Chapter 2

Literature Review
2.1 *Helicobacter pylori*

*Helicobacter pylori* (*H. pylori*) are a Gram-negative microaerophilic, spiral shape ε-proteobacteria that specifically colonizes the gastric mucosa, and it is the most common bacterial infection worldwide. Typically they are acquired during childhood and the infection can persist in the gastric ecosystem throughout the life span of the host, if untreated (Everhart 2000).

**Nomenclature and Classification:**

The bacteria were initially named Campylobacter pyloridis, and then were renamed as *C. pylori* to fix a Latin grammar error. However, when 16S rRNA gene sequencing and other research showed in 1989 that the bacterium did not belong in the genus Campylobacter, therefore it was placed in to its own genus, Helicobacter. Here pylori of *H. pylori* specifies pyloric valve of human digestive track.

*Figure 2.1.1 Helicobacter pylori*

**Scientific Classification:**

- **Kingdom:** Bacteria
- **Phylum:** Proteobacteria
- **Class:** Epsilon-Proteobacter
- **Order:** Campylobacterales
- **Family:** Helicobacteraceae
- **Genus:** *Helicobacter*
- **Species:** *H. pylori*

**Binomial Name:** *Helicobacter pylori*
History

Dr. Robin Warren and Dr. Barry Marshall discovered *Helicobacter pylori* in 1983. Before that, the German scientists found spiral-shaped bacteria in the lining of the human stomach in 1875, but they were unable to culture it and the results were eventually forgotten. Evidences of same spiral-shaped bacteria were found in 1893 and 1899 in the stomach of dogs and humans respectively but this work had a very little impact.

Interest in understanding the role of bacteria in stomach diseases was rekindled in the 1970s with the visualization of bacteria in the stomach of gastric ulcer patients. They were observed in 1979 by Australian pathologist Robin Warren, who did further research on it with Australian physician Barry Marshall beginning in 1981. After numerous unsuccessful attempts at culturing the bacteria from the stomach, they finally succeeded in visualizing colonies in 1982 when they unintentionally left their Petri dishes incubating for 5 days over the Easter weekend. In their original paper, Warren and Marshall contended that most stomach ulcers and gastritis were caused by infection by this bacterium and not by stress or spicy food, it was assumed. To demonstrate that *H. pylori* caused gastritis, Marshall drank a beaker of *H. pylori*. He became ill several days later with nausea and vomiting. An endoscopy report ten days after inoculation revealed signs of gastritis and the presence of *H. pylori*. These results proved that *H. pylori* were the causative agent of gastritis. Marshall and Warren went on to show that antibiotics are effective in the treatment of many cases of gastritis. In 1987, the Sydney gastroenterologist Mr. Thomas Borody invented the first triple therapy for the treatment of duodenal ulcers. In 1994, the National Institutes of Health (USA) published an opinion stating that most recurrent duodenal and gastric ulcers were caused by *H. pylori* and recommended that antibiotics be included in the treatment regimen. Dr. Robin Warren and Dr. Barry Marshall was awarded the Nobel Prize in
**Medicine in year 2005** for their work on *Helicobacter pylori* (http://www.helico.com/).

### 2.2 *Helicobacter pylori* Genome

The first strain of *H. pylori* to be sequenced in 1997 was strain 26695. The chromosome of strain 26695 is circular and composed of 16, 67,867 base pairs. The average GC content is approximately 39%. In 1999, strain J99 was sequenced which was isolated from an American patient with a duodenal ulcer. Compared to strain 26695, it has a slightly smaller circular chromosome with 16, 43,831 base pair. The overall genomic organization, gene order and predicted proteomes of the sequenced strains were very similar. In 2006, a chronic atrophic gastritis *H. pylori* strain HPAG1 with 15, 96,366 base pair, was sequenced (Oh *et al.* 2006). Similar to the sequenced strains 26695 and J99, HPAG1 is a type-1 strain that contains CagA and a virulent allele of VacA (Enroth *et al.* 2000). The *H. pylori* strain G27 was sequenced recently (Baltrus *et al.* 2009). It was originally isolated from an Italian patient (Covacci *et al.* 1993) and has been used widely in *H. pylori* research. The G27 genome has a similar size to the other three sequenced strains. It is 16, 52,983 base pairs long and have a GC content of 38.9%. In addition, G27 also contains one 10032 base pairs AT rich (65.2%) plasmid resembling that found in strain HPAG1. The plasmid encodes 11 genes. There are similarities and the variations in the size of genomes several strains. The disease causing factor Pathogenic Island is present in all most all the strains but yet the strains cause different diseases.

