ISSN NO. 2320-5407

Journal homepage: http://www.journalijar.com

INTERNATIONAL JOURNAL OF ADVANCED RESEARCH

RESEARCH ARTICLE

In-silico analysis of Clp protease catalytic subunit-2 of *Mycobacterium tuberculosis*: Modeling and docking analysis.

Mamta Garg, Navdeep Kaur Sidhu, *Kavita K Kakarala

Centre for Biotechnology and Bioinformatics (CBB), School of Life Sciences, Jawaharlal Nehru Institute of Advanced Studies (JNIAS), 6th Floor, Buddha Bhawan, M. G. Road, Secunderabad - 500003, Andhra Pradesh

Manuscript Info

Manuscript History:

Received: 12 February 2014
Final Accepted: 15 March 2014
Published Online: April 2014

Key words:
Mycobacterium tuberculosis, Caseinolytic protease, Homology modeling, Docking, Virtual Screening, binding affinity

Corresponding Author

Kavita K Kakarala

Abstract

Tuberculosis (TB) is global. The World Health Organization (WHO) estimates that one third of the world's population is infected with Mtb and India accounts for one-fifth of the global TB incident cases. With the emergence of Multi Drug Resistance (MDR), Extreme Drug Resistance (XDR) and Total drug resistance (TDR), the need for identification of novel targets and inhibitors has become necessary to control TB pandemic. The proteases belong to the class of druggable enzymes and have already been validated as therapeutic targets in the treatment of HIV, hepatitis, and cancer. ClpP1P2 protease is very recently proved to be essential for growth and infection of Mycobacterium tuberculosis (Mtb) and thus makes it an attractive target for drug development. However, ClpP2 was reported earlier as one of the genes essential for growth of Mtb and the crystal structure of it is unavailable yet, We had modeled ClpP2 protease of the Mtb using crystal structure of Clp protease of E. coli (1TYF-PDB ID). The computed model's energy was minimized and validated using PROCHECK, ProSA, Verify 3D and ERRAT to obtain a stable model structure. Stable model was further investigated for the binding affinity of drugs already in use, phytoconstituents, and with compounds of zinc database using GLIDE of Virtual screening workflow of Schrödinger Suite 2012. The docked complexes were validated based on GLIDE score. Our investigation indicates that ClpP2 inhibition could be the mechanism of action of already available drugs, and interestingly, it was observed that phytoconstituents like Luteolin-7-O-glucoside, Quercetin, Aloe emodin, Naphthoquinone and Xanthone showed binding energy comparable to the second line and advanced drugs, which we believe is an important outcome as most of the second line drugs have acute toxicity. We have also performed virtual screening using ligands from zinc database and identified that amides are the most potent inhibitors of ClpP2 protease of Mycobacterium tuberculosis. Despite their potential importance in the survival and virulence of Mtb, to our knowledge, functional studies of these peptidases have not yet been reported. To our knowledge this work on structural modelling of ClpP2 of Mycobacterium is unique bringing out information on possible inhibitors of ClpP2 protease to combat Mycobacterium tuberculosis.

Introduction

Tuberculosis (TB) is a global pandemic and a major public health problem in India. India accounts for one-fifth of the global TB incident cases. Each year, nearly 2 million people in India to develop TB, of which around 0.87

million are infectious cases. It is estimated that annually around 330,000 Indians die due to TB (http://www.whoindia.org/en/section3/section123.htm). Treatment of Tuberculosis still poses a challenge because of various reasons, the most important being multi drug resistance (MDR) and extreme drug resistance (XDR). Furthermore, adding to the complexity of drug resistance, the emergence of Total drug resistance (TDR) [1] in India has created an urgent intervention in terms of identifying novel targets and inhibitors to control the TB pandemic. According to WHO estimates, in 2010, there were 8.8 million new cases of tuberculosis (TB) and 1.5 million deaths [2]. Unfortunately, Tuberculosis treatment has not seen much progress as Mycobacterium tuberculosis (Mtb) is a stubborn pathogen. The available treatment options are few depending on a relatively small set of chemotherapeutic agents, which includes the widely used front-line drugs like isoniazid, Ethambutol, rifampicin, and Pyrazinamide [3]. The need for the development of novel targets and drugs targeting them has gained much attention as these organisms are endowed with the ability to avoid elimination even with antibiotic intervention possibly due to their restricted metabolism and replication processes that are essentially targeted by the current antibiotics. The Clp protease is one of the major cellular proteases responsible for degrading misfolded or damaged proteins and plays a central role in maintaining protein function, especially under stress conditions [4]. Its role in the virulence of numerous other pathogen bacteria, including Salmonella Typhimurium [5], Listeria monocytogenes [6], Streptococcus pneumonia [7], Staphylococcus aureus [8], and Helicobacter pylori [9] is already reported. The complete genome sequencing of M. tuberculosis strain (H37Rv) has revealed the presence of two paralog genes encoding putative ClpP peptidase, ClpP1 (Rv2461c) and ClpP2 (Rv2460c) (Pasteur Institute TubercuList, http://genolist.pasteur.fr/TubercuList), organized as an operon [10, 11]. This serine protease was first discovered and characterized in Escherichia coli [12, 13]. Interestingly, the ClpP proteolytic subunit is an unusual drug target, where both inhibition and activation are lethal events [14]. ClpP2 is required for mycobacterial growth as demonstrated elegantly by Sasseti et al (2003) [15] and from the therapeutic point of view, the main advantage of targeting ClpP2, is the absence of this enzyme in the eukaryotic cytoplasm [16]. Recently, it has been demonstrated that both ClpP1 and ClpP2 are essential for growth and infection of Mtb [16]. Thus, essential nature of ClpP1P2 protease makes it an attractive target for antibiotic development [14], particularly because the proteases as a class are druggable enzymes and have already been validated as therapeutic targets in the treatment of HIV, hepatitis, and cancer [17].

We modeled ClpP2 as overexpression of ClpP2, independently of ClpP1, was apparently toxic to the cells [14] and was one of genes required for Mycobacterial growth defined by high density mutagenesis [15] and the ATPase subunits ClpX and ClpC ATPase complexes are not an obligatory requirement for its peptidase activity [18]. However the combination of chemical and genetic validation suggests that the Clp proteases are attractive new targets [14]. The investigation also included an evaluation of binding affinity of available anti-tuberculosis drugs, lead like molecules from Zinc database and phytoconstituents.

Material and Methods

Model Building

The protein sequence of ATP-dependent Clp protease catalytic subunit 2 (ClpP2) of Mtb (SWISSPROT accession: P63783) was retrieved from Protein Knowledgebase UniProt KB (http://www.uniprot.org/uniprot/P63783) and saved in FASTA format. Templates for homology modeling were identified using BLAST [www.ncbi.nlm.nih.gov/blast/]. The 2.3Å resolution X-ray structure of the proteolytic component of the caseinolytic Clp protease (ClpP) from E.coli was selected and obtained from the Protein Data Bank (PDB; http://www.rcsb.org/pdb):PDB code: 1TYF [19]. The selection of 1TYF was based on identity, query coverage and E value. Protein modeling was done using PRIME of Schrödinger Suite 2012 [20]. Prime provides powerful tools for template identification, alignment, model building, and refinement for the full range of sequence identity. Prime utilizes a hierarchical, multi-scale technique to exhaustively sample loop conformations. A clustering algorithm removes redundant candidate structures and efficiently reduces the number of loops for complete energy minimizations [21]

Model validation

The obtained model was validated with the help of PROCHECK [22], Verify 3D [23], and ProSA z-score [24] and ERRAT protein verification [25].

Ligands

The docking studies were performed using available drugs (inhibitors), belonging to First line, second line and advanced compounds. The selection of the drugs were based on a recent review on potential inhibitors [3]. Virtual

screening was also performed using lead like molecules of Zinc database (http://zinc.docking.org/). Finally, docking studies were performed, using phytoconstituents of natural antitubercular plants [26]. The 3D structures of available drugs/Phytoconstituents were downloaded from the PubChem database (http://pubchem.ncbi.nlm.nih.gov/) and the Zinc database (http://zinc.docking.org/) in a structured data format (sdf) format.

Ligand Preparation

All the ligand structures obtained from PubChem database were prepared using LigPrep module of Maestro software using the Lipinski filter. The LigPrep process consists of a series of steps that perform conversions, apply corrections to the structures, generate variations on the structures, eliminate unwanted structures, and optimize the structures.

Protein structure Preparation

The validated homology model of ClpP2 protease was prepared using protein preparation wizard of Schrödinger Suite 2012 which is designed to ensure chemical correctness and to optimize protein. The modeled structure was preprocessed and prepared by adding the missing hydrogen, correcting the bond orders followed by optimization and energy minimization with force field: OPLS 2005 (optimized potential for liquid simulations) [27].

Docking

Docking with Glide: Docking studies were performed by Glide module of Schrodinger 9.3 [28,29].The Glide (Grid-Based Ligand Docking With Energetics) algorithm approximates a systematic search of positions, orientations, and conformations of the ligand in the receptor binding site using a series of hierarchical filters. Glide approximates a complete systematic search of the conformational orientation and positional space of the docked ligand. It performs grid-based ligand docking with energetic and searches for favorable interactions between one or more typically small ligand molecules and a typically larger receptor molecule, usually a protein.

Virtual Screening

The virtual screening was also performed using lead like molecules of Zinc database (http://zinc.docking.org/) with prepared homology model of ClpP2 protease using virtual screening workflow (VSW) module of the Schrödinger Suite 2012. The Virtual Screening Workflow is designed to run an entire sequence of jobs for screening large collections of compounds against a particular target. The workflow includes ligand preparation using LigPrep, filtering using propfilter on QikProp properties or other structural properties, and Glide docking at the three accuracy levels, High throughput virtual screening (HTVS), standard precision (SP), and extra precision (XP). HTVS and SP modes were used for a large set of ligands and XP docking is more accurate than the above two method. After ensuring that the protein and ligands were in the correct form for docking, the receptor-grid files were generated using a Receptor grid-receptor generation program. The grid was generated at the centroid of the active site residues consisting of residues Ser111, His136& Asp187 forming a catalytic triad. The ligands were docked with the active site using Glide of Schrödinger Suite 2012 [28, 29]. Glide generates conformations internally and passes these through a series of filters. This first places the ligand centre at various grid positions of a 1 Å grid and rotates it around the three Euler angles. At this stage, crude score values and geometrical filters weed out unlikely binding modes. The next filter stage involves a grid-based force field evaluation and refinement of docking solutions including torsional and rigid body movements of the ligand. The final energy evaluation is done with Glide Score, and a single best pose is generated as the output for a particular ligand.

G Score = a * vdW + b * Coul + Lipo + Hbond + Metal + BuryP + RotB + Site where, vdW => van der Waal energy; Coul => Coulomb energy; Lipo => lipophilic contact term; HBond => hydrogen-bonding term; Metal => metal-binding term; BuryP => penalty for buried polar groups; RotB => penalty for freezing rotatable bonds; Site => polar interactions at the active site; and the coefficients of vdW and Coul are: a =0.065, b = 0.130.

