cytotoxicity and mutagenicity. Bioavailability assessment results for top 3 compounds identified as best binders with renin are provided in (**Table 5**). Bioavailability assessment results for all compounds present in the dataset is provided in Supplementary Table S4.

By considering multiple parameters such as free energy of binding, torsional alerts, significant interaction with target protein (renin) and bioavailability, S8000024 was identified as potential lead molecule for exhibiting renin inhibitory activity. Even though S8000024 possessed good binding energy, atom-based free energy contributions, significant interactions with catalytic residues (Asp32 and Asp215) and good bioavailability, compound was prone to exhibit hERG liability. Hence, derivative compounds were identified for the compound S8000024 to identify highly potent lead molecules possessing better bioavailability than parent compound (S8000024).

The structural features of compound S8000024 depicting the physicochemical feature, **hERG pIC50** (predicted by StarDrop's glowing molecule module) is provided in (**Figure 10**). The

hERG liability of the compound S8000024 is attributed to the presence of the presence of nitrogen and oxygen atoms attached to the ring and alkyl chain.

The structural features of compound S8000024 contributing for all other physicochemical parameters encoding bioavailability is provided in the Supplementary Figure S6.

3.5 Identification of derivatives

New derivatives were identified for S8000024 and focusing on the unfavourable atoms in the binding affinity studies. Identification of derivatives was accomplished by utilizing inSight tool, which aids in the prediction of synthetically feasible compounds by navigating chemical spaces like Enamine, GalaXi, etc. The parent compound S8000024 was provided in SDF format and pharmacophore features required for making specific interactions with catalytic aspartate residues were defined. GalaXi, (chemical space created in collaboration with WuXi LabNetwork) comprising of 2 billion synthetically feasible chemical compounds was utilized for navigating derivative molecules.

3.6 Molecular docking studies for derivative compounds

The binding affinity as well as significant binding interactions of derivative molecules with target protein (renin) was analysed by executing molecular docking studies. Active site definition was followed by molecular docking as well as HYDE assessment studies. (**Table 6**) depicts top 6 compounds which showed excellent binding potential with renin.

The binding interactions of D1 is provided in (**Figure 11**). D1 exhibited hydrogen bond interactions with the catalytic aspartate (Asp215) residue. Out of top 6 derivative molecules

identified, compound D1 showed exceptional binding affinity of -35 kJ/mol and its binding affinity was identified in Nanomolar range. Compound D1 exhibited higher binding energy compared to parent compound (S8000024). But binding energy of all other derivative molecules were significantly lower and their binding affinity was predicted in millimolar range. The binding interactions of D1 with renin is represented in the (**Figure 11**).

The MFTA studies concluded the significance of charge distribution associated with the electronegative groups such as NH, and O (oxygen atom) in establishing stable hydrogen bond interactions with the target protein (renin). In coherent with the observations from MFTA studies, the electronegative groups (NH and O) present in the derivative compound D1 also exhibited hydrogen bond interactions with the active site residues of renin. Hence, the presence of these electronegative groups possessing specific charge distributions are significant for making stable interactions with the active site residues of renin.

3.7 Evaluation of torsional alerts and atom-based free energy contributions

The torsional alerts and atom-based free energy contributions were evaluated for the identified derivative compounds. The atom-based free energy contributions of compound D1, which demonstrated significantly higher binding energy values compared to all other derivative compounds, is provided in the (**Figure 12**).

Derivative compound D1 displayed a total of 9.9 kJ/mol atom-based unfavourable contribution, which is comparable with the parent compound S8000024, which revealed atom-based unfavourable contributions of 9.6 kJ/mol. D1 exhibited favourable torsion alerts (green colour) similar to the parent compound, S8000024.