### 2.3 General mechanism of *H. pylori* in Gastritis/Stomach Ulcer and Cancer

Colonization of *H. pylori* in the stomach causes chronic gastritis. The initial infection during the decades can remain silent, due to the dynamic equilibrium between the bacterium and its human host, or evolve into more severe
diseases, such as atrophic gastritis, peptic ulcer, lymphoma of the mucosa-associated lymphoid tissue or gastric adenocarcinoma (Zarrilli et al. 1999). The organism due to presence of its pathogenic-island causes chronic persistent and is of serious concern for developing countries. Studies indicate that infected individuals have 2 - 6 fold increased risk of developing gastric cancer and mucosal associated lymphoid tissue lymphoma compared to their uninfected counterparts (Dutta et al. 2006; Cendron and Zanotti 2011). Some clinical work in past at Japan suggests that *H. pylori* eradication reduces the risk of new gastric carcinomas in patients with a history of the disease (Kawai et al. 2011).

*Helicobacter pylori* mode of infection is still unknown, some believe this bacterium integrates like viruses and some do not. But after infection it colonizes in to mucus layer of host stomach (Figure 2.3.1). To colonize the stomach, they must survive the acidic pH of the lumen and burrow into the mucus to reach its niche, close to the stomach's epithelial cell layer. The bacterium has flagella to move through the stomach lumen to the mucus lining of the stomach. To avoid being carried into the lumen, *H. pylori* senses the pH gradient within the mucus layer by chemotaxis and swims away from the acidic contents of the lumen towards the more neutral pH environment of the epithelial cell surface.

*H. pylori* need basic pH to survive. They produce certain enzymes like adhesions and ureases, which helps them in above mechanism. Adhesins which bind to membrane-associated lipids and carbohydrates and help it adhere to epithelial cells. A bacterium also produces large amounts of the enzyme urease, which are localized inside and outside of the bacterium. The survival of *H. pylori* in the acidic stomach is dependent on urease, and it would eventually die without the enzyme. Urease breaks down urea to carbon dioxide and ammonia, which neutralizes gastric acid and helps bacteria to survive. The
ammonia that is produced is toxic to the epithelial cells and, along with the other products of \textit{H. pylori} including protease, Vacuolating cytotoxin A (VacA), and certain phospholipases damages epithelial cells and cause Ulcer (Figure 2.3.2).

![Colonization of H. pylori](image1)

**Figure 2.3.1 Colonization of \textit{H. pylori}**

The type IV secretion apparatus secreted by genes of cagPAI of \textit{H. pylori} also injects the cag PAI-encoded protein CagA into the stomach’s epithelial cells of host, where it disrupts the host cytoskeleton, its adherence of adjacent cells, intracellular signaling, and other cellular activities.

![Ulcer caused by H. pylori](image2)

**Figure 2.3.2 Ulcer caused by \textit{H. pylori}**
This colonization results in chronic gastritis, an inflammation of the stomach lining. The severity of the inflammation is likely to underlie *H. pylori*-related diseases. Duodenal and stomach ulcers result when the consequences of inflammation allow the acid and pepsin in the stomach lumen to overwhelm the mechanisms that protect the stomach and duodenal mucosa from these caustic substances. However chronic inflammation induced by the bacteria causes’ further reduction of acid production and, eventually, atrophies of the stomach lining, which may lead to Gastric Ulcer and increases the risk for Gastric Cancer in host (Hassell *et al.* 2001).