Post Docking Analysis

Ligands were ranked based on the glide and dock score. For analyzing the interactions of docked protein-ligand complexes, the Ligand interaction diagram of Schrödinger Suite 2012, was used to visualize the bonds between receptor and ligand atoms within a range of 4 Å.

Note: The amino acid numbering is according to the numbering generated by software, which is one more than Uniprot numbering.

Result and Discussion

According to WHO statistics, tuberculosis remains one of the leading causes of death [2]. With the first cases of TDR reported in India, during January 2012 and the mortality rate of MDR, XDR and TDR due to TB reported to be 30%, 60% and 100% respectively (http://www.dnaindia.com/mumbai/report_first-cases-of-totally-drug-resistant-tb-in-india-one-dead_1634439) research in this area assumes great importance. These alarming effects of bacterial pathogens, points to the urgent need to identify novel targets and to develop new drugs to combat infection. Intracellular protein degradation is critical for maintaining cellular homeostasis through protein quality control and regulation of numerous biological pathways [30, 31, 32, 33].

ClpP1P2 of Mtb was identified very recently as an ideal potential target for the development of drugs as it is required for normal growth [16]. In mycobacteria, however, two different protein species contribute to protease activity. Although Mtb ClpP1 forms a tetradecameric complex, a crystal structure of Mtb ClpP1 lacks appropriate active site geometry to support proteolysis [36].

The model was built based on the crystal structure of 2.3Å resolution X-ray structure of Caseinolytic protease from E. coli (PDB ID-1TYF) as it shares 49% sequence identity and a query coverage of 90%. The sequence identity of the ClpP1, closely related protein in the same species showed 48% identity with around 80% query coverage thus we have used Caseinolytic protease from E. coli (PDB ID-1TYF) as a template [19]. Prime module of Schrödinger [20] was used for homology modeling. The alignment used for modeling is shown in Figure 1. The loops of the resultant model were refined followed by energy minimization using OPLS2005 force fields with default settings: until no further change in the energy was observed, to obtain a stable model suitable for docking analysis. This energy minimized structure was used for Glide (grid-based ligand docking with energetic) docking. The energy minimized model thus obtained validated with the help of PROCHECK [22], Verify 3D [23], and ProSA z-score [24], ERRAT protein verification [25]. PROCHECK program checks the detailed residue-by-residue stereochemical quality of protein structure. Procheck result shows that altogether >90% of the residues in the homology models are in favored and allowed regions (Figure.2). ProSA calculates an overall quality score for a specific input structure. As the Z score obtained is -6.9, which is within the scores obtained for groups of structures from different sources (X-ray, NMR) are distinguished by different colors (Figure. 3), it indicates the overall model quality as the range falls within a characteristic range of native proteins. It can be used to check whether the z-score of the input structure is within the range of scores typically found in native proteins of similar size. Verify 3D analysis of the models found no confirmation errors, and there were no values less than 0.09, further indicating that all of the residue was located in favorable structural environments (Figure.4). The ERRAT program checks the "Overall quality factor" that works by analyzing the statistics of non-bonded interactions between different atom types, with higher scores depicting higher quality structures. In our model, the ERRAT score for all of the model was >85%, well within the range of high quality model (Figure.5). The PROCHECK, ProSA z-score,Verify 3D and ERRAT results together show that the representative structure was satisfactory and can thus be considered as a reliable source for further analysis.

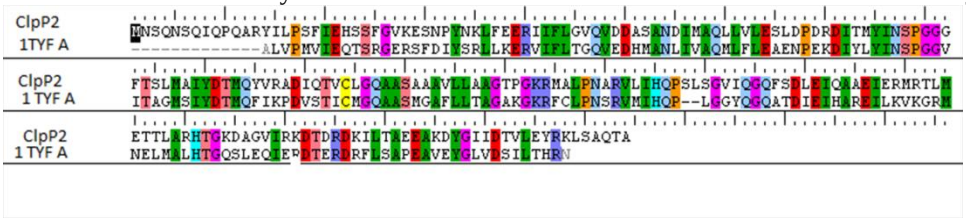


Figure 1. ClustalW alignment of ClpP2 of Mycobacterium tuberculosis with 1TYF (PDB ID of proteolytic component of caseinolytic Clp Protease from E.Coli at 2.3 Angstrom resolution).

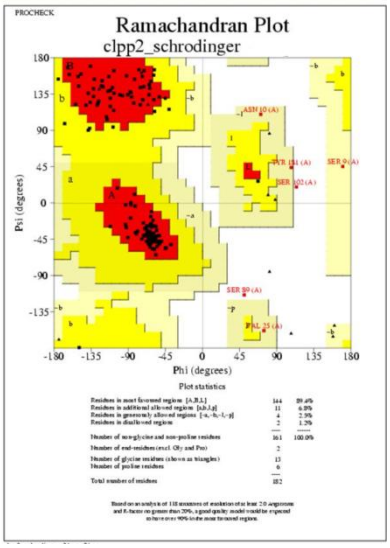


Figure 2. Procheck results of Modeled structure of ClpP2 of Mycobacterium tuberculosis using 1TYF (PDB ID of proteolytic component of caseinolytic Clp Protease from E.Coli at 2.3 Angstrom resolution) as template. Results indicate that 93.3% residues are in most favoured regions.

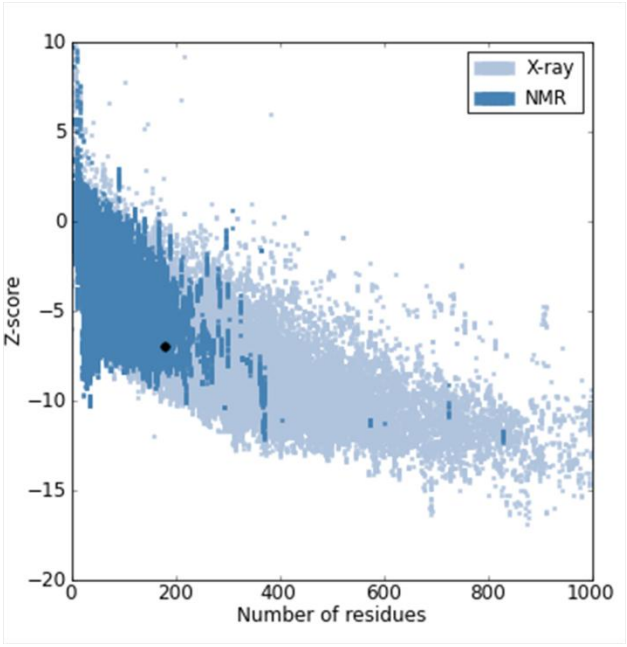


Figure 3. Investigation of the modeled structure of ClpP2 using the ProSA-web service. Z score obtained is -6.9, which is within the scores obtained for groups of structures from different sources (X-ray, NMR). It indicates overall model quality as the range falls within a characteristic range of native proteins.

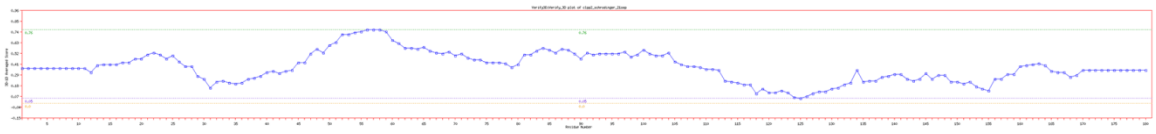


Figure 4. Verify 3D plot of the modeled structure of ClpP2. Verify 3D analysis of the model found no conformational errors, and there were no values less than 0.09, further indicating that all of the residues were located in favorable structural environments.

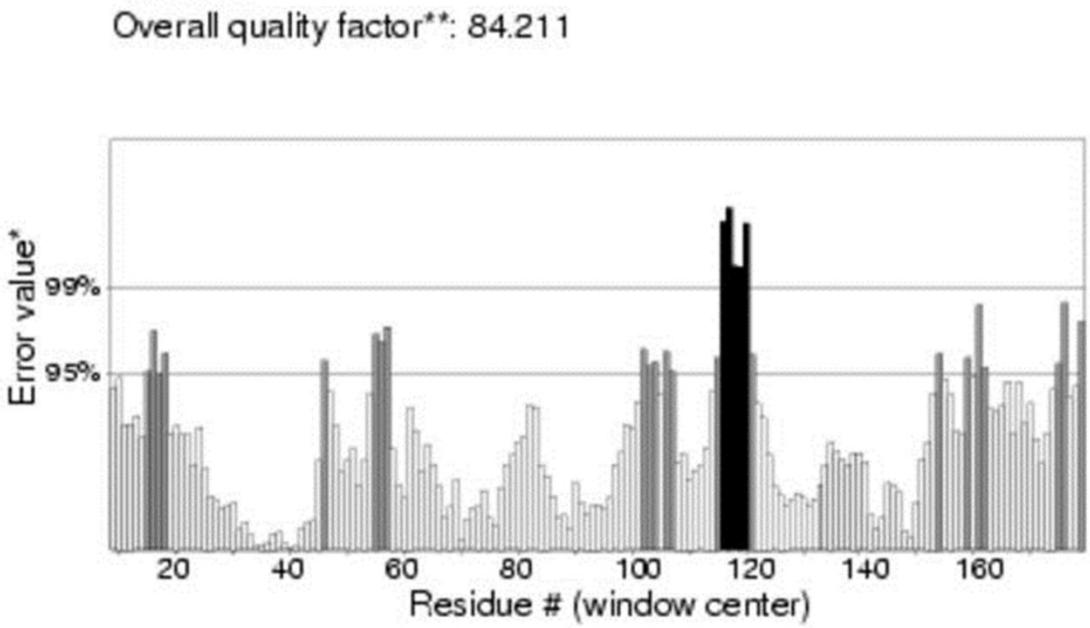


Figure 5. Showing ERRAT structural quality factor of the modeled structure of ClpP2.The ERRAT score for all of the models was >85%, well within the range of a high quality model.