3.8 Bioavailability assessment – derivative compounds

Bioavailability assessment were done for all the 6 derivative compounds identified by chemical space exploration. Physicochemical properties such as aqueous solubility, lipophilicity, hERG (human Ether-a-go-go-Related Gene) liability, Human Intestinal Absorption (HIA), and affinity towards cytochrome p450 enzyme isoform were predicted for derivative molecules. (**Table 7**). Depicts details on physicochemical properties associated with bioavailability assessment of derivative compounds.

Among 6 derivative compounds identified, D1 showed good bioavailability scores compared to all other derivatives. D1 exhibited improved scores for solubility (logS), partition coefficient (logP), lower affinity towards cytochrome p450 isoform, inactive for toxicity profiles and lower values for hERG liability compared to its parent compound S8000024. The improved bioavailability profile of compound D1 compared to its parent compound S8000024 makes it as an efficient drug candidate for eliciting renin inhibition. The structural features of derivative compound D1 contributing for reduced hERG pIC₅₀ is provided in (**Figure 13**).

The structural features of compound D1 contributing for all other physicochemical parameters encoding bioavailability is provided in the Supplementary figure S7.

This compound demonstrated excellent binding energy, significant interactions as well as favourable torsion alerts compared to parent compound S8000024. Hence, we propose compound D1 as a potential renin inhibitor with improved potency and bioavailability.

4. Conclusion

45 indole-3-carboxamide compounds reported for renin inhibitory activity were considered for this study. MFTA study revealed the significance of charge distribution as well as molecular size of chemical species in eliciting renin inhibition. Molecular docking, atom-based free energy contributions, torsion alerts and bioavailability assessment studies was performed to identify the most potent lead molecule among 45 compounds. Even though, S8000024 displayed higher binding energy, favourable torsion alerts, significant interactions with renin as well as good bioavailability, compound was prone to exhibit hERG liability. Hence, derivatives were identified for S8000024 by employing chemical space exploration. Among 6 derivatives identified, compound D1 showed good binding energy, atom-based free energy contributions, favourable torsion alerts as well as bioavailability scores compared to parent compound (S8000024).

5. Acknowledgements

GGP thank Alexander Oliferenko and N.S Zefirov for providing access to MFTA tool. Authors thank Zastra Innovations for providing facilities, access to BioSolveIT GmbH, Germany, Optibrium Ltd, UK software licenses and mentoring. AUP also thank Zastra Innovations for the financial support for the completion of this research study.

6. Funding

Authors thank Vellore Institute of Technology for providing financial support for carrying out this research work.

7. Declaration of competing interest

Authors report no potential conflict of interest.

8. Ethics Statement

The present study does not involve human/animal participants.

9. CRediT author statement

GGP aided in the conceptualization of the study, provided resources and ensured accuracy of the study. AUP carried-out experiments, formal analysis and wrote manuscript. SLS, GGP analysed and validated data together.

Supplementary Data

All datasets, descriptor values, QSAR models for predictions usage and plots will be available in <https://github.com/giribio/agingdata>

References

- [1] S.S. Khan, B.D. Singer, D.E. Vaughan, Molecular and physiological manifestations and measurement of aging in humans, *Aging Cell*. 16 (2017) 624–633. <https://doi.org/10.1111/accel.12601>.
- [2] A. Yazdanyar, A.B. Newman, The Burden of Cardiovascular Disease in the Elderly: Morbidity, Mortality, and Costs, *Clin. Geriatr. Med.* 25 (2009) 563–577. <https://doi.org/10.1016/j.cger.2009.07.007>.
- [3] G.A. Roth, C. Johnson, A. Abajobir, F. Abd-Allah, S.F. Abera, G. Abyu, M. Ahmed, B. Aksut, T. Alam, K. Alam, F. Alla, N. Alvis-Guzman, S. Amrock, H. Ansari, J. Ärnlöv, H. Asayesh, T.M. Atey, L. Avila-Burgos, A. Awasthi, A. Banerjee, A. Barac, T. Bärnighausen, L. Barregard, N. Bedi, E. Belay Ketema, D. Bennett, G. Berhe, Z. Bhutta, S. Bitew, J. Carapetis, J.J. Carrero, D.C. Malta, C.A. Castañeda-Orjuela, J. Castillo-Rivas, F. Catalá-López, J.Y. Choi, H. Christensen, M. Cirillo, L. Cooper, M. Criqui, D. Cundiff, A. Damasceno, L. Dandona, R. Dandona, K. Davletov, S. Dharmaratne, P. Dorairaj, M. Dubey, R. Ehrenkranz, M. El Sayed Zaki, E.J.A. Faraon, A. Esteghamati, T. Farid, M. Farvid, V. Feigin, E.L. Ding, G. Fowkes, T. Gebrehiwot, R. Gillum, A. Gold, P. Gona, R. Gupta, T.D. Habtewold, N. Hafezi-Nejad, T. Hailu,