### 2.4 Essential Genes

The completion of human genome project and the completion of genome sequences of pathogenic bacteria have increased the momentum of field of drug discovery against threatening human pathogens. The sequencing of pathogenic bacteria has provided a lot amount of raw material for *in silico* analysis (Dutta *et al.* 2006). Identification of bacterial genes that are nonhomologous to human genes and important for the survival of bacteria is one of the promising means to identify novel drug targets. Availability of genome sequences of pathogens has provided a tremendous amount of information that can be useful in drug target and vaccine target identification (Sarangi *et al.* 2009). The target should be essential for growth and viability of the organism, should provide selectivity, and should yield a drug which is highly selective against pathogen with respect to human host. Essential genes are those important for the survival of an organism, and therefore considered a foundation of life. A subtractive genomics approach and bioinformatics provide opportunities for finding the optimal drug targets (Reddy and Satpathy 2009). A **subtractive genomics** has been successfully used by authors to locate novel drug targets in *Pseudomonas aeruginosa* (Sakharkar *et al.* 2004). The work has been effectively complemented with the compilation of the **Database of Essential Genes** (DEG) for a number of pathogenic micro-organism (Dutta *et al.* 2006; Sarangi *et al.*
The whole approach is built on the assumption that the target should play an important role in the survival of the pathogen and it should not have any conserved homolog in the human host. Non-human homologous can eradicate possibilities of cross contamination that might be harmful to the human host. The subtractive genomics approach is subtractive because we focus on the complement of the genome of the pathogen that is essential for the viability of the pathogen but is not present in the human. For prokaryotes, the DEG database contains essential genes in more than 10 bacteria, such as *E. coli*, *B. subtilis*, *H. pylori*, *S. pneumoniae*, *M. genitalium* and *H. Influenza* (Dutta et al. 2006; Bhasin et al. 2005; Galperin and Koonin 1999) whereas for eukaryotes, the DEG database contains those in yeast, humans, mice, worms, fruit flies, zebra fish and the plant *A. thaliana*, by sequence comparison of query sequence with the prokaryotes sequence in the DEG database we can find out whether the query sequence is essential or not.

### 2.4.1 MurA transferase

According to NCBI enzyme commission, one of the pathogenic genes product UDPN- acetylglucosamine enolpyruvyl transferase (MurA transferase, MurA, EC 2.5.1.7) catalyses enolpyruvyl transfer from phosphoenolpyruvate to UDP-Nacetylglucosamine (UDP-GlcNAc) and is converted to UDP-N-acetyl-3-O-(1-carboxyvinyl)-D-glucosamine/UDP-N acetylmuramic acid [UDP-MurNAC] by the addition of a lactyl group to the glucosamine. This is an important reaction involved in the first stage of peptidoglycan biosynthesis, which is the major component of the bacterial cell wall. Although it was often regarded as an inert structure surrounding bacteria, the peptidoglycan layer is a highly dynamic and tightly regulated macromolecule that is constantly remodeled to allow cell growth and division (Boneca 2005).
Further literature study and investigation of KEGG results reported MurA transferase (EC 2.5.1.7) an important target involved in metabolic pathways mainly amino sugar and nucleotide sugar metabolism and peptidoglycan biosynthesis; which indirectly plays a major role in cell wall synthesis, strengthening, cell division and plays active role in balancing the osmotic pressure of cytoplasm of cell. This MurA could be targeted to control the proliferation of pathogen in host. Because according to our studies it is both, essential to pathogen and non-homologous to human host. According to available literature we found that activity of MurA present in \textit{E. coli} is inhibited by naturally occurring antibiotic fosfomycin and Fosfomycin (C$_3$H$_7$O$_4$P) and a fungal product terreic acid (C$_7$H$_6$O$_4$) (Han \textit{et al.} 2010).

### 2.5 Virulent factors of \textit{Helicobacter pylori} Genome

Three major virulence factors of \textit{H. pylori} have been described: 1) The adhesion protein BabA2; 2) The Vacuolating toxin (VacA) and 3) The Cytotoxin-associated gene product (CagA).