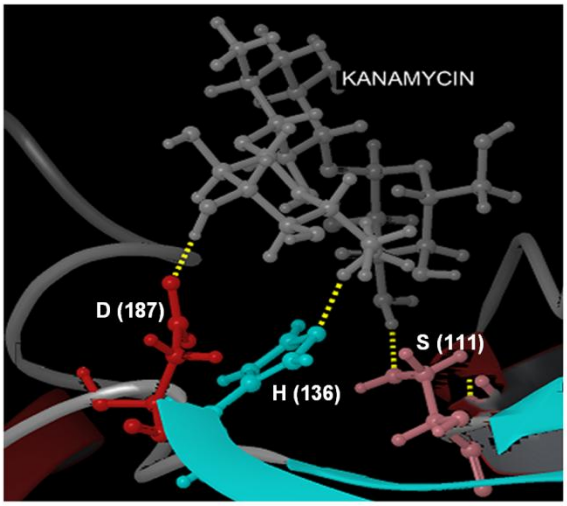


Figure 6 (a)

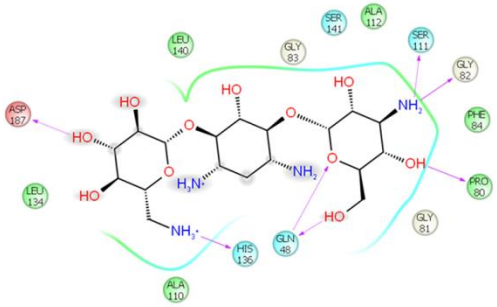
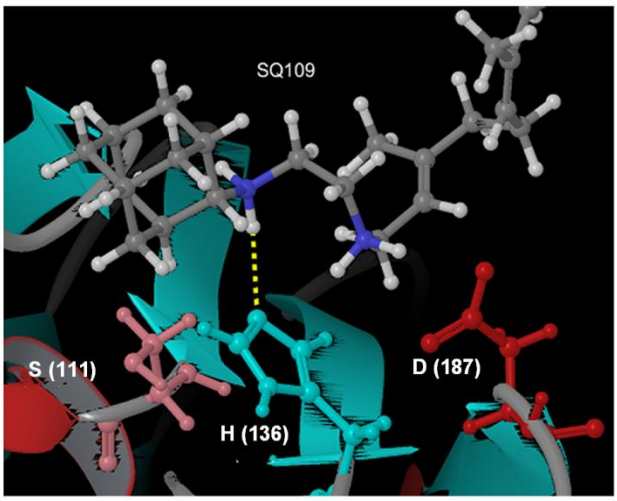


Figure 6 (b)

Figure 6. (a) Docked pose of Kanamycin. Kanamycin is shown to make hydrogen bond contacts with ClpP2 protease. (b) Ligand interaction map showing residues within 4A distance. S (111), D (187), H (136) of catalytic triad is shown to interact. Apart from this it is shown to interact with G (82), Pro (80), Gln (48). Pink arrows indicates the hydrogen bonding between ligand and residues.



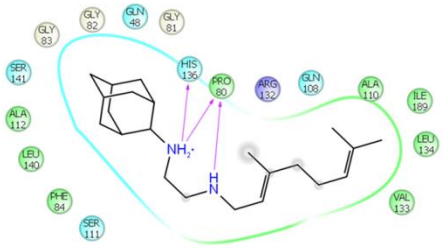
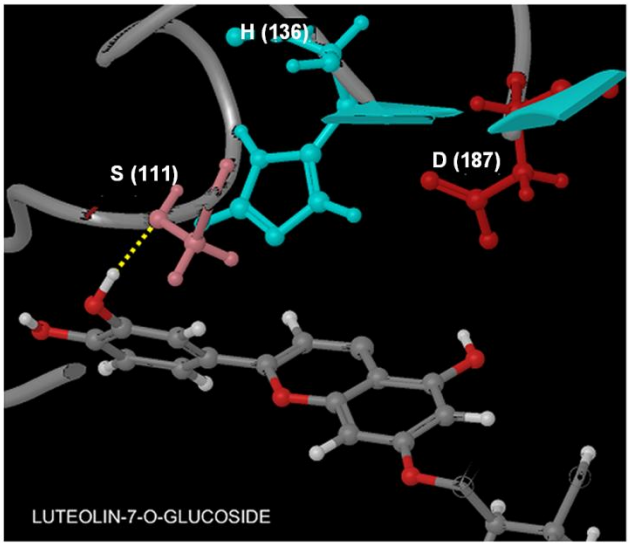


Figure 7 (a)

Figure 7 (b)

Figure 7. (a) Docked pose of SQ109 (Ethylenediamines). ClpP2 protease is shown to make hydrogen bond contacts with SQ109 (b) Ligand interaction map showing residues with in 4A distance. Side chain of D (187) of catalytic triad is shown to be involved in the Hydrogen bond. The other interactions shown are with P (80).



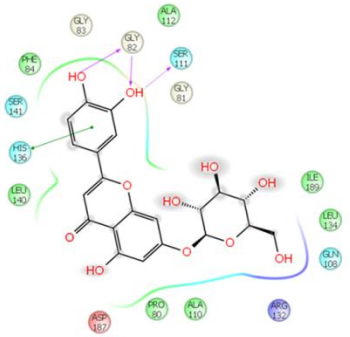
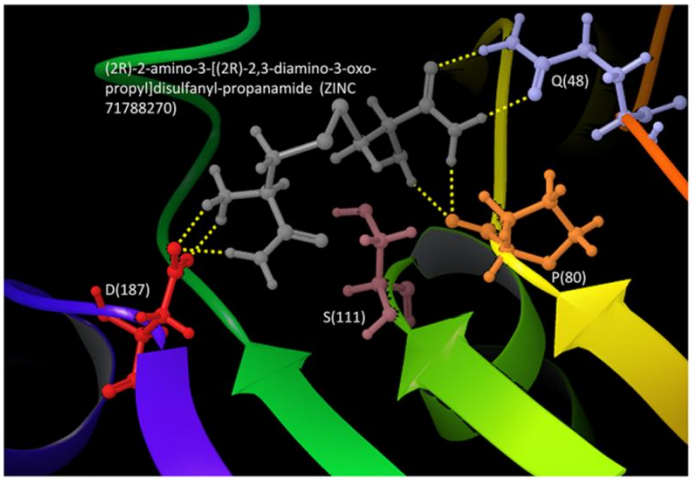


Figure 8 (a)

Figure 8 (b)

Figure 8. (a) Docked pose of Luteolin-7-O-glucoside. ClpP2 protease is shown to make hydrogen bond contacts with Luteolin-7-O-glucoside. (b) Ligand interaction map showing residues within 4A distance. Main chain interaction with G (82). Side chain interaction possible with S (111).



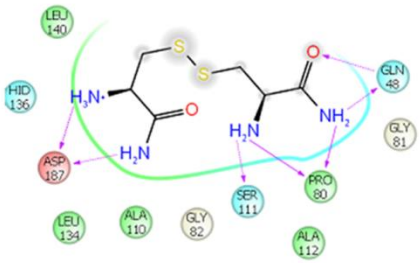
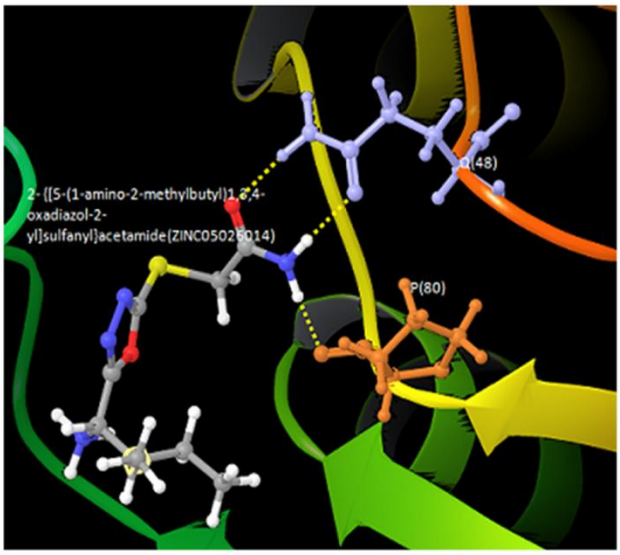


Figure 9 (a) **Figure 9 (b)**

Figure 9. (a) Docked pose of ZINC71788270 {(2R)-2-amino-3-[(2R)-2,3-diamino-3-oxo-propyl]disulfanyl-propanamide}. ClpP2 protease is shown to make hydrogen bond contacts with ZINC71788270. (b) Ligand interaction map showing residues within 4A distance showing interactions with D (187), Q (48) and P (80).



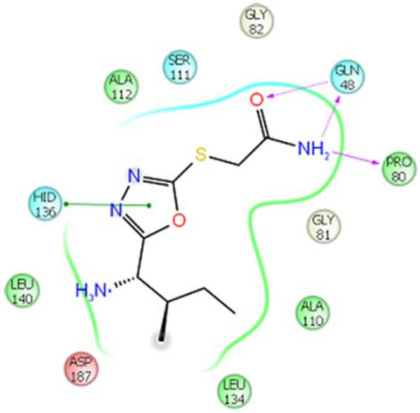
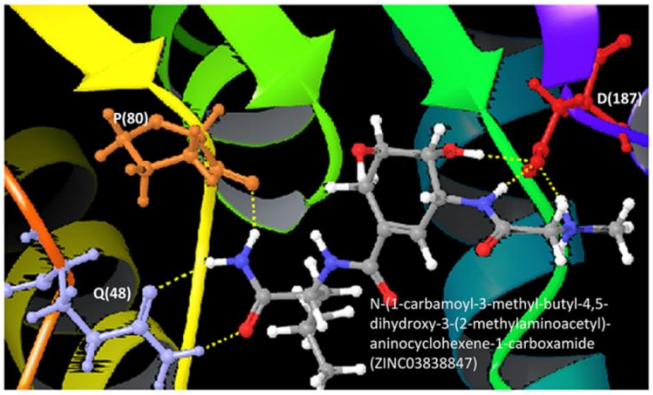


Figure 10 (a) **Figure 10 (b)**

Figure 10. (a) Docked pose of ZINC05026014 (2-{[5-(1-amino-2-methylbutyl) -1,3,4-oxadiazol-2-yl] sulfanyl} acetamide). ClpP2 protease is shown to make hydrogen bond contacts with ZINC05026014. (b) Ligand interaction map showing residues within 4A distance showing interactions with P (80) and Q (48).



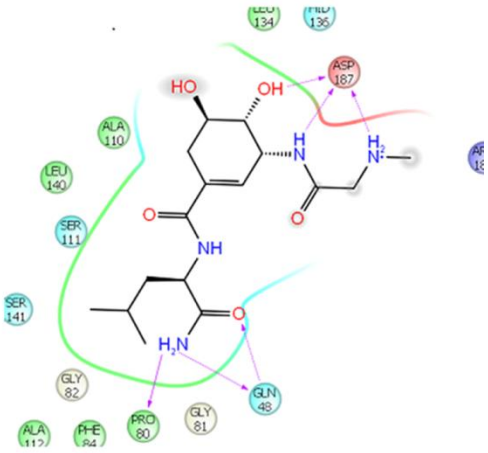
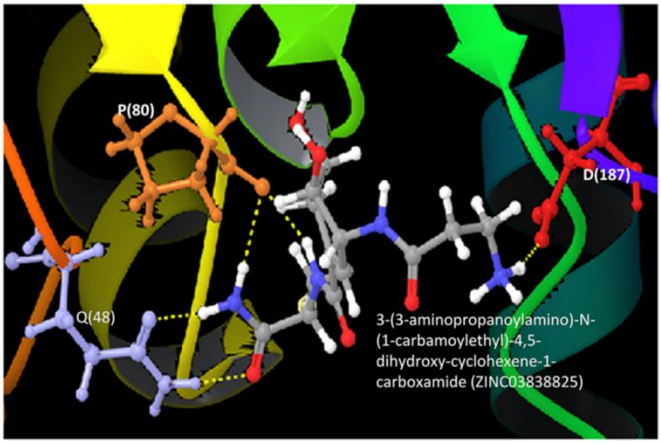


Figure 11 (a) **Figure 11 (b)**

Figure 11. (a) Docked pose of ZINC03838847 {N-(1-carbamoyl-3-methyl-butyl) -4,5-dihydroxy-3-(2-methylaminoacetyl) amino-cyclohexene-1-carboxamide}. ClpP2 protease is shown to make hydrogen bond contacts with ZINC03838847. (b) Ligand interaction map showing residues within 4A distance showing interactions with D (187), P (80) and Q (48).



