Helicobacter pylori – Multitalented Adaptation of Binding Properties

Sara Henriksson
The years of anxious searching in the dark, with their intense longing, their intense alternations of confidence and exhaustion and the final emergence into the light—only those who have experienced it can understand it.

*Albert Einstein*

To My Family!
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>III</td>
</tr>
<tr>
<td>LIST OF PAPERS</td>
<td>V</td>
</tr>
<tr>
<td>ABBREVIATIONS</td>
<td>VII</td>
</tr>
<tr>
<td>ENKEL SAMMANFATTNING</td>
<td>VIII</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td><strong>GENERAL FACTS ON H. PYLORI</strong></td>
<td>1</td>
</tr>
<tr>
<td>Infection and colonization of the stomach</td>
<td>2</td>
</tr>
<tr>
<td>Motility and chemotaxis</td>
<td>3</td>
</tr>
<tr>
<td>Urease</td>
<td>3</td>
</tr>
<tr>
<td><strong>TRANSMISSION, EPIDEMIOLOGY, AND DISEASE OUTCOME</strong></td>
<td>4</td>
</tr>
<tr>
<td>Transmission and epidemiology</td>
<td>4</td>
</tr>
<tr>
<td>Disease progression and outcome</td>
<td>5</td>
</tr>
<tr>
<td><strong>GENETIC AND GEOGRAPHICAL DIVERSITY</strong></td>
<td>7</td>
</tr>
<tr>
<td><strong>VIRULENCE FACTORS</strong></td>
<td>9</td>
</tr>
<tr>
<td>cagPAI</td>
<td>10</td>
</tr>
<tr>
<td>VacA</td>
<td>10</td>
</tr>
<tr>
<td><strong>OUTER MEMBRANE PROTEINS</strong></td>
<td>11</td>
</tr>
<tr>
<td><strong>BACTERIAL ADHERENCE</strong></td>
<td>12</td>
</tr>
<tr>
<td>Receptor specificity</td>
<td>13</td>
</tr>
<tr>
<td>Detachment</td>
<td>13</td>
</tr>
<tr>
<td>H. pylori adherence and detachment</td>
<td>14</td>
</tr>
<tr>
<td>BabA</td>
<td>16</td>
</tr>
<tr>
<td>SabA</td>
<td>19</td>
</tr>
<tr>
<td>Mucin interactions</td>
<td>20</td>
</tr>
<tr>
<td><strong>THE RECEPTORS FOR H. PYLORI</strong></td>
<td>21</td>
</tr>
<tr>
<td>The ABH antigens</td>
<td>21</td>
</tr>
<tr>
<td>Lewis antigens</td>
<td>23</td>
</tr>
<tr>
<td>AIMS OF THE THESIS</td>
<td>24</td>
</tr>
<tr>
<td><strong>RESULTS</strong></td>
<td>25</td>
</tr>
<tr>
<td><strong>PAPER I</strong></td>
<td>25</td>
</tr>
<tr>
<td><strong>PAPER II</strong></td>
<td>27</td>
</tr>
<tr>
<td><strong>PAPER III</strong></td>
<td>30</td>
</tr>
<tr>
<td><strong>PAPER IV</strong></td>
<td>31</td>
</tr>
<tr>
<td><strong>DISCUSSION AND CONCLUDING REMARKS</strong></td>
<td>33</td>
</tr>
<tr>
<td><strong>ACKNOWLEDGEMENTS</strong></td>
<td>36</td>
</tr>
<tr>
<td><strong>REFERENCES</strong></td>
<td>38</td>
</tr>
</tbody>
</table>
Abstract

*Helicobacter pylori* infects and persists within the stomach of most of the world's population causing chronic inflammation of the stomach that can result in gastritis, peptic ulcer disease and cancer. Adherence to gastric epithelial cells is important for the establishment of infection and is mainly mediated by the two adhesins BabA (Blood group Antigen Binding Adhesin) and SabA (Sialic Acid Binding Adhesin). BabA binds the ABO/Leb blood group antigens (Leb), and SabA binds to sialylated antigens. Both adhesins demonstrate high-affinity to their receptors. This very strong protein-glycan type of interaction is both complex and also seemingly contradictory within the dynamic environment of the stomach. This is so because the high turnover rate and exfoliation of the gastric epithelium would otherwise have managed to shed and discard of attached bacteria into the acidic and bactericidal stomach lumen.

We have found that BabA mediated binding is regulated by acid and binding to Leb is inactivated at low pH, but recovers when pH is neutralized. The SabA adhesin behaves similarly, which suggests that this reciprocity in binding properties is a unique and important feature of *H. pylori*. The acid lability of the binding interaction was traced to specific amino acid positions in the BabA protein. These mutations were also found to correspond to differences in local acidity of different parts of the stomach and to human populations with different patterns of gastric disease. Mutations in similar amino acid positions were identified in BabA also from a 5-year animal infection study. Our hypothesis is that when BabA is exposed to a pH-gradient the binding is gradually decreased and lost at lower pH, i.e., when the epithelial cells are discarded and shed into the acidic lumen. Such acid regulation of binding property allows the bacteria (and hence the infection) to detach from the dying cells and by their flagella bacteria can swim back to the intact stomach wall and hence, recycle the infection. Thus, *H. pylori* can take advantage of pH-regulated binding affinity of its adhesins for functional recycling and continuous adaptation to the stomach environment in order to maintain persistent life-long infection.

*H. pylori* strains from southern Europe have a lower prevalence of Leb-binding ability. Seventy-five percent of strains in Sweden but only 40% of strains in Spain bind Leb. Among the Spanish isolates, we identified one strain, 812, that binds poorly to Leb but adheres strongly to histo-tissue sections of gastric mucosa. This binding was not perturbed by removal of sialylated carbohydrate, and, therefore, SabA binding sites. These novel results suggest alternative BabA-mediated binding properties and receptors
for *H. pylori*. We first identified this new receptor as a glycolipid that was independent of secretor phenotype, i.e., Leb, in secretions such as saliva and milk. However, glycan arrays demonstrated stronger binding of 812 to blood group A and B antigens compared to blood group O, the Leb/ALeb ratio on glycan array was 0.65 compared to 1.18 for the reference strain 17875/Leb. These results suggest that some *H. pylori* strains have adapted their binding properties to individuals of blood groups A and B.

To investigate the impact of secretor status in BabA-positive *H. pylori* infections, we studied a large group of volunteers from Porto, Portugal, for factors that could affect adherence. We found a significant correlation between BabA expression and other virulence factors such as CagA and VacA, pre-dysplastic appearance, and positive secretor status. BabA-mediated *in vitro* adherence of *H. pylori* to the gastric tissue correlated with positive secretor status.

Detailed investigations of *H. pylori* BabA-mediated binding to cells in culture have been difficult because *H. pylori* is a human-specific pathogen and most cell lineages are derived from immortalized cells with dysplastic glycosylation expression patterns. We transfected cells with glycosyltransferases necessary for expression of the Leb or Lea blood group antigens; Leb-expression levels were correlated to binding of *H. pylori*. Thus, these cell lineages could be used for detailed studies of BabA-mediated binding and what effect this interaction has on the host cell.
This thesis is based on the following papers:


pH Regulated H. pylori Adherence: Implications for Persistent Infection and Disease.

*These authors contributed equally to this work


Clinical isolates of Helicobacter pylori demonstrate alternative BabA-mediated binding properties for adherence to the gastric mucosa.

Manuscript in progress


Please note the change in name from S. Eriksson to S. Henriksson
Papers not included in this thesis

### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BabA</td>
<td>Blood group antigen binding adhesin</td>
</tr>
<tr>
<td>CagA</td>
<td>Cytotoxin-associated gene A</td>
</tr>
<tr>
<td>CBD</td>
<td>Carbohydrate binding domain</td>
</tr>
<tr>
<td>FITC</td>
<td>Fluorescein isothiocyanate</td>
</tr>
<tr>
<td>FucT</td>
<td>Fucosyltransferase</td>
</tr>
<tr>
<td>Gal</td>
<td>Galactose</td>
</tr>
<tr>
<td>GalNAc</td>
<td>N-Acetylgalactosamine</td>
</tr>
<tr>
<td>GlcNAc</td>
<td>N-Acetylgalactosamine</td>
</tr>
<tr>
<td>HC</td>
<td>High cancer</td>
</tr>
<tr>
<td>Hop</td>
<td>Helicobacter outer membrane protein</td>
</tr>
<tr>
<td>Hor</td>
<td>Hop-related</td>
</tr>
<tr>
<td>HU</td>
<td>High ulcer</td>
</tr>
<tr>
<td>Le</td>
<td>Lewis antigen</td>
</tr>
<tr>
<td>Lea</td>
<td>Lewis a</td>
</tr>
<tr>
<td>Leb</td>
<td>Lewis b</td>
</tr>
<tr>
<td>MALT</td>
<td>Mucosa-associated lymphoid tissue</td>
</tr>
<tr>
<td>OMP</td>
<td>Outer membrane protein</td>
</tr>
<tr>
<td>RIA</td>
<td>RadioImmunoAssay</td>
</tr>
<tr>
<td>SabA</td>
<td>Sialic acid binding adhesin</td>
</tr>
<tr>
<td>Se</td>
<td>Secretor</td>
</tr>
<tr>
<td>SPR</td>
<td>Surface plasmon resonance</td>
</tr>
<tr>
<td>T4SS</td>
<td>Type 4 secretion system</td>
</tr>
<tr>
<td>VacA</td>
<td>Vacuolating cytotoxin A</td>
</tr>
</tbody>
</table>
Enkel sammanfattning

Redan som barn blir ungefär hälften av jordens befolkning infekterad av *Helicobacter pylori*, en kronisk infektion som kan botas med antibiotikabehandling. De flesta bär på infektionen utan tydliga besvär, men ungefär 10–20% kommer få magsår och 1-2% drabbas av magcancer. I utvecklingsländer är förekomsten av *H. pylori* infektion ännu högre där 80-90% kan vara infekterade.


Syftet med denna avhandling var att studera hur *H. pylori* anpassar sitt vidhäftningsprotein BabA till att binda till ABO-antigena trots de pH- förändringar som *H. pylori* utsätts för under infektion. Jag ville även studera stammar från olika befolkningar och undersöka deras bindningsegenskaper då vår grupp tidigare visat att många isolat inte tycks binda till blodgruppsantigenen.