**I. The BabA adhesin** of \textit{H. pylori} is an outer membrane protein that binds to the fucosylated histo-blood group antigens on the surface of gastric epithelial cells (Boren \textit{et al.} 1993; Ilver \textit{et al.} 1998). It has been reported that \textit{H. pylori}
strains possessing babA2 gene, which encodes active BabA adhesin, are associated with increased gastric inflammation (Prinz et al. 2001) and increased risk for duodenal ulcer and adenocarcinoma (Gerhard et al. 1999).

II. The Vacuolating cytotoxin gene vacA is polymorphic, varying in the signal and middle regions. The main signal region alleles are s1 and s2, whereas the middle region alleles are m1 and m2 (Cover et al. 1994; Atherton et al. 1995). VacA is a toxin that binds to several epithelial receptors (Seto et al. 1998; Padilla et al. 2000; Yahiro et al. 2003) and forms hexameric pores (Czajkowsky et al. 1999), which later are endocytosed and converted in large vacuoles (Papini et al. 1994). VacA is initially synthesized in the bacteria as a single polypeptide of about 140 kDa, in most cases comprising 1,287 residues (Cover et al. 1994; Telford et al. 1994; Schmitt and Haas 1994; Phadnis et al. 1994). The protein is subsequently released by type V-secretion system. In this context, the N-terminal leader sequence of 33 residues and the C-terminal autotransporter domain of about 33 kDa are cleaved, yielding a mature toxin of 88.2 kDa (Nguyen et al. 2001). The polypeptide can be further processed into an N-terminal fragment of 33.4 kDa (named p34 or p37, residues 1-311) and a C-terminal fragment of 54.8 kDa (named p55 or p58, comprising residues 320-821), both parts stay associated by non-covalent interactions. The relevance of the cleavage is not clear; it is not required for VacA activity. However, many data have shown that the N-terminal domain is essential in the toxic activity of VacA while the C-terminal domain is essential in binding to target membranes (Montecucco and Rappuoli 2001; Cover and Blanke 2005).

III. CagA, a Cytotoxin-associated gene A coded antigen. CagA protein (Antigen) is virulence-associated factors coded by cagA gene of cagPAI and is approximately of 145-KD. (Covacci et al. 1993; Tummuru et al. 1993). Protein CagA is translocated into gastric epithelial cells cytoplasm via type IV secretion system, encoded by the cag pathogenicity island (cag PAI) (Odenbreit et al. 2001).
The translocation further triggers tyrosine phosphorylation and alters host cell signaling and stimulates release of Interleukin 8 (IL-8). It also causes apoptosis, cell proliferation and inflammation of host cells. Upon delivery into epithelial cells via the t4ss, CagA anchors to the inner surface of cell membrane. At these sites, it recruits and modifies the junctional adhesion molecule (JAM), which is normally located at cell junctions. By interaction with JAM, CagA alters the structure and function of the apical-junctional. Apical junctional dysfunction may result in loss of control over cytoskeleton architecture, cell polarity, proliferation and differentiation, which is characteristic of oncogenic transformation. CagA contains repeated 5 amino acid sequences (Glu-Pro-Ile-Tyr-Ala), designated as EPIYA motifs, and are located at the C terminus of the protein. Functionally, EPIYA motifs are the binding targets of many host cell proteins.

**Figure 2.5.1 Mechanism of invasion of CagA**

After delivery, CagA is then phosphorylated by several Src family kinases (SFKs) on the tyrosine residues of every EPIYA motif. CagA activated SHP-2 also stimulates Extracellular signal-Regulated Kinases (Erk) through both Ras-dependent and Raslin dependent pathways. Corresponding to its stronger binding affinity for SHP-2, it has greater ability to induce cell growth. Activation
of the Ras/Erk pathway, also serves as an upstream event of CagA-induced Nuclear Factor kappa-light-chain-enhancer of activated B cells (NF-κB) stimulation. The explicit involvement of CagA in the NF-κB/IL-8-activating pathway signifies its direct role in the chronic mucosal inflammation that precedes, and is thought to be linked with, cancer (Momoyo et al. 2000).