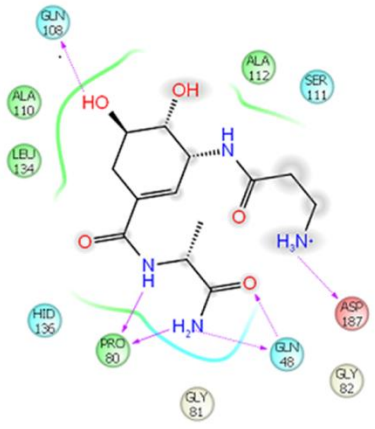


Figure 12 (a) **Figure 12 (b)**

Figure 12. (a) Docked pose of ZINC03838825 {3-(3-aminopropanoylamino) -N-(1-carbamoylethyl) -4,5-dihydroxy-cyclohexene-1-carboxamide}. ClpP2 protease is shown to make hydrogen bond contacts with ZINC03838825. (b) Ligand interaction map showing residues in 4A distance showing interactions with D (187), P (80) and Q (48).

Docking studies

We studied the binding affinity of available drugs, some of the phytoconstituents with reported antitubercular activity and compounds from zinc database (http://zinc.docking.org/). Docking studies were performed using ligands belonging to First line drugs (Isoniazid, Rifampin, Pyrazinamide and Ethambutol), second line drugs and advanced drugs belonging to the class of Aminoglycosides (Kanamycin and Amikacin), polypeptides (capreomycin, viomycin and enviomycin), Fluroquinolones (ciprofloxacin), Thioamides (Ethionamide, Propionamide), D-Cycloserine, p-aminosalicylic acid, Quinolones (Moxifloxacin, Gatifloxacin), Oxazolidinones (Linezolid), Nitroimidazoles (OPC-67683, PA 824). The "Extra Precision" (XP) mode of Glide docking [34] (version 9.3, Schrodinger Inc.) was used to perform all docking calculations using the OPLS-AA 2005 force field. In this work the bounding box of size 10 Å ×10 Å ×10 Å was defined in ClpP2 Protease of Mtb by selecting residues of a catalytic triad namely S (111), H (136) and D (187). The catalytic triad of ClpP2 of Mtb was identified from the alignment of Mtb proteins onto E. coli ClpP and locating the catalytic triad of Asp-His-Ser, which is characteristic of serine proteases [35, 36]. These docking calculations provide insight into the interactions of ligands with amino acids in the binding pocket of the target and to predict the corresponding binding affinities of ligands [37]. Table-S1 (see supplementary material) presents the Glide XP docking results. Their Glide score values among the ligands vary in between -10.181449 and 0.18402 kcal/mol. Of the first line drugs, Ethambutol showed highest binding affinity with a Glide score of (-3.224532) in comparison to other drugs. Thus, apart from inhibiting the synthesis of arabinogalactan [3], Ethambutol may also inhibit ClpP1P2 protease. The second-line drugs studied in the present investigation included Kanamycin, amikacin, capreomycin, ciprofloxacin, ethionamide, cycloserine and p-aminosalicylic acid. Of the second line drugs Kanamycin showed the highest glide score (-10.181449) (Table.S1) Kanamycin was observed to interact with all the three active site residues namely S (111), H (136) and D (187) Figure 6). Amikacin showed a docking score of (-6.25064) 5), Capreomycin (--5.709099), p-aminosalicylic acid (-3.95398), ethionamide (-3.403275), cycloserine (-3.472878) and ciprofloxacin (-4.005264).

Thus, it appears that the ClpP2 protease inhibition could be one of the mechanisms of action of these drugs. Our results are in accordance with previous reports where it was shown that the synergistic nature of ClpP1P2 protease depletion with aminoglycosides, a class of drugs already used to treat tuberculosis, points to a potential combination therapy against Mtb [16].

In addition, ClpP1P2 protease is necessary for degrading abnormal proteins, such as those produced in the presence of certain antibiotics. The accumulation of such non-functional misfolded proteins might result in cellular stress in the absence of an effective system for their removal [33]. In fact, it was recently discovered that the natural product cyclomarin kills Mtb by targeting the ClpC1 ATPase and presumably increasing Clp-mediated proteolysis, as demonstrated in the whole cell fluorescence-based assay [38].

Advanced Drugs

SQ109 was identified as anti-Mtb leads among analogues of ethambutol prepared through combinatorial chemistry. SQ109 is implicated in affecting cell wall biosynthesis, but it is also active against ethambutol-resistant strains, and its precise mechanism of action has not been elucidated and is potentially novel. It nonetheless produces a remarkable synergy in vitro and animal models when combined with rifampin, isoniazid, or TMC207. Our results demonstrate that the inhibition of ClpP2 protease of Mtb appears to be one of its mechanism of inhibition of SQ109 (Figure. 7), as it showed a Glide score of -3.946659 (Table.S1).

A combination of clavulanic acid, a β-lactamase inhibitor, and meropenenm, a carbapenenm antibiotic, was shown to have potent activity in vitro killing XDR Mtb under aerobic as well as anaerobic conditions [3]. Docking of Meropenenm & clavulanic acid showed a glide score of -3.710785, indicating it's efficacy in inhibiting ClpP2 protease of Mtb. Earlier studies have shown that d-transpeptidase is the target of carbapenems in Mtb. In comparison to older quinolones, moxifloxacin and gatifloxacin are more potent in vitro. These days they are targeted against TB as anti-TB agents. It has been reported that they exclusively target Mtb. From our docking studies, it has been observed that the two quinolones, moxifloxacin could inhibit ClpP2 protease to the greatest extent in comparison to Gatifloxacin as docking studies yielded a Glide score of -3.818629 and -3.576391 (Table.S1).

Linezolid, which represents the oxazolidinone class of antibiotics. It targets Mtb ribosome and exhibits potent activity against MDR-TB clinical isolates. Linezolid has been used for MDR-TB (and XDR-TB), and additional Phase II clinical trials are ongoing [3]. Docking studies have shown that ClpP2 is another target for Linezolid (glide score -3.189251) (Table. S1).

OPC-67683 has a similar mechanism of action to that of PA-824, which might be due to inhibition of mycolic acid biosynthesis. OPC-67683 is 4–16 times more potent than PA-824 in vitro and it does not show cross-resistance to other first-line TB drugs [3]. It is currently in Phase II clinical trials for MDR-TB patients [3]. Our results of

docking agree with the reports as evident by higher glide score for OPC-67683 (-3.524075) in comparison to PA-824 (-1.860373) (Table.S1) whereas, BT043 showed no interaction with a catalytic triad.

Binding affinity of Anti-Tubercular phytoconstituents

Adverse effects of antitubercular drugs include skin rash, hepatitis, Abdominal pain, nausea, vomiting, hepatitis, renal damage, Dizziness, headache, depression, psychosis, convulsions.

Phytoconstituents responsible for anti-tubercular activity includes alkaloids, glycosides, tannins, phenolics, xanthones, quinones, sterols, triterpenoids, etc. These phytoconstituents present in plant exert a desired pharmacological effect on the body and thus act as natural anti-tubercular agents. The various phytoconstituents investigated for binding affinity in the present investigation are quercetin, urosalic acid, oleanolic acid, aloe emodin, naphthoquinone, lignans, α-Amyrin, lupeol, luteolin-7-O-glucoside, coumarin, gallic acid, xanthone, piperine (Table.S1). They were tested for their binding affinity with ClpP2 protease, which is a novel target in developing antituberculosis drugs. Based on the Glide docking for Luteolin-7-O-glucoside (-6.087934) (Figure-8) highest binding energy followed by Quercetin (-5.856694), Aloe emodin (-4.527424), Naphthoquinone (3.34881), Xanthone (-2.889023), Lignan (-1.815696), Leupol (-1.324373), piperine (-1.103315), Oleanolic acid -0.200838) urosalic acid (-0.184021 (Table-S1)

Virtual Screening

The virtual screening of substrate database with prepared protein was performed with OPLS2005 force field using virtual screening workflow (VSW) module of the Schrodinger Suite. VSW uses glide docking to rank the best compound which utilizes the scoring functions, High throughput virtual screening (HTVS), standard precision (SP) and extra precision (XP). HTVS and SP modes are used for a large set of ligands and XP docking is more accurate than the above two methods. It uses the ligand poses that have a high score from SP docking. The XP GlideScore scoring function was used to order the best ranked compounds and the specific interactions as pi-cation and pi-pi stacking were analyzed using XP visualizer in Glide module. Based on results obtained after the screening of ligands from zinc database using virtual screening workflow, Out of 1,25000 compounds screened, 126 molecules were selected from the XP output based on its good interaction, forming hydrogen bonds with the active site residues S (111), D (187) and H (136) of ClpP2 of Mtb (Table S2; Figure.3) Hydrogen bonding measures the intermolecular interaction between the protein and ligands. The top ranking compounds based on GLIDE scores are listed in Table-S2. The GLIDE scores of these compounds have ranges of -8.338 to -4.615Kcal/mol. The top ranking molecule (ZINC71788270 (2R)-2-amino-3-[(2R)-2,3-diamino-3-oxopropyl]disulfanyl-propanamide) has a Glide score of -8.338 Kcal/mol with 8 hydrogen bonds and 7 rotatable bonds. The details of other potential inhibitors are given in Table-S2 (Supplementary Table). The top ligands and their hydrogen bond interactions are depicted in Figure 9, 10, 11 and 12 respectively. It is evident from this analysis that the best compounds are located in the center of the active site formed by catalytic triad. Since computational screenings need experimental validation in order to confirm the accurate drug molecule(s), the proposed LEAD molecules need to be optimized in further studies. Thus, the significance of this work lies in providing a rapid approach to screen compounds that are likely to inhibit the ClpP2 protease of Mtb, as the therapeutic efficacy of Clp protease inhibition is already reported in Staphylococcus aureus [39]. Our study has identified some of the inhibitors which are currently used drugs like Etambutanol, Kanamycin and Amikacin apart from giving structural basis to the binding affinity of some of the phytoconstituents with reported antitubercular activity [26]. We have also identified 126 molecules from Virtual screening workflow of Schrodinger software (Table S-2) as potential inhibitors with ClpP2 of Mtb. However, experimental studies are needed to validate these in silico findings.

Conclusions

ClpP of Mtb is one of the most recent drug target for tuberculosis. In this work, we have constructed a 3D model of ClpP2 and the refined model was further used for docking studies and high throughput virtual screening. Docking results indicate that, out of the various classes of drugs tested, drugs belonging to aminoglycosides showed highest binding affinity as evident by docking score and glide score. Furthermore, we have observed that some of the antitubercular phytoconstituents have a higher binding affinity than second line drugs. As it is known that second line and advanced drugs have more adverse effects, our work suggests that phytoconstituents could be an alternative treatment option or should be given as adjunct to reduce the adverse effects and increase the antibacterial activity. High throughput Virtual screening results have indicated ligands belonging to amide group as potential inhibitors of ClpP2 protease which is one of the components of ClpP1P2 complex, essential for growth and infection of Mtb. Even though, these findings will be useful in rational drug development, but further experimental evaluation is required for complete understanding the mechanism of inhibition.