Sammanhanget är att även dessa förändringar uppkommit genom sekvensutbyte med ett annat BabA-liknande protein.

Ungefär 20 % av Europas befolkning är s.k. icke-sekretor och saknar därför blodgruppsantigener i tårvätska, saliv och bröstmjölk, de har inte heller blodgruppsantigener i sin magslemhinna. Istället har de en kortare kolhydrat som *H. pylori* inte kan binda till och som kallas Lewis a (Lea). De övriga 80 %, vilka har ABo blodgruppsantigener i magen, kallas ”sekretor positiva”. Vi undersökte *H. pylori* infektion hos ”sekretor positiva” för att se om de oftare blev infekterade av *H. pylori* som har BabA. När vi studerade vidhäftning *H. pylori* med BabA till vävnadssnitt från magslemhinna kunde vi se att sekretorindivider uppvisade högre bindning. Vi kunde också se att individer som har blodgruppsantigenerna i magen är i större utsträckning infekterade med *H. pylori* som har BabA. Infektion med BabA-uttryckande stammar påvisade fler bakterier i nära kontakt med våra celler i epitellagret, medan de *H. pylori* stammar som saknar BabA var på större avstånd från epitelcellerna. Vi fann ett samband mellan förekomst av BabA och två andra sjukdomsassocierade faktorer hos *H. pylori*, CagA och VacA-protein Förekomst av BabA var också vanligare hos individer med mer avancerad sjukdom såsom intestinal metaplasi, ett förstadium till magcancer.


Genom att bättre förstå hur bakterien vidhäftar och därigenom hur man skulle kunna förhindra att de kommer i kontakt med våra celler så ökar möjligheten att hitta nya principer för läkemedelsbehandling mot magsår och cancer.
Introduction

History

*Helicobacter pylori* infects the stomachs of more than 50% of the world’s population and has lived in such close association with modern humans since before they migrated from East Africa more than 58,000 years ago (Linz et al. 2007). Before the discovery of *H. pylori* in the early 1980s, stomach disorders such as gastritis and peptic ulcers were ascribed to bad diet, too much coffee, or a stressful lifestyle. The disorders were treated accordingly with drugs such as antacids and proton pump inhibitors to reduce the acidity of the stomach and thus eliminate the symptoms. Bacteria had been seen in the stomach as early as 1874, but these findings were ignored because nothing was believed to survive the acidic environment in the stomach. In the early 1980s, *H. pylori* was isolated from the antrum of patients with gastritis and ulcer disease, and later experiments fulfilled Kosh’s postulates and, importantly, antibiotic treatments got rid of the infection and the inflammation disappeared (Marshall & Warren 1984). For this finding, Marshall & Warren were awarded with the Nobel Prize in Medicine and physiology in 2005. Hence, *H. pylori* was confirmed as the cause of gastritis and the more serious peptic ulcer diseases. In epidemiological studies, *H. pylori* infection was also found to correlate with gastric cancer, i.e. *H. pylori* is considered as an onco-pathogen (Parsonnet et al. 1991). In 1994, the World Health Organization listed *H. pylori* infection as a carcinogen (IARC 1994). The understanding that *H. pylori* infection is the causative agent of overt gastric disease opened a paradigm shift that has completely changed the treatment of stomach disorders, which are now considered as infectious diseases.

General facts on *H. pylori*

*H. pylori* is Gram negative, helical shaped bacteria, and has a tuft of 5–7 polar flagella located at its distal end. *H. pylori* is microaerophilic and has its optimal growth conditions at 5–19% O₂, 5–10% CO₂, 37° C, and high humidity; similar to conditions that are found in the gastric mucosa. Although being a gastric pathogen, the high acidity of the stomach juice is lethal for the bacteria. *H. pylori*, therefore, resides in the slimy mucus layer that lines and protects the epithelial cells from the acidic gastric juice.
Infection and colonization of the stomach

*H. pylori* colonizes the stomach and, in particular, the less acidic antrum (Figure 1). During disease progression, pH initially decreases due to hypersecretion and *H. pylori* might move into the first parts of the intestine, the duodenum. This region is less resistant to infection and peptic ulcer could develop. Long-term hypersecretion can cause atrophic gastritis to the mucosa and even loss of the acid producing parietal cells and higher stomach pH. Atrophy can also result in a gastric ulcer formation, which is sometimes a precursor to gastric cancer (described in the section on Disease progression and outcome). Within the mucus layer, *H. pylori* is mainly confined to the 100 µm of mucus closest to the epithelial cells where pH is more neutral. Thirty percent are found within the first 5 µm and around 20% are found tightly attached to the cells (Schreiber et al. 2004; Hessey et al. 1990). The colonization of *H. pylori* is restricted to the superficial epithelial cells, colocalizing with the expression of the mucin MUC5AC (Van den Brink et al. 2000; Hidaka et al. 2001). The strict colonization of the superficial zone might relate to the glandular mucin MUC6 that possesses terminal α1,4-GlcNAc. This structure inhibits cell wall synthesis in *H. pylori* thereby making these glandular regions toxic to *H. pylori* (Kawakubo et al. 2004). In addition to the extracellular habitat, *H. pylori* has also been found between cells, deeper in the tissue, and in intracellular vesicles of both cultured gastric epithelial cells and in gastric biopsies (Amieva et al. 2002; Necchi et al. 2007; Aspholm et al. 2006). These invasive bacterial cells can repopulate the extracellular environment suggesting that the intracellular lifestyle might be a way for *H. pylori* to escape the immune system as well as antibiotic treatment (Amieva et al. 2002 and (reviewed in Dubois & Borén 2007).

![Figure 1. Simplified anatomical illustration of the stomach](image)

Figure 1. Simplified anatomical illustration of the stomach
Motility and chemotaxis

For chemotactic orientation, *H. pylori* senses the pH and bicarbonate gradient in the mucus and swims away from the gastric juice and towards the neutral environment close to the epithelial cells (Schreiber et al. 2004).

The spiral shape of the bacteria, caused by specific crosslinking of the peptidoglycan and coiled-coil-rich proteins, together with flailing movements of the flagella, enables *H. pylori* to penetrate the slimy mucus layer in a cork screw-like motion, a function that is essential for survival (Eaton et al. 1992; Eaton et al. 1996; Sycuro et al. 2010; Specht et al. 2011; Ottemann & Lowenthal 2002). The flagella are composed of three structures: the basal body, the hook, and the filament that ends with a bulb (Chevance & Hughes 2008). The hook and filament protrude from the basal body and are covered in a membranous sheath that is contiguous with the outer membrane, a sheath that probably functions to protect the filament from low pH or against the host immune response (C. S. Goodwin et al. 1985) and (reviewed in Lertsethtakarn et al. 2011). The flagella filaments are mainly composed of the two proteins, FlaA and FlaB (Kostrzynska et al. 1991; Suerbaum et al. 1993). During assembly of the filament, both FlaA and FlaB are heavily glycosylated with pseudaminic acid, a carbohydrate that resembles human sialic acid (Josenhans et al. 2002; Schirm et al. 2003).

Urease

In addition to the chemotactic and swimming capabilities that *H. pylori* utilizes to escape the acidic gastric juice, it also possesses the cytoplasmic enzyme urease. Urease is a large 1.1-MDa complex that neutralizes the local environment by converting urea to NH₃ and HCO₃⁻. NH₃ neutralizes protons in the periplasm and HCO₃⁻ acts as a buffer to maintain a pH of 6.1 (Scott et al. 2007; Boyanova et al. 2011). Lysis of bacterial cells in the surrounding releases urease and buffers into the microenvironment of the live bacteria, thus the bacterial cells are buffered both internally and externally (Dunn et al. 1997; Marcus & Scott 2001). This neutralization of the mucin also changes its rheological properties and makes it easier for *H. pylori* to penetrate (Celli et al. 2009).
Transmission, epidemiology, and disease outcome

Transmission and epidemiology

*H. pylori* infects around 50% of the population, is acquired during childhood, and typically persists in the stomach throughout life unless treated with antibiotics. Overcrowding, poor sanitation, and low economic standards are the risk factors for acquiring *H. pylori* infection, which is inversely correlated with socioeconomic status (reviewed in Khalifa et al. 2010). Because of the improved living standards in developed compared to developing countries, the infection rate is still much higher in developing countries where it can sometimes reach 80-90%. Related to this is the birth cohort effect that is seen in developed countries where the *H. pylori* prevalence is much higher in the older generation because of a lower acquisition rate in the younger generation and wide spread use of antibiotics. In a study of Peruvian children, 71% were infected by the age of 6 months (determined by urea-breath test) compared to 1.2% of 6-month-old Swedish children (determined by serological test) (reviewed in Delport & van der Merwe 2007). The increased growth rate of developing countries proposes that the total number of infected individuals in the world will increase even further (reviewed in Khalifa et al. 2010). Consequently, gastric disease related to *H. pylori* infection will also increase.

No natural reservoir has been found to be responsible for the spread of *H. pylori*, and it is probably transmitted from person to person through oral-oral transmission because bile in the fecal environment is toxic for *H. pylori*. The most plausible route of transmission is during events of gastroenteritis where *H. pylori* can be isolated particularly from vomitus but also from the air. During such disease, *H. pylori* can also survive and spread through diarrhea (Perry et al. 2006); and (reviewed in Delport & van der Merwe 2007). In developed countries, *H. pylori* is mainly transmitted in an intra-familial manner from parents to children but also between siblings. In the developing world with poor sanitation and other family arrangements, extra-familial spread is more common (Schwarz et al. 2008); and (reviewed in Delport & van der Merwe 2007).

In 10–20% of infected individuals, the infection progresses towards peptic ulcer disease, 1–2% acquire gastric cancer, and <1% acquire mucosa-associated lymphoid tissue (MALT) (Figure 2). Because of the high prevalence of *H. pylori* infection, gastric cancer has become the second most common form of cancer in the world, second only to lung cancer. In Sweden there were 9323 deaths due to gastric cancer between 1991-2000 (Sonnenberg 2011). However, during the same period, 4373 individuals died
due to bleeding peptic and gastric ulcers that can also be associated with *H. pylori* infection. Considering that these numbers represent Sweden, a country with a relatively low prevalence of *H. pylori*, they are expected to be even higher worldwide. However, infection does not necessarily lead to these disorders and many infected carriers have no symptoms. Because of the long relationship with humans, some scientists almost consider this a commensal relationship, and *H. pylori* infection has suggested to protect against diseases such as gastroesophageal reflux disease (reviewed in de Vries & Kuipers 2010).