Figure 2.5.2 CagA Signaling Pathway
2.6 Gastric cancer

Cancer is a group of diseases characterized by unregulated cell growth and spread of cells from the site of origin to other sites in body. Over 100 types of cancers have been classified. The issue of origin gives the cancer its distinguishing characteristics. Cancer originates almost anywhere in the body and they are different kinds of cancers, like:

I. **Sarcomas:** are cancers arising from cells found in the supporting tissues of the body such as bone, cartilage, fat, connective tissue, and muscle.

II. **Lymphomas:** are cancers that arise in the lymph nodes and tissues of the body’s immune system.

III. **Leukemia:** are cancers of the immature blood cells that grow in the bone marrow and tend to accumulate in large numbers in the bloodstream.

IV. **Carcinomas:** are the most common types of cancer arising from the cells that cover external and internal body surfaces. Oral, Lung, Stomach & breast are the most frequent cancers of this type.

Hallmarks of cancer:

Hanahan and Weinberg (2011) have defined six hallmarks of cancer. An effective cancer cell must be able to:

- Grow when it is not supposed to
- Ignore orders to stop growing
- Do this forever
- Avoid self-destructing
- Bring in nutrients
- Mount an invasion and to get this far it probably needs
The study is focused on the most common types of carcinoma Gastric Cancer/Ulcer.

**Gastric Cancer/Ulcer:**
Gastric Cancer is the breach occurred in Gastric and Duodenal Mucosa. They belong to type carcinoma cancers. *H. pylori* alter acid production and lead to increased acid outputs, which causes peptic disease. *H. pylori* cause gastric cancer; involve alterations in gastric epithelial cell responses that are perturbed within the context of a chronic gastric inflammatory infiltrate, which can persist for decades.

**Symptoms and Diagnosis:**
Early stomach cancer often has no symptoms or merely causes a stomachache. However, as it becomes more advanced, it can cause internal inflammation,
nausea, vomiting, severe stomach pain and weight loss. Indigestion (dyspepsia) is a very common symptom, and a high proportion of people with dyspepsia will have *Helicobacter pylori* infection causing gastritis.

Infection with *Helicobacter pylori* can be diagnosed by biopsy with pathological examination and the urea breath test; both are accurate. A blood antigen test can indicate previous exposure, but not necessarily active infection (Hanahan and Weinberg 2011).

**Gastric cancer**, despite its declining incidence rate, remains the **fourth most common cancer**, the **second leading cause of cancer-related death**, and the **14th most common cause of death overall worldwide**, which kills > 700 000 people each year (Herszenyi and Tulassay 2010; Herrera and Parsonnet 2009). Early stages of the disease are often clinically silent, with patients having advanced stage disease at the time of diagnosis, and reported 5-year survival rates are approximately 20% (Correa et al. 2004). *H. pylori* infection is the strongest known risk factor for gastrointestinal malignancies that arise within the stomach, and epidemiological studies have determined that the attributed risk for gastric cancer conferred by *H. pylori* is approximately 75% (Peek and Blaser 2002). While *H. pylori* infection increases the risk of developing both types of gastric cancer (i.e. diffuse and intestinal), chronic inflammation is not a prerequisite for development of diffuse-type cancer, thus suggesting that different mechanisms underlie the ability of *H. pylori* to induce gastric malignancies. Also, it is likely that *H. pylori* influences early stages in gastric carcinogenesis, as suggested by the demonstration that eradication of the infection significantly decreases the incidence of gastric cancer only in patients without premalignant lesions at the time of diagnosis (Wong et al. 2004).
2.7 Current Treatments and Drawbacks:

I. Bismuth subsalicylate, Metronidazole, and Tetracycline Treatment

Bismuth subsalicylate is a mild antibiotic. Likewise, metronidazole and tetracycline are also antibiotics. They fight against bacteria in our body. Together, bismuth subsalicylate, metronidazole, and tetracycline are used to treat *Helicobacter pylori*, a bacterial infection involved in causing stomach ulcers. A drug Ranitidine bismuth citrate is also used to decrease the amount of acid in the stomach and to treat *Helicobacter pylori*, a bacterial infection involved in causing stomach ulcers. Ranitidine bismuth citrate is most commonly used with clarithromycin (Biaxin), an antibiotic, to treat this infection ([http://www.nlm.nih.gov/medlineplus/druginfo](http://www.nlm.nih.gov/medlineplus/druginfo)).

