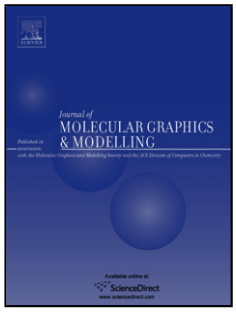


Title: Structure-based virtual screening as a tool for the identification of novel inhibitors against *Mycobacterium tuberculosis* 3-dehydroquinate dehydratase

Author: Guilherme O. Petersen Shalini Saxena Janupally Renuka Vijay Soni Perumal Yogeeswari Diogenes S. Santos Cristiano V. Bizarro Dharmarajan Sriram

PII: S1093-3263(15)00079-0
DOI: http://dx.doi.org/doi:10.1016/j.jmgm.2015.05.001
Reference: JMG 6541

To appear in: *Journal of Molecular Graphics and Modelling*

Received date: 17-12-2014
Revised date: 10-4-2015
Accepted date: 1-5-2015

Please cite this article as: G.O. Petersen, S. Saxena, J. Renuka, V. Soni, P. Yogeeswari, D.S. Santos, C.V. Bizarro, D. Sriram, Structure-based virtual screening as a tool for the identification of novel inhibitors against *Mycobacterium tuberculosis* 3-dehydroquinate dehydratase, *Journal of Molecular Graphics and Modelling* (2015), http://dx.doi.org/10.1016/j.jmgm.2015.05.001

Research Highlights

- Structure-based Pharmacophore models of Mtb DHQase protein was established
- Pharmacophore model was validated and used as a 3D search query
- Identification of top five lead molecules after virtual screening.
- *In vitro* enzymatic studies leads identification of top lead molecule.

Structure-based virtual screening as a tool for the identification of novel inhibitors against *Mycobacterium tuberculosis* 3-dehydroquinate dehydratase

Guilherme O Petersen[a][b][c], Shalini Saxena[c], Janupally Renuka [c], Vijay Soni[c], Perumal Yogeeswari[c], Diogenes S Santos[a][b], Cristiano V Bizarro*[a][b], Dharmarajan Sriram* [c].

[a] Centro de Pesquisas em Biologia Molecular e Funcional (CPBMF), Instituto Nacional de Ciência e Tuberculose (INCT-TB), Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS), Av. Ipiranga 6681, 90619-900 Porto Alegre, Brazil

[b] Programa de Pós-Graduação em Biologia Celular e Molecular, PUCRS, Porto Alegre 90619-900, Brazil

[c] Department of Pharmacy, Birla Institute of Technology & Science-Pilani, Hyderabad Campus, Hyderabad 500078, India

*For correspondence: D. Sriram, Telephone: +91-4066303506; Fax: +91-4066303998;
Email: dsriram@hyderabad.bits-pilani.ac.in

*Cristiano V Bizarro Telephone: +5551333203629
Email: cristiano.bizarro@pucrs.br

Abstract

3-Dehydroquinate dehydratase (DHQase), the third enzyme of the shikimate pathway, catalyzes the reversible reaction of 3-dehydroquinate into 3-dehydroshikimate. The aim of the present study was to identify new drug-like molecules as inhibitors for *M. tuberculosis* DHQase employing structure-based pharmacophore modeling technique using an in house database consisting of about 2500 small molecules. Further the pharmacophore models were validated using enrichment calculations, and finally three models were employed for high-throughput virtual screening and docking to identify novel small molecules as DHQase inhibitors. Five compounds were identified, out of which, one molecule (**Lead 1**) showed 58% inhibition at 50 μM concentration in the Mtb DHQase assay. Chemical derivatives of the **Lead 1** when tested evolved top two hits with IC_{50}s of 17.1 and 31.5 μM as well as MIC values of 25 and 6.25 μg/mL respectively and no cytotoxicity up to 100 μM concentration.

Keywords: Tuberculosis, Drug discovery, Shikimate pathway, DHQase, Virtual screening.

1. Introduction

Tuberculosis (TB) is the second greatest killer worldwide due to a single infectious agent being responsible for 1.3 million deaths and 8.6 million people infected in 2012 [1]. According to WHO (World Health Organization), approximately 2 billion people are latently infected and about 10% of those person could develop active disease during their lifetime. Moreover, 95% of TB deaths occur in low- and middle-income countries, and it is among the top three causes of death for women aged between 15 and 44. It was estimated that 530,000 children became ill and 74,000 HIV-negative children died of TB in 2012. Also, TB is a leading killer of people living with HIV, causing one fifth of all deaths of HIV-infected people. *M. tuberculosis* is the main causative agent of TB in humans and the dissemination of multi-drug-resistant strains challenges the existing treatments for TB [1].

Targeting essential metabolic pathways of micro-organisms absent in humans is an attractive strategy for the development of new therapies. The shikimate pathway produces an important precursor of aromatic compounds in bacteria, fungi, plants and apicomplexan parasites [2-5]. The pathway comprises of seven enzymes each catalyzing a separate step that converts erythrose-4-phosphate and phosphoenolpyruvate into chorismate, the precursor for the synthesis of aromatic amino acids, folic acid, ubiquinone, and many other aromatic compounds [2, 6-7]. As reported from gene disruption studies, enzymes from the shikimate pathway were found to be essential for *M. tuberculosis* survival [8].

3-Dehydroquinate dehydratase (DHQase) is the third enzyme of the shikimate pathway catalyzing the reversible dehydration of 3-dehydroquinate into 3-dehydroshikimate. There are two forms of DHQase (type I and type II), with distinct structures, but both catalyze the overall reaction [9]. *M. tuberculosis* uses type II DHQase (Mtb DHQase), a 17.79 kDa protein containing 147 amino acids. Crystallographic data revealed Mtb DHQase as a homododecamer formed from a tetramer of trimers with the active site found near the C-

terminal region of strands β1 and β3 of the parallel β-sheet [9,10]. The present study involved structure based design of Mtb DHQase inhibitors using energy optimized pharmacophore as a query to small molecule database searching and retrieved suitable hits based on pharmacophore fitness, followed by docking in the active site of DHQase protein. Finally, five compounds were short-listed and further evaluated for their *in vitro* DHQase assay.

2. Materials and methods

2.1 Computational details

All the computational studies were carried out on an Intel Core 2 Duo E7400 2.80 GHz capacity processor with memory of 2 GB RAM running with RHEL 5.2 operating system. PHASE 3.3 implemented in the Maestro 9.3 software package (Schrodinger, LLC) was used to generate e-pharmacophore models [11]. Glide energy grid was generated for the prepared protein complex. Binding site was defined by a rectangular box surrounding the x-ray ligand. Ligand was refined using the "Refine" option in Glide, and the option to output Glide XP descriptor information was chosen (Glide v5.7, Schrodinger, LLC, New York, NY). For refinement and docking calculations, the default settings as available in the software package were used.

2.2 Protein and ligand preparation

The crystal structure of Mtb DHQase bound to inhibitor (PDB ID: 2Y71) with a resolution of 2.0 Å was retrieved from PDB (protein data bank) and was prepared using the Protein Preparation Wizard of Schrödinger Suite 9.3 [12]. The protein was subjected to many steps such as addition of hydrogens, bond order and formal charge corrections, adjustment of tautomerization and ionization states of protein etc. The water molecules were removed from the protein and the hydrogen bonding network was optimized by reorienting the hydroxyl and thiol groups in the protein residues. Finally the protein was subjected to energy minimization using OPLS_2005 (Optimized Potential for Liquid Simulations) force field. The reference

ligand was also retrieved from PDB which had a reported K_i value of approximately 45 nM [13] and was subjected to energy minimization using impact of Schrodinger suite version 9.3. The ligand was minimized using 500 cycles each of SD (Steepest Descent) and CG (Conjugate Gradient). The interactions of the ligand with the protein residues in the active site were visualized using ligand interaction diagrams in Schrodinger suite version 9.3

2.3 Glide XP (extra-precision) docking

The prepared crystal structure of protein was used for grid generation, specifying the refined x-ray ligand as a reference to identify the active site. The docking was done by Glide XP (extra precision) protocol where the output gives a docking pose with XP glide score with rmsd (root mean square deviation) to crystal ligand at active site. The score can be considered as a total of binding energies of ligand with the protein (bonding and non-bonding interactions). For pharmacophore hypothesis generation, the energies of hydrophobic interactions, cation-π and π-π interactions were given out in XP descriptor file. The XP Glide scoring function was used to order the best ranked compounds and the specific interactions like π-cation and π-π stacking were analyzed using XP visualizer in Glide module.

2.4 E-pharmacophore generation

The e-pharmacophore hypothesis was created for the ligand by using the xpdes results of the Glide XP output in the docking post processing tool of the scripts module. Starting with the refined crystal ligand, pharmacophore sites were automatically generated with Phase (Phase, v3.0, Schrodinger, LLC, New York, NY) [14-15] using the default set of six chemical features: hydrogen bond acceptor (A), hydrogen bond donor (D), hydrophobic (H), negative ionizable (N), positive ionizable (P), and aromatic ring (R). Hydrogen bond acceptor sites were represented as vectors along the hydrogen bond axis in accordance with the hybridization of the acceptor atom. Hydrogen bond donors were represented as projected points located at the corresponding hydrogen bond acceptor positions in the binding site. The

ligand was docked with Glide XP and the pose was refined. The Glide XP scoring terms were computed, and the energies were mapped onto atoms. The pharmacophore sites were generated, and the Glide XP energies from the atoms that comprised each pharmacophore sites were summed up. These sites were then ranked based on the individual energies, and the most favorable sites were selected for the pharmacophore hypothesis [16]. This pharmacophore model was then used as query for virtual screening.