### Disease progression and outcome

During the primary, acute infection, the stomach responds by releasing mucins and sheds the cells of the superficial cellular layer in an attempt to remove the bacteria. This is accompanied by hypochlorhydria (low acid secretion) that can continue for weeks until the acid secretion returns to normal (Schubert & Peura 2008). For very few individuals the infection is spontaneously cleared, and it is more common that the bacteria will persist in the stomach for the individual’s lifetime unless treated with antibiotics.

**Figure 2.** Disease progression of *H. pylori* infection. Adapted from (Suerbaum & Josenhans 2007)

**Hyperchlorhydria-related diseases**

All *H. pylori* infected individuals develop gastritis, primarily of the antrum where the bacteria mainly colonizes. Gastritis is characterized by an increased infiltration of mononuclear leucocytes and PMN cells that causes an inflammation of the mucosa (reviewed in Correa & Piazuelo 2012). During this stage of disease, the acid secretion from the parietal cells is
increased because of an increase in gastrin production by the G-cells in the antrum. For most *H. pylori* infected individuals, the disease progression will stop here and they might carry the infection unnoticed for the rest of their lives. Some infected individuals, however, are more sensitive to the high acidity (low pH) and will develop duodenal ulcers (reviewed in El-Zimaity 2007). Eradication of *H. pylori* at this step will allow the ulcer to heal (Graham et al. 1992).

**Hypochlorhydria-related diseases**

If the infection of the antrum progresses, it will eventually degrade the G-cells and gastrin secretion will decrease. This causes a reduction in the flow of gastric juice and the corpus becomes less acidic. *H. pylori* can then migrate towards the corpus and cause a corpus-predominant gastritis or pangastritis, further degrading the parietal cells and eventually causing corpus atrophy and loss of normal glandular tissue (reviewed in Correa & Piazuelo 2012). During this step, some individuals develop gastric ulcers, that, unlike duodenal ulcers, occur in an environment of low acidity (high pH) (reviewed in El-Zimaity 2007). Intestinal metaplasia is the primary step towards cancer development and is characterized by a progressive spread and replacement of the stomach cells by intestinal cells accompanied by a loss of the normal mucins of the stomach, MUC5ac and MUC6. These are then replaced by more intestinal-like sialylated and sulphated mucins, such as MUC2 (reviewed in Correa & Piazuelo 2012). These mucins are secreted by large vacuoles of the intestinal Goblet cells, and are easily detected with Alcian blue staining, a typical hallmark of intestinal metaplasia. Intestinal metaplasia advances into dysplasia that almost irreversibly progresses towards gastric cancer. During the cancer stage, very few *H. pylori* have remained in the stomach due to unfavorable growth conditions. The prognosis for gastric cancer is poor, usually because it is detected at a late stage. However, better prognosis is seen for patients receiving anti-*H. pylori* therapy simultaneously with surgery (Fukase et al. 2008).

The rate at which this entire process of disease progression occurs is slow, and if it progresses all the way to gastric cancer this will occur later in life, usually >60 years of age. The exact mechanism that triggers the progression towards corpus-predominant gastritis and gastric cancer is not known. Some host factors such as polymorphisms of IL-1β, TNF-α, and IL-10 might be involved. IL-1β is particularly important for gastric cancer development because of its inhibitory effects on acid secretion. Additional host factors involve high salt consumption and cigarette smoking (reviewed in Wroblewski et al. 2010). Bacterial virulence factors such as a prevalence of
CagA and the VacAs1 allele are also correlated with gastric cancer (described in the section on virulence factors).

**Genetic and geographical diversity**

To persist in the gastric niche, *H. pylori* must quickly and constantly adapt to new environments. *H. pylori* demonstrates a substantial heterogeneity and allelic diversity making it one of the most diverse bacterial species known. No two isolates are alike and, despite being classified as the same strain, differences are found even in isolates from the same individual at different time points (Kuipers et al. 2000; Kersulyte et al. 1999; Israel et al. 2001; Falush et al. 2001; Gressmann et al. 2005; Salama et al. 2007). This diversity was noticed early using simple techniques, and modern developments of more elaborate genome sequencing have enabled detailed studies on the mechanisms that confer this variability. Sequence comparison of seven different genomes identified 4.9 to 12.4% of the total gene pool to be strain specific, and microarray analyses have found this number to be as high as 27% (Fischer et al. 2010; Gressmann et al. 2005). Diversification of *H. pylori* is conferred by two main processes. First, during replication the genome is subjected to changes such as point mutations, rearrangements, and slipped-strand mispairing. Second, these changes are further distributed between genetic variants within the strain, or between strains in a mixed infection, by natural transformation and recombination (Figure 3). Because of inversions, many strains also demonstrate different syntenies (gene orders) (Lara-Ramírez et al. 2011). The outcome of this diversity is a quasi-panmictic population structure where, in a changing environment, one genetic variant is superior and could surpass the other clones when they are challenged by changes in the environment (Suerbaum et al. 1998); and (reviewed in Dorer et al. 2009).

The microenvironment in which *H. pylori* lives is a constant source of mutagens such as free radicals and low pH. In addition, *H. pylori* also uses a polymerase that does not possess proof reading activity and results in a hypermutator phenotype when it is overexpressed (García-Ortíz et al. 2011). In line with having a small genome, the mechanisms to cope with such high mutation rate are limited and no homolog for the mismatch repair system has been identified (reviewed in Dorer et al. 2011). Because of the lack of many conventional repair mechanisms, homologous recombination is important for DNA-repair. This is supported by the findings of Dorer et al. (2010) that upon DNA damage, *H. pylori* upregulates a lysozyme and genes involved in natural competence such as the unique comB Type IV secretion system to acquire new genes from lysed cells in the microenvironment.
(Hofreuter et al. 2001; Dorer et al. 2010). The average length of imported gene fragments is estimated to be 417 base pairs and is much shorter than that seen in other bacteria (Falush et al. 2001). Interestingly, it appears that the uptake of external DNA consists of larger pieces that are then digested to smaller fragments before being imported into the chromosome to further increase the genetic diversity of the organism (Kennemann et al. 2011; Lin et al. 2009; Kulick et al. 2008).

Figure 3. Mechanisms of genetic variability in H. pylori. Allelic change can be the result of a point mutation in an existing gene (1) or acquire of a new gene from other strains by natural transformation, which could either replace a gene (2) or add a new gene to the genome (3). H. pylori can turn off expression of genes by phase variation and slipped strand mispairing (4). Intra-genomic recombination can create either mosaic genes (5) or complete gene conversion creating two identical genes (6).

In a genome-wide analysis, the average mutation rate was $2.5 \times 10^{-5}$ per site and year and this can be compared to the recombination rate, $5.5 \times 10^{-5}$ per initiation site and year, measured on the same set of samples (Kennemann et al. 2011). This mutation frequency is even higher than that seen in the Escherichia coli mutator phenotype (Björkholm et al. 2001). In a previous analysis of related samples, it has been demonstrated that only 1 out of 650 recombination events gives rise to a loss or gain of a gene, and that recombination is responsible for three times as many substitution mutations making it the most important factor in generating diversity (Morelli et al. 2010; Kennemann et al. 2011). Deletion mutants of RecA, the H. pylori recombinase, are sensitive to both UV and DNA-damaging agents and are also defective in successful colonization demonstrating the necessity for

Related to recombination are the numerous restriction enzymes (R-M system) that *H. pylori* utilizes to digest and restrict the import of DNA from other species or strains. Humbert and colleagues (2011) indicated in their discussion that the recognition sites for the R-M system are not evenly distributed in the genome indicating a selection for recombination of specific loci (Humbert et al. 2011). In line with this, it has recently been shown that the outer membrane proteins are among the highest frequency genes involved in recombination (Kennemann et al. 2011; Yahara et al. 2012).

The quasi-panmictic population of *H. pylori* and its long association with humans since before they migrated out of Africa, has resulted in a diversity that expands globally (Linz et al. 2007; Falush 2003). Human migrations have limited the exchange of alleles between strains and isolated groups of *H. pylori* ancestral strains have evolved into seven population types that reflect the migrations of humans: hpEurope, hpEastAsia, hpAfrica1, hpAfrica2, hpAsia2, hpNEAfrica and hpSahul (Falush 2003; Linz et al. 2007; Moodley et al. 2009). Interestingly, the cag pathogenicity island (cagPAI) was acquired after the populations split in Africa as it is completely missing in the hpAfrica2 strain (Gressmann et al. 2005). These different population classes thus reflect different genotypes that, to some extent, are responsible for the dissimilar prevalence of *H. pylori*-related diseases worldwide. An example of this is gastric cancer being more common in Japan but duodenal ulcers being more common in India (reviewed in R. Suzuki et al. 2012).

**Virulence factors**

Virulence is defined as the ability of a bacterial species to induce disease. Consequently, the outcome of a bacterial infection is highly dependent on the prevalence and status of its virulence factors. The genetic diversity and variability of *H. pylori* is mirrored in the wide range of virulence factors that vary by disease, age, country, and ethnicity (Boyanova et al. 2011). To be defined as an *H. pylori* virulence factor, the protein must be correlated with disease both in vitro and in vivo and with epidemiological disease patterns (reviewed in Lu et al. 2005). Three main virulence factors of *H. pylori* are the cytotoxin-associated gene pathogenicity island (cagPAI), the vacuolating cytotoxin (VacA), and the outer membrane proteins (OMPs). Many of the OMPs are proposed to be involved in disease-associated mechanisms such as adherence and manipulation of the immune response. VacA and CagA are,
together with BabA, associated with the more severe cases of gastric disease (described in the section babA and pathogenesis).