However, these drugs have many side effects like:

- Difficulty breathing
- A severe headache and Vision changes
- Closing of your throat
- Kidney disease, as bismuth of drug is absorbed in kidney which is very toxic
- Swelling of your lips, tongue, or face
- Liver damage
- Blood problems
- Fever, fatigue, easy bruising or bleeding, etc.

Due to plethora of side effects many of the drugs are withdrawn from the markets, like Ranitidine bismuth citrate which was withdrawn from the U.S. market in 1998.

II. Triple therapy

Triple therapy, including two antibiotics, amoxicillin and clarithromycin, and a proton pump inhibitor taken simultaneously for 7-10 days. However, this treatment may fail for several reasons and the main reasons for failure are
compliance, side effects and *H. pylori* resistance to the antibiotics like clarithromycin, metronidazole, tetracycline, fluoroquinolones, and rifamycins (Megraud 2004). Considering the probiotic treatment of *H. pylori* infection, a medicine named A. pylori is available to control the infection of *H. pylori*. But according to the local survey analysis medicine A. pylori is not efficient enough to be considered as first line of treatment and thus, it is given with other antibiotics as second line of treatment. Therefore, it can be considered that *H. pylori* resistance to the antibiotics has become an emerging issue and have rise a need for better medication options.

### III. Chemotherapy

The chemotherapy is widely used in cancer treatment. This treatment currently uses DNA as primary target. Several aspects of DNA are targeted, including the structure of individual nucleotides; the integrity of nucleotides or their bases within DNA; the main enzymes active in the synthesis phase (DNA polymerase and topoisomerases, which are active in DNA replication and in DNA unwinding, respectively); and the structures and enzymes active in the mitosis phase. By acting on these targets, it prevents completion of the cell cycle. The actions of treatment do not target cancer cells specifically but inhibit the proliferation of any cell in the cell cycle. It means that normal cells in the cell cycle, such as hair cells, immune cells, and cells of the gastrointestinal lining, are also harmed frequently along with cancer cells. So the normal body cells also suffer along the cancerous cells.

### 2.8 Proteins

#### 2.8.1 Protein, its importance and property

A gene is a functional hereditary unit (Brandt et al. 2009) and each gene codes for at least one polypeptide (Protein), except for non-coding genes. Proteins are the functional unit of cell. They are the macromolecules found in all biological systems, from lower prokaryotes to higher eukaryotes. Proteins are the most
abundant class of bio-molecules since they represent over 50% of the dry weight of the cells. Along with the quantitative parameter, functionally also proteins are required in all most all biological processes. Either it is the central dogma of life that is transcription and translation or it may be inter and intra cellular cell signaling proteins are major regulatory factors. They also regulate the cellular compartmental transport. The major chemical reactions occurring in the natural system are catalyzed by proteins called enzymes. The immune system in any biological system is also governed by proteins. The immune globulins used by system for defense are proteins. In a nut shell proteins are necessary functional units of any biological system (Cozzone 2002). Proteins are basically made of 20 different amino acids commonly having an alpha-carbon centre with carboxylic, amino and hydrogen group attached to it. All alpha carbon has a fourth group attached called R-group. All 20 amino acids have 20 different R-groups known as side chains. Physicochemical property of amino acid depends on the type of the side chain present in it. Depending on the quantitative prevalence of each amino acid the physicochemical properties of proteins also changes. Physicochemical properties of unknown proteins are important to study as they give incite about their function. According to research in past, the three dimensional structures and biological activity of proteins depend on the physicochemical properties of their constituent amino acids (Cozzone 2002). Current development in the sequencing technology have provided plethora of data for analysis, and also have increased the scope to analyze unknown and hypothetical proteins. Due to easy sequencing and annotation, the number of predicted and hypothetical proteins has been reported so high in the database, which arise a need to analyze their physicochemical, structural and functional properties. There are number of protein sequence databases exist, ranging from simple sequence repositories, which store data with little or no manual intervention in the creation of the records, to expertly curated universal databases that cover all species and in which the original sequence data are enhanced by the manual addition of
further information in each sequence record (Apweiler et al. 2004). Some widely used protein sequence repositories are NCBI’s Entrez Protein (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?dbProtein); Reference sequence (http://www.ncbi.nlm.nih.gov/RefSeq); Protein Information Resource Protein Sequence Database (PIRPSD) (http://pir.georgetown.edu/); The leading universal curated protein sequence database Swiss-Prot (http://www.ebi.ac.uk/swissprot/index.html) and UniProt (http://www.uniprot.org). Focusing on prediction of physicochemical properties of proteins a number of online software facilities are available. The Supercomputing Facility for Bioinformatics & Computational Biology, IIT Delhi provides software for calculating the properties of protein (http://www.scfbio-itd.res.in/bioinformatics/bioinformaticssoftware.htm) but this facility have separate software for each property so user has to use number of software for single protein analysis which is inconvenient and time consuming task. AACompIdent is a tool provided by Uniprot (Expasy server) which allows the identification of a protein from its amino acid composition (Wilkins et al. 1996). Other Software provided by Expay is Protparam (http://expasy.org/tools/protparam.html); it estimates many basic physicochemical properties of a polypeptide on the basis of its sequence. There is a standalone software ProPAS (Songfeng and Yunping 2012) for computing physicochemical properties of protein but it consists of only three components. There are many such web based tools available beyond cited here.

