Identification of novel thymidylate synthase therapeutics for treatment of H. pylori
Sharon Priya Alexander¹, Charan Kumar M², Shaik Parveen³, Natarajan Pradeep⁴ and Amineni Umamaheswari⁵

ABSTRACT: Helicobacter pylori is a helix-shaped, Gram-negative, urease-positive, curved, microaerophilic bacterium found in the stomach, which majorly causes gastric ulcers, stomach cancer, stomach inflammation (gastritis). H. pylori depict multi drug resistance to antibiotics such as clarithromycin, metronidazole, tetracycline, amoxicillin and furazolidone. The drug resistance of pathogen to the existing drug molecules necessitates the implementation of alternative strategy through in silico techniques. sRNAs are non-coding, small RNAs that regulate the metabolic function for forming metabolic complexes in the bacteria. Ten sRNA candidates were predicted in H. pylori using sRNAPredict. Thymidylate synthase (thyX) is an enzyme non-homologous to Homo monophosphate (dTMP). It has critical role in various biological activities such as DNA damage, DNA replication, and pyrimidine metabolism necessary for proliferation of H. pylori. Inhibitors having 1040 structural analogs. Library of thyX inhibitors were subjected to sRNA homologous to Homo monophosphate (dTMP). It has critical role in various biological activities such as DNA damage, DNA replication, and pyrimidine metabolism leads to growth, multiplication of the pathogens and hence, it was selected as putative drug target for H. pylori. Using PHASE v.3.8, four existing inhibitors were subjected to geometry based similarity search against in-house library of one million small molecules, which resulted a library of thyX inhibitors having 1040 structural analogs. Library of thyX inhibitors was considered for rigid receptor docking (RRD) studies through Maestro v.9.6. Tertiary structure of thyX was retrieved from the Protein databank (PDB) with ID: 3AH5, co-crystallized with FAD. The residues present within 4 Å around the FAD were defined as allosteric site residues and a grid of 10x10x10 Å was generated. RRD of Glide v.6.3 consists of a three-tier docking pipeline with high throughput virtual screening (HTVS), standard precision (SP) and extra precision (XP) docking to selectively filter the ligands at every stage from lower stringency to higher stringency, which disclosed 15 leads. Binding free energy calculations were performed to the resulted 15 lead complexes using Prime-MM/GBSA. On comparison, five best leads were found to have least docking scores and binding free energies than the existing inhibitors considered in the present study. Lead1 showed least docking score with XP G Score as -10.66 kcal/mol and better binding free energy (ΔG) as -117.20 kcal/mol with better binding orientation towards thyX. As the five proposed leads are found to obey the ADMET properties and having better binding affinity with thyX, these five leads were adequate to hinder the biological activity of thyX. Thus, hindering the thyX activity decreases the DNA replication, pyrimidine metabolism necessary for proliferation of H. pylori, in turn reduces the progression of pathogenesis leading to stomach cancer.

Keywords: Helicobacter pylori, sRNAPredict, ADMET, Thymidylate synthase, dUMP, dTMP, docking, binding free energy.

Introduction
Helicobacter pylori (campylobacter pylori) [1] is a Gram-negative, urease-positive, microaerophilic, non-fermenting, single polar flagellum bacterium found in the human gastric mucosa and stomach. H. pylori are a helix/spiral shaped pathogenic bacterium which is about 3µm in length. H. pylori infection is usually acquired from intake of contaminated food/water and spreads through person to person. The infection is commonly observed in crowded living conditions with poor sanitation. Study of the H. pylori is centered on attempts to understand pathogenesis, the ability of this organism to cause diseases. The analysis also given approximately 29% of the loci is in the “pathogenesis” category of the genome data base.