References

1. **Zarir FU, Rohit A A, Kanchan K A, and Camilla R (2011)** Totally Drug-Resistant Tuberculosis in India. Clin Infect Dis. first published online December 21: 2011 doi:10.1093/cid/cir889.
2. **Juan-Pablo M, Antonio M, Laia F, Lucía del B, Angels Ou, et al. (2012)** Factors that influence current tuberculosis epidemiology. European spine journal, DOI: 10.1007/s00586-012-2334-8.
3. **Takushi Kaneko, Christopher Cooper, Khisimuzi Mdluli (2011)** Challenges and opportunities in developing novel drugs for TB. Future Med. Chem 3: 1373–1400.
4. **Kress W, Maglica Z, Weber-Ban E. (2009)** Clp chaperone-proteases: structure and function. Res. Microbiol. 160:618–628.
5. **Webb C, Moreno M, Wilmes-Riesenberg M, Curtiss R, Foster JW (1999)** Effects of DksA and ClpP protease on sigma S production and virulence in Salmonella typhimurium. Mol Microbiol. 34:112–123.
6. **Gaillot O, Pellegrini E, Bregenholt S, Nair S, Berche P (2000)** The ClpP serine protease is essential for the intracellular parasitism and virulence of Listeria monocytogenes. Mol Microbiol. 35:1286–1294.
7. **Robertson GT, Ng WL, Foley J, Gilmour R, Winkler ME (2002)** Global transcriptional analysis of clpP mutations of type 2 Streptococcus pneumoniae and their effects on physiology and virulence. J Bacteriol. 184:3508–3520.
8. **Frees D, Qazi SN, Hill PJ, Ingmer H (2003)**. Alternative roles of ClpX and ClpP in Staphylococcus aureus stress tolerance and virulence. Mol Microbiol. 8:1565–1578.
9. **Loughlin MF, Arandhara V, Okolie C, Aldsworth TG, Jenks PJ (2009)** Helicobacter pylori mutants defective in the clpP ATP-dependant protease and the chaperone clpA display reduced macrophage and murine survival. Microb Pathog. 46:53–57.
10. **Sherrid AM, Rustad TR, Cangelosi GA, Sherman DR (2010)** Characterization of a Clp protease gene regulator and the reaeration response in Mycobacterium tuberculosis. PLoS One. 5(7):e11622.
11. **Cole ST, et al. (1998)** Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. Nature 393:537–544.
12. **Katayama-Fujimura, Y Gottesman, S Maurizi, M.R (1987)** A multiple-component, ATP-dependent protease from Escherichia coli. J Biol Chem 262: 4477–4485.
13. **Hwang BJ, Woo KM, Goldberg AL, Chung CH (1988)** Protease Ti, a new ATP-dependent protease in Escherichia coli, contains protein-activated ATPase and proteolytic functions in distinct subunits. J Biol Chem 263: 8727-34.
14. **Juliane Ollinger, Theresa O'Malley, Edward A. Kesicki, Joshua Odingo and Tanya Parish (2012)** Validation of the Essential ClpP Protease in Mycobacterium tuberculosis as a Novel Drug Target. J. Bacteriol 194:663-668.
15. **Sassetti CM, Boyd DH, Rubin EJ (2003)** Genes required for Mycobacterial growth defined by high density mutagenesis. Mol Microbiol 48:77–84.
16. **Ravikiran M. Raju et al. (2012)** Mycobacterium tuberculosis ClpP1 and ClpP2 Function together in protein degradation and are required for viability in vitro and during Infection. PLoS Pathogen 8(2): e1002511. doi:10.1371/journal.ppat.100251.
17. **Drag M, Salvesen GS (2010)** Emerging principles in protease-based drug discovery. Nat Rev Drug Discov 9: 690–701.
18. **Nadia Benaroudj et al (2011)** Assembly and proteolytic processing of mycobacterial ClpP1 and ClpP2. BMC Biochemistry 12:61.
19. **Jimin Wang, James A. Hartling, John M. Flanagan (1997)** The Structure of ClpP at 2.3 A Resolution Suggests a Model for ATP-Dependent Proteolysis. Cell 91: 447–456.
20. Prime, version 3.1, Schrödinger, LLC, New York, NY, 2012.
21. **M. P. Jacobson, D. L. Pincus, C. S. Rapp, T. J. F. Day, B. Honig, D. E. Shaw, and R. A. Friesner(2004)** A Hierarchical Approach to All-Atom Loop Prediction. Proteins 55:351-367.
22. **Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993)** PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Cryst 26: 283-291.
23. **Bowie JU, Lüthy R, Eisenberg D (1991)** A method to identify protein sequences that fold into a known three-dimensional structure. Science 253:164-70.
24. **Wiederstein, M. & Sippl, M.J. (2007)** ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Research 35:W407-W410.
25. **Colovos C, Yeates TO (1993)** Verification of protein structures: patterns of nonbonded atomic interactions. Protein Sci. 2:1511-1519.

26. **Vikrant Arya (2011)** A Review on Anti-Tubercular Plants. International Journal of PharmTech Research 3: 872-880.
27. **Banks JL, Beard HS, Cao Y, Cho AE, Damm W, Farid, R, Felts A.K, Halgren TA, Mainz DT, Maple JR, Murphy R. Philipp DM, Repasky MP, Zhang LY, Berne BJ, Friesner RA, Gallicchio E. and R.M. Levy (2005)** Integrated Modeling Program, Applied Chemical Theory (IMPACT). J. Comp. Chem. 26:1752
28. **Thomas A. Halgren, Robert B. Murphy, Richard A. Friesner, Hege S. Beard, Leah L. Frye, W. Thomas Pollard, and Jay L. Banks (2004)** Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening. J. Med. Chem. 47:1750-1759.
29. **Richard A. Friesner, Robert B. Murphy, Matthew P. Repasky, Leah L.Frye, Jeremy R. Greenwood, Thomas A. Halgren, Paul C. Sanschagrin,and Daniel T. Mainz (2006)** Extra Precision Glide: Docking and ScoringIncorporating a Model of Hydrophobic Enclosure for Protein-LigandComplexes. J. Med. Chem. 49:6177-6196.
30. **Ingmer H, Brøndsted L (2009)** Proteases in bacterial pathogenesis. Res Microbiol 160: 704–710.
31. Goldberg AL (2003) Protein degradation and protection against misfolded or damaged proteins. Nature 426: 895–899.
32. **Glickman MH, Ciechanover A (2002)** The ubiquitin-proteasome proteolytic pathway: destruction for the sake of construction. Physiol Rev 82: 373–428.
33. **Goldberg AL (1972)** Degradation of abnormal proteins in Escherichia coli (protein breakdown-protein structure-mistranslation-amino acid analogs-puromycin). Proc Natl Acad Sci USA 69: 422–426.
34. **RA Friesner, et al. (2004)** Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. J Med Chem 47:1739-49.
35. **Maurizi, M.R., Clark, W., Kim, S, H., and Gottesman, S (1990).** Clp P represents a unique family of serine proteases. J.Biol.Chem 265:12546-12552.
36. **Ingvarsson H, Mate´ MJ, Ho¨gbom M, Portnoı¨ D, Benaroudj N, et al. (2007)** Insights into the inter-ring plasticity of caseinolytic proteases from the X-ray structure of Mycobacterium tuberculosis ClpP1. Acta Crystallogr D Biol Crystallogr 63: 249–259. doi:10.1107/S0907444906050530.
37. **E.M Krovat, T.Steindl and T.Lange (2005)** Recent advance in docking and scoring. Current computer-aided drug design 1:93-102.
38. **Schmitt EK, Riwanto M, Sambandamurthy V, Roggo S, Miault C, et al. (2011)** The Natural Product Cyclomarin Kills Mycobacterium Tuberculosis by Targeting the ClpC1 Subunit of the Caseinolytic Protease. Angew Chem Int Ed Engl 50:5889–5891.
39. **Böttcher T, Sieber SA (2008)** Beta-lactones as specific inhibitors of ClpP attenuate the production of extracellular virulence factors of Staphylococcus aureus. J Am Chem Soc 130:14400–14401.

Table.S1 (Supplementary data)

Entry ID	Compound ID_Entry Name	docking score	glide gscore	glide evdw	glide ecoul	glide energy	XP HBond
1.	CID_6032_Kanamycin	-7.908949	--10.181449	-18.551383	-24.985783	-43.537166	-5.44162
2.	CID_37768_Amikacin	-6.15844	-6.25064	-17.548005	-27.534527	-45.082532	-3.48669
3.	CID_5280637_Luteolin-7-O-glucoside	-6.071434	-6.087934	-27.982371	-11.188833	-39.171204	-3.38895
4.	CID_5280343_quercitin.1	-5.849594	-5.856694	-20.577932	-8.416987	-28.994919	-3.28755
5.	CID_370_Gallic acid	-5.190873	-5.190873	-12.33501	-9.024444	-21.359454	-2.88
6.	CID_3034236_Capreomycin IIA	-4.962999	-5.709099	-22.182919	-17.995193	-40.178112	-3.33952
7.	CID_3037981_Viomycin	-4.862172	-5.272872	-17.692519	-20.237088	-37.929608	-2.84413
8.	CID_10207_Aloe emodin	-4.527424	-4.527424	-19.66682	-9.587741	-29.254561	-2.07494
9.	CID_4649_Aminosalicylic Acid	-3.95398	-3.95398	-9.060618	-5.460405	-14.521023	-2.16859
10.	CID_152946_Moxifloxcin	-3.707229	-3.818629	-25.891373	-6.519847	-32.41122	0
11.	CID_23665637_Meropenenm plus clavulanic acid	-3.578185	-3.710785	-20.405555	-15.020099	-35.425654	-1.18651
12.	CID_2761171_ ethionamide	-3.403275	-3.403275	-16.98628	-6.124187	-23.110466	-1.26045
13.	CID_8530_Naphthoquinone_isoniazed	-3.34881	-3.34881	-14.207608	-7.18389	-21.391498	-1.35714
14.	CID_6234_d-cycloserine	-3.287978	-3.472878	-12.022442	-5.44543	-17.467871	-1.25722
15.	CID_441401_Linezolid	-3.18	-3.189251	-22.567516	-6.767812	-29.335328	-0.7
16.	CID_14052_ethambutol	-3.12	-3.224532	-13.367152	-13.77941	-27.146562	-1.10186
17.	CID_5274428_SQ109	-3.081659	-3.946659	-15.288194	-6.944295	-22.232489	-0.63902
18.	CID_2764_ciprofloxacin	-2.991464	-4.005264	-20.477068	-10.462637	-30.939705	-1.82493
19.	CID_7020_Xanthone	-2.889023	-2.889023	-11.900819	-4.021612	-15.922431	-1.29714
20.	CID_6480466_OPC-67683	-2.725875	-3.524075	-28.043351	-8.051299	-36.09465	-1.30316
21.	CID_5379_gatifloxacin	-2.495091	-3.576391	-15.034498	-16.468992	-31.50349	-2.10992
22.	CID_6578_propionamide	-2.345441	-2.345441	-8.725306	-6.934172	-15.659478	-1.33
23.	CID_6323497_Rifapentine	-2.071247	-2.783047	-28.94276	-0.547571	-29.490331	-1.44
24.	CID_456199_PA-824	-1.860373	-1.860373	-29.016039	-3.274754	-	