- G.B. Hailu, G. Hankey, H.Y. Hassen, K.H. Abate, R. Havmoeller, S.I. Hay, M. Horino, P.J. Hotez, K. Jacobsen, S. James, M. Javanbakht, P. Jeemon, D. John, J. Jonas, Y. Kalkonde, C. Karimkhani, A. Kasaeian, Y. Khader, A. Khan, Y.H. Khang, S. Khera, A.T. Khoja, J. Khubchandani, D. Kim, D. Kolte, S. Kosen, K.J. Krohn, G.A. Kumar, G.F. Kwan, D.K. Lal, A. Larsson, S. Linn, A. Lopez, P.A. Lotufo, H.M.A. El Razek, R. Malekzadeh, M. Mazidi, T. Meier, K.G. Meles, G. Mensah, A. Meretoja, H. Mezgebe, T. Miller, E. Mirrakhimov, S. Mohammed, A.E. Moran, K.I. Musa, J. Narula, B. Neal, F. Ngalesoni, G. Nguyen, C.M. Obermeyer, M. Owolabi, G. Patton, J. Pedro, D. Qato, M. Qorbani, K. Rahimi, R.K. Rai, S. Rawaf, A. Ribeiro, S. Safiri, J.A. Salomon, I. Santos, M. Santric Milicevic, B. Sartorius, A. Schutte, S. Sepanlou, M.A. Shaikh, M.J. Shin, M. Shishehbor, H. Shore, D.A.S. Silva, E. Sobngwi, S. Stranges, S. Swaminathan, R. Tabarés-Seisdedos, N. Tadele Atnafu, F. Tesfay, J.S. Thakur, A. Thrift, R. Topor-Madry, T. Truelsen, S. Tyrovolas, K.N. Ukwaja, O. Uthman, T. Vasankari, V. Vlassov, S.E. Vollset, T. Wakayo, D. Watkins, R. Weintraub, A. Werdecker, R. Westerman, C.S. Wiysonge, C. Wolfe, A. Workicho, G. Xu, Y. Yano, P. Yip, N. Yonemoto, M. Younis, C. Yu, T. Vos, M. Naghavi, C. Murray, Global, Regional, and National Burden of Cardiovascular Diseases for 10 Causes, 1990 to 2015, *J. Am. Coll. Cardiol.* 70 (2017) 1–25. <https://doi.org/10.1016/j.jacc.2017.04.052>.
- [4] H.S. Buttar, T. Li, N. Ravi, Prevention of cardiovascular diseases: Role of exercise, dietary interventions, obesity and smoking cessation, *Exp. Clin. Cardiol.* 10 (2005) 229–249.
- [5] T.W. Buford, Hypertension and Aging HHS Public Access, *Ageing Res Rev.* 26 (2016) 96–111. <https://doi.org/10.1016/j.arr.2016.01.007>.
- [6] H.E. Yim, K.H. Yoo, Renin-angiotensin system - Considerations for hypertension and

