CHAPTER IV

SUMMARY

4.1 INTRODUCTION

*Helicobacter pylori* (previously known as *Campylobacter pylori*) is about 3 µm long and 0.5 µm in diameter, microaerophilic (requires oxygen), neutralophilic (adapting to highly acidic environment) helix-shaped (curved rod), gram negative bacteria colonizing upper gastrointestinal tract.

*H. pylori* infection is coordinated in a cascade manner that helps it to colonize in the host. Colonization of bacterium starts by adapting itself to the harsh acidic environment in the stomach. *H. pylori* have the necessary machinery to neutralise the pH of its surroundings. It also has the ability to sense the pH of its surroundings and move towards the less acidic region. *H. pylori’s* next hurdle is gastric mucosal barrier in the stomach and it has the capability to overcome this gastric mucosal barrier. Once the gastric mucosal barrier is weakened, pathogen uses different adhesion molecules to adhere to the epithelial lining of the stomach. Pathogen then establishes interaction with the host using several toxins that indirectly leads to development of inflammation or gastritis. Prolonged inflammation damages epithelial cells leading to ulcers in the stomach. Genetic changes in the host cell due to *H. pylori* infection leads to development of gastric cancer. Thus *H. pylori* infection induces gastric inflammation, ulcer, and cancer.

Gastric cancer is the most common cancer in the world caused by *H. pylori* infection leading to damage of gastric mucosa and gastric glands. Drugs relieve from gastritis or pain but are not specific to *H. pylori*. Hence, there is a need for discovery of drug targets and drugs for *H. pylori*. The advances in silico drug discovery methods made drug target discovery efficient, inexpensive and less time consuming.

4.2 OBJECTIVE

An objective of the current study is to identify drug targets and drugs for *H. pylori* using various in silico approaches.
4.3 IDENTIFICATION OF DRUG TARGETS

Subtractive genome analysis, metabolic pathway analysis, pathogenic island analysis and meta-analysis of microarray data identified nearly one hundred twenty three potential drug targets in various metabolic categories. Current study categorized 26 drug targets as novel drug targets, whereas 97 drug targets are already experimentally validated lending credence to our approach and rest 41 targets are designated as a critical drug targets.

4.3.1 Subtractive Genome Analysis

Genes unique to *H. pylori* were filtered by comparing genomes between *H. pylori* and *Homosapien sapiens*; *H. pylori* and other *Helicobacter* species (*H. acinonychis, H. hepaticus, H. mustalae*); and 23 *H. pylori* strains using RAST. Bacterial genes which are non-homologous to humans and essential for pathogen are identified using BLASTp. These genes later were subjected to gene property analysis to identify 29 potential drug targets. Eleven genes of the 29 drug targets are already experimentally validated. This method enabled rapid identification of drug targets with possible therapeutic implications for gastric cancer.

4.3.2 Metabolic Pathway Analysis

Comparative metabolic pathway analysis resulted in the identification of complete and partial unique metabolic pathways. There are 35 metabolic pathways and around 3,178 reactions in sub class of 23 *H. pylori* strains. Comparative metabolic pathway analysis between *H. pylori* and *Homosapien sapiens*; *H. pylori* and other *Helicobacter* species (*H. acinonychis, H. hepaticus, H. mustalae*); and 23 *H. pylori* strains using MetaCyc resulted in 1084, 1783 and 34 complete unique metabolic pathways respectively; whereas 1489, 1018, and 30 partial unique metabolic pathways are identified respectively. Enzymes/proteins of complete and partial unique metabolic pathways were subjected to BLASTp to identify non-homologous proteins. These non-homologous enzymes/proteins are designated as potential drug targets. Metabolic reactions in the pathways revealed twelve metabolic pathways with forty two drug targets.
Comparative analysis of complete and partial metabolic reactions between *H. pylori* and *Homosapien sapiens* identified twelve and fifteen potential drug targets respectively. Comparative analysis of complete and partial metabolic reactions between *H. pylori* and *H. pylori* and other *Helicobacter* species (*H. acinonychis, H. hepaticus, H. mustalae*) identified five and eight potential drug targets respectively. Comparative analysis of complete and partial metabolic reactions among the *H. pylori* strains identified two and one potential drug targets respectively.