Bioinformatics Centre, Department of Bioinformatics, Sri Venkateswara Institute of Medical Science University, Tirupati, Andhra Pradesh – 517507; Ph: +91-877-2287727 Email: svims.btisnet@nic.in

The infections of H. Pylori are also linked with chronic antral gastritis and peptic ulceration [2] and also causes duodenal/gastric ulcer disease, chronic superficial gastritis, gastric adenocarcinoma, and mucosa-associated tissue lymphoma (MALT) and primary B-cell gastric lymphoma, gastric adenocarcinoma, which is a type of stomach cancer. Four clinical H. pylori isolates with high-level resistance to ß-lactams exhibited low to moderate level resistance to the structurally and functionally unrelated antibiotics ciprofloxacin, chloramphenicol, metronidazole, rifampin, and tetracycline (triple therapy). This pattern of multidrug resistance was transferable to susceptible H. pylori by natural transformation using naked genomic DNA from a clinical multidrug-resistant isolate. Acquisition of the multidrug resistance was also associated with a change in the genotype of the transformed multidrug-resistant H. pylori. But this triple therapy treatment is high cost and every seven days one time taken, repeated for 3 times and also has some side effects like vomiting, headache and skin rash. Hence with an aim to design a novel inhibitor against this infectious pathogen, a number of bioinformatics tools have been developed in this current work. In recent years, the expression of small RNAs (sRNAs) has been found to be widespread among bacteria. These regulatory RNAs play a major role in the gene expression and also essential for their diversified roles. Even in the best-studied bacterial transcriptomes, identities and functions of sRNAs are not fully understood. Recently, research has been started to find

Copyright-IJCSME

DOI: 10.18645/IJCSME SPB.0005
the role and importance of sRNAs in Gram-negative pathogens such as Salmonella typhimurium and Pseudomonas aeruginosa [3]. S. typhimurium has showed genetic islands showed the host induced expression in macrophages and thus contributed to virulence [4]. While experimental methods are critical for functional characterization of sRNAs, computational methods for prediction of sRNAs, owing to their efficiency, can be a useful complement to experimental approaches. Multidrug resistance nature of H. Pylori and adverse effects of existing treatment paved a new challenge for designing novel lead molecules for treating the H. Pylori mediated stomach cancers. In the present study, sRNA candidates were predicted from the H. Pylori genome and subjected to non-homology followed by metabolic pathway analysis. Systematic protocol was followed to target the defined thymidylate synthase (thyX) for designing inhibitor against the pathogen H. Pylori.

**Material and methods**

**1. sRNA prediction**

Available whole genome sequence of H. pylori ATCC 700392/26695 was retrieved from the National center for Biotechnology information (NCBI). sRNA candidates present in genome sequence of H. pylori were identified using sRNAPredict tool which rapidly identifies the putative intergenicRNA candidates [6]. sRNAPredict tool works on co-ordinate-based algorithm which incorporates the predictive features of sRNA’s at their relevant positions [7]. Comparative metabolic pathway analysis was performed to the predicted sRNA candidates [8, 9].

**2. Non-homology analysis**

The predicted sRNA candidates were further subjected to BLAST-P with parameters such as threshold e-value as 0.0001 and bit score cutoff as 100 against Homo sapiens, so as to define the non-homology with the host [10].

**3. Metabolic pathway analysis**

In order to trace the critical role of the targets in the survival of the pathogen, the identified sRNA sequences were subjected to metabolic pathway analysis [11] using KEGG pathway database [12]. Pathogen specific (Unique) enzymes that are non-homologous to humans were further characterized and verified their role for the presence of alternative pathway for the synthesis of the product and essential for the survival of the pathogen in the host.

**4. Protein structure preparation**

Co-crystal structure of thymidylate synthase (thyX) crystallized with FAD and 2'-Deoxy Uridine 5'- monophosphate was retrieved from the protein database. The protein structure was prepared using protein preparation wizard of Schrodinger by assigning bond orders to hydrogen, recreating the disulfide bonds and by filling up with selenomethionine to methionine. Missing side chains and loops in the structure was filled, further subjected to optimization of structure using OPLS_2005 force field at neutral pH. [13] In the present study, for inhibiting the thyX, allosteric site residues were targeted and was defined by the co-factor interactions. These residues were cross validated using PDBsum.

**5. ASINEX database**

Four reported inhibitors such as Rapamycin [14], fluourourasil [15] Raltitrexed [16] pemtrexed [17] and two co-crystal ligands (co-factor (FAD) and the substrate (UMP)) were considered as ligand dataset for generating the thyX inhibitors library. The ligands were subjected to shape based similarity screening against the prepared ASINEX 3D platinum database having 4.5 lakhs of small molecules. The obtained thyX inhibitor library was optimized using OPLS-AA_2005 force field using conjugate gradient algorithm method. Optimized thyX inhibitors library was then prepared using LigPrep and Epik of Maestro v.9.6. Reactive filters and Lipinski filters were applied to refine the generated tautomers [18]. The ligands disobeying the Lipinski’s rule of five were removed [19].