- kidney, *Electrolyte Blood Press.* 6 (2008) 42–50.
<https://doi.org/10.5049/EBP.2008.6.1.42>.
- [7] M.A. Sparks, S.D. Crowley, S.B. Gurley, M. Mirotso, T.M. Coffman, Classical renin-angiotensin system in kidney physiology, *Compr. Physiol.* 4 (2014) 1201–1228.
<https://doi.org/10.1002/cphy.c130040>.
- [8] H.E. Yoon, E.N. Kim, M.Y. Kim, J.H. Lim, I.-A. Jang, T.H. Ban, S.J. Shin, C.W. Park, Y.S. Chang, B.S. Choi, Age-Associated Changes in the Vascular Renin-Angiotensin System in Mice, *Oxid. Med. Cell. Longev.* 2016 (2016).
<https://doi.org/10.1155/2016/6731093>.
- [9] A.M. Garrido, K.K. Griendling, NADPH oxidases and angiotensin II receptor signaling, *Mol. Cell. Endocrinol.* 302 (2009) 148–158.
<https://doi.org/10.1016/j.mce.2008.11.003>.
- [10] F. Turnbull, B. Neal, C. Algert, J. Chalmers, M. Woodward, S. MacMahon, C. Baigent, J. Cutler, R. Fagard, P. Whelton, S. Yusuf, N. Chapman, L. Agodoa, H. Black, J.P. Boissel, B. Brenner, M. Brown, C. Bulpitt, R. Byington, R. Collins, B. Dahlof, B. Davis, J. Dens, R. Estacio, K. Fox, L. Hansson, R. Holman, L. Hunsicker, J. Kostis, K. Kuramoto, E. Lewis, L. Lindholm, J. Lubsen, E. Malacco, G. Mancina, C. Pepine, M. Pfeffer, B. Pitt, P. Poole-Wilson, G. Remuzzi, A. Rodgers, P. Ruggenenti, R. Schrier, P. Sever, P. Sleight, J. Staessen, K. Teo, R. Turner, L. Wing, Y. Yui, A. Zanchetti, Effects of different blood-pressure-lowering regimens on major cardiovascular events: Results of prospectively-designed overviews of randomised trials, *Lancet.* 362 (2003) 1527–1535. [https://doi.org/10.1016/S0140-6736\(03\)14739-3](https://doi.org/10.1016/S0140-6736(03)14739-3).
- [11] A. Stanton, C. Jensen, J. Nussberger, E. O'Brien, Blood Pressure Lowering in Essential

- Hypertension with an Oral Renin Inhibitor, Aliskiren, *Hypertension*. 42 (2003) 1137–1143. <https://doi.org/10.1161/01.HYP.0000101688.17370.87>.
- [12] A. Stanton, C. Jensen, J. Nussberger, E. O'Brien, Blood pressure lowering in essential hypertension with an oral renin inhibitor, aliskiren., *Hypertens. (Dallas, Tex. 1979)*. 42 (2003) 1137–43. <https://doi.org/10.1161/01.HYP.0000101688.17370.87>.
- [13] A. Unni P, S.L. Sudhakaran, G.G. Pillai, Review on druggable targets of key age-associated properties regulated by therapeutic agents, *Chem. Biol. Drug Des.* (2020) cbdd.13759. <https://doi.org/10.1111/cbdd.13759>.
- [14] E.G. Da Mota, M.H. Duarte, E.F.F. Da Cunha, M.P. Freitas, Theoretical design of new indole-3-carboxamide derivatives as renin inhibitors, *Med. Chem. Res.* 24 (2015) 3097–3106. <https://doi.org/10.1007/s00044-015-1362-4>.
- [15] V. Palyulin, E. Radchenko, A. Melnikov, N. Zefirov, Molecular field topology analysis and structure generation, *Chem. Cent. J.* 2 (2008) 2007. <https://doi.org/10.1186/1752-153x-2-s1-p8>.
- [16] R.S. Vařeková, Z. Jiroušková, J. Vaněk, Š. Suchomel, J. Koča, Electronegativity Equalization Method: Parameterization and Validation for Large Sets of Organic, Organohalogen and Organometal Molecule, *Int. J. Mol. Sci.* 8 (2007) 572.
- [17] E. V. Radchenko, V.A. Palyulin, N.S. Zefirov, Molecular field topology analysis (MFTA) in the design of neuroprotective compounds, in: *Neuromethods*, Humana Press Inc., 2018: pp. 139–159. https://doi.org/10.1007/978-1-4939-7404-7_5.
- [18] H. Abdi, Partial least squares regression and projection on latent structure regression (PLS Regression), *Wiley Interdiscip. Rev. Comput. Stat.* 2 (2010) 97–106. <https://doi.org/10.1002/wics.51>.