### 4.3.3 Pathogenic Islands Analysis

Island Pick, SIGI-HMM and Island Path methods of Islands viewer predicted 31 pathogenic islands with genes/virulence factors for 22 *H. pylori* strains. No pathogenic islands were detected in *H. pylori* strain HpSAT464. Nearly 642 genes/virulence factors associated with pathogenic islands were identified in 22 *H. pylori* strains. Of them 282 were known with known functions and rest 361 were hypothetical proteins. Analysis of 642 bacterial genes identified 101 genes which are non-homologous to humans and are essential for pathogen. Gene property analysis of 101 genes identified 31 potential drug targets.

### 4.3.4 Meta-analysis of the Microarray Data

Meta-analysis of the microarray data identified twenty one drug targets (non-homologous proteins) under metabolic categories such as motility and chemotaxis; two-component system; membrane transport; cell structures biosynthesis; stress response; amino acids biosynthesis; peptidoglycan biosynthesis; generation of precursor metabolites and energy; potassium metabolism; fatty acids, lipids, and isoprenoids; and virulence factors and other biosynthesis.

### 4.4 IDENTIFICATION OF HITS

Virtual screening of chemical libraries such as Drugbank, Therapeutic Target Database, Chembl and also mining of literature identified 63018 hits. Drugbank, TTD, Chembl and literature mining identified 142, 62839, 31, and 6 hits respectively. The following drug targets and drugs are shortlisted from 123 drug targets and 63018 hits.
**Drug Targets**

- D-alanine-D-alanine ligase B
- Molybdopterin-guanine dinucleotide biosynthesis protein MobA
- Shikimate dehydrogenase
- Thiamin phosphate pyrophosphorylase/hydroxethylthiazole kinase
- 4–diphosphocytidyl-2-c-methyl–D-erythritol kinase
- NADH ubiquinone oxidoreductase subunit
- UDP-N-acetylmuramoyl-L-alanine-D-glutamate ligase

**Hits/Ligands**

- Quercetin
- 6- N-hydroxylaminopurine (HAP)
- N4-hydroxycytidine
- p-Chloromercuribenzoate
- 4-fluoro-3,5-dihydroxy-4-methylpent-1-enylphosphonic acid
- Piericidin A
- Iodoacetate
- Curcumin
- Butyl-2-\{[3-(2-naphtholyxy)-4-oxo2-(trifluoromethyl)-4H-chormen-7-yl-]oxy\}propanoate
- Phosphinate

**4.5 HIT TO LEAD OPTIMIZATION**

Hit to lead optimization was carried out by screening for analogues related to eight hits mentioned above. PubChem identified 28 lead molecules. Lead optimization was carried out by docking studies, Lipinski rule of five and ADMET property analysis.

**4.6 DOCKING STUDIES**

Twenty eight analogues were docked with their respective drug targets. Drug target D-alanine-D-alanine ligase B was docked with five analogues. Based on the binding affinities and hydrogen bonding’s quercetin 7-glucoside (-8.3 Kcal/mol and 11 hydrogen bonds), rutin (-8.3 Kcal/mol and 12 hydrogen bonds), and querectrinion
(-8.5 Kcal/mol and 11 hydrogen bonds) proved to be the best analogues for the drug target D-alanine-D-alanine ligase B.

Drug target molybdopterin-guanine dinucleotide biosynthesis protein mobA was docked with three analogues. Based on the binding affinity and hydrogen bondings N4-hydroxycytidine; uridine (-7.1 Kcal/mol and 10 hydrogen bonds) proved to be the best analogue for the drug target molybdopterin-guanine dinucleotide biosynthesis protein mobA.