2.5 Dataset preparation

BITS-Pilani *in house* database containing 2500 compounds was considered for the e-pharmacophore based screening. All the compounds were prepared using LigPrep module [17] of Schrodinger 9.3 so as to generate high quality structures with appropriate ionization states, tautomers, and ring conformations and stereochemistry. Epik was employed for adjustment of tautomerization and ionization states of molecules at pH 7.0±2.0. Ligand conformers were generated for all the database molecules using ConfGen [17]. All the molecules were energy minimized and filtered using the OPLS_2005 force field, by means of a distance dependent dielectric constant with a prefactor of 4. A maximum of 32 conformations for every ligand were generated and carried forward for virtual screening.

2.6 E-Pharmacophore validation

For the fraction of known actives recovered when the database was screened, enrichment factor (EF) was employed. In this study, decoy set consisted of 1000 molecules with an average molecular weight of 400 kDa which were available for download (http://www.schrodinger.com/glidedecoyset). In decoy set further 15 known active molecules of DHQase inhibitors were included for validation. For this, we focused primarily on EF (1%), the enrichment in the top 1% of the decoys, and along with goodness of fit (GH), % actives, % yield were calculated using the following equations,

$$EF = \frac{(H_s \times D)}{(H_t \times A)} \qquad (1)$$

$$GH = \left(\left(\frac{H_a}{4H_tA}\right) \times (3A + H_t)\right) \times \left(1 - \left(\frac{H_t - H_a}{D - A}\right)\right) \quad (2)$$

$$\% \text{ Yield} = \left[\left(\frac{Ha}{Hb}\right) \times 100\right] \quad (3)$$

$$\%A = \left[\left(\frac{Ha}{A}\right) \times 100\right] \quad (4)$$

Where

'H_t' was total number of compounds in the hit list

'H_a' was the total number of active molecules in the hit list

'A' was the total number of actives in the decoy set

'D' was the total number of molecules in the decoy set

2.7 E-pharmacophore based screening and docking

For the e-pharmacophore approach, explicit matching was required for the most energetically favorable site (scoring better than -1.0 kcal mol^{-1}). For filtering the database molecules, a minimum of 4 sites were required to match for hypotheses with 5 or more sites [18]. The above criterion was followed in the present work to screen the *in house* database. In order of their fitness score, database hits were ranked to measure how well the aligned ligand conformer matched the hypothesis based on RMSD, site matching, vector alignments and volume terms [19]. Database ligands after e-pharmacophore filter were docked into the binding sites of the protein with Glide 5.7 (Glide, version 5.7, Schrodinger, LLC, New York, NY, 2011), utilizing the high-throughput virtual screening (HTVS) scoring function to estimate protein-ligand binding affinities. Ligands filtered from HTVS were subjected to Glide SP (standard precision) docking. Compounds with best docking and Glide scores were then finally subjected to Glide XP (extra precision) docking.

2.8 Molecular Docking using GOLD

The Gold 4.1.2 program (Genetic Optimization for Ligand Docking) from the Cambridge Crystallographic Data Centre, UK, was used to analyze the binding orientations of the top

identified compound obtained from the *in house* databases. The binding orientations of these compounds were compared with the Glide docking results for final validation. The crystal structure of Mtb DHQase (PDB: 2Y71) with a resolution of 1.50 Å was retrieved from PDB (protein data bank) and was prepared using the Protein Preparation Wizard of Schrödinger Suite 9.3. The energy minimized structure was used for further docking analysis using the Gold docking software/program with the default parameters: population size (100); selection-pressure (1.1); number of operations (10,000); number of islands (1); niche size (2); and operator weights for migrate (0), mutate (100) and crossover (100). The active site was defined within 10 A ˚and the ligand binding sites were analysed. The docked poses were scored using the scoring functions GoldScore (GS) to find the required docking pose.

2.9 Cloning and purification of Mtb DHQase

All enzymes were purchased from New England Biolabs. Oligos and other chemicals were procured from Sigma. pQEII vector was obtained from Qiagen. 3-Dehydroquinate dehydratase (AroD) was PCR amplified, using specific forward primer (5'CACCCATATGAGCGAACTGATCGTGAACGTG3') and reverse primer (5'AGCTAAGCTTTCACGTCCCGACATGCTCAGCTAG3'), from *M. tuberculosis* genomic DNA. Pfu polymerase (NEB) was used with 60°C of primer annealing temperature. PCR products were digested with NdeI and HindIII followed by cloning into NdeI and HindIII sites of pQEII vector. Sequence confirmed clones were further transformed into BL21 (DE3) CODON PLUS cells for expression and purification. Transformants were grown in LB broth at 37°C with constant aeration in the presence of 100 μg/mL ampicilin. Exponentially growing cultures (A_{600} of 0.6) were induced with 1.0 mM IPTG and further grown for 4–6 h at 37°C. Cells were then harvested and lyzed by sonication in lysis buffer. The cell lysate containing His6-fusion proteins were mixed with equilibrated Ni-NTA affinity resins and the tagged proteins were eluted with buffer containing 750 mM imidazole.

2.10 In vitro Mtb DHQase enzyme assay

Enzymatic assay for Mtb DHQase was performed using Molecular Devices® SpectraMax M4 Multiplate and Cuvette reader. Photometric determination of DHQase activity was accomplished by measuring the rate of production of 3-dehydroshikimate at 234 nm in a 1 mL quartz cuvette. The assays were performed in duplicate at 25°C in 50 mM Tris/HCl (pH 7.0). A final concentration of 50 nM of Mtb DHQase was used. The assay was initiated by adding the enzyme after incubation of buffer and substrate (3-dehydroquinate) at 25°C. The values for K_m and V_{max} were obtained by measuring the reaction over a range of different substrate concentrations. The data was processed using GraphPad Prism® (GraphPad Software).

2.11 Antibacterial activity

The compounds were further screened for their *in vitro* antimycobacterial activity against *M. tuberculosis* H37Rv strain by microplate alamar blue assay method. Briefly, the inoculum was prepared from fresh LJ medium, re-suspended in 7H9-S medium (7H9 broth, 0.1% casitone, 0.5% glycerol) supplemented with OADC (oleic acid, albumin, dextrose, and catalase), adjusted in a McFarland tube number 1, and diluted 1:20 and 100 µl was used as inoculum [20]. Each drug stock solution was thawed and diluted in 7H9-S at four-fold the final highest concentration tested. Serial two-fold dilutions of each drug were prepared directly in a sterile 96-well microtiter plate using 100 µl 7H9-S. A growth control containing no antibiotic and a sterile control were also prepared for each plate. In order to avoid evaporation during incubation for 7 days, sterile water was added to all perimeter wells. The plate was covered, sealed in plastic bags and incubated at 37°C in normal atmosphere. After 7 days of incubation, 30 µl of alamar blue solution was added to each well, and the plate was further re-incubated overnight. A change in color from blue (oxidized state) to pink (reduced) indicated the growth of bacteria, and the minimal inhibitory concentration (MIC) was defined

as the lowest concentration of drug that prevented this change in color.

2.12 Cytotoxicity assay

The cytotoxic activity against HEK293 (human embryonic kidney cells) cellswas measured by incubating the test compounds in 96-well plate containing 5×10^5 cells at different concentrations at 37°C, with 5% CO_2 and 95% O_2 atmosphere for 48 h [21]. About 4 h before the end of incubation period, 10 µL of MTT reagent (10 mg mL^{-1}) was added and the 96-well plate was centrifuged at 1200 rcf for about 3 min and the supernatants were removed. Subsequently to each well, 200 µL of DMSO was added. The absorbance was measured at a wavelength of 560 nm on Perkin Elmer Victor X3 microplate reader against the blank. The assay was performed in triplicates for each concentration of drug to minimize the error rate. The cytotoxicity of each compound was expressed as % inhibition.

2.13 Biophysical characterization

Differential scanning fluorimetry (DSF) is a screening method used to identify ligands that bind and stabilize purified proteins. The ability of the compound to stabilize Mtb DHQase was assessed by measuring the fluorescence of the native protein and the protein-ligand complex in the presence of a fluorescent dye whose fluorescence increased when exposed to non-polar residues of the protein and reached a maximum when the protein denatured [22,23]. Native Mtb DHQase was compared to Mtb DHQase in the presence of ligand by heating from 25 to 95°C in steps of 0.1 °C in the presence of the fluorescent dye sypro orange. As the protein denatured, non-polar residues are exposed to the dye, increasing its fluorescence. A right side positive shift of melting temperature (T_m) in comparison to native protein indicated higher stabilization of the protein-ligand complexes, which was a consequence of the inhibitor binding.

3. Results and discussion

A major goal of rational drug design would be to identify and develop new ligands

with high inhibitory activity towards a cellular target. The characterization of the binding groups involved in protein-ligand interactions is crucial for understanding an inhibitor's binding mechanism, a prerequisite to further improve the inhibitory activity in a rational way. A very useful technique for achieving this goal was reported to be the structure based pharmacophore modeling [24]. In the present study, the reference ligand bound to the protein was re-docked with Mtb DHQase to generate a pharmacophore hypothesis to screen the *in house* database compounds. The identified compounds were further evaluated for *in vitro* enzymatic inhibition and antimycobacterial potential.