**6. High throughput virtual screening**

A grid of 10x10x10 Å was generated around the centroid of the allosteric site residues. Three tier rigid receptor docking (RRD) was performed with the prepared library of thyX inhibitors towards the receptor grid using Glide v.6.0 module of Maestro v.9.6 [13]. Three tier docking was implemented for the refinement based on the position, binding orientation and conformations of the ligands. RRD consists of HTVS (high throughput virtual screening) and SP (standard precision) followed by XP (extra precision) docking by eliminating false positives with the increasing stringency in order to get the best leads favorable binding orientation and binding affinity [13].

**7. Free binding energy calculations**

The binding free energy (ΔG) of thyX-lead complexes were calculated using molecular mechanics/generalized Born surface area (MM/GBSA) by Prime approach possibly brought closer to the experimental value by adding the entropic contributions to the binding free energy (ΔG). The ΔG calculations are much more precise than the XPG Score [20]. ΔG calculations were given as: ΔGbind = ΔE + ΔGsolv+ ΔGSA

Where, ΔE: minimized energies; ΔGsolv: solvation free energies; ΔGSA: difference in surface area energy of the thyX-lead complex and sum of the surface energies of thyX and leads respectively.

Using MM-GBSA approach with OPLS_2005 along with molecular energies (EMM), polar solvation through surface generalized Born solvation model (GSGB) and a nonpolar solvation term (GNP) composed of nonpolar solvent accessible and vdW interactions were also calculated.
8. ADME/T screening
The best leads were further subjected to analyze the pharmacological properties. Absorption, distribution, metabolism, excretion / toxicity (ADME/T) screening of obtained the best ranked leads were carried out using QikProp.v4.7 module of Schrodinger [13]. QikProp v.4.7 predicts the different principal descriptors and pharmaceutically relevant 44 properties along with a detailed analysis molecular weight, number of rotatable bonds, number of hydrogen bond acceptors, number of hydrogen bond donors, log S, ClogS, PCaco, log BB, log KP values, globularity, QP%, log HERG, Lipinski’s rule of five and Jorgensen rule of three were predicted through QikProp module of Schrodinger 2014 for the best leads.

Results and discussion
1. Prediction of sRNA’s
The H. pylori genome sequence is of 1.67 Mb size and comprising of 38.9% GC content, 1,563 genes and 1,469 proteins. The 1,469 proteins obtained from NCBI were further subjected to sRNA predict tool and obtained ten sRNA candidates. Among the 10 sRNA candidates, five were proteins and remaining five were enzymes. Enzymes are biologically significant than non-enzymes as they catalyze numerous biochemical reactions that aids for the survival of the pathogens. As the active site and the allosteric sites of the enzymes were predefined and thus enzymes aids in designing of novel inhibitors specific to the target.

2. Non-homology analysis
Present drug discovery and development focused to identify and optimizing the drug candidates that may be act through the inhibition of specific enzyme targets. In order to avoid the unwanted cross reactivity towards the host (Homo sapiens), the five enzymes were analyzed for the non-homology of the potential drug. Among the five enzymes, thymidylate synthase (thyX) is pathogen specific or unique and non-homologous to the host, (Homo sapiens) and these five enzymes were subjected to metabolic pathway analysis.

3. KEGG analysis
The sRNA candidates were subjected to KEGG analysis such that the drug target should be pathogen specific. The metabolic pathways of the five enzymes were analyzed and compared to the host metabolic pathways revealed that, many pathogen specific metabolic pathways were absent in host but four enzymes were having alternative pathways. Those pathways that were absent in the host and present in the pathogens were selected as unique pathways. The sRNA candidates having alternative pathways to synthesis the product were not considered, because blocking of these drug targets would be ineffective as the product is synthesized by alternate way. Thymidylate synthase (thyX) was found to be unique to pathogen with no alternative pathway to form the product thymidine. ThyX was found to be crucial for the survival of the pathogen as it is essential for the DNA replication, DNA repair mechanism along with the pyrimidine metabolism of the pathogen. Thus thyX was selected as potential drug target against H. Pylori.