25.	CID_10607_Lignans.	-1.815696	-1.815696	-19.334844	-4.287432	-23.622276	-0.7
26.	CID_5388906_Bedaquiline	-1.616177	-1.616177	-23.245986	-6.74219	-29.988176	-1.30918
27.	CID_42609849_BTZ043	-1.37608	-1.93168	-25.219378	-0.409086	-25.628464	-0.39528
28.	CID_73145_lupeol	-1.324373	-1.324373	-25.886035	-2.152901	-28.038936	-0.46832
29.	CID_638024_Piperine	-1.103315	-1.103315	-22.428424	-4.418711	-26.847136	-0.7
30.	CID_1049_ Oleanolic acid	-0.192238	-0.200838	-25.773503	0.662432	-25.111071	0
31.	CID_64945_Urosalic acid	2.304579	-0.184021	-17.270969	-8.323981	-25.594949	-1.15282

Table-S2 (Supplementary Table)

S.NO	COMPOUND ID	COMPOUND	DOCKING SCORE	GLIDE GSCORE	pH range	xlogP	Apolar desolvation (kcal/mol)	Polar desolvation (kcal/mol)	H-bond donors	H-bond acceptors	Net charge	tPSA (Å²)	Molecular weight (g/mol)	Rotatable bonds
1	ZINC71788270	(2R)-2-amino-3-[(2R)-2,3-diamino-3-oxo-propyl]disulfanyl-propanamide	-8.338	-8.338	Mid (pH 6-8	-2.35	-6.01	-12.26	8	6	0	138	238.338	7
2	ZINC05026014	2-{[5-(1-amino-2-methylbutyl)-1,3,4-oxadiazol-2-yl]sulfanyl}acetamide	-6.721	-6.721	pH 7	-1.99	-6.9	-50.29	5	6	1	109	245.328	6
3	ZINC03838847	N-(1-carbamoyl-3-methyl-butyl)-4,5-dihydroxy-3-(2-methylaminoacetyl)amino-cyclohexene-1-carboxamide	-6.325	-6.325	pH 7	-0.97	-10.15	-52.12	8	9	1	158	357.431	8
4	ZINC03838825	3-(3-aminopropanoylamino)-N-(1-carbamoylethyl)-4,5-dihydroxy-cyclohexene-1-carboxamide	-6.261	-6.261	pH 7	-2.51	-2.51	-53.35	10	11	1	201	384.413	7
5	ZINC03838818	N-(1-carbamoylethyl)-4,5-	-6.204	-6.204	pH 7	-2.45	-10.76	-45.29	8	9	1	158	329.377	6

		dihydroxy-3-(2-methylaminopropanoylamino)cyclohexene-1-carboxamide												
6	ZINC05499882	2-[[(4-ethoxyphenyl)amino-sulfanyl-methylene]aminomethylene]propanediamide	-6.175	-6.175	pH 7	-0.30	-0.49	-33.32	6	7	0	121	309.371	8
7	ZINC00492320	3-phenyloxirane-2,2-dicarboxamide	-6.127	-6.127	pH 7	-0.74	-2.55	-11.25	4	5	0	99	206.201	3
8	ZINC03838821	N-(1-carbamoylethyl)-4,5-dihydroxy-3-(2-methylaminoacetyl)amino-cyclohexene-1-carboxamide	-6.107	-6.107	pH 7	-2.28	-11.09	-48.36	8	9	1	158	315.35	6
9	ZINC03995579	2,4-DIHYDROXYPHENYLACETYL-L-ASPARAGINE	-6.084	-6.084	pH 7	-1.34	-3.32	-61.86	5	8	-1	153	281.244	6
10	ZINC72154561	N-[(1S,2R)-2-aminocyclobutyl]-2-oxo-1,2-dihydrobenzo[cd]indole-6-sulfonamide	-6.082	-6.082	pH 7	1.02	0.3	-59.17	5	6	1	107	318.378	3
11	ZINC03838841	N-(2-carbamoylethyl)-4,5-dihydroxy-3-(2-methylaminoacetyl)amino-cyclohexene-	-6.033	-6.033	pH 7	2.90	-11.75	-56.05	8	9	1	158	315.35	7

		1-carboxamide												
12	ZINC03838850	3-(2-amino-3-methyl-butanoyl)amino-N-(1-carbamoyl-3-methyl-butyl)-4,5-dihydroxy-cyclohexene-1-carbox	-5.984	-5.984	pH 7	-1.27	-11.04	-53.63	9	9	1	169	385.485	8
13	ZINC03838832	N-[3-(1-carbamoylethylcarbamoyl)-5,6-dihydroxy-1-cyclohex-2-enyl]-4-hydroxy-pyrrolidine-2-carboxamid	-5.964	-5.964	pH 7	-3.31	-14.59	-46.71	9	10	1	178	357.387	5
14	ZINC12763227	(2S)-2-[[4-(3-amino-3-oxo-propyl)-5-(2-furyl)-1,2,4-triazol-3-yl]sulfanyl]-N-(ethylcarbamoyl)propana	-5.885	-5.885	pH 7	0.13	3.47	-23.32	4	10	0	145	380.43	8
15	ZINC02601319	methyl 2-(3,4-dihydroxyphenyl)-3,3,3-trifluoro-2-hydroxypropanoate	-5.816	-5.816	pH 7	1.15	-1.12	-9.78	3	5	0	86	266.171	4
16	ZINC20463689	1-[4-[(1S,2S,3S,4R,5R)-2-amino-3-hydroxy-6,8-dioxabicyclo[3.2.1]octan-4-yl]piperazin-1-yl]ethanone	-5.815	-5.815	pH 6-8	-2.55	-1.1	-11.91	3	7	0	88	271.317	1
17	ZINC03838761	3-amino-N-(carbamoylme	-5.795	-5.795	pH 7	-3.69	-11.85	-52.68	8	7	1	140	230.244	3

		thyl)-4,5-dihydroxy-cyclohexene-1-carboxamide												
18	ZINC03838829	N-(carbamoylmethyl)-4,5-dihydroxy-3-(2-methylaminoacetyl)amino-cyclohexene-1-carboxamide	-5.740	-5.740	pH 7	-3.17	-11.69	-47.97	8	9	1	158	301.323	6
19	ZINC55200276	3-(3-{[(2-hydroxy-3-phenylpropyl)amino]methyl}-1H-indol-1-yl)propanamide	-5.716	-5.716	pH 7	1.88	6.13	-52.91	5	5	1	85	352.458	9
20	ZINC04179503	5-(hydroxymethyl)-2-[(2-nitrophenyl)amino]oxolane-3,4-diol	-5.713	-5.713	pH 7	0.55	-8.76	-12.84	4	8	0	127	270.241	4
21	ZINC03838823	N-[3-(carbamoylmethylcarbamoyl)-5,6-dihydroxy-1-cyclohex-2-enyl]-4-hydroxy-pyrrolidine-2-carboxamide	-5.709	-5.709	pH 7	-2.38	-15.18	-46.46	9	10	1	178	343.36	5
22	ZINC03118398	N1-(2-amino-1-{[(4,6-dimethoxypyrimidin-2-yl)amino]methyl}-2-oxoethyl)benzamide	-5.686	-5.686	pH 7	1.29	-0.38	-15.72	4	9	0	128	345.359	8
23	ZINC03112227	[[(5E)-2-methyl-5-(thiocarbamoylhydrazono)-	-5.684	-5.684	pH 7	1.78	4.61	-29.04	8	8	0	130	387.494	5

		1,4-dihydroindeno[1,2-b]pyridine-3-carbonyl]amino]thioure												
24	ZINC47028706	N-[(1S)-2-imidazol-1-yl-1-methyl-ethyl]-4-pyrimidin-2-yl-piperazine-1-carboxamide	-5.679	-5.679	pH 6-8	0.86	9.27	-18.4	1	8	0	79	315.381	4
25	ZINC25491351	N4,N4-dimethylpyrimidine-2,4,5,6-tetramine	-5.625	-5.625	pH 6-8	-0.72	-0.09	-32.65	7	6	1	108	169.212	1
26	ZINC67756698	N~4~-[2-(4,5-dimethyl-1,3-thiazol-2-yl)ethyl]-5,6,7,8-tetrahydropyrido[3,4-d]pyrimidine-2,4-diamine	-5.493	-5.493	pH 7	0.69	5.07	-50.64	5	6	1	93	305.431	4
27	ZINC39736865	(2R)-N-(2,5-dimethylphenyl)-2-(imidazo[1,2-a]pyridin-2-ylmethylamino)-2-phenyl-acetamide	-5.491	-5.491	pH 7	4.31	12.74	-105.19	4	5	2	64	386.499	6
28	ZINC03838836	N-(carbamoylmethyl)-4,5-dihydroxy-3-(3-methyl-2-methylamino-butanoyl)amino-cyclohexene-1-carboxamide	-5.437	-5.437	pH 7	-2.56	-10.89	-43.99	8	9	1	158	343.404	7
29	ZINC32629443	(4,6-dimethoxypyri	-5.389	-5.389	pH 6-8	-0.30	-2.84	-5.7	2	5	0	70	169.184	3

		midin-5-yl)methylamine												
30	ZINC21984993	5-Amino-3-pyridinemethanamine	-5.376	-5.376	pH 7	-0.49	-1.45	-48.48	5	3	1	67	124.167	1
31	ZINC55246064	4-[[(5-chloro-1-methyl-imidazol-4-yl)sulfonylamino]methyl]tetrahydropyran-4-carboxamide	-5.366	-5.366	pH 7	-0.66	-1.39	-18.21	3	8	0	116	336.801	5
32	ZINC01065700	5-(6-amino-9H-purin-9-yl)-N-ethyl-3,4-dihydroxytetrahydro-2-furancarboxamide (non-preferred name)	-5.301	-5.301	pH 7	-0.74	-10.12	-21.79	5	10	0	148	308.298	3
33	ZINC03838831	N-(carbamoylmethyl)-4,5-dihydroxy-3-(4-methyl-2-methylamino-pentanoyl)amino-cyclohexene-1-carboxamid	-5.276	-5.276	pH 7	-2.03	-10.77	-45.06	8	9	1	158	357.431	8
34	ZINC14755862	(3S,7S,8aS)-3-(4-aminobutyl)-7-[(3-methoxybenzyl)amino]hexahydropyrrolo[1,2-a]pyrazine-1,4-dione	-5.264	-5.264	pH 6-8	-0.21	1.38	-53.17	5	7	1	98	361.466	8
35	ZINC55036332	3-[3-({[(6-methoxy-1H-benzimidazol-2-yl)methyl]ami	-5.261	-5.261	pH 7	1.75	5.79	-57.35	5	7	1	103	378.456	8