- [19] R. Kiralj, M.M.C. Ferreira, Basic validation procedures for regression models in QSAR and QSPR studies: Theory and application, *J. Braz. Chem. Soc.* 20 (2009) 770–787. <https://doi.org/10.1590/S0103-50532009000400021>.
- [20] M.R. Andrea Volkamer, Daniel Kuhn, Friedrich Rippmann, DoGSiteScorer: a web server for automatic binding site prediction, analysis and druggability assessment | *Bioinformatics* | Oxford Academic, <https://doi.org/10.1093/Bioinformatics/Bts310>. (2012) 2074–2075. <https://academic.oup.com/bioinformatics/article/28/15/2074/236684> (accessed 4 June 2020).
- [21] E. de Castro, C.J.A. Sigrist, A. Gattiker, V. Bulliard, P.S. Langendijk-Genevaux, E. Gasteiger, A. Bairoch, N. Hulo, ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins., *Nucleic Acids Res.* 34 (2006) W362-5. <https://doi.org/10.1093/nar/gkl124>.
- [22] M. Rarey, B. Kramer, T. Lengauer, G. Klebe, A Fast Flexible Docking Method using an Incremental Construction Algorithm, *J. Mol. Biol.* 261 (1996) 470–489. <https://doi.org/10.1006/jmbi.1996.0477>.
- [23] N. Schneider, G. Lange, S. Hindle, R. Klein, M. Rarey, A consistent description of HYdrogen bond and DEhydration energies in protein-ligand complexes: methods behind the HYDE scoring function., *J. Comput. Aided. Mol. Des.* 27 (2013) 15–29. <https://doi.org/10.1007/s10822-012-9626-2>.
- [24] P. Schyman, R. Liu, V. Desai, A. Wallqvist, vNN Web Server for ADMET Predictions, *Front. Pharmacol.* 8 (2017) 889. <https://doi.org/10.3389/fphar.2017.00889>.
- [25] T. Hoffmann, M. Gastreich, The next level in chemical space navigation: going far beyond enumerable compound libraries, *Drug Discov. Today.* 24 (2019) 1148–1156.

- <https://doi.org/10.1016/j.drudis.2019.02.013>.
- [26] X. Du, Y. Li, Y.L. Xia, S.M. Ai, J. Liang, P. Sang, X.L. Ji, S.Q. Liu, Insights into protein–ligand interactions: Mechanisms, models, and methods, *Int. J. Mol. Sci.* 17 (2016). <https://doi.org/10.3390/ijms17020144>.
- [27] G. Bitencourt-Ferreira, M. Veit-Acosta, W.F. de Azevedo, Electrostatic energy in protein–ligand complexes, in: *Methods Mol. Biol.*, Humana Press Inc., 2019: pp. 67–77. https://doi.org/10.1007/978-1-4939-9752-7_5.
- [28] P. Kuki, J.E. Nielsen, Electrostatics in proteins and protein-ligand complexes, *Future Med. Chem.* 2 (2010) 647–666. <https://doi.org/10.4155/fmc.10.6>.
- [29] H.X. Zhou, X. Pang, Electrostatic Interactions in Protein Structure, Folding, Binding, and Condensation, *Chem. Rev.* 118 (2018) 1691–1741. <https://doi.org/10.1021/acs.chemrev.7b00305>.
- [30] A. V Brethomé, Brethomé, S.P. Fletcher, R.S. Paton, Conformational Effects on Physical-Organic Descriptors: The Case of Sterimol Steric Parameters, (2019). <https://doi.org/10.1021/acscatal.8b04043>.
- [31] P. Bultinck, W. Langenaeker, P. Lahorte, F. De Proft, P. Geerlings, M. Waroquier, J.P. Tollenaere, The electronegativity equalization method I: Parametrization and validation for atomic charge calculations, *J. Phys. Chem. A.* 106 (2002) 7887–7894. <https://doi.org/10.1021/jp0205463>.
- [32] B. Scheiper, H. Matter, H. Steinhagen, U. Stolz, Z. Böcskei, V. Fleury, G. McCort, Discovery and optimization of a new class of potent and non-chiral indole-3-carboxamide-based renin inhibitors, *Bioorganic Med. Chem. Lett.* 20 (2010) 6268–6272. <https://doi.org/10.1016/j.bmcl.2010.08.092>.