3.1 e-Pharmacophore generation of crystal ligand

The crystal structure of Mtb DHQase (PDB: 2Y71) bound with an inhibitor was used for structure-based pharmacophore modeling. The protein was considered for grid generation using crystal ligand as a reference to identify the active site. The ligand was optimized for hydrogen network and refined using OPLS_2005 force field. Originally this reference ligand was showing hydrogen bonding interaction with Arg112, His81, Ser103, Ile102, Asn75, His101 and Tyr24 amino acid residues. To validate the active site cavity reference ligand was re-docked with the active site residues of the Mtb DHQase protein. The crystal ligand was also retrieved from PDB and was subjected to energy minimization using impact of Schrödingersuiteversion9.3. The ligand was minimized by using 500 cycles each of SD (steepest descent) and CG (conjugate gradient).

The ligand exhibited a Glide score of -11.12 kcal mol^{-1} and was found in the vicinity of important amino acids like Arg112, His81, Ser103, Ile102, Asn75, His101 and Tyr24. Re-docking results showed that the compound exhibited interactions similar to the ones found in the original crystal structure and the RMSD was found to be 0.340 Å (Fig. 1). The result of this docking was used to ascertain pharmacophoric features for Mtb DHQase inhibition.

Insert Fig. 1.

The initial number of pharmacophore sites was set up to 10 for the crystal structure. The pharmacophore generation protocol applied for the ligand yielded in a hypothesis with seven features as shown in Fig 2. The e-pharmacophore features consisted of two hydrogen donor (D) moieties, two hydrogen acceptor (A) groups, two aromatic ring (R) interactions and one negatively (N) charged group. The results of the e-pharmacophore features of the reference ligand with their respective energetics are given in Table 1. The energy contribution for binding of ligand to the protein was the key to derive pharmacophoric features in structure-based design.

Insert Fig. 2.

The A3 and D6 features corresponded to the third hydroxyl group of ligand which was involved in H-bond with His81 and Arg112. A2 and D5 corresponded to carboxylic group of ligand which was involved in H-bond interactions with Ser103, Asn75 and Ile102; while the N9 feature was found interacting with Asn75 and His101 in the protonated state amino acid residues. The two aromatic ring features (R10 and R11) were mapped to the two aromatic rings in the ligand which were mainly involved in π-π interactions.

Insert Table 1

To validate the pharmacophore model, we made combinations of pharmacophores having 6, 5, 4 and 3 point features totaling 36 combinations on the basis of their energy scores as represented in Table 2.

3.2 Validation of constructed pharmacophore models

The enrichment results for all the 36 pharmacophoric hypotheses were compared based on enrichment factor (EF) and % of actives based on recovery rate against the ranked decoy database. This helped us to eliminate the pharmacophore models that lacked significant interactions as well as allowed to prioritize the sites for further virtual screening, leading to good enrichment as well as diversity in hits. The EF reflected the capability of a screening

application to detect active ligands (true positives) compared to random selection. Thus, the value was expected always to be greater than 1 and the higher it was, the better the enrichment performance for virtual screening [25]. The recovery rate of actives was performed against the total 1015 compounds in which 15 were known inhibitors of Mtb DHQase and 1000 were the decoy set which represented inactives. The 6 point (A2, A3, D5, D6, N9, and R11), 5 point (D5, D6, N9, R10, R11) and 4 point (D5, N9, A2, R11) pharmacophore models showed the highest enrichment at 1% (EF1%) values (Fig.3). Also, the overall % active, and % yield and GH were found to be better for all these selected pharmacophore models. The overall results indicated that the three pharmacophore models generated from the crystal structure of Mtb DHQase could differentiate effectively the actives from the inactives. The validated e-pharmacophore hypothesis was then utilized for the virtual screening of *in house* database following a protocol as summarized in Fig 4.

Insert Table 2

Insert Fig. 3.

Insert Fig. 4.

3.3 Virtual screening of database compounds

The basic goal of any virtual screening protocol would be to reduce the enormous virtual chemical space of small organic molecules to a manageable number of compounds that could inhibit the protein with a higher chance so as to identify a candidate drug [26]. The top pharmacophore models after validation were subjected to virtual screening of commercially available database (Asinex) following a protocol as summarized in Fig. 4. The option implying "find matches" to the hypothesis in Maestro was employed for this step. Compounds retrieved by each e-pharmacophore models using phase with a fit value of above 1.0 were regarded as potential hits. A fit value was a measure of how well the ligand matched the e-pharmacophore model. Top hits retrieved by the e-pharmacophore model with a fit

value above 1.0 were carried forward for HTVS where the top 10% compounds were forwarded to SP (standard precision) and XP docking. Top hits from HTVS resulting in a score of ≥ -6.0 kcal mol^{-1} were subjected to another round of docking by Glide SP. The survivors of these preliminary screening steps were taken up to Glide XP docking. The XP Glide combines accurate, physics-based scoring terms and thorough sampling and the results showed scores ranging from -8.14 to -6 $kcalmol^{-1}$.The results of the virtually screened ligands using Glide XP docking showed the hydrogen-bond interactions, electrostatic interaction, hydrophobic enclosure, and π–π stacking interactions.

3.4 Molecular docking studies

Hit molecules obtained from Glide SP were further evaluated with Glide XP and GOLD 4.1.2 [27] to understand how these ligands bind to the enzyme and also to confirm their potency. Compounds selected from the SP were subjected to Glide XP and GOLD docking. All the compounds were found to retain key interactions at the active site like hydrogen bonding with Arg112, His81, Ser103, Ile102, Asn75, His101 and Tyr24. The XP glide score of the selected compounds ranged between -6.011 to -9.118 $kcalmol^{-1}$ and their GOLD score between 50.89 to 76.12. The pose obtained from both the docking studies (Glide and GOLD) are quite similar as they able to retain the important interaction with the proteins.

Thus we selected top five compounds from the Glide XP docking study with the best Glide scores (-7.03 to -8.14 $kcal.mol^{-1}$) and GOLD scores (56.89 to 76.12), suggesting strong protein-ligand interactions. The chemical structures of these lead compounds (Lead 1 to Lead 5) are illustrated in Fig. 5. All these top hits showed good docking scores and interactions with important amino acids such as Arg112, His81, Ser103, Ile102, Asn75, His101 and Tyr24. Fig. 6 presents the binding modes of all five lead molecules and their interacting pattern in the active site of the protein using ligand interaction diagram in Schrodinger suite version 9.3. All the compounds were found well fit in the active site cavity of the protein.

Insert Fig. 5.

Insert Fig. 6.

Further to prove the design concept, the final hits were subjected to biological assays. The docking score, H-bond and important interactions of these hits are presented in Table 3.

Insert Table 3

3.5 Enzyme kinetics of Mtb DHQase and in vitro enzyme inhibition studies.

The activity of enzyme concentration followed first order kinetics. The apparent constants K_m, V_{max} and K_{cat} were determined and their values were found to be 66.08 μM, 0.156 U/mg and 0.041 s^{-1} respectively which were determined using GraphPad (GraphPad Prism software, San Diego, CA) (Fig. 7).

Insert Fig. 7.

Enzymatic inhibition assays for Mtb DHQase were performed by measuring the rate of production of 3-dehydroshikimate at 234 nm. The top five hits were screened initially at 50 μM concentration. The **Lead 1** showed 58% inhibition at 50 μM, presenting the highest inhibition percentage of all hit molecules and was chosen as lead compound (Fig. 8). Sixty five chemical derivatives of the lead obtained from our *in house* chemical library were screened at 50 μM. Top two lead derivatives (**lead1a** and **lead1b**) (Fig. 8) were found very promising and hence were tested at lower concentrations to estimate IC_{50} and found to be 17.1 and 31.5 μM for **lead1a** and **lead1b** respectively (Table 4). To validate this *in vitro* enzyme inhibition results we have superimposed the lead compounds (**lead1**, **lead1a** and **lead1b)** on the pharmacophore. The superimposition of lead compounds on the pharmacophore is represented in Fig. 9.

Insert Fig. 8.

Insert Fig. 9.

3.6 Antibacterial activity and cytotoxicity assay

The top two leads were evaluated for both antibacterial and cytotoxic activities (Table 4). The antibacterial activity of both the compounds was evaluated against *M. tuberculosis* H37Rv while cytotoxicity was evaluated using HEK293 cells. MIC values of 25 µg/mL and 6.25 µg/mL were obtained for **lead1a** and **lead1b** respectively. Both compounds presented low cytotoxicity, with 14.3% of cell death at 100 µM for **lead1a** and 11.9% for **lead1b** (Table 4).

Insert Table 4

3.7 Biophysical characterization

The top active inhibitor against Mtb DHQase (**lead1a**) was evaluated for its stabilization of the protein using DSF. This was measured by comparing the thermal stability of native Mtb DHQase over gradual increase in temperature with that of the protein-ligand complex. The positive shift in melting temperature (T_m) over the native Mtb DHQase T_m accounted to the increased stability of protein due to ligand binding which implied the stabilization of target protein Mtb DHQase by the inhibitor. Fig 10 shows the melting temperature of active compound **lead1a**. The protein Mtb DHQase T_m was found to be 45.1°C whereas melting temperature for protein with **lead1a** was 48.3°C. This observation explained the stabilization effect of active inhibitor (**lead1a**) over Mtb DHQase accounting for their activity. The lead derivative, **lead1a**, displayed a T_m shift of 3.2°C, showing affinity with the protein and correlating with its IC_{50} value of 17.1µM.

Insert Fig. 10.