		no}methyl)-1H-indol-1-yl]propanamide												
36	ZINC03838803	3-amino-N-(1-carbamoyl-3-methyl-butyl)-4,5-dihydroxy-cyclohexene-1-carboxamide	-5.257	-5.257	pH 7	-1.49	-10.31	-53.77	8	7	1	140	286.352	5
37	ZINC04980610	1-[2-(dimethylamino)ethyl]-4-[(2,4-dimethyl-1,3-thiazol-5-yl)carbonyl]-5-(2-fluorophenyl)-3-hydroxy-1,5-dihydro-2H-pyrrol-2-one	-5.252	-5.252	pH 6-8	1.05	0.27	-55.73	1	6	1	71	404.487	6
38	ZINC33821171	(1R,2S)-1-[(6R)-2,4-diamino-5,6,7,8-tetrahydropteridin-6-yl]propane-1,2-diol	-5.217	-5.217	pH 7	-0.99	-3.79	-26.49	9	8	1	144	241.275	2
39	ZINC20268663	C-(3-Benzo[1,3]dioxol-5-yl-[1,2,4]oxadiazol-5-yl)-methylamine	-5.205	-5.205	pH 7	1.12	-0.18	-57.38	3	6	1	85	220.208	2
40	ZINC72135436	4-{[(5-fluoro-2-pyrimidinyl)amino]methyl}-4-azepanol	-5.203	-5.203	pH 7	0.45	3	-40.44	4	5	1	75	241.29	3
41	ZINC55246027	4-[[(1-methylpyrazol-4-yl)sulfonylamino]methyl]tetr	-5.195	-5.195	pH 7	-1.43	-3.34	-17.44	3	8	0	116	302.356	5

		ahydropyran-4-carboxamide												
42	ZINC13440025	2,8-diamino-9-[(2R,3R,4S,5R)-3,4-dihydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl]-1H-purin-6-one	-5.192	-5.192	pH 7	-2.21	-5.54	-23.28	8	11	0	186	298.259	2
43	ZINC39399027	6-(aminomethyl)-N-[3-(methylthio)benzyl]pyrimidin-4-amine	-5.151	-5.151	pH 7	1.44	5.01	-48.71	4	4	1	65	261.374	5
44	ZINC71951078	(2S)-1-[(2-phenyltriazol-4-yl)methyl]pyrrolidine-2-carboxamide	-5.127	-5.127	pH 7	0.85	4.12	-43.45	3	6	1	78	272.332	4
45	ZINC67514516	3,5-dimethyl-N-[1-methyl-2-(3-methyl-1H-pyrazol-5-yl)ethyl]-1H-indole-2-carboxamide	-5.105	-5.105	pH 7	3.37	6.5	-13.42	3	5	0	74	310.401	4
46	ZINC03838839	N-(1-carbamoyl-3-methyl-butyl)-4,5-dihydroxy-3-(4-methyl-2-methylamino-pentanoyl)amino-cyclohexene-1	-5.087	-5.087	pH 7	0.17	-9.2	-48.93	8	9	1	158	413.539	10
47	ZINC45108176	(2R,3S,4R,5R,6R)-5-fluoro-6-(fluoromethyl)tetrahydropyran-2,3,4-triol	-5.084	-5.084	pH 7	-1.22	-3.98	-11.6	3	4	0	70	184.138	1
48	ZINC12729228	2-[(3-methoxy-4-	-5.079	-5.079	pH 7	0.14	-3.3	-17.38	5	7	0	125	265.269	5

		methyl-benzoyl)amino]propanediamide												
49	ZINC65453096	2-(2-aminoethyl)-N-(1-methyl-2-pyridin-4-ylethyl)quinazolin-4-amine	-5.075	-5.075	pH 7	1.45	6.43	-49.28	4	5	1	78	308.409	6
50	ZINC71951076	(2S)-1-[[3-(2-pyridyl)-1,2,4-oxadiazol-5-yl]methyl]pyrrolidine-2-carboxamide	-5.063	-5.063	pH 7	0.39	1.03	-44.2	3	7	1	99	274.304	4
51	ZINC03838763	N-[3-(1-carbamoylethylcarbamoyl)-5,6-dihydroxy-1-cyclohex-2-enyl]thiophene-2-carboxamide	-5.054	-5.054	pH 7	-0.95	-10.29	-17.99	6	8	0	141	353.4	5
52	ZINC72143772	N-[(1S,2R)-2-aminocyclobutyl]-5-methoxy-1H-indole-2-carboxamide	-5.053	-5.053	pH 7	0.88	1.97	-57.53	5	5	1	82	260.317	3
53	ZINC12729222	2-(naphthalene-2-carbonylamino)propanediamide	-5.037	-5.037	pH 7	0.89	-1.06	-17.54	5	6	0	115	271.276	4
54	ZINC15417502	N-(4-chlorophenyl)-2,3-dimethyl-1H-indole-5-carboxamide	-5.028	-5.028	pH 7	4.30	8.25	-11.47	2	3	0	45	298.773	2
55	ZINC65534161	(3R,6R)-6-methyl-1-[(E)-3-phenylprop-2-enoyl]piperidi	-5.021	-5.021	pH 7	1.85	5.19	-12.37	2	4	0	63	272.348	3

		ne-3-carboxamide												
56	ZINC42783881	N-(sec-Butyl)-4,5,6,7-tetrahydro-1H-pyrazolo-[4,3-c]pyridine-3-carboxamide hydrochloride	-5.014	-5.014	pH 7	0.73	1.89	-41.42	4	5	1	74	223.3	3
57	ZINC03838868	3-(2-bromo-5-fluoro-phenoxy)-N-(1-carbamoylethyl)-4,5-dihydroxy-cyclohexene-1-carboxamide	-5.013	-5.013	pH 7	1.13	-8.35	-16.18	5	7	0	121	417.231	5
58	ZINC05163496	3-[(3-chloro-4-methyl-phenyl)aminoiminomethyl]benzene-1,2-diol	-5.005	-5.005	pH 7	5.05	4.22	-9.13	3	4	0	65	276.723	3
59	ZINC01995835	3-Amino-3-Ureido-Butyric Acid Ethyl Ester	-4.995	-4.995	pH 7	-1.15	-5.68	-43.25	6	6	1	109	190.223	5
60	ZINC05751362	3-amino-2-methylpropanenitrile	-4.987	-4.987	pH 7	-0.35	-0.4	-54.29	3	2	1	51	85.13	1
61	ZINC67688325	6-(3-hydroxyazetidin-1-yl)-N-[3-(2-isopropyl-1H-imidazol-1-yl)propyl]nicotinamide	-4.987	-4.987	pH 7	0.65	5.06	-38.3	3	7	1	85	344.439	7
62	ZINC67473171	({6-[4-(ethylsulfonyl)piperazin-1-yl]-2-methylpyrimidin-4-yl}methyl)ami	-4.986	-4.986	pH 7	-0.39	1.44	-57.27	3	7	1	94	300.408	4

		ne												
63	ZINC04808038	1-carbamoylethyl	-4.966	-4.966	pH 7	2.01	-2.43	-12.19	3	5	0	89	288.097	4
64	ZINC72168905	N-(1,2-diethyl-1H-benzimidazol-5-yl)-2-(3-pyrrolidinyl)benzamide	-4.964	-4.964	pH 7	3.24	9.71	-45.36	3	5	1	64	363.485	5
65	ZINC55215205	(3R)-1-[(E)-2-(p-tolyl)vinyl]sulfonylpiperidine-3-carboxamide	-4.960	-4.960	pH 7	1.68	2.99	-13.67	2	5	0	80	308.403	4
66	ZINC55246072	4-[[(5-chloro-2-thienyl)sulfonylamino]methyl]tetrahydropyran-4-carboxamide	-4.958	-4.958	pH 7	0.94	-1.46	-13.58	3	6	0	98	338.838	5
67	ZINC08036154	4-(2-pyridylaminoiminomethyl)benzene-1,2,3-triol	-4.955	-4.955	pH 7	1.30	0.11	-13.53	4	6	0	98	245.238	3
68	ZINC45108180	(2R,3S,4R,5S,6R)-5-fluoro-6-(fluoromethyl)tetrahydropyran-2,3,4-triol	-4.953	-4.953	pH 7	-1.22	-4.05	-8.04	3	4	0	70	184.138	1
69	ZINC02584876	2-Amino-4-(2,4-dichlorophenyl)-5-methylthiophene-3-carboxamide	-4.940	-4.940	pH 7	3.18	-4.18	-7.71	4	3	0	69	301.198	2
70	ZINC71479701	5-[3-hydroxy-5-(2-piperidin-2-ylethyl)phenyl]nicotinamide	-4.914	-4.914	pH 7	2.29	3.3	-51.93	5	5	1	93	326.42	5
71	ZINC71782960	(2S)-2-	-4.908	-4.908	pH 7	-1.27	-4.2	-11.29	6	6	0	118	251.286	6

		amino-N-[(1R)-2-amino-1-(hydroxymethyl)-2-oxo-ethyl]-3-phenyl-propanamide												
72	ZINC02773433	N-(4-{[(carbamoylcarbonylamino)amino]sulfonyl}phenyl)acetamide	-4.894	-4.894	pH 7	-1.96	-11.18	-18.34	5	9	0	147	300.296	5
73	ZINC12446815	2-{[(6-methyl-3-oxo-2H,4H-benzo[e]1,4-oxazin-7-yl)sulfonyl]amino}acetamide	-4.893	-4.893	pH 7	-0.86	-9.09	-17.47	4	8	0	128	299.308	4
74	ZINC52781587	2-[(2R)-6-chloro-3-oxo-4H-1,4-benzothiazin-2-yl]-N-(4-fluorophenyl)acetamide	-4.888	-4.888	pH 7	3.24	6.76	-10.82	2	4	0	58	350.802	3
75	ZINC04557026	2-(hydroxymethyl)-5-(6-imino-1-methylhydropurin-9-yl)oxolane-3,4-diol	-4.878	-4.878	pH 7	-4.76	-2.39	-36.34	5	9	1	131	282.28	2
76	ZINC69659476	1-[(3S)-3-imidazol-1-yl-1-piperidyl]-3-(2-naphthyl)propan-1-one	-4.878	-4.878	pH 7	3.24	13.33	-37.55	1	4	1	39	334.443	4
77	ZINC67784267	N~4~-[(5-isobutylisoxazol-3-yl)methyl]-5,6,7,8-tetrahydropyrido[3,4-	-4.870	-4.870	pH 7	1.44	5.63	-48.2	5	7	1	106	303.39	5

		d]pyrimidine-2,4-diamine												
78	ZINC04814145	N-(3,5-dimethoxyphenyl)-2-(3-oxopiperazin-2-yl)acetamide	-4.836	-4.836	pH 7	1.39	-0.05	-45.08	4	7	1	93	294.331	5
79	ZINC47074323	2-[[2-(2-oxo-1H-quinolin-3-yl)acetyl]amino]propanediamide	-4.824	-4.824	pH 7	-1.05	-3.6	-22.19	6	8	0	148	302.29	5
80	ZINC00519077	4-(4-(2,3-dihydrobenzo[b][1,4]dioxin-6-yl)-3-methylisoxazol-5-yl)benzene-1,3-diol	-4.822	-4.822	pH 7	3.02	3.03	-11.4	2	6	0	85	325.32	2
81	ZINC06232586	5-({[5-(1-aminoethyl)-1,3,4-oxadiazol-2-yl]sulfanyl}methyl)-6-methylpyrimidine-2,4(1H,3H)-dione	-4.812	-4.812	pH 7	-2.84	-4.85	-55.99	5	8	1	132	284.321	4
82	ZINC63869110	3-[(4-hydroxybenzoyl)amino]-4-(2-methylpiperidin-1-yl)benzamide	-4.806	-4.806	pH 7	2.44	3.88	-18.09	4	6	0	96	353.422	4
83	ZINC71479700	5-[3-hydroxy-5-(2-piperidin-2-ylethyl)phenyl]nicotinamide	-4.804	-4.804	pH 7	2.29	3.3	-54.04	5	5	1	93	326.42	5
84	ZINC21984991	5-(aminomethyl)	-4.800	-4.800	pH 7	-0.34	0.41	-50.2	3	3	1	58	137.162	2