**cagPAI**

The cagPAI is a pathogenicity island in the *H. pylori* genome and encodes numerous genes that, upon cell contact, are expressed and assembled into the needle-like type 4 secretion system (T4SS) (Rohde et al. 2003). The T4SS is evolutionarily conserved among many Gram-negatives such as *Agrobacterium*, *Bordetella*, and *Legionella*, but differs in different organisms in terms of what substrates are transferred. *H. pylori* T4SS binds the integrin β1 receptor that is located on the basal membrane and transfers the cytotoxin associated gene A (CagA) which is also encoded by the cagPAI (Odenbreit et al. 2000; Backert et al. 2000; Stein et al. 2000; Kwok et al. 2007; Jiménez-Soto et al. 2009). Once inside the cell, CagA is phosphorylated on specific EPIYA motifs by host kinases, and phosphorylated CagA goes on to manipulate the cell by interacting with numerous host cell proteins. In addition, injected non-phosphorylated CagA manipulates proliferation and immune response of host cells (M. Suzuki et al. 2009). Cultured epithelial cells respond by forming the characteristic “hummingbird phenotype” that is the effect of both cell scattering and elongation (reviewed in Tegtmeyer et al. 2011). CagA is not considered as a virulence factor (although the cagPAI is), but it is considered an oncoprotein and is associated with development of gastric adenocarcinoma (Parsonnet et al. 1991). *H. pylori* infections of Mongolian gerbils resulted in more gastric adenocarcinomas in a CagA-dependent manner and so did mice that were transgenic for CagA expression (Ohnishi et al. 2008; Franco et al. 2005).

**VacA**

Vacuolating cytotoxin A (VacA) is a multifunctional secreted cytotoxin. The vacA gene is found in all *H. pylori* isolates though there are differences among the alleles. The s1 allele, especially in combination with the m1 allele, is highly associated with the risk of developing peptic ulcers and gastric cancer (reviewed in Palframan et al. 2012). The VacA toxin forms large vacuoles in gastric cells; however such vacuoles are not seen in biopsies. VacA localizes to, and exerts effects on, the mitochondria where it triggers the apoptotic cascade and induces cell death by mitochondrial fission. The detailed molecular mechanisms for this, however, are not known (Jain et al. 2011) and (reviewed in Palframan et al. 2012). In addition, VacA has been
found to bind the integrin subunit CD18 on T-cells and suppressing their activities (Gebert et al. 2003; Sewald et al. 2008).

**Outer membrane proteins**

The OMPs of *H. pylori* comprise about 4% of the genome, more than in any other bacterial species. They are divided into five families with the largest group being the Hop family. The Hops are comprised of 21 genes that all encode a conserved N-terminal motif (A↓EX[D,N]G) (Alm et al. 2000). Related to the Hop family are the 11 Hor genes, which are homologous in sequence to the Hop family but lack the N-terminal motif. Together the Hop and Hor families form the major OMP family (Alm et al. 2000). The C-terminal domains of the OMPs are homologous and have alternating hydrophobic and hydrophilic domains that are proposed to form anti-parallel amphipatic β-sheets that assemble into a β-barrel in the outer membrane (Bina et al. 2000; Tomb et al. 1997). Phylogenetic analyses based on the homologous C-terminal domains reveal clustering of certain genes. High nucleotide similarities in the 5’ and 3’ regions of the genes indicate a preference for recombination between these alleles and this has indeed been show for babA, babB, and babC, and more recently for sabA, sabB, and hopQ as well (Bäckström et al. 2004; Solnick et al. 2004; Talarico et al. 2012). Located between the N- and C-terminal domains, and exposed to the extracellular environment, is the hypervariable region that is unique for each OMP and is proposed to determine their functions. This variable region not only varies between the OMPs, but also varies for the same OMP between different strains (Alm et al. 1999). Because this region is exposed to the outside of the cell, it is part of the ‘face’ towards the immune response and is, therefore, under constant selective pressure to avoid recognition.

In relation to the diversity seen in the variable region, some genes (sabA, sabB, hopZ, oipA, babB, and babC) are also proposed to have CT repeats that can be regulated by slipped-strand mispairing to undergo phase variation (Alm et al. 1999; Tomb et al. 1997; Oh et al. 2006). By readily turning proteins on and off, the bacterial community is more dynamic and able to prepare for sudden changes that can occur during the course of the infection.

Although functions are still not assigned to all OMPs, their cellular localization hints that these proteins are involved in adherence or transport of nutrients. The first five characterized OMPs were classified as porins, and two of the OMPs, BabA (HopS) and SabA (HopP), are defined and characterized as adhesins (Ilver et al. 1998; Mahdavi et al. 2002). Many others are proposed to be involved in adhesion such as HopZ, HopQ and
AlpAB (HopB and HopC) but no receptors for these proteins have yet been identified (Peck et al. 1999; Loh et al. 2008; Senkovich et al. 2011).

The OMPs are thought to have been acquired as a single gene event that underwent duplication and divergence before the speciation of *H. pylori* because they are also present in their closest relative, *Helicobacter acinonychis*, though the gene repertoire is somewhat dissimilar (Gressmann et al. 2005). The evolution of the OMPs has probably been a continuous process because some genes are still found as duplicates (hopJ/K and hopM/N) indicating a recent duplication event (Alm et al. 2000). The geographic variation that can be seen for other virulence factors of *H. pylori* such as CagA and VacA is also seen for the OMPs. Kawai et al. has recently published a comparison between 20 complete genomes from East Asia, Europe, West Africa, and among Amerindian populations and found distinct differences in the prevalence of *babC* (HopU), which only seems to be present in the European strains (Kawai et al. 2011).

**Adherence**

Bacterial cells that are attached to the host cells have a benefit over free-swimming, planktonic bacteria because of the higher supply of essential nutrients, especially iron, close to the host cells. A tight adherence to the host cells also enables the bacteria to manipulate and control the immune response through delivery of their effector molecules and to promote invasion for an intracellular lifestyle. Adhesins are proteins expressed by the bacteria that mediate attachment to specific receptors on the host cell surface. FimH and PapG are the two most described bacterial adhesins are so called fimbrial adhesins that are found at the tip of fimbriae. Afimbrial adhesins, such as in *H. pylori*, are found integrated in the membrane. Such adhesins are proposed to promote more cellular damage because of their tighter association with the host cell, and have been demonstrated to be important for the delivery of effector molecules (reviewed in Kline et al. 2009). The most common receptors for bacteria and viruses are carbohydrates on glycoproteins and glycolipids. Such interactions are typically weak, but by combining adhesins and adhering to glycosylated domains in the host membrane, such as lipid rafts, multivalency increases the affinity and is commonly known as the Velcro™ effect (reviewed in Moran et al. 2009).
Receptor specificity

The interactions between adhesins and their cognate receptors determine, to some extent, the host specificity and tissue tropism (Klemm & Schembri 2000). Some bacteria, like *H. pylori*, colonize a narrow niche whereas other bacteria can colonize and spread through multiple niches. A shift in niche is sometimes preceded by a switch in receptor phenotype that benefits an adherent clone over the non-adherent clones, and consequently leads to outcompitition by the adherent clone. A rapid way of acquiring new receptor specificity is to express multiple adhesins that can easily be switched on and off to fit the local receptor repertoire. Such multiple adhesins are present in many bacterial species e.g., *E. coli*, where different pathogenic strains possess distinct adhesin-receptor repertoires (reviewed in Kline et al. 2009). Additionally, allele changes could arise resulting in a change of the protein structure, either directly in the binding site or at a distant site causing secondary effects to the conformation or orientation of the binding site. Such small changes can be acquired by point mutation, recombination, allelic exchange, or gene rearrangement (Ofek & Doyle 1994). PapG is the adhesin of the P-fimbriae in uropathogenic *E. coli* and has been demonstrated to regulate its tissue tropism by allelic exchange. Initial receptor specificity studies by Strömberg et al. identified three allelic variants of the adhesin based on differential recognition of isoreceptors for the globoseries glycosphingolipids (GbO) (Strömberg et al. 1991). All alleles are specific for the GbO determinant Gal(α1-4)Gal epitope, but different specificities for the extended GbO structures discriminate among the alleles. The PapGI adhesin binds GbO₃, expressed on uroepithelial cells, though this allele is uncommon. PapGII, which is correlated with pyelonephritis and bacteremia, prefers GbO₄ that is also found on uroepithelial cells. PapGIII is correlated with human cystitis and genitourinary infections in cats and dogs and prefers GbO₃, Forssman, or GloboA, the latter of which is found in the human urinary tract (Strömberg et al. 1991; Lindstedt et al. 1989) (reviewed in Lane & Mobley 2007). Such drifting in receptor specificity is not restricted to bacteria as viruses are also known to possess such capability. Recently, point mutations in the hemagglutinin receptor binding site of avian influenza virus (H5N1) were shown to promote airborne transmission between ferrets (Herfst et al. 2012; Morens et al. 2012).
Detachment

Despite the importance of attachment for colonization, it is also an obstacle that needs to be overcome under certain circumstances. A tight adherence could trigger a tremendous immune response that could be risky for the attached bacteria. A strong permanent attachment would also limit the possibility to inhabit other surfaces. An obvious way for a bacteria to cope with this is to simply turn off expression of the adhesin, but not all situations support this type of regulation. Such an environment is in the urinary tract, where a loss of adherence during conditions of low urinary flow would present a risk for the bacteria of being flushed out during urination. Both commensal and uropathogenic E. coli express the FimH adhesin, the tip protein of Type I fimbriae. The most common receptor for FimH is trismannose and enhanced binding is accomplished by multivalency. Uropathogenic strains, especially those causing pyelonephritis, are found with a different allele that, in addition to the binding of tri-mannose, also can bind mono-mannose that is common in the urinary tract, binding to mono-mannose had 10 times higher affinity (Sokurenko et al. 1997). Such FimH adhesins increase the affinity for mono-mannose under conditions of high shear stress, such as during urination, and enable the bacteria to withstand 10-fold higher flow rates. The bacteria can then revert back to the lower affinity state when the flow has ceased (Thomas et al. 2002). Consequently, nearly permanent high affinity binding is limited to just those situations in which it is required. However, this type of regulation requires changes in flow conditions that not all colonized sites experience. Neisseria meningitidis uses Type IV pili for colonization of the nasopharynx epithelium where it multiplies to microcolonies and then detaches and spreads to a new host or causes an invasive disease. It has recently been shown that detachment of N. meningitidis from epithelial surfaces is phosphorylation-dependent and caused by increased expression of the pptB gene during the multiplication phase. Phosphorylation of a serine on the pili causes destabilization of the pilus fiber and subsequent detachment, thus enabling the bacteria to restart the infectious cycle (Chamot-Rooke et al. 2011).