4. Conclusion

The main objective of the present study was to identify the drug like molecules as inhibitors for Mtb DHQase by using structure-based modeling. Structure-based virtual

screening and pharmacophore modeling have emerged as corresponding methods to high throughput screening of large chemical databases. The crystal structure of Mtb DHQase bound with an inhibitor was used to construct pharmacophore models and virtual screening of *in house* database retrieved five top hit compounds with good docking score and interaction pattern. *In vitro* enzymatic inhibition studies of these five ligands yielded one molecule (**Lead 1**) which showed 58% inhibition at 50 μM concentration in Mtb DHQase assay. Chemical derivatives of the lead compound were tested and top two hits presented IC_{50} of 17.1 and 31.5 μM for **lead1a** and **lead1b** respectively. Further the tight binding with the protein was biophysically confirmed by DSF. It was satisfying to see that the best inhibitory compound **lead1a** also showed a positive shift in DSF, indicating an increase in thermal stability of the inhibitor-protein complex that matched with its *in vitro* Mtb DHQase enzyme activity as well. These compounds were tested for antibacterial activity against *M. tuberculosis* H37Rv and for cytotoxicity against HEK293 cells. The MIC values for **lead1a** and **lead1b** were found to be 25 μg/mL and 6.25 μg/mL, respectively. Both compounds presented low cytotoxicity, with less than 15% of cell death at 100 μM Therefore, the structure-based pharmacophore modeling provided useful information required for proper understanding of the important structural and binding features for designing novel Mtb DHQase inhibitors. The present study thus reported two new molecules as prototypical leads uses against TB.

Acknowledgement

This project was financially supported by the National Counsel of Technological and Scientific Development (CNPq), Ministry of Science, Technology and Innovation of Brazil.

References

[1] World Health Organization, Global Tuberculosis Control: Surveillance and Financing, WHO Report, 2012.

[2] C. Abell in Enzymology and molecular biology of the shikimate pathway. In Comprehensive Natural Products Chemistry, Vol. 1 (Ed. Sankawa, U.), Elsevier, Amsterdam, 1999 pp. 573-607.

[3] J. R. Coggins, C.Abell, L. B. Evans, M. Frederickson, D. A. Robinson, A. W. Roszak, A. P. Lapthorn, Experiences with the shikimate-pathway enzymes as targets for rational drug design, Biochem. Soc. Trans. 2003, 31, 548-552.

[4] F. Dosselaere, J. Vanderleyden, A metabolic node in action: chorismate-utilizing enzymes in microorganisms, Crit. Rev. Microbiol. 2001, 27, 75-131.

[5] F. Roberts, C. W. Roberts, J. J. Johnson, D. E. Kyle, T. Krell, J. R. Coggins, G. H. Coombs, W. K. Milhous, S. Tzipori, D. J. Ferguson, D. Chakrabarti, R. McLeod, Evidence for the shikimate pathway in apicomplexan parasites, Nature 1998, 393, 801-805.

[6] J. Harris, C. Kleanthous, J. R. Coggins, A. R. Hawkins, C. Abell, Different mechanistic and stereochemical courses for the reactions catalysed by type II dehydroquinases, J. Chem. Soc., Chem. Commun. 1993, 13, 1080-1081.

[7] E. Haslam in Shikimic Acid: Metabolism and Metabolites, (Ed. Wiley and Sons) Chichester, UK, 1993

[8] T. Parish, N. G. Stoker, The common aromatic amino acid biosynthesis pathway is essential in *Mycobacterium tuberculosis*, Microbiology 2002, 48, 3069-3077.

[9] D. G. Gourley, A. K. Shrive, I. Polikarpov, T. Krell, J. R. Coggins, A. R. Hawkins, N. W. Isaacs, L. Sawyer, The two types of 3-dehydroquinase have distinct structures but catalyze the seme overall reaction, Nat. Struct. Biol. 1999, 6, 521-525.

[10] M.V.B. Dias, W.C. Snee, K.M. Bromfield, R.J. Payne, S.K. Palaninathan, A. Ciulli, N.I.

Howard, C. Abell, J.C. Sacchettini, T.L. Blundell. Structural investigation of inhibitor designs targeting 3-dehydroquinate dehydratase from the shikimate pathway of *Mycobacterium tuberculosis*, Biochem. J. 2011, 436, 729-739.

[11] Phase, version 3.4, Schrödinger, LLC, New York, NY, 2012; ConfGen, version 2.3, Schrödinger, LLC, New York, NY, 2012

[12] Schrödinger Suite 2012 Protein Preparation Wizard; Epik version 2.2, Schrödinger, LLC, New York, NY, 2012; Impact version 5.7, Schrödinger, LLC, New York, NY, 2012.

[13] L. Tizón, J.M. Otero, V.F. Prazeres, A.L. Llamas-Saiz, G.C. Fox, M.J. van Raaij, H. Lamb, A.R. Hawkins, J.A. Ainsa, L. Castedo, C. González-Bello, A Prodrug Approach for Improving Antituberculosis Activity of Potent *Mycobacterium tuberculosis* Type II Dehydroquinase Inhibitors, J. Med. Chem. 2011, 54, 6063-6084.

[14] S.L. Dixon, A.M. Smondyrev, E.H. Knoll, S.N. Rao, D.E. Shaw, R.A. Friesner, PHASE: a new engine for pharmacophore perception, 3D QSAR model development, and 3D database screening: methodology and preliminary results, Comput. Aided Mol. Des. 2006, 20, 647-671.

[15] S.L. Dixon, A.M. Smondyrev, S.N. Rao, A novel approach to pharmacophore modeling and 3D database searching, Chem. Biol. Drug Des. 2006, 67, 370-372.

[16] N.K. Salam, R. Nuti, W. Sherman, Novel method for generating structurebased pharmacophores using energetic analysis, J. Chem. Inf. Model. 2009, 49, 2356-2368.

[17] LigPrep, version 2.5, Schrödinger, LLC, New York, NY, 2012 and ConfGen, version 2.3, Schrödinger, LLC, New York, NY, 2012.

[18] R. Palakurti, D. Sriram, P. Yogeeswari, R. Vadrevu1, Multiple e-pharmacophore modeling combined with high-throughput virtual screening and docking to identify potential inhibitors of β-secretase, Mol. Inf. 2013, 32, 385-398.

[19] S. Nagamani, C. Kesavan, K. Muthusam, e-Pharmacophore mapping and docking

studies on Vitamin D receptor (VDR), Bioinformation. 2012, 15, 705-710.

[20] S.G. Franzblau, R.S. Witzig, J.C. McLaughlin, P. Torres, G. Madico, A. Hernandez, M.T. Degnan, M.B. Cook, V.K. Quenzer, R.M. Ferguson, R.H. Gilman, Rapid, low-technology MIC determination with clinical *Mycobacterium tuberculosis* isolates by using the microplate Alamar Blue assay, J. Clin. Microbiol. 1998, 36, 362.

[21] D. Gerlier, N. Thomasset, Use of MTT colorimetric assay to measure cell activation, J. Immunol. Methods. 1986, 94, 57-63.

[22] V.U. Jeankumar, J. Renuka, P Santosh, V. Soni, J. P. Sridevi, P. Suryadevra. P. Yogeeswari, D. Sriram, Thiazole-aminopiperidine hybrid analogues: design and synthesis of novel *Mycobacterium tuberculosis* GyrB inhibitors, Eur. J. Med. Chem. 2013, 70, 140-153.

[23] F.H. Niesen, H. Berglund, M. Vedadi, The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability, Nat. Protoc. 2007, 2, 2212-2221.

[24] V. Aparna, N. Mohanalakshmi, W. Hopper, Structure-based pharmacophore and virtual screening for bacterial MexB EffluxPump Inhibitors, International Conference on Bioscience, Biochemistry and Bioinformatics（ICBBB 2011). 2011, 458-461.

[25] O. Dror, D. Schneidman-Duhovny, Y. Inbar, R. Nussinov, H.J.Wolfson, Novel approach for efficient pharmacophore-based virtual screening: method and applications, J. Chem. Inf. Model, 2009, 49, 2333-2343

[26] V. Vyas, A. Jain, A. Jain, A. Gupta, Virtual screening a fast tool for drug design, Sci. Pharm. 2008, 76, 333-360.

[27] Gold, Version 4.1.2, Cambridge Crystallographic Data Centre, Cambridge, UK.

List of Figures with Captions

Table 1 . Scores of hypothesis in e-pharmacophore.

Rank [a]	Feature label	Score	Type
1	D5	-1.60	D[b]
2	D6	-1.60	D
3	N9	-1.32	N[c]
4	A2	-0.80	A[d]
5	A3	-0.80	A
6	R10	-1.41	R[e]
7	R11	-1.07	R

[a] Rank: Impor
[b] Donor group
[c] Negative Ionizable group
[d] Acceptor group
[e] Ring aromatic group

Table 2: Validation of e-pharmacophore models.