4.3D structure and allosteric site analysis
To design an inhibitor towards thyX, 3D structure is necessary to check the structural complementarity with better binding orientation and favorable conformation. Tertiary structure of thyX was retrieved from the Protein databank3AH5) as homo 6-mer having six chains namely B, C, D, E and F with chain length of 207 amino acids and chain A of 3AH5 has 208 amino acids in length [20]. ThyX was co-crystallized with the co-factor flavin-adenine dinucleotide (FAD) and substrate 2’-Deoxyuridin 5’-monophosphate (UMP) at 2.5 Å resolution.

Results present around 4 Å region round the co-factor FAD was defined as allosteric site residues (Fig. 1A), as co-factor binding sites are crucial in the reaction that are tightly bound to the enzyme for the synthesis of product. Thus, for inhibition studies these allosteric site residues were targeted for hindering the activity of thyX and these residues were cross checked with PDBsum. Charged / polar amino acids such as Arg {(Arg–70 of B chain), (Arg–97 of A chain), (Arg–99 of A,B,C chains), (Arg–188 of A,B chains) and (Arg–197 of A chain)}, Glu–76 of A,B chains, Ser–189 of A,B chains, His–98 of A chain and Asn–192 of A chain; Hydrophobic amino acids such as, Phe–69 of B chain, Ile–100 of A,B,C chains and Leu–196 of A chain were found to be present in the allosteric site.

5. Virtual screening
Structurally similar compounds exhibits similar activity, hence structurally and geometrically similar compounds were retrieved. Shape based and geometry based similarity screening was performed to the four published inhibitors and two co-crystal ligands using PHASE module of Schrodinger against the prepared ASINEX platinum 3D database. The retrieved hits were subjected to preparation using LigPrep module of Schrodinger. Conformers obtained from post LigPrep were further subjected to Epik analysis. To avoid the false positives, all the conformers were subjected to reactive filters and Lipinski’s filters to generate thyX inhibitors library of 1040 compounds.

6. Docking studies and free energy calculations
To analyze the molecular interactions between thyX and ligand compounds from the library of thyX inhibitors were docked in to thyX grid. 1040 ligands were docked into thyX (3AH5) grid generated around the allosteric site residues resulted 200 compounds from HTVS and 51 compounds from SP docking. Further filtering with XP docking with higher stringency resulted 15 ligands. The
obtained 15 ligands were compared to the existing inhibitors, which revealed five the best leads (TABLE 1) and were reconsidered for binding free energy calculations. Analyzing the docking results revealed the molecular interactions of the leads with the receptor thyX with better binding affinity and better binding orientation. The leads were having similar binding orientation and high binding affinity towards thyX of H. Pylori to form a stable complex and competitively block the allosteric site of thyX and stop the activity of thyX. Lead1 showed least XPG score of -10.66 kcal/mol and binding free energy (ΔG) of -117.20 kcal/mol. Analysis of thyX-lead1 complex revealed that, chlorobenzene moiety of lead1 formed of π-π stacking and π-Cation interactions with allosteric residue Arg-188 of A chain of thyX. Chloro-ethoxy phenol hydrate moiety of lead1 involved in π-π stacking with B chain of thyX. Side chain atoms of allosteric site residue Arg-188 of B chain was involved in hydrogen bond with oxygen moiety of sulfanylethyl acetate moiety of lead1 (Fig. 1B). Remaining residues such as Arg-70 of B chain, Arg-97 of A chain, Arg-99 of A,B,C chains and Arg-197 of A chain, Glu-76 of A,B chains, Ser-189 of A,B chains, His-98 of A chain and Asn-192 of A chain along with the hydrophilic residues such as, Phe-69 of B chain, Ile-100 of A,B,C chains and Leu-196 of A chain were involved in van der Waals interactions to stabilize the thyX-lead1 dock complex.