H. pylori adherence and detachment

During infection, around 20% of the H. pylori infectious load bacterial cells are found adherent to human gastric mucosa (Hessey et al. 1990). The adhesins of H. pylori are membrane-bound, and this enables a more intimate contact with the host cells compared to fimbrial adhesins.
Adherence of *H. pylori* is beneficial for the colonization, establishment, and propagation of infection due to three factors (reviewed in Amieva & El-Omar 2008): 1) adhesion enables an intimate contact with host cells that can induce cellular damage and inflammation that will lead to an increased release of nutrients by the host cell; 2) adhesion helps the bacteria to avoid mechanical clearance and supports invasion and persistence; and 3) the host cell surface could be used as a site for *H. pylori* replication as recently demonstrated by Tan et al. 2009. Nonetheless, the majority of *H. pylori* cells are found free-swimming, which can be rather challenging for the bacteria.

In the outer most part of the mucin layer, the immune response is less active but the bacteria are at increased risk of swimming towards the acidic lumen. Therefore, attachment would appear to be a solution to this problem. However, close to the epithelial cells peristaltic movements and immune responses are more active and the bacteria are at risk of being cleared. Also, epithelial cells have a short lifespan of around 3 days, and they carry any attached bacteria with them as they shed (Hidaka et al. 2001 and reviewed in Moore et al. 2011). Thus, for *H. pylori*, the attachment dilemma is particularly evident, and tight regulation of attachment and detachment is necessary for the bacteria to persist in the proper environment. In addition to phase variation, that disassembles the bacterial adherence machinery, we have recently identified a functional regulation of *H. pylori* adhesins by affinity reduction in an acidic environment. This helps the bacteria to detach from the epithelial cells before they reach the acidic lumen enabling them to use chemotaxis for proper orientation back towards the epithelium (Paper I).

Adherence by *H. pylori* must also be flexible enough to adapt to changes in the receptor repertoire that occurs during infection and is triggered by both the bacteria and the host cells. To date only two *H. pylori* adhesins are characterized; BabA and SabA. BabA binds almost constitutively to ABO/Leb blood group antigens whereas the more phase-variable SabA binds inflammation-associated sialylated structures (Borén et al. 1993; Ilver et al. 1998; Mahdavi et al. 2002; Aspholm-Hurtig et al. 2004). In addition to adherence to the epithelial cells, BabA and SabA are known to interact with their carbohydrate receptors when they are expressed on the mucins that serves to protect the gastric mucosa from acidity (described in the section Mucin interactions) (Van de Bovenkamp et al. 2003). Numerous additional OMPs have been demonstrated to be adherence-related, such as HopZ, AlpAB, HorB, and HopQ, though no clear mechanism or receptor has been described (Snelling et al. 2007; Loh et al. 2008; Peck et al. 1999; Odenbreit et al. 2002; Senkovich et al. 2011).
**BabA**

The Blood group Antigen Binding Adhesin, BabA, mediates binding to the ABO/Leb blood group antigens. The first hint of the existence of an adhesin was provided by application of FITC-labeled *H. pylori* to paraffin-embedded tissue sections of human gastric mucosa and observed adherence to the foveolar epithelial cells (Falk et al. 1993). Inhibition with various substrates such as human colostrum from secretors and non-secretors, antibodies, and glycoconjugates identified the receptor as the H1 and Lewis b (Leb) blood group antigens (Borén et al. 1993). H1/Leb are terminal carbohydrate structure that defines blood group O. They are found on red blood cells and on gastro-intestinal (GI) epithelial linings such as in the stomach (described in the section The receptors for *H. pylori*). Related to H1 and Leb structures are the A and B blood group antigens, but neither of these structures were identified as receptors, nor did *H. pylori* bind the related Lea structure, demonstrating specificity for a fucose moiety (Borén et al. 1993). The cognate adhesin, Blood group Antigen Binding Adhesin (BabA), was identified by use of the re-tagging technique (Ilver et al. 1998). This technique utilizes a cross-linker attached to the receptor glycoconjugate. Binding of this receptor exposes the adhesin for cross-linking and enables detection by streptavidin-biotin via SDS-PAGE and mass spectrometry to identifies the cognate adhesin. The strain used in these studies, CCUG17875, was found to contain two babA alleles of which one is silent due to defects in the translational sequence and signal peptide (Bäckström et al. 2004; Ilver et al. 1998). This allele was called babA1 and the allele encoding a functional BabA protein was called babA2. Deletion of babA2 confirmed that the BabA protein is the functional adhesin for binding to H1/Leb. During these studies, another gene was identified and called babB. It was homologous to babA at the 3’ and 5’ ends but divergent in the middle (Ilver et al. 1998). Both of these proteins were later classified as belonging to the Hop family of OMPs (Alm et al. 2000).

H1 and Leb were identified as receptor structures for BabA, whereas ALeb and BLeb did not demonstrate any binding to the Peruvian isolate that was used in the original studies (Borén et al. 1993). When Aspholm-Hurtig et al. investigated the worldwide receptor specificity for BabA, they identified strains that, in addition to Leb/H1, could also bind the ALeb and BLeb structures. These strains were termed ‘generalists’ while those that were restricted to Leb/H1 were termed ‘specialists’ (Aspholm-Hurtig et al. 2004). Interestingly, the specialist strains originated from parts of the world, such as South America, where O is the predominant blood group. The affinity of BabA for its receptors is high, with values in the μM to nM range, though
they vary substantially between clinical isolates (Aspholm-Hurtig et al. 2004). Interestingly, the affinity for H1/Leb is stronger than for ALeb or BLeb, and this could explain the higher risk of peptic ulcer disease in blood group O individuals (Aspholm-Hurtig et al. 2004 and reviewed in Anstee 2010). In addition to the BabA-mediated attachment to the epithelial cells, BabA also interacts with the mucins MUC5AC and MUC1 (described in the section Mucin interactions).

Clinical isolates have exceedingly diverse babA sequences, with the highest variability in the middle domain. In addition, not all strains express BabA, and not all expressed BabA proteins are functional (Paper II) (Aspholm-Hurtig et al. 2004). Thus, the presence of a babA gene is not necessarily evidence of a functional BabA protein. The mechanism by which this diversity is obtained is not known, but phylogenetic analyses of the BabA variable region in different populations reveals heterogeneous selective pressures, such as escape from host immune response, receptor specificity, and affinity, that act on the protein (Aspholm-Hurtig et al. 2004). During H. pylori infection studies in Rhesus macaques, mice, and gerbils, expression of BabA was frequently lost (Solnick et al. 2004; Ohno et al. 2011; Styer et al. 2010). The loss of BabA in Rhesus macaques might be because of a higher inflammatory response to BabA-expressing adherent bacterial cells. However, it could also be because of the lower prevalence of the ABO/Leb blood group antigens in gastric mucosa during infection and hence selection against BabA-expression (Lindén et al. 2008). Interestingly, BabA is not immunogenic in humans indicating an evolved function to resist human immune responses (Kimmel et al. 2000; Haas et al. 2002).

*babA, babB, and babC recombination*

Phylogenetic analysis of the OMP C-terminal regions reveals a close genetic relationship between BabA, BabB, and BabC where BabA and BabB are most similar. Despite this similarity, only BabA has been assigned a function (Alm et al. 2000). BabA and BabB are most prevalent, and BabC is less so and was recently associated with strains of European origin (Kawai et al. 2011). The three alleles are commonly found at three different loci, A, B, or C, that are in each third of the genome (Tomb et al. 1997). The existence of three similar OMPs suggested that H. pylori utilizes them to switch antigenic appearance to avoid the immune response (Alm et al. 2000). Indeed, after the first genome comparison between strains 26695 and J99, it was clear that babA and babB were found at inverted loci. In addition, 26695 has the babC gene at the C locus, which J99 is missing (Alm et al. 1999). Comparisons of clinical isolates demonstrated that these alleles frequently recombine and switch loci to form duplicates, deletions, and chimeric genes at a frequency...
of about $3 \times 10^{-6}$ gene conversions per cell division (Colbeck et al. 2006; Pride & Blaser 2002; Hennig et al. 2006; Amundsen et al. 2008). Interestingly, the C-terminal regions of BabA and BabB are more homologous within a genome/strain than between genomes/strains, and this suggests a concerted evolution between BabA and BabB (Pride & Blaser 2002). No such analysis has been done for BabC, but the close relationship between these genes suggests similar results. The Leb-binding function of BabA is not affected by the loci from which it is expressed (Hennig et al. 2006). However, expression of BabA from the B or C loci could result in a gain of function because BabA could then make use of their CT-repeats for faster switching of expression (Bäckström et al. 2004; Colbeck et al. 2006). The recombination frequency is very high, and a seemingly homogeneous population from a single clone can retain a subpopulation of single clones that demonstrate an altered arrangement of the bab genes caused by recombination (Colbeck et al. 2006). Such rearrangements can also revive lost BabA expression demonstrating that loss of a functional BabA is reversible (Bäckström et al. 2004). Recombination is also seen during experimental infections, such as in the Rhesus macaque model, where the expression of BabA is sometimes switched off by babA-babB recombination.

BabA and pathogenesis

The pathogenic importance of an ABO/Leb-mediated attachment by BabA was demonstrated in H. pylori-infected, Leb-transgenic mice that had higher inflammatory scores compared to their non-transgenic littermates (Falk et al. 1995; Guruge et al. 1998). This study was later confirmed in Mongolian gerbils where the animals infected with a babA mutant had lower infiltration of inflammatory cells and a reduced cytokine response (Sugimoto et al. 2011). In 1999, Gerhard et al. updated the old Type 1 (CagA+, VacA+) and Type 2 (CagA-, VacA-) classifications of H. pylori by adding a third denominator, babA2 (Gerhard et al. 1999). Genotypic studies demonstrated that babA2 was correlated with cagA and the more pathogenic vacAs1 allele and was significantly associated with adenocarcinoma and better discriminated against gastritis compared to the conventional Type 1 vs. 2 definition. These results indicated that BabA expression has an important role in the disease process, and these strains were termed ‘triple-positive strains’ (Gerhard et al. 1999). Many groups tried to repeat the task of correlating babA2 with gastric disease but obtained inconclusive results (reviewed in Yamaoka 2008). Because the prevalence of babA2 is not equivalent to expression of BabA, additional studies have investigated the correlation with BabA expression. Such studies demonstrated a correlation between BabA, CagA, and VacA, and also to more severe gastric disease such as intestinal metaplasia (Odenbreit et al. 2009; Yamaoka et al. 2002;
Yamaoka et al. 2006; Azevedo et al. 2008) (Paper III). Interestingly, Fujimoto et al. demonstrated a stronger correlation with duodenal ulcer and gastric cancer for strains with low levels of BabA expression than for those with high or no expression (Fujimoto et al. 2007).