E-pharm features	EF(manual)[a]	GH[b]	%yield[c]	%A[d]
D5,D6,N9,A2,A3,R10,R11	50	0.7103	75	60
D5,D6,N9,A2,R10,R11	28.57	0.4656	42.85	60
D5,D6,N9,A3,R10,R11	21.42	0.3835	32.14	60
D5,D6,N9,A2,A3,R10	10.72	0.3277	16.09	93.33
D5,D6,N9,A2,A3,R11	18.30	0.4227	27.45	93.33
D5,N9,A2,A3,R10,R11	14.63	0.3044	21.95	60
D6,N9,A2,A3,R10,R11	14.81	0.3214	22.22	66.66
D5,D6,A2,A3,R10,R11	6.80	0.2214	10.20	66.66
D5,N9,A2,R10,R11	5.37	0.2008	8.06	66.66
D6,N9,A3,R10,R11	6.00	0.2102	9.0	66.66
D5,D6,N9,R10,R11	13.63	0.2926	20.45	60
D5,N9,A3,R10,R11	5.12	0.1849	7.69	60
D6,N9,A2,R10,R11	6.60	0.21	9.90	66.66
D5,D6,N9,A2,A3	3.66	0.20	5.49	93.33
D5,D6,A2,A3,R10	2.61	0.17	3.92	93.33
D5,D6,A2,A3,R11	3.35	0.19	5.03	93.33
D5,N9,A2,R10	1.83	0.12	2.75	93.33
D5,N9,A3,R10	1.70	0.116	2.56	93.33
D6,N9,A2,R10	1.70	0.11	2.56	93.33
D6,N9,A3,R10	1.82	0.12	2.73	93.33
D5,N9,A2,R11	2.76	0.17	4.14	93.33
D5,N9,A3,R11	2.55	0.16	3.82	93.33
D6,N9,A2,R11	2.80	0.17	4.20	93.33
D6,N9,A3,R11	3.05	0.19	4.58	100
D5,D6,A2,A3	1.68	0.11	2.53	100
D5,A2,R10,R11	1.17	0.07	1.76	66.66
D6,A3,R10,R11	1.26	0.08	1.89	66.66
D5,N9,A2	1.32	0.06	1.99	100
D5,N9,A3	1.16	0.037	1.74	100
D6,N9,A2	1.17	0.04	1.76	100

D6,N9,A3	1.30	0.062	1.95	100
D5,A3,R11	1.06	0.03	1.60	93.33
D6,A2,R10	1.08	0.02	1.63	100
D6,A3,R11	2.21	0.12	3.31	53.33
D5,N9,R10	1.14	0.04	1.71	93.33
D6,N9,R11	1.87	0.12	2.80	100

[a] Enrichment Factor
[b] Goodness of hit

Table 3: The fitness, docking Score, gold score and hydrogen bond interaction of top five ligands.

Ligand	Fitness	Docking score	H-Bond	GOLD score	Interaction
Lead 1	1.34	-8.14	5	72.17	Tyr24, Val23, Ser103, Asn75, Ile102
Lead 2	1.56	-7.18	3	60.41	Asn75, Ile102, Ser103
Lead 3	1.67	-7.03	6	62.57	Asn75, Asn12, Ile102, His101(protonated state), Arg112
Lead 4	1.23	-7.09	4	76.12	Asn75, Ile102, Arg15, His101(protonated state)
Lead 5	1.43	-7.50	3	56.89	2Ser103, Ile102

Table 4: Activity table showing IC_{50} value, MIC and cytotoxicity.

Compounds	IC_{50} (µM)	MIC (µg/mL)	Cytotoxicity (% at 100 µM)
Lead1a	17.1	25.0	14.3
Lead1b	31.5	6.25	11.9

Figure 1

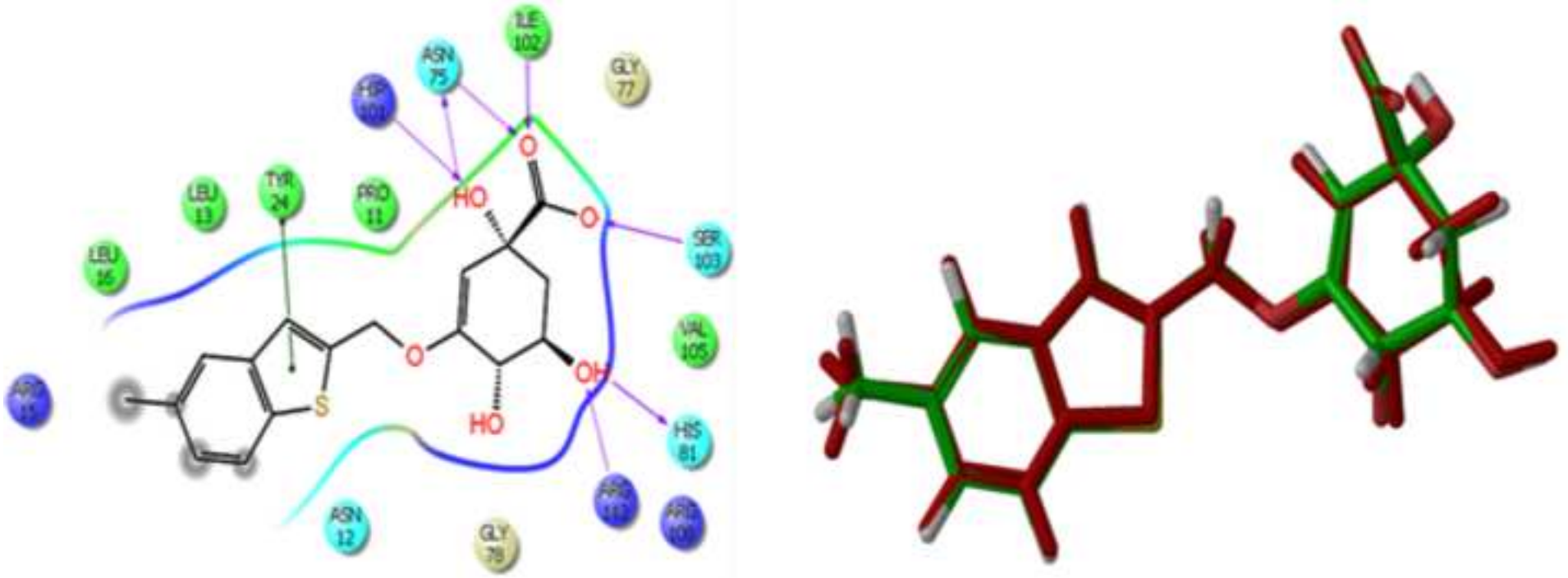


Figure 2

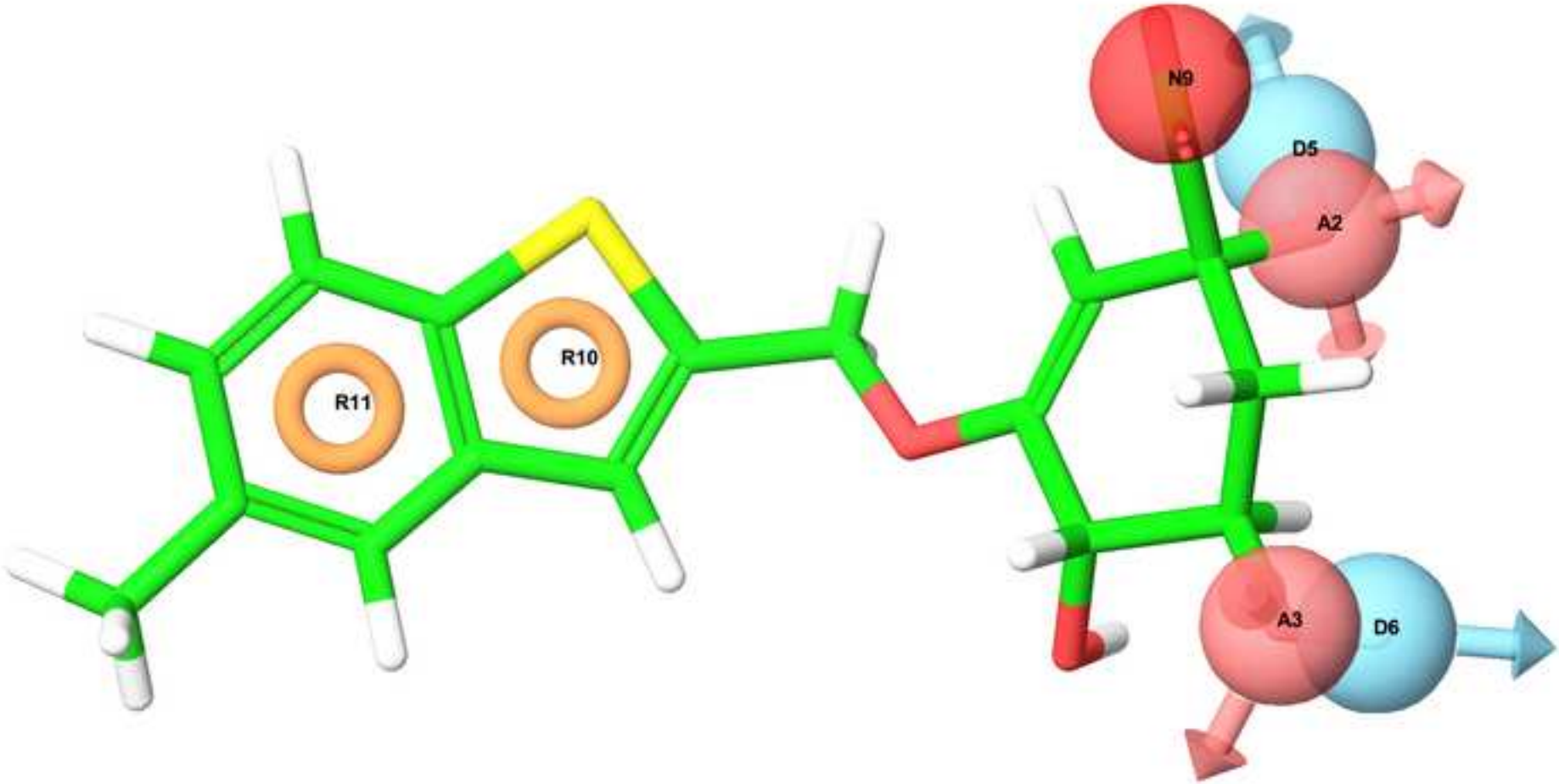
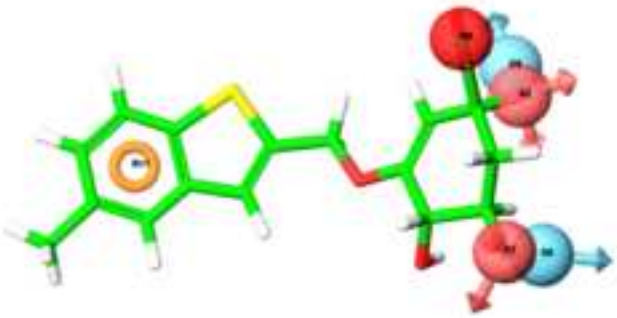
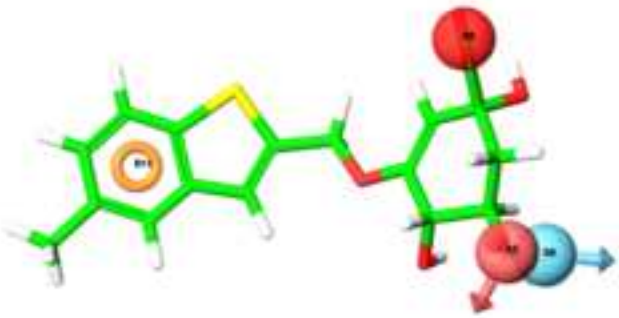