2.8.2 Protein Modeling

One of the most exciting challenges for computational biology has been to predict the function of uncharacterized proteins for which an experimental derived structure is not available (Gangrade and Panchal 2011). According to the present condition of protein sequence and structure repositories, the number of protein sequences excels higher than presence of their relative structures. Protein structure modeling deals with predicting the secondary and
tertiary structure of a particular protein from its primary structure; which is amino acid sequence of proteins. The tertiary structure of protein can reveal much more information on functional aspect of proteins. As the wet-lab methods to predict structure X-ray crystallography, NMR (Chasse et al. 2001) and spectroscopy are too expensive for small scale research, one way to bridge this gap between sequence and structure is to use computer-generated structure models of proteins (Malmstrom and Goodlett 2010). There are various methods to generate computed models. Major methods are ab initio based, threading (or fold recognition) and comparative modeling.

The first approach is de novo or ab initio methods. This method predicts the structure from sequence alone, without relying on similarity at the fold level between the modeled sequence and any of the known structures (Bonneau and Baker 2001). Threading and Comparative (or Homology) modeling methods make use of experimental protein structures to build models for evolutionary related proteins. Experimental structural biology and homology modeling thereby complement each other in the exploration of the protein structure space. There are many types of software and servers available for comparative modeling and collection of some better comparative modeling servers are available at Protein Model Portal (PMP) (http://www.proteinmodelportal.org); a collaborative initiative named Protein Structure Initiative Knowledgebase (PSI KB) (Arnold et al. 2009).

PMP provides access to M4T (http://www.fiserlab.org/servers/m4t) (Fernandez-Fuentes et al. 2007), Swiss model (http://swissmodel.expasy.org/workspace) (Guex and Peitsch 1997) and Mod Web (http://salilab.org/modweb) (Eswar et al. 2003) servers for comparative modeling. As all of them are template based modeling methods, according to their algorithms, target input and usage of single or multiple templates differ in the quality of output structure they provide.
In *silico* modeling is a multidisciplinary method integrating mathematical models with experimental (in vitro and in vivo) and clinical data. (Sanga *et al.* 2007) Homology or evolutionary relatedness represents a key concept in studying protein sequence, structure, and function. Homologs can be inferred by sequence similarity search tools such as the popular sequence-profile comparison method PSI-BLAST. (Bong-Hyun *et al.* 2009) Basic Local Alignment Search Tool (BLAST) provides an "expect" value, statistical information about the significance of each alignment (Christiam *et al.* 2009) **Comparative, or homology, modeling** of protein structures is the most widely used prediction method when the target protein has homologues of known structure (Piedra *et al.* 2008). This study is aimed at the genomic and proteomic characterization, modeling and evaluating the verified structure of proteins as a potential drug target in gastric cancer therapy. **Docking** is a method which predicts the preferred orientation of one molecule to a second when bound to each other to form a stable complex (Zachairias 2003). A variety of experimental and computational techniques can be used to identify possible protein binding partners of protein. The prediction of putative protein-ligand interaction studied by computational docking methods is of increasing importance in the field of structure based drug designing (Lyskov and Gray 2008).