		pyridine-3-carbaldehyde												
85	ZINC40535547	4-acetyl-3,5-dimethyl-N-[2-(pyrimidin-2-ylamino)ethyl]-1H-pyrrole-2-carboxamide	-4.800	-4.800	pH 7	0.79	4.89	-16.64	3	7	0	100	301.35	6
86	ZINC21953304	(4,5-dimethyl-1H-benzimidazol-2-yl)methylamine	-4.795	-4.795	pH 7	1.52	2.47	-49.64	4	3	1	56	176.243	1
87	ZINC55246041	4-[[[(2R)-2-methyl-1-piperidyl]sulfonylamino]methyl]tetrahydropyran-4-carboxamide	-4.791	-4.791	pH 7	-0.10	-0.99	-14.05	3	7	0	102	319.427	5
88	ZINC65341632	2-chloro-5-(3-cyano-5-hydroxy-phenyl)benzamide	-4.791	-4.791	pH 7	2.36	2.75	-14.41	3	4	0	87	272.691	2
89	ZINC06657470	N-[4-(1H-benzoimidazol-2-yl)phenyl]cyclobutanecarboxamide	-4.781	-4.781	pH 7	3.77	8.03	-15.71	2	4	0	58	291.354	3
90	ZINC67770325	6-(1-benzofuran-2-yl)-1,3-benzodioxole-5-carboxamide	-4.770	-4.770	pH 7	2.82	5.3	-13.51	2	5	0	75	281.267	2
91	ZINC25491340	N-(2,4,6-triaminopyrimidin-5-yl)benzamide	-4.761	-4.761	pH 7	0.07	1.81	-16.19	7	7	0	133	244.258	2
92	ZINC03838769	N-(1-carbamoylethyl)-3-(cyclohexylcarbamoylamino)-4,5-	-4.757	-4.757	pH 7	-0.33	-11.97	-18.34	7	9	0	153	368.434	5

		dihydroxy-cyclohexene-1-carboxamide												
93	ZINC00162531	3-(3-chlorophenyl)-2-(2-pyridyl)oxirane-2-carboxamide	-4.748	-4.748	pH 7	1.61	-3.73	-11.07	2	4	0	68	274.707	3
94	ZINC40285023	N-(5-chloro-2-hydroxyphenyl)-2-oxo-2-(3-oxopiperazin-1-yl)acetamide	-4.747	-4.747	pH 7	0.32	-1.15	-11.46	3	7	0	99	297.698	2
95	ZINC05347659	2-((7-chloro-1-hydroxy-9-oxo-9H-xanthen-3-yl)oxy)acetamide	-4.737	-4.737	pH 7	2.70	-2.33	-17.03	3	6	0	102	319.7	3
96	ZINC00542681	1-(2-hydroxy-2-(1H-indol-3-yl)ethyl)piperidine-4-carboxamide	-4.735	-4.735	pH 7	0.81	-4.73	-47.35	5	5	1	83	288.371	4
97	ZINC39351821	3-N-Cbz-amino-2,6-Dioxo-piperidine	-4.731	-4.731	pH 7	-1.08	-0.8	-48.38	3	4	1	63	143.166	1
98	ZINC08690210	2-[[5-(5-methyl-1H-benzoimidazol-2-yl)-2-pyridyl]sulfanyl]-N-(4-methylthiazol-2-yl)-acetamide	-4.716	-4.716	pH 7	3.92	8.08	-14.1	2	6	0	84	395.513	5
99	ZINC03838843	N-[3-(2-carbamoylethylcarbamoyl)-5,6-dihydroxy-1-	-4.714	-4.714	pH 7	-3.93	-15.24	-54.55	9	10	1	178	357.387	6

		cyclohex-2-enyl]-4-hydroxy-pyrrolidine-2-carboxamid												
100	ZINC71745814	4-{2-[5-(hydroxymethyl)-2-furyl]-1H-imidazol-1-yl}benzenesulfonamide	-4.710	-4.710	pH 7	0.80	0.46	-17.59	3	7	0	111	319.342	4
101	ZINC69457068	(2S)-1-[(1-phenylpyrazol-3-yl)methyl]pyrrolidine-2-carboxamide	-4.709	-4.709	pH 7	1.39	5.06	-37.41	3	5	1	65	271.344	4
102	ZINC05631131	2-Deoxy-D-ribose	-4.708	-4.708	pH 7	-1.52	-7.88	-7.54	3	4	0	69	134.131	0
103	ZINC20357536	2-((4-amino-7-methyl-5,6,7,8-tetrahydropyrido[4',3':4,5]thieno[2,3-d]pyrimidin-2-yl)thio)acetamide	-4.702	-4.702	pH 7	1.12	0.26	-18.61	4	6	0	98	309.42	3
104	ZINC40543261	2-[3-cyclopropyl-5-[(1R)-tetralin-1-yl]sulfanyl-1,2,4-triazol-4-yl]acetamide	-4.701	-4.701	pH 7	1.93	8.74	-18.54	2	5	0	74	328.441	5
105	ZINC67674831	(2S)-1-(3-{[(1,3-dimethyl-1H-1,2,4-triazol-5-yl)amino]carbonyl}benzyl)pyrrolidine-2-carboxamide	-4.699	-4.699	pH 7	0.22	7.17	-54.75	4	8	1	107	343.411	5
106	ZINC40817262	(3,4-dihydroxyphenyl)-[(3R)-3-([1,2,4]triazol	-4.692	-4.692	pH 7	0.87	5.2	-17.51	2	7	0	91	338.367	2

		o[4,3-a]pyridin-3-yl)-1-piperidyl]methanone												
107	ZINC00506647	3-((4-ethylphenyl)amino)-6,7-dimethoxyisobenzofuran-1(3H)-one	-4.686	-4.686	pH 7	3.74	0.31	-13.01	1	5	0	56	313.353	5
108	ZINC18068845	1-((1E)-2-{2-[(1E)-2-(2-hydroxynaphthyl)-1-azavinyl]cyclohexyl}-2-azavinyl)naphthalen-2-ol	-4.682	-4.682	pH 7	6.82	11.32	-103.04	4	4	2	68	424.544	4
109	ZINC65341631	2-chloro-5-(4-cyano-3-hydroxy-phenyl)benzamide	-4.675	-4.675	pH 7	2.83	2.79	-16.17	3	4	0	87	272.691	2
110	ZINC19851028	Ethyl 1-[4-(aminomethyl)pyridin-2-yl]piperidine-4-carboxylate	-4.671	-4.671	pH 7	1.29	4.72	-54.49	3	5	1	70	264.349	5
111	ZINC50274403	(3S)-2-[2-(4-methoxyphenyl)ethyl]-3,4-dihydro-1H-isoquinoline-3-carboxamide	-4.662	-4.662	pH 7	2.88	4.92	-10.07	2	4	0	56	310.397	5
112	ZINC02361465	1,3-diamino-1,3-dioxopropan-2-yl diethylcarbamodithioate	-4.660	-4.660	pH 7	-0.52	-1	-13.93	4	5	0	89	249.361	7
113	ZINC23478381	3-({4-[4-(aminomethyl)-1H-1,2,3-triazol-1-yl]-1-piperidinyl}methyl)-7-	-4.659	-4.659	pH 7	0.67	5.41	-105.53	5	7	2	96	354.458	4

		methyl-2(1H)-quinolinone												
114	ZINC69563860	2-chloro-N-[4-[[(1S)-1-(hydroxymethyl)propyl]carbamoyl]phenyl]pyridine-4-carboxamide	-4.659	-4.659	pH 7.	1.99	2.83	-14.88	3	6	0	91	347.802	6
115	ZINC71941455	3-[(3-chloro-4-ethoxy-phenyl)carbamoylamino]-2,2-dimethyl-propanamide	-4.645	-4.645	pH 7	2.37	1.89	-17.29	4	6	0	93	313.785	6
116	ZINC71747698	4-[4-(isopropylamino)-5,6,7,8-tetrahydropyrido[3,4-d]pyrimidin-2-yl]benzamide	-4.642	-4.642	pH 7	1.58	4.27	-53	5	6	1	98	312.397	4
117	ZINC63869112	3-[(4-hydroxybenzoyl)amino]-4-(2-methylpiperidin-1-yl)benzamide	-4.641	-4.641	pH 7	2.44	3.84	-18	4	6	0	96	353.422	4
118	ZINC16886381	(2S)-2-[2-[(3-aminophenyl)amino]-4-methyl-6-oxo-1H-pyrimidin-5-yl]propanoic	-4.637	-4.637	pH 7	0.94	5.85	-62.03	4	7	-1	124	287.299	4
119	ZINC03649942	(3S)-3,6,9-trihydroxy-8-methoxy-3-methyl-2,4-dihydroanthracen-1-one	-4.637	-4.637	pH 7	1.88	1.35	-11.66	3	5	0	87	288.299	1
120	ZINC71853644	N-(3,5-dimethyl-1-phenyl-pyrazol-4-yl)-2-ethyl-5-methyl-pyrazole-3-	-4.637	-4.637	pH 7	2.21	8.21	-14.59	1	6	0	65	323.4	4

		carboxamide												
121	ZINC36395220	N-[2-(2-amino-2-oxo-ethyl)phenyl]-3-fluoro-4-methyl-benzamide	-4.637	-4.637	pH 7	2.23	4.85	-13.62	3	4	0	72	286.306	4
122	ZINC17080350	(3-iminobenzo[f]chromen-2-yl)-N-(4-sulfamoylphenyl)carboxamide	-4.622	-4.622	pH 7	1.59	1.78	-43.63	5	7	1	128	394.432	3
123	ZINC55246042	4-[[[(2S)-2-methyl-1-piperidyl]sulfonylamino]methyl]tetrahydropyran-4-carboxamide	-4.617	-4.617	pH 7	-0.10	-0.97	-14.79	3	7	0	102	319.427	5
124	ZINC03838849	3-(3-aminopropanoylamino)-N-(1-carbamoyl-3-methyl-butyl)-4,5-dihydroxy-cyclohexene-1-carboxamide	-4.615	-4.615	pH 7	-2.18	-12.06	-61.2	9	9	1	169	357.431	8
125	ZINC04010946	4-chlorophenyl hexopyranoside	-4.615	-4.615	pH 7	0.35	-8.64	-9.73	4	6	0	99	290.699	3
126	ZINC42789676	2-[[(E)-3-(1-naphthyl)prop-2-enoyl]amino]propanediamide	-4.615	-4.615	pH 7	1.32	0.01	-18.23	5	6	0	115	297.314	5