Figure 3



6 point pharmacophore

5 point pharmacophore



4 point pharmacophore

Figure 4

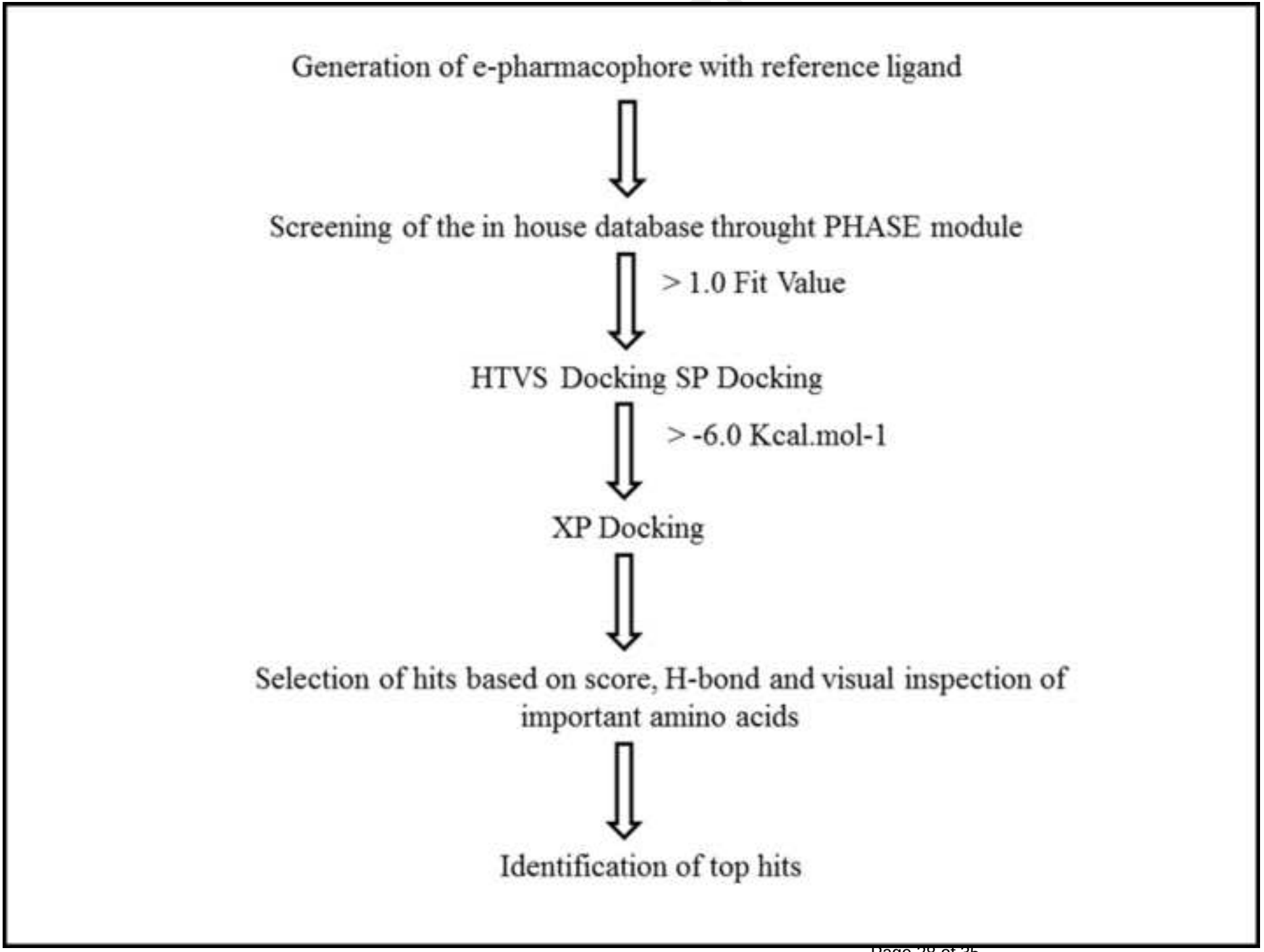
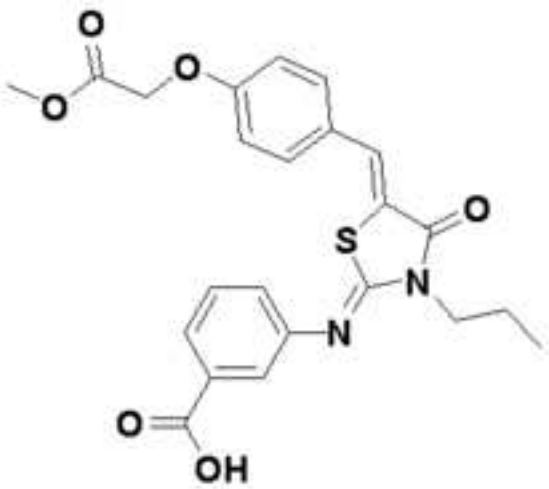
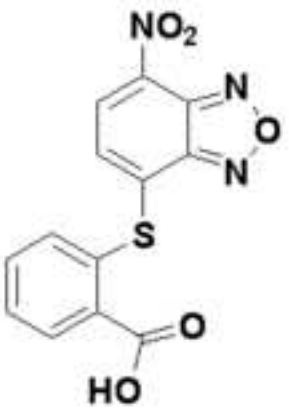

Generation of e-pharmacophore with reference ligand
Screening of the in house database throught PHASE module
> 1.0 Fit Value
HTVS Docking SP Docking
> -6.0 Kcal.mol-1
XP Docking
Selection of hits based on score, H-bond and visual inspection of important amino acids
Identification of top hits