- [33] C. Schärfer, T. Schulz-Gasch, H.C. Ehrlich, W. Guba, M. Rarey, M. Stahl, Torsion angle preferences in druglike chemical space: A comprehensive guide, *J. Med. Chem.* 56 (2013) 2016–2028. <https://doi.org/10.1021/jm3016816>.
- [34] L. Sultatos, Drug absorption, in: *XPharm Compr. Pharmacol. Ref.*, Elsevier Inc., 2007: pp. 1–2. <https://doi.org/10.1016/B978-008055232-3.60020-0>.
- [35] U.M. Zanger, M. Schwab, Cytochrome P450 enzymes in drug metabolism: Regulation of gene expression, enzyme activities, and impact of genetic variation, *Pharmacol. Ther.* 138 (2013) 103–141. <https://doi.org/10.1016/j.pharmthera.2012.12.007>.
- [36] M.E. Curran, I. Splawski, K.W. Timothy, G.M. Vincent, E.D. Green, M.T. Keating, A molecular basis for cardiac arrhythmia: HERG mutations cause long QT syndrome., *Cell*. 80 (1995) 795–803. [https://doi.org/10.1016/0092-8674\(95\)90358-5](https://doi.org/10.1016/0092-8674(95)90358-5).

Figure captions

Figure 1: Scatter plot depicting correlation between experimental and predicted values.

Figure 2: **(2a)** Depicts the molecular supergraph of S8000024. **(2b)**. Represents the distribution of local descriptor charge Q on S8000024. The regions marked with blue coloured circles in activity maps illustrates that an increase in descriptor value tend to decrease the activity, while positions marked with red coloured circles indicate that increase in descriptor value tend to increase the activity. **(2c)** : Represents the distribution of local descriptor VdW.Re on S8000024. The positions marked with red coloured circles indicate that increase in descriptor value (VdW.Re) tend to increase the activity.

Figure 3: Binding interactions of S8000024 with renin. The ligand (S8000024) is depicted in stick model.

Figure 4: Binding interactions of S8000005 with renin. Ligand (S8000005) is depicted in stick model.

Figure 5: Two dimensional (2D) and three dimensional (3D) depiction of binding interactions of the compound S8000022. Ligand S8000022 is depicted in stick model.

Figure 6: Binding interactions of the co-crystal ligand (depicted in stick model) with renin protein.

Figure 7: Depicts atom-based free energy contributions of S8000024 interacting with renin (target protein). Positive values of free energy of binding is represented by red coloured corona (6.4 kJ/mol 2.00 kJ/mol and 1.2 kJ/mol).

Figure 8: Illustrates atom-based free energy contributions of S8000005 interacting with renin (target protein). Positive values of free energy of binding is represented by red coloured corona (10.9 kJ/mol 4.6 kJ/mol and 2.2 kJ/mol).