### 2.9 Antibiotics

**Antibiotics** are chemicals that kill or inhibit the growth of bacteria and are used to treat bacterial infections. They are produced in nature by soil bacteria and fungi. This gives the microbe an advantage when competing for food, water and other limited resources in a particular habitat, as the antibiotic kills off their competition.

**How do antibiotics work?**

Antibiotics take advantage of the difference between the structure of the bacterial cell and the host’s cell. They either prevent the bacterial cells from multiplying so that the bacterial population remains the same, allowing the
host’s defense mechanism to fight the infection or kill the bacteria, for example stopping the mechanism responsible for building their cell walls.

An antibiotic can also be classified according to the range of pathogens against which it is effective. Penicillin G will destroy only a few species of bacteria and is known as a narrow spectrum antibiotic. Tetracycline is effective against a wide range of organisms and is known as a broad spectrum antibiotic.

**Antibiotic resistance**

Bacteria are termed drug-resistant when they are no longer inhibited by an antibiotic to which they were previously sensitive. The emergence and spread of antibacterial-resistant bacteria has continued to grow due to both the over-use and misuse of antibiotics.

Treating a patient with antibiotics causes the microbes to adapt or die; this is known as ‘selective pressure’. If a strain of a bacterial species acquires resistance to an antibiotic, it will survive the treatment. As the bacterial cell with acquired resistance multiplies, this resistance is passed on to its offspring. In ideal conditions some bacterial cells can divide every 20 minutes; therefore after only 8 hours in excess of 16 million bacterial cells carrying resistance to that antibiotic could exist.

**How is resistance spread?**

Antibiotic resistance can either be inherent or acquired. Some bacteria are naturally resistant to some antibiotics due to their physiological characteristics. This is inherent resistance. Acquired resistance occurs when a bacterium that was originally sensitive to an antibiotic develops resistance. For example, resistance genes can be transferred from one plasmid to another plasmid or chromosome, or resistance can occur due to a random spontaneous chromosomal mutation ([http://www.microbiologyonline.org.uk/](http://www.microbiologyonline.org.uk/)).
Antibiotic Fosfomycin

Fosfomycin is an antibiotic produced by *Streptomyces fradiae* and is currently accepted widely as approved drug. Fosfomycin is considered as anti-bacterial agent. Fosfomycin is also known as Fosfocina, Phosphomycin, Phosphonomycin. Considering the pharmaceutical industry Fosfomycin is known as Monurol or Veramina. Fosfomycin is a broad spectrum antibiotic that concentrates in kidney and bladder and is used to treat uncomplicated urinary tract infections. Fosfomycin also reduces nephrotoxicity and ototoxicity of platinum-containing anti-tumor agents. Fosfomycin is a phosphoenolpyruvate analogue produced by Streptomyces that irreversibly inhibits enolpyruvate transferase (MurA), which prevents the formation of N-acetylmuramic acid, an essential element of the peptidoglycan cell wall. The 3 salts of Fosfomycin are

I. Fosfomycin calcium monohydrate with average Mass: 196.152 Dalton
II. Fosfomycin disodium with Average Mass: 182.0227 Dalton
III. Fosfomycin tromethamine with Average Mass: 259.1941 Dalton

Fosfomycin tromethamine is rapidly absorbed following oral administration and converted to fosfomycin. Oral bioavailability under fasting conditions is 37%. When given with food, oral bioavailability is reduced to 30% ([http://www.drugbank.ca/drugs/DB00828](http://www.drugbank.ca/drugs/DB00828)).