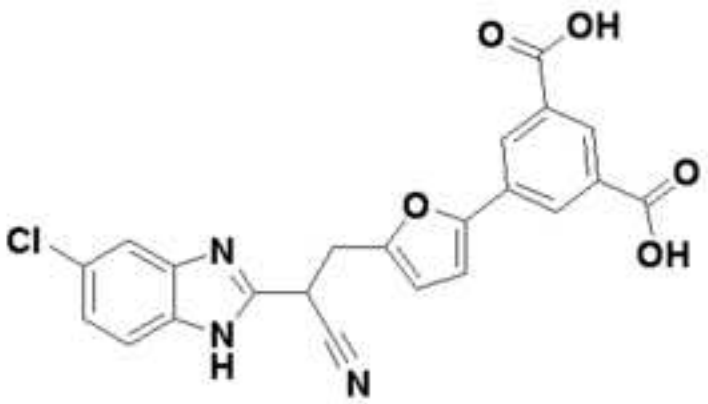
Figure 5



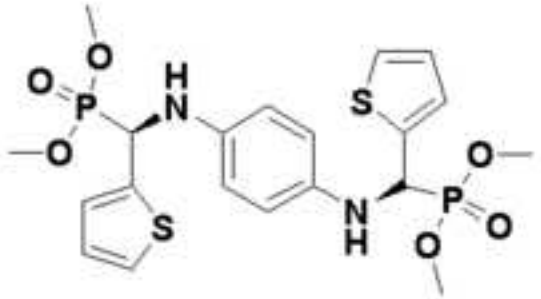
Lead 1



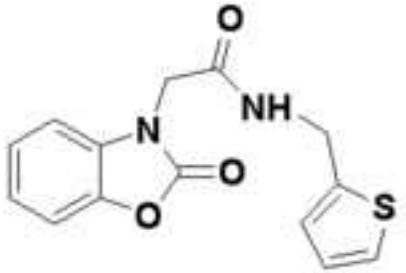
Lead 2



Lead 3

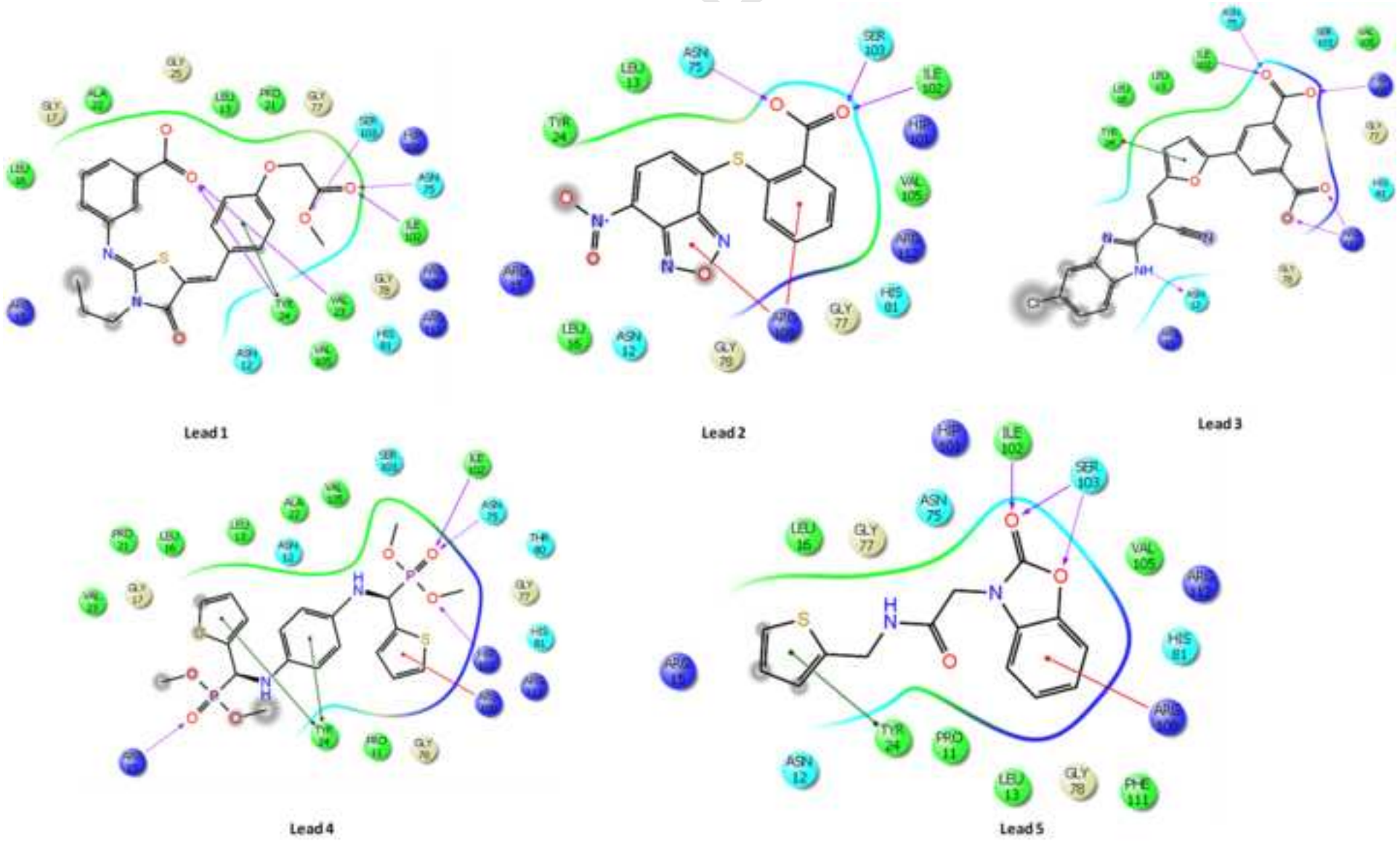


Lead 4

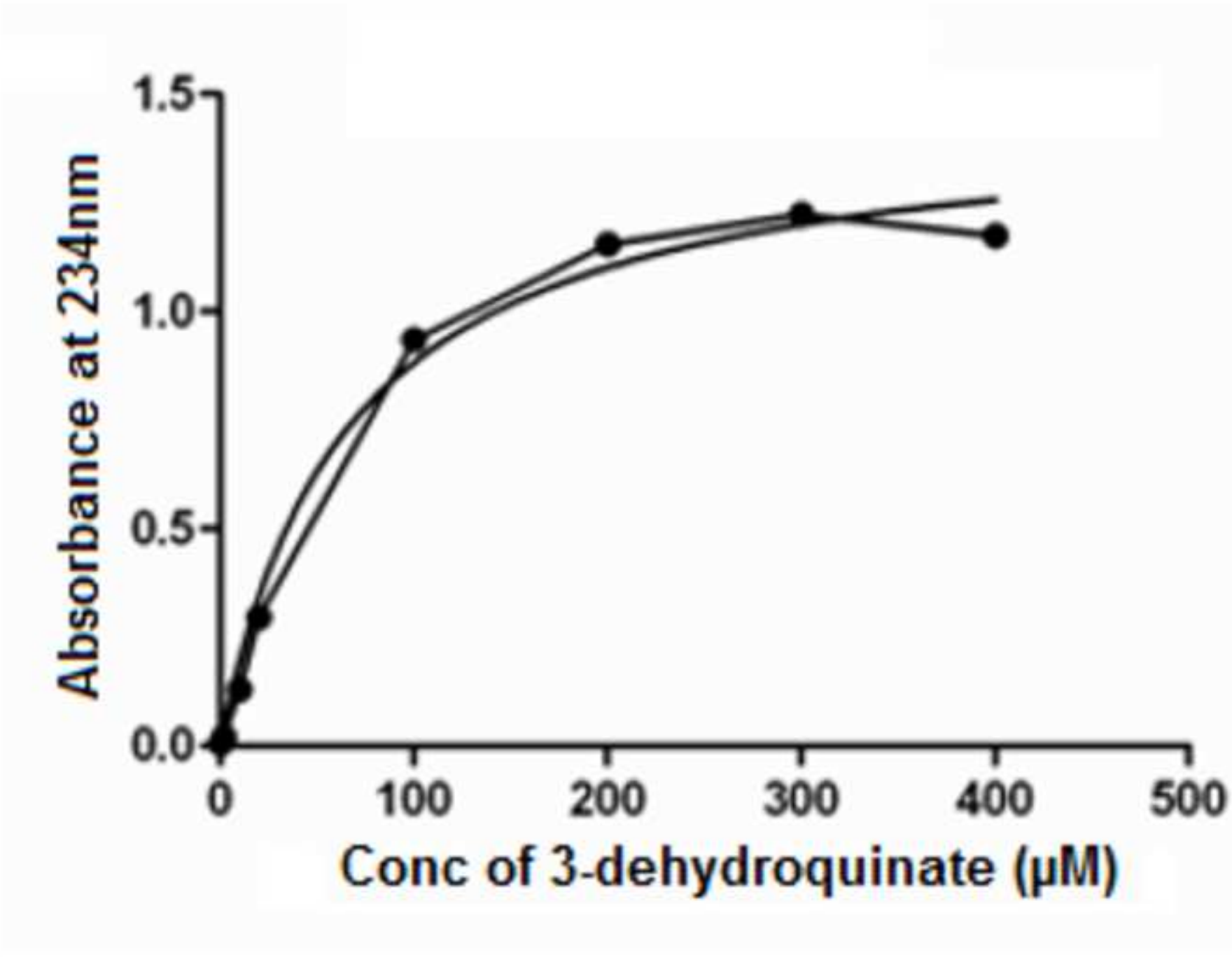


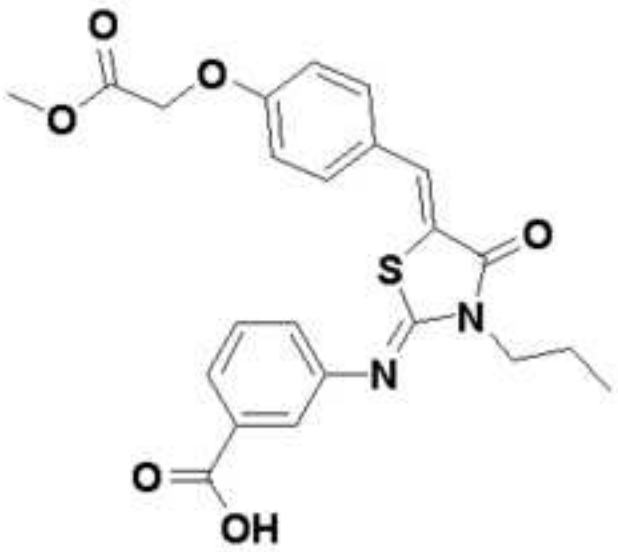
Lead 5

Figure 6

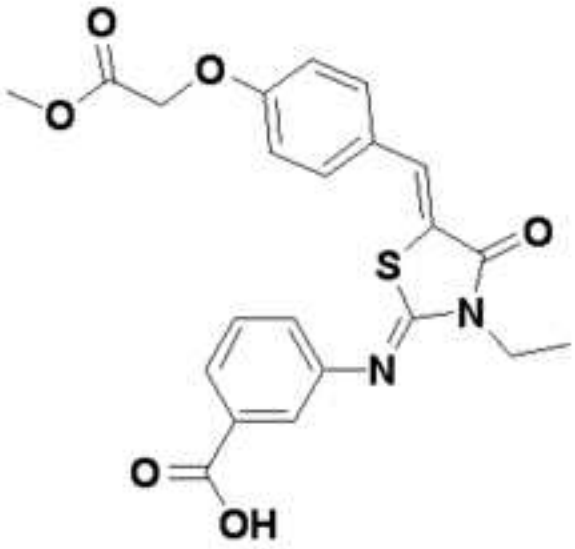


Figure(s)