Figure 9: Illustrates atom-based free energy contributions of S8000022 interacting with renin (target protein). Positive values of free energy of binding is represented by red coloured corona (9.2, kJ/mol, 2.6 kJ/mol 4.2 kJ and 3.2 kJ/mol).

Figure 10: Represents the contribution of structural features for enhancing hERG liability. Regions represented in red colour shows positive contribution and regions marked in blue colour indicate least contribution for eliciting each of the respective physicochemical property.

Figure 11: Represents binding interactions of D1 with renin. Compound D1 is shown in stick model.

Figure 12: Portrays atom-based free energy contributions of derivative compound D1 interacting with renin (target protein). Positive values of free energy of binding is represented by red coloured corona (6.6 kJ/mol, 1.3 kJ/mol, 1.3 kJ/mol and 0.7 kJ/mol).

Figure 13: Portrays contribution of structural features of D1 for the bioavailability feature, hERG liability. Regions represented in red colour shows positive contribution and regions marked in blue colour indicate least contribution for eliciting each of the respective physicochemical property.

Sl.no	Descriptor codes
1	Charge.Q, VdW.Re
2	Charge.Q, VdW.Re, HBond.Hd, HBond.Ha
3	Charge.Q, VdW.Re, Lipo.Lg
4	Charge.Q, VdW.Re, HBond.Hd, HBond.Ha, Lipo.Lg

Tables

Table 1: Represents descriptor combinations for executing MFTA calculations.

Table 2: Represents significant statistical parameters related to the model.

Statistical parameters	Values
r^2	0.81
Q^2	0.65

Root Mean Square Error (RMSE)	0.17
No. of factors	3

Table 3. Visually scrutinized best poses from docking studies.

Sl.no	Compound ID	Score	Energy (kJ/mol)	Ligand efficiency	Binding affinity
					range
1	S8000024	-32.37	-34	0.22	nM
2	S8000005	-35.33	-31	0.22	mM- μ M
3	S8000022	-34.87	-29	0.23	mM- μ M

Table 4: Illustrates structural parameters such as torsion, interclash and intraclash for top ranked compounds S8000024, S8000005 and S8000022.

Compound	Torsion	Intra-clash	Inter-clash
S8000024	Good	Moderate	Moderate
S8000005	Moderate	Good	Good
S8000022	Good	Good	Good

Table 5. Physicochemical properties to predict bioavailability for lead compounds.

Sl.no	ID	logS	logP	2C9 pKi	hERG pIC5	Human Intestinal Absorption	Cytotoxicity	Mutagenicity
-------	----	------	------	------------	--------------	-----------------------------------	--------------	--------------

0 (HIA)								
1	S800002	1.88	3.69	5.8	6.5	+	No	No
2	S800000	1.2	4.32	5.9	6.4	+	No	No
3	S800002	1.44	2.52	5.4	6	+	No	No

Table 6. Depicts dock scores, HYDE assessment scores for top binding compounds.

Sl.no	Compound ID	Dock score	Binding energy (kJ/mol)	Binding affinity range
1	D1	-27.13	-35	nM
2	D2	-22.63	-6	mM-μM
3	D3	-18.28	-10	mM-μM
4	D4	-11.41	-10	mM-μM
5	D5	-10.54	-6	mM-μM
6	D6	-19.04	-6	mM-μM

Table 7. Depicts bioavailability scores for top derivative compounds.

Sl.no	ID	logS	logP	2C9 pKi	hERG pIC50	HIA	Cytotoxicity	Mutagenicity
1	D1	2.586	3.369	5.565	4.9	+	No	No
2	D2	1.46	3.433	5.647	5.7	+	No	Yes

3	D3	2.055	3.226	5.377	5.4	+	No	No
4	D4	2.454	2.88	5.535	5.4	+	No	Yes
5	D5	0.6352	4.957	6.088	6.1	+	No	No
6	D6	0.3239	4.474	5.774	6.1	+	No	No