Evidence for the functional correlation between BabA, CagA, and VacA is now being elucidated. Adherent *H. pylori* are frequently found associated with the intercellular junctions where they would have immediate access for penetration of the mucosa (Amieva et al. 2003). Such tight adherence, mediated by BabA (or other adhesins), simplifies the secretion and delivery of VacA to host cells that could trigger separation of the cellular junctions and facilitate penetration of *H. pylori* through the intercellular space. On the basal side, these bacteria have access to the T4SS integrin β1 receptor that is located where the bacteria can deliver CagA (reviewed in Wessler & Backert 2008). Indeed, CagA-positive bacteria are found more tightly associated with epithelial cells during infection (Camorlinga-Ponce et al. 2004). This hypothesis has recently been supported by Ishijima et al. who demonstrated that the BabA-Leb interaction increases T4SS-mediated induction of mRNAs for pro-inflammatory cytokines and precancerous factors in cell cultures, and stimulates intracellular levels of phosphorylated CagA (Ishijima et al. 2011). BabA, SabA, and CagA have also recently been demonstrated to share regulatory mechanism upon interaction with mucins (Skoog et al. 2012). Although not all studies were able to demonstrate a correlation between BabA, CagA, and VacA in disease progression, there is a strong indication that BabA itself can cause damage to the epithelial cells. A recent finding has suggested that BabA is involved in induction of double-strand breaks and such grave DNA damage could tilt the disease state towards more aggressive pathogenesis (Toller et al. 2011). However, it is difficult to determine if these effects are mediated by direct binding to the receptor, or are secondary and caused by the tighter association with host cells.

**SabA**

SabA, the Sialic-acid binding adhesin, was discovered some years after BabA. Mahdavi et al. observed adherence of the CCUG17875 babA2 mutant to human gastric mucosa from a gastritis patient (Mahdavi et al. 2002). The receptor was characterized to be sialyl-dimeric-Lewis x antigen (sdiLex), but more detailed analysis of the binding specificity has identified the minimal binding epitope to be NeuAcα2-3Gal with a polymorphism for the core chain (Mahdavi et al. 2002; Aspholm et al. 2006). Such sialylated antigens are expressed by inflamed tissues to recruit neutrophils and are, therefore, triggered by *H. pylori* infection. SabA expression is highly variable with both
a PolyT tract in the promoter and CT-repeats in the coding region, and it has also recently been shown to be controlled by the acid-responsive ArsRS two-component system that also regulates urease and carbonic anhydrase (A. C. Goodwin et al. 2008). Similarly to the recombination between babA and its related alleles, sabA can also recombine with its related allele sabB, and to some extent with hopQ (Talarico et al. 2012). This occurs at a frequency of $1.4 \times 10^{-9}$ gene conversions per cell generation, which is lower than that seen for babA and babB recombination.

**Mucin interactions**

Mucins are large, heavily O-glycosylated proteins and represent the major component of the mucus layer. This layer covers and protects mucosal surfaces from the external environment such as in the stomach where it also serves to maintain the pH gradient (Bhaskar et al. 1992). Healthy gastric mucosa contains the two secreted gel-forming mucins, MUC5AC and MUC6, and the cell surface-bound MUC1. MUC5AC is the major constituent of the mucus and is secreted from the surface foveolar cells and pit cells and is reflective of their glycosylation pattern. MUC6 is secreted from the glands in lower amounts (Phillipson et al. 2008). *H. pylori* specifically interacts with MUC5AC via the BabA adhesin, and this interaction triggers proliferation and faster division of the bacterial cells (Van de Bovenkamp et al. 2003; Skoog et al. 2012). Contradictory reports have been written about changes in distribution of these mucins during *H. pylori* infection. Byrd et al. reported a downregulation of MUC5AC and upregulation of MUC6, but later studies by Teixeira et al. and Van de Bovenkamp et al. saw no such changes (Byrd et al. 1997; Teixeira et al. 2002; Van de Bovenkamp et al. 2003). Though the results are inconsistent, similar effects are seen in acute infections in Rhesus macaques (Cooke et al. 2009; Marcos et al. 2008). A downregulation of MUC5AC could be triggered by the bacteria, as it would thin the mucin layer and diminish the Leb structures in the mucus that otherwise might hinder the bacterial cells from reaching the cell surface. The increase in MUC6 could reflect a host response because, as described previously, *H. pylori* is unable to colonize MUC6-producing regions due to the toxic effects of the α1,4-GlcNAc moieties (Kawakubo et al. 2004) and (reviewed in Moran et al. 2011).

MUC1 is a membrane-bound mucin found on the surface epithelium and glandular tissues. Both BabA and SabA can mediate binding to Muc1 and Muc1 could potentially mediate cell signaling through its cytoplasmic tail (Lindén et al. 2009) and (reviewed in Moran et al. 2011). However, the interaction of *H. pylori* with MUC1 triggers shedding of its highly
glycosylated extracellular domain, thus acting as a releasable decoy. MUC1 also blocks bacterial cells that are unable to bind the mucin and acts as a steric hindrance at the cell surface (Lindén et al. 2009; McGuckin et al. 2007). This is supported by findings that alleles coding for shorter versions of MUC1 can enable a closer contact between *H. pylori* and host cells and are associated with more severe gastritis and gastric cancer (Carvalho et al. 1997; Vinall et al. 2002).

**The receptors for *H. pylori***

Surrounding the surface of all cells is the glycocalyx, a halo-like structure that is a result of substantial post-translational glycosylation that takes place in the cell. Glycosylation is the most diverse post-translational modification because it is not encoded in the genome but is accomplished by the action of several glycosyl transferases that polymerize monosaccharides into long chains of polysaccharides. These chains can be found as glycoconjugates, connected to proteins or lipids, but also as free glycans. The polysaccharide chains vary in linkages, are linear or branched, and can have different endings such as acetylation or sulfation. It is estimated that 1–2% of the human genome is involved in the synthesis of glycans, and that more than 70% of human proteins are glycosylated (Apweiler et al. 1999). In addition, glycosylation is tissue and cell specific, and glycosylation levels change in response to intrinsic or extrinsic factors and during disease progression. One such disease is cancer where several specific glycans are involved in the pathophysiological steps that are characteristic of malignancy (reviewed in Fuster & Esko 2005). Recent advances in new technologies for studying glycans has resulted in high sensitivity mass spectrometry, high throughput HPLC, and glycan arrays (Adamczyk et al. 2012; Rillahan & Paulson 2011; Blixt et al. 2004). This has made glycobiology an emerging field in not only the diagnosis and treatment of cancer, but also, with the use of glycan arrays, in the study of pathogens and their receptor specificities.

**The ABH antigens**

The ABH-blood group antigens are the terminal carbohydrate structures that define the human blood groups. Despite the name, most of the ABH antigens are not found on erythrocytes but instead in other tissues such as gastric mucosal cells and their secretions, saliva, milk, and tear fluid where they are designated 'histo-blood group antigens'. These terminal structures are found on four different poly-N-acetyllactosamine core chains: Type 1 (Galβ1-3GlcNAcβ1-R) and Type 2 (Galβ1-4GlcNAcβ1-R), which are N-linked, Type 3
(Galβ1-3GalNAcα1-R), which is O-linked, and Type 4 (Galβ1-3GalNAcβ1-R), which is found on glycolipids. The majority of the antigens are found on Type 1 and 2 core chains. Differential expression of these core chains together with the substrate specificity of the glycosyl transferase determines what terminal structure will be formed. In the stomach, the ABH antigens are found on the Type 1 chain in epithelium and foveolar cells and on the Type 2 chain in the deeper glands, though the border between the two is not always obvious. Type 3 is found on stomach mucins. BabA-mediated binding of *H. pylori* is restricted to the terminal structures of the Type 1 and Type 3 chains, and there is no binding to the Type 2 chain or its derivatives.

There are three antigens defining the blood groups A, B, and H (blood group O) in which H is a disaccharide and the precursor of the A and B trisaccharides. These antigens are synthesized by enzymes encoded in three loci (H, Secretor (Se), and ABO). The H-determinant is formed by an α1-2 fucosyl transferase (FucT) encoded by the H or Se locus. The H α1-2FucT is almost 100% prevalent in humans and forms the H-antigen on Type 2 or Type 4 core chains on erythrocytes. However, in epithelial cells and in secretions the Se α1-2FucT is the active enzyme and forms the H-antigen on Type 1, 2, and 3 core chains.

The ABO locus is the blood group-determining locus because it contains the A and B alleles as well as the O allele that encodes a non-functional enzyme. The A and B alleles each encode the A or B transferase that strictly modifies the fucosylated galactose of the H-antigen with a GalNAc (A) or Gal (B). The A and B transferases are not 100% efficient, resulting in residual levels of A and B structures in the tissues.

**Secretors and non-secretors**

Around 20% of Caucasians are homozygotes for a non-functional Se α1-2FucT, mainly caused by a G428A point mutation in the gene that results in an early stop-codon and a non-functional protein (Kelly et al. 1995). Because the A and B transferases are specific for the H-determinant, this results in an almost complete lack of all histo-blood group antigens on epithelial cells and in secretions. Such individuals are termed ‘non-secretors’. Another point mutation in the Se α1-2FucT, A385T, renders the enzyme weaker and fewer histo-blood group antigens are formed. This phenotype is more common in Asia and these individuals are termed ‘weak secretors’ (Henry et al. 1996; Lindén et al. 2008 and reviewed in Anstee 2010). Neither of these phenotypes shows any obvious negative side effects. On the contrary, non-secretors are protected from infection by some Noroviruses that use the ABO
antigens as receptors (Thorven et al. 2005; Carlsson et al. 2009). In very rare cases, there is a complete lack of both the H and Se transferases, termed the ‘Bombay’ phenotype, and such individuals are completely devoid of ABO antigens but they do not demonstrate any obvious side effects.