The proposed leads were further subjected to analyze the pharmacological properties. Principal descriptors (TABLE 2) and pharmaceutically relevant properties (TABLE 3) such as Molecular volume, LogP o/w, SASA, IP (ev) and Log Khsa etc., were analyzed. Molecular volume defines the total solvent-accessible volume in cubic angstroms using a probe with a 1.4 Å radius, high solubility reduces the binding affinity. Log P for octane to water (LogP o/w) measures the hydrophobicity of a compound that was measured by its distribution coefficient as a major determinant of how the drug must pass specifically through lipid bilayers for the transcellular transport. Solvent-accessible surface area (SASA) is the surface area of a biomolecule that can be accessible by a solvent, often used to calculate the free energy transfer required to move a drug molecule from aqueous solvents to a non-polar solvents. IP (ev), calculated ionization potential of the drug molecule. Log Khsa defines the drug binding affinity towards human serum albumin for the distribution of the drug molecule [13]. The overall ADMET properties of proposed five leads were well within the normal ranges without any violations and biologically active without any toxic functional groups [13].

Table: 1. XPG Score and Binding free energies of proposed leads and published inhibitors.
Table 2: Principle descriptors of proposed inhibitors of thyX and published inhibitors.

<table>
<thead>
<tr>
<th>Compound</th>
<th>MW</th>
<th>Rot</th>
<th>SA</th>
<th>FO</th>
<th>W</th>
<th>PI</th>
<th>Vol</th>
<th>Do</th>
<th>Ac</th>
<th>GL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lead1</td>
<td>57.6</td>
<td>87</td>
<td>6.0</td>
<td>40</td>
<td>8.0</td>
<td>16</td>
<td>1.8</td>
<td>19</td>
<td>4.4</td>
<td>4.2</td>
</tr>
<tr>
<td>Lead2</td>
<td>56.2</td>
<td>83</td>
<td>3.0</td>
<td>40</td>
<td>8.0</td>
<td>17</td>
<td>1.1</td>
<td>21</td>
<td>0.4</td>
<td>4.2</td>
</tr>
<tr>
<td>Lead3</td>
<td>46.2</td>
<td>77</td>
<td>1.1</td>
<td>9</td>
<td>21</td>
<td>0.6</td>
<td>14</td>
<td>34</td>
<td>4.7</td>
<td>10</td>
</tr>
<tr>
<td>Lead4</td>
<td>40.8</td>
<td>66</td>
<td>5.2</td>
<td>25</td>
<td>4.8</td>
<td>19</td>
<td>4.6</td>
<td>28</td>
<td>9.2</td>
<td>9.2</td>
</tr>
<tr>
<td>Lead5</td>
<td>42.4</td>
<td>71</td>
<td>2.1</td>
<td>25</td>
<td>3.7</td>
<td>29</td>
<td>1.3</td>
<td>12</td>
<td>46</td>
<td>1.1</td>
</tr>
</tbody>
</table>

Range 95% of drugs.

MW, molecular weight = (130.0/725.0); Rot, No. of rotatable bonds = (0.0/15.0), SASA, total solvent accessible surface area = (300.0/1000.0), FOSA, hydrophobic solvent accessible surface area = (0.0/750.0), PISA, carbon pi solvent accessible surface area = (0.0/450.0), WPISA, weakly polar solvent accessible surface area = (0.0/175.0), volume, molecular volume (A3) = (500.0/2000.0), Donor, donor – hydrogen bonds = (0.0/6.0), Accept HB, acceptor – hydrogen bonds = (2.0/20.0), IP (eV), ionization potential = (7.9/10.5), EA (eV), electron affinity = (-0.9/1.7), Glob, globularity = (0.75/0.95).

Range 95% of drugs.

LogP o/w, log P for octavo/water = (-2.0/6.5); Logs, log S for aqueous solubility = (-6.5/0.5), CIlogS, log S – conformation independent = (-6.5/0.5); LogBB, log BB for brain/blood = (-3.0/1.2), Log KP, log KP for skin permeability = (KP in cm/h); Log Kh, log K hsa serum protein binding = (-2.5/1.5); Lipinski rule of five violations = (maximum is 4); Jorgensen rule of three violations = (maximum is 3).