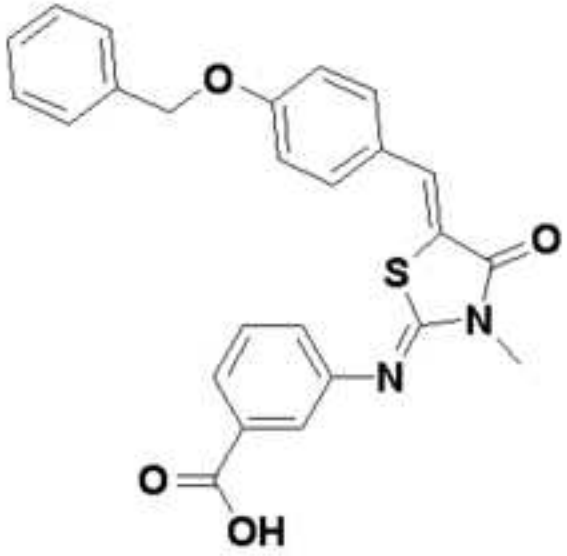

Absorbance at 234nm
1.5
1.0
0.5
0.0
0
100
200
300
400
500
Conc of 3-dehydroquinate (μM)



Lead 1
58% inhibition at 50μM



Lead 1a
IC_{50}=17.1 μM



Lead 1b
IC_{50}=31.5 μM

Figure 9

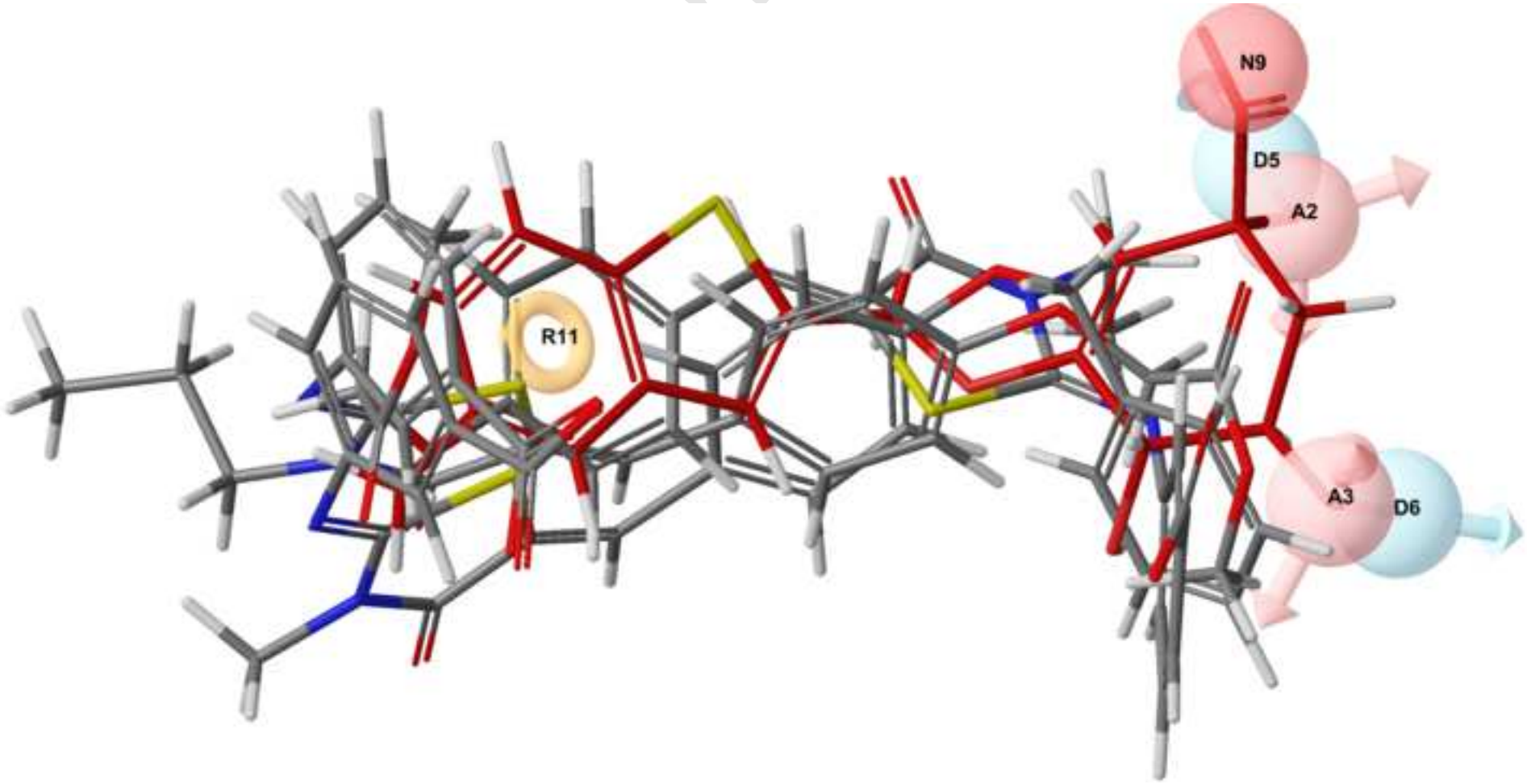
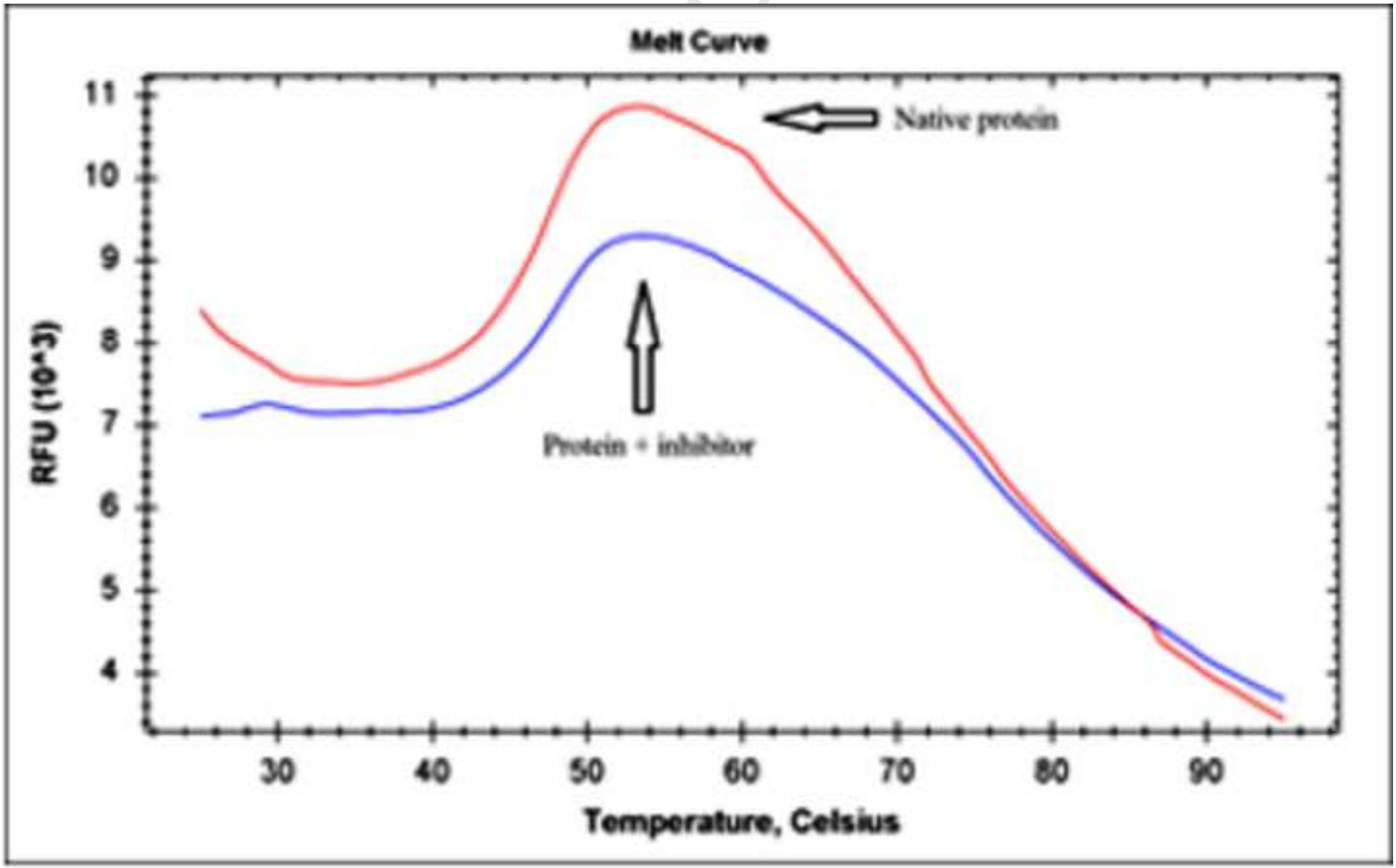


Figure 10


Melt Curve
Native protein
Protein + inhibitor
RFU (10^3)
11
10
9
8
7
6
5
4
30
40
50
60
70
80
90
Temperature, Celsius

Graphical Abstract (for review)