Figure 4 Schematic illustration of the synthesis of the ABO/Leb-antigens

**Lewis antigens**

Related to the ABH blood group antigens are the Lewis antigens (Figure 4). These are the result of concerted action between the Se transferase and the Lewis (Le) transferase, which is encoded by the Le locus. The Le transferase is a fucosyl transferase that modifies the ABO antigens (H1, A1, and B1) by adding a fucose in an α1-3 or α1-4 linkage to the sub-terminal GlcNAc
thereby forming Leb, ALeb, and BLeb, respectively. The concurrent expression of the Le transferase and Se transferase converts almost all ABO antigens expressed on the stomach epithelium into Le antigens. In non-secretors, Lea is formed from the unmodified Type 1 core chain and is, therefore, the main antigen found in the stomach mucosa of non-secretors. Lea is not a substrate for the Se-transferase, thus minute levels of Lea will, therefore, be found in secretors. The Le transferase also modifies Type 2 chains generating Lex (corresponding to Lea) and Ley (corresponding to Leb). Because of the location of these core chains, such antigens are predominantly found in the glandular region. Le antigens can be detected on erythrocytes but, despite the specificity of the Le transferase for Type 2 chains, it does not modify ABH blood group antigens on erythrocytes. These Le antigens are instead glycolipids that have passively adsorbed onto the membrane.

Ten percent of the European population is Lewis-negative in that they lack a functional Le transferase. Together with the phenotypes of the secretor transferase (secretor, weak secretor, or non-secretor), this generates four different phenotypes designated by the antigens that are expressed: Lea-b-, Lea-b+, Lea+b-, and Lea+b+.

The lack of terminal structures on the Gal of Lea and Lex makes them targets for modifications such as sialylation and sulfation. Such antigens are common in healthy gastric mucosa of non-secretors, but not of secretors. However, their expression is triggered during inflammation to recruit neutrophils and is also triggered by *H. pylori* during infection (Sakamoto et al. 1989; Mahdavi et al. 2002; Lee et al. 2006; Lindén et al. 2008; Marcos et al. 2008).

**Aims of the Thesis**

1. To investigate how *H. pylori* adherence to human gastric mucosa is affected by low pH.
2. To investigate adherence mediated by the European non-Leb binding strains.
3. To investigate how secretor status affects the *H. pylori* infection.
4. To create transgenic cell-lineages for detailed investigation into BabA-Leb interactions.
Results

Paper I

pH Regulated H. pylori Adherence: Implications for Persistent Infection and Disease

This study focused on the attachment dilemma of H. pylori, where a tight adhesion to the gastric cells during infection could be detrimental because permanently attached bacteria would be lost when dead cells shed into the acidic lumen.

We initially observed that in vitro adhesion of FITC-labeled H. pylori strain 17875 was lowered if the adhesion was carried out at a pH below 7. The loss of adherence was attributed to a reduction in affinity for the Leb receptor, and, interestingly, the ability to bind Leb was fully restored when pH was reconditioned and returned to neutral. We compared this property of BabA with other lectins, and despite these other lectins also losing binding at low pH, none of them had the ability to recover their binding properties when the pH was neutralized. The stability of complexes of bacterial cells, Leb conjugates, and BabA at low pH indicates that the acid-lability is attributed to a functional property of the BabA adhesin. We detected a similar property for the SabA adhesin, although SabA was more sensitive to low pH compared to BabA.

The polymorphism of BabA in terms of both affinity and specificity for the receptor is also reflected in its acid-labile property. In a large-scale screen of clinical isolates, we found that while some strains still bound Leb at pH 2, other strains had lost most of the binding already at pH 4. We concluded that the acid lability of the BabA-Leb interaction is reflective of their diversity.

Purified BabA from H. pylori demonstrated a specific binding to the human gastric mucosa at neutral pH, and low pH caused detachment of BabA. However, pretreatment of BabA by low pH followed by neutralization restored binding. Pure BabA from two bacterial isolates with different acid-lability were compared in a Leb-binding ELISA. The acid-lability that each protein demonstrated correlated well with the acid-labile binding that the bacterial cells of each strain demonstrated. These results confirmed the finding that the BabA protein has a unique ability to restore its binding ability after exposure to low pH.
To determine how the structure of BabA is affected by low pH, we analyzed BabA protein stability at lower pH by circular dichroism. We only detected subtle changes, probably because of rearrangements in the α-helices, and no further changes to the global structure occurred below pH 2.5. Crossover experiments of BabA from an acid-tolerant isolate to an acid-sensitive isolate showed that the recipient improved in Leb binding and became almost as acid-tolerant as the donor. This demonstrates that most of the unique acid-sensitivity in BabA-mediated binding is encoded within the babA gene.

We further tested pH tolerance/sensitivity of isolates from the same individual, but collected from different regions of the stomach that are characterized by low vs. high acidity (antrum and corpus). The aim was to test if such putative differences could relate to the distinct acidity of their gastric niche. One patient, ‘Örebro 2’, demonstrated differences in acid-lability between isolates from the antrum vs. corpus tissue. Sequencing revealed close to identical babA sequences with the exception of two amino acid positions. In the antrum strain, we detected Leu in the putative carbohydrate binding domain (CBD), i.e. the Leb-binding site, and Glu in an area distant from the binding site, the C-terminal domain. In the corpus strain, these amino acids were, instead, Pro and Gly. We cloned these babA genes and expressed them as recombinant BabA. In addition, we also constructed recombinant BabA where the two amino acids were swapped giving a mosaic BabA genotype. These recombinant proteins were tested for Leb binding at lower pH by surface plasmon resonance (SPR). The recombinant proteins reflected the acid-sensitivity of the antrum isolate and the more tolerant corpus isolate. In the mosaic BabA, we detected no differences in acid-lability when exchanging Glu for Gly, but exchanging of Leu for Pro revealed a shift towards higher acid tolerance suggesting that the Pro residue is important for acid tolerance.

H. pylori-related diseases are reflective of the different acid profiles of the stomach: ulcers are typically related to high acid secretion and cancer to low acid secretion. A comparison of acid-lability from high peptic ulcer (HU) vs. high cancer (HC) prevalent regions revealed that strains from HU regions were more acid sensitive in binding compared to HC regions. These findings were somewhat contradictory but might reflect the different niches that these strains colonize. Duodenal ulcers are associated with high acid-secretion that typically arises from an antrum-predominant colonization where bacteria in the antrum are adapted to the acidic antrum conditions. Sequencing of babA from these regions showed that almost all BabA from HC regions have the Pro-position in the potential CBD. In addition, sequences from HU regions had several amino acids missing in this region as
they lacked the Pro and contextual amino acids. Instead HU strains had required the antrum-associated Glu position in the C-terminal. These results further emphasize the importance of these amino acids for adapting to local pH and regional gastric patterns of disease.

To investigate evolution in BabA acid-lability during chronic infection, we analyzed *H. pylori* from Rhesus macaques that had been experimentally infected for five years. From one monkey, we found clones from the antrum and corpus that differed in acid-lability. Antrum isolates and the original infecting strain were similar in acid-lability whereas the corpus clone was more acid-sensitive. We sequenced *babA* from these clones and found identical sequences in input and output bacteria from the antrum except for some mutations in the β-barrel domain. However, in the corpus clone we identified several amino acid differences in the same region as the Glu and Gly in the Örebro 2 antrum and corpus. These mutations were confirmed to affect acid-lability by SPR experiments, which also confirmed that this corpus strain was similar to the corpus strain from Örebro 2. We could trace these mutations in the macaque corpus BabA as coming from a recombination event between *babA* and a *babA* homolog.

In conclusion, we propose that the unique method by which *H. pylori* regulates attachment and detachment is vital during infection. By keeping the binding activity of the adhesins acid-regulated, *H. pylori* is able to detach from a dead cell that moves towards the toxic acidic lumen, orient itself by chemotaxis towards the epithelial cells, and start a new cycle i.e. biopanning for the most fitting phenotypes and then recycle the infection.

**Paper II**

**Clinical isolates of *Helicobacter pylori* demonstrate alternative BabA-mediated binding properties for adherence to the gastric mucosa**

This project was initiated during the large screen for ABO/Leb binding among different populations (Aspholm-Hurtig et al. 2004). The phenotype of Leb binding in European populations decreased towards southern Europe. We were interested to test if these strains utilize alternative mechanisms for adherence to gastric mucosa.

We initially screened strains from Spain for Leb binding properties and *in vitro* adhesion to human gastric mucosa. Most of the strains that were low in Leb binding were similarly low in adherence to gastric mucosa with some
exceptions. One isolate, 812, was low in Leb-binding as measured by RIA (RadioImmunooassay) but demonstrated strong adherence to the human gastric tissue sections. We expanded the screen to include strains from northern Scandinavia, Sweden, Spain, and Portugal. We correlated Leb binding activity with expression of BabA and identified a set of strains that were low in Leb binding but high in BabA expression, indicating additional functions of BabA. We tested these strains in the in vitro adherence assay, i.e. binding to the human gastric mucosa, and found a group of strains with binding patterns similar to the previously identified 812, although only one of those strains demonstrated adherence as strong as 812. We concluded that although many strains from Europe are high in expression of BabA they still don’t adhere to gastric mucosa. The functional reason for the high BabA expression but low Leb binding in these strains needs further investigation.

We then focused our study on the adherence properties of the strain 812. Periodate treatment and trypsinization of sections from human mucosa suggests the receptor for 812 to be glycolipids. Strain 812 also showed a different adherence phenotype when we tested adherence to porcine gastric mucosa and Leb transgenic mice; the reference strain 17875/Leb (possessing ABO/Leb-mediated binding) adhered strongly to all tissues but 812 adhered weakly if at all. Interestingly, also adherence to Rhesus macaque gastric mucosa seemed to be complex; 17875/Leb could adhere to all macaque mucosal samples, 812 only bound to every second Rhesus macaque gastric mucosa.