Drug target thiamin phosphate pyrophosphorylase / hydroxyl ethyl thiazole kinase was docked with two analogues. Based on the binding affinity and hydrogen bondings p-hydroxymercuribenzoate (-5.7 Kcal/mol and 4 hydrogen bonds) proved to be best analogue for the drug target thiamin phosphate pyrophosphorylase / hydroxyl ethylthiazole kinase.

Drug target 4–diphosphocytidyl-2-c-methyl–D-erythritol kinase is docked with [(1e,3r,4s)-4-fluoro-3,5-dihydroxy-4-methyl-pent-1-enyl]phosphonic acid. Based on the binding affinity and hydrogen bondings [(1e,3r,4s)-4-fluoro-3,5-dihydroxy-4-methyl-pent-1-enyl]phosphonic acid (-5.3 Kcal/mol and 9 hydrogen bonds) was proved to be best analogue for drug target 4–diphosphocytidyl-2-c-methyl–D-erythritol kinase.

Drug target NADH ubiquinone oxidoreductase subunit was docked with two analogues. Based on the binding affinity and hydrogen bondings glucopiericidin B (-6.4 Kcal/mol and 3 hydrogen bonds) was proved to be best analogue for drug target NADH ubiquinone oxidoreductase subunit.

Drug target shikimate dehydrogenase was docked with ten analogues. Based on the binding affinity and hydrogen bondings bisdemethoxy curcumin (-8.4 Kcal/mol and 6 hydrogen bonds) curcumin monoglucoside (-8.9 Kcal/mol and 15 hydrogen bonds) proved to be the best analogues for the drug target shikimate dehydrogenase.

Drug target UDP-N-acetylmuramoyl-L-alanine-D-glutamate ligase is docked with three analogues. Based on the binding affinity and hydrogen bondings RC4 (-5.7 Kcal/mol and 1 hydrogen bond) proved to be the best analogue for the drug target UDP-N-acetylmuramoyl-L-alanine-D-glutamate ligase. Docking studies helped to identify ten potential analogues from the pool of twenty eight analogues.
4.7 SCREENING OF LEAD MOLECULES USING LIPINSKI RULE OF FIVE

Lipinski rule of five estimated the permeability and solubility of leads. It filtered unfavorable leads and eliminated lead candidates with poor physicochemical properties. Lipophilicity (logP) of a compound is considered as one of the important parameter with respect to its biological pH values. Twenty eight lead molecules are screened based on mass, hydrogen bond donor, hydrogen bond acceptors, logP value, and molar refractivity.

Leads N4-hydroxycytidine, guanine (oxime), amorphigenin, [(1e,3r,4s)-4-fluoro-3, 5-di hydroxy-4-methyl-pent-1-eny] phosphonic acid, bisdemethoxycurcumin, 4-nitrophenyl iodoacetate, curcumin 4, 4'-diacetate, hexyl iodoacetate, tert-butyl 2-iodoacetate, iodoacetic acid hydroxysuccinimide ester, methyl bis (ethylenimido) phosphinate satisfied all five descriptors.

Leads 6-N-hydroxylaminopurine, p-hydroxymercuribenzoate, [18FP]-curcumin, 2, 3-Diiodo-2-buten-1,4-diol iodoacetate, 2-[(2-chloroethyl)(phenyl) amino] ethyl iodoacetate, 4-(benzyloxy)phenyl bis(aziridin-1-yl)phosphinate and RC4 satisfied four descriptors out of five.

4.8 SCREENING OF LEAD MOLECULES USING ADMET PROPERTIES

ADMET analysis was performed for twenty eight analogues. Based on the properties like absorption, distribution, metabolism, excretion and toxicity one analogue was identified in ADMET analysis. P-hydroxymercuribenzoate is the potential analogue according to ADMET analysis.

4.9 CONCLUSION

Extensive study helped to find out the best lead molecule P-hydroxymercuribenzoate. Docking studies, Lipinski rule of five and ADMET properties identified analogue p-hydroxymercuribenzoate as the best drug candidate. This potential drug candidate can further be subjected to invitro and invivo for therapeutic intervention against gastric cancer.