The proposed leads showed better binding affinity with favorable binding orientation in forming a stable complex with no violations were observed for Lipinski’s rule of five and observed to possess beneficial pharmacological properties against thyX of H. Pylori. The conversion of dUMP to dTMP was catalyzed by thyX, which is necessary for thymine synthesis required for DNA biosynthesis and DNA repair mechanism [21, 22]. The proposed five leads were enough to bind thyX, which hinders the thymine synthesis by blocking the allosteric site competitively and exerts an allosteric effect to change the shape of the thyX binding site. Thus, in turn results in hindering the DNA multiplication or DNA repair mechanism mediated by thyX of pathogen H. Pylori. Thus by hampering the pyrimidine synthesis, proliferation and multiplication of the pathogen can be reduced by treating with the proposed leads against H. Pylori mediated stomach cancers.

Table 3: ADME properties of proposed inhibitors thyX and published inhibitors.

<table>
<thead>
<tr>
<th>Compound</th>
<th>LogP o/w</th>
<th>Log S</th>
<th>CI logS</th>
<th>Log BB</th>
<th>Rule of 5</th>
<th>Rule of 3</th>
<th>Log KP</th>
<th>Log Kh</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lead1</td>
<td>1.79</td>
<td>9.73</td>
<td>10.3</td>
<td>-0.78</td>
<td>2</td>
<td>2</td>
<td>-1.95</td>
<td>5.61</td>
</tr>
<tr>
<td>Lead2</td>
<td>1.8</td>
<td>10.2</td>
<td>10.0</td>
<td>0.68</td>
<td>2</td>
<td>2</td>
<td>1.88</td>
<td>4.97</td>
</tr>
<tr>
<td>Lead3</td>
<td>1.43</td>
<td>-</td>
<td>-</td>
<td>-1.86</td>
<td>0</td>
<td>1</td>
<td>0.022</td>
<td>-</td>
</tr>
<tr>
<td>Lead4</td>
<td>1.40</td>
<td>4.12</td>
<td>4.71</td>
<td>2.12</td>
<td>0</td>
<td>0</td>
<td>-0.193</td>
<td>-</td>
</tr>
<tr>
<td>Lead5</td>
<td>1.18</td>
<td>4.20</td>
<td>4.09</td>
<td>-0.20</td>
<td>0</td>
<td>0</td>
<td>-0.244</td>
<td>-</td>
</tr>
<tr>
<td>Rapa piricin</td>
<td>2.05</td>
<td>6.89</td>
<td>7.23</td>
<td>-3.01</td>
<td>3</td>
<td>0</td>
<td>-2.77</td>
<td>-</td>
</tr>
<tr>
<td>Fluor ouaci l</td>
<td>1.56</td>
<td>0.87</td>
<td>0.83</td>
<td>2.15</td>
<td>2</td>
<td>1</td>
<td>4.98</td>
<td>1.786</td>
</tr>
<tr>
<td>Kahlitr ened</td>
<td>1.89</td>
<td>6.43</td>
<td>6.43</td>
<td>1.45</td>
<td>0</td>
<td>0</td>
<td>5.78</td>
<td>5.32</td>
</tr>
<tr>
<td>Pentet rased</td>
<td>2.42</td>
<td>7.29</td>
<td>7.29</td>
<td>3.78</td>
<td>3</td>
<td>0</td>
<td>1.32</td>
<td>-1.245</td>
</tr>
</tbody>
</table>
pathway analysis to define thyX as a drug target against H. Pylori. ThyX plays a major role in thymine synthesis by catalyzing the conversion of dUMP to dTMP required for DNA replication as well as DNA repair mechanism. Comparative metabolic pathway analysis revealed the absence of alternative pathway, which makes thyX as a potential drug target against H. Pylori. Based on the rigid receptor docking and binding free energy analysis by screening 1040 compounds, revealed five the best leads. These five leads are having better binding affinity and binding free energy than the existing inhibitors. Thus proposed five leads are having favorable ADMET properties and enough to hinder the pyrimidine synthesis of the pathogen H. Pylori without interfering the host’s system. Thus the proposed five leads would bring new possibilities in developing potential drug moieties in treating the H. Pylori mediated infections that leads to stomach cancers.

Acknowledgment
The author SA is highly thankful for DBT, Ministry of Science and Technology, Govt. of India for the support and providing the necessary infrastructure facilities to carry out the research work at BIF (No.BT/BI/25/001/2006).

References