We attempted to characterize the receptor of 812 by inhibition tests and found no inhibition by use of free Leb glycan, which argues that Leb is not the main receptor for 812. In support, both human secretor and non-secretor milk inhibit adherence of 812, which taken together suggests that 812 binds to carbohydrate that lack α1,2-fucose alt. to fucosylated antigens that are not dependent and expressed due to individual secretor phenotypes.

We found that adherence of 812 to human gastric mucosa was mediated by BabA because an isogenic BabA deletion failed to bind and because crossover experiments transferred binding phenotype of 812. However, recombinant 812 BabA demonstrated a specific but much lower binding to human gastric mucosa (Figure 10b) than was expected comparing with the strong adherence of the 812 bacterial cells (Figure 10b).

To expand the search for the cognate 812BabA adhesin binding site, i.e. the best receptor, we applied 812 to a glycan array with ~350 different carbohydrates but detected no adherence to any of these carbohydrates. We expanded the array to also include the albumin-glycoconjugates used for
RIA-assay that might present the ABO/Leb-antigens further away from the array surface. By this array, 812 binds to the series of ABO/Leb blood group antigens, and a stronger binding to ALeb and BLeb than to Leb and H1. We also analyzed binding of 812 to glycosphingolipids by TLC found affinity for A-hexa structures indicating that 812 might have shifted its receptor binding epitope towards GalNAc.

We further analyzed single colonies of 812 for Leb binding properties and found a marked heterogeneity with a mixture of three Leb-binding phenotypes, which we named “red”, “circle”, and “darker” according to their appearance on the membrane after the Leb-binding colony screening test. The dominant “red” colonies were very low in Leb binding, “circle” colonies bind similar to Leb as “red” but appeared red in the middle with a darker ring, and the “darker” colonies were stained because they bound stronger to Leb, although the frequency of “darker” colonies were as low as \( \approx 1/500 \). In addition to Leb binding, the “darker” clone also showed stronger binding to ALeb and A-hexa, as confirmed by RIA. Such an increase in Leb binding could come from higher BabA expression, but immunoblots actually revealed decreased BabA expression in the “darker” clone. Sequencing demonstrated identical BabA sequences in “red” and “circle” clones, but the “darker” one had two amino acids replaced in the C-terminal region: Asn to Ala and Arg to Lys. We could trace these amino acid replacements to a recombination event with a babA homolog. Interestingly, the deletion of this homolog had no effect on adherence.

Glycosylation is a common post-translational modification and often involves Asn, such as the Asn we found replaced to Ala in the “darker” clone. By staining carbohydrates we identified a different glycosylation pattern in the “red” clone. These experiments revealed that the “red” and “darker” phenotypes exhibit additional differences accompanying previously observed amino acid substitutions in BabA. We proceeded to study the clones that adhered to human gastric mucosa to determine the phenotype they belonged to. We made use of tissue sections from gastric mucosa as a bio-panning tool to select and to isolate a population of adherent clones. Colony screening of these clones by printing membranes and incubating with Leb as bait and further babA sequencing showed that they were of the “red”, low Leb binding genotype.

In conclusion, we have found that many of the European strains express BabA but don’t bind Leb, and that the majority of such strains cannot bind sections of human gastric mucosa. Further investigations are needed to determine if these BabA proteins have modified binding properties. Our studies focused on strain 812 that demonstrates low-affinity binding to
ABO/Leb but fully maintained strong binding to gastric mucosa. Glycan array analysis suggested that 812 has shifted its binding epitope towards the Gal/GalNAc epitope and displays increased preference for the ALeb and BLeb structures. We also detected single clones with higher Leb binding due to amino acid replacements acquired by recombination event with a BabA homolog.

**Paper III**

**Infection by Helicobacter pylori expressing the BabA adhesin is influenced by the secretor phenotype**

This study focused on investigating how the secretor phenotype, i.e. expression of ABO-antigens in secretions and the GI-tract epithelium, influences *H. pylori* infection and adherence. We hypothesized that secretors are more likely to be infected with a BabA-expressing strain.

For these types of studies, it is crucial to select non-biased subjects. We screened volunteers from a gastric pathology survey at a shipyard in northern Portugal, which could be considered a cross section of the population. BabA status was determined from biopsies by both PCR and immunofluorescence. In accordance with previous findings, the prevalence of the *babA*-gene and the BabA-protein did not correlate. Therefore, to assure that we investigated expressed BabA, we assessed BabA by histotissue immunodetection.

First, we analyzed BabA for correlation with other virulence factors. Positive BabA status was correlated with both CagA and VacAs1. This result is in accordance with the definition of ‘triple-positive strains’. We next analyzed how BabA status affects the pathological state of the stomach. We found BabA to be correlated with a higher *H. pylori* density. BabA was also found to be correlated with intestinal metaplasia and degenerative alterations. The histopathology found in BabA-positive cases is possibly related to the correlation between BabA, CagA, and VacAs1 (Gerhard et al. 1999). Of particular relevance, BabA was recently demonstrated to be important for delivery of CagA (Ishijima et al. 2011).

We analyzed secretor status in biopsies and saliva by Ulex staining, determination of Lewis status by immunofluorescence, and by PCR for the G428A mutation in the secretor fucosyltransferase FUT2. We could only detect a correlation of BabA with Ulex staining of biopsies, not with secretor status determined from saliva. It was also not correlated with Lewis
phenotypes of either gastric biopsies or saliva or with the FUT2 G428A polymorphism.

To test how BabA-mediated adherence is influenced by the phenotype of the tissues, we applied FITC-labeled 17875/Leb to sections from 61 of the gastric biopsies. A significantly higher binding was observed between Lewis phenotypes a-b (probably secretor) and a-b+ (secretor) compared to a-b- (non-secretor) and a-b+ (weak secretor), i.e. Lea staining suggested that the secretor expression is reduced. Significantly higher binding was also detected in those tissues that were strongly positive for Ulex staining and normal/heterozygous for the FUT2 G428A polymorphism, the marker for secretor positive phenotype.

In conclusion, we have confirmed a functional correlation between BabA, CagA, and VacA, and most importantly, BabA-positive strains demonstrated higher binding to tissues of positive secretor phenotype. Thus, we demonstrated that secretors are more likely to be infected with a BabA-positive strain that enables closer contact with epithelial cells and thus causes more of tissue damage.

**Paper IV**

**Studies of Lewis antigens and *H. pylori* adhesion in CHO cell lines engineered to express Lewis b determinants**

In this study, we constructed Leb-expressing recombinant Chinese hamster ovary (CHO) cells. We further analyzed how these cells presented Leb and if recombinant Leb could support BabA-mediated adhesion of *H. pylori*.

CHO cells were transfected with expression vectors containing four different genes that are necessary for Leb expression on core 3 O-glycans, GlcNAc- and Gal-transferases for the backbone and the secretor- and Lewis-fucosyl transferases for addition of the fucose. Two Leb-expressing clones (1c5 and 2c2) were selected for further analysis. Flow cytometry revealed differences in expression of Leb: 2c2 with high levels of both Leb and Ley expression and 1c5 having lower levels of Leb expression and no Ley expression. However, both clones express Lea. By co-expressing the reporter proteins PSGL-1/mIgG2b and AGP/mIgG2b that are glycosylated by N- and O-glycans, respectively, we detected Leb on O-glycans on the surface of both clones but lower expression in 1c5. The 2c2 clone also expressed Leb and Ley in N-glycans. In addition, 2c2 demonstrated Leb and Ley expression also in glycolipids whereas 1c5 was low in Leb and with no Ley glycolipid.
We further investigated whether the Leb expression of these cells was suitable to support BabA-mediated adherence of *H. pylori*. We applied FITC-labeled 17875/Leb and specific adherence was detected to the 2c2 and 1c5 clones, \( \approx 7600 \) and \( \approx 2200 \) *H. pylori* were found, respectively, compared to \( \approx 400 \) for non-recombinant CHO cells.

To summarize these results, we could detect a distinct difference between the CHO cells clones in terms of Leb expression. The 1c5 clone expressed Leb on O-glycans and a very poorly in glycolipids. In contrast, the 2c2 clone displayed strong expression in both N- and O-glycans and in addition in glycolipids. 2c2 also expressed Ley both on N-glycans and in glycolipids. The binding results demonstrate that O-glycan expression of Leb supports BabA-mediated adherence to Leb. These results, however, do not reveal if N-glycans or glycolipids could support adherence. However, biochemical tests of fractionated glyco-conjugates might provide for information about the receptor proportion of N-glycans and glycolipids.
Discussion and Concluding Remarks

The glycosylated mucin layer that lines the entire gastro-intestinal tract is one of the first lines of defense against gastrointestinal pathogens and, therefore, displays individual diversity in its carbohydrate repertoire. This polymorphism is presumably the net result of an arms race between the pathogens and their host. Changes of a pathogen in response to a new receptor, and potentially a new niche, demand that the host follow with additional changes to avoid being overtaken by the pathogen (Bishop & Gagneux 2007). In addition, diversity among individuals in carbohydrate patterns limits the horizontal spread of infections, such as that seen for the resistance against disease by Norovirus among non-secretors.

For most enteropathogens, the acidity of the stomach is lethal. However, *H. pylori* has cleverly adapted its tissue tropism to this unique niche in many ways and is, therefore, the (almost) exclusive persistent pathogen of human stomach. The quasi-panmictic population structure i.e. resembling an unstructured population, of *H. pylori* facilitates persistence of the infection by ensuring that a subpopulation of at least a single clone is capable of surviving and pursuing infection in response to environmental changes. The persistence of a bacterial clone is highly dependent on adherence capability. A good example of an adhesin that is capable of adaptation is the *E. coli* FimH protein. From 133 sequenced *fimH* genes there are 63 unique alleles (Sokurenko 2004). Such a high diversity of the FimH adhesin would be beneficial as it can shift in binding affinity and target other mono or multivalent mannose receptors. In comparison, BabA is even more polymorphic and each gene represents a unique allele. *H. pylori* strains are very different in BabA-mediated Leb binding and display a >1500-fold spectra of high binding affinities (Aspholm-Hurtig et al. 2004). Moreover, as it has been shown in the present study, some strains are very low in binding affinity and many strains don’t bind at all and have lost all affinity even though they still express BabA. This indicates that there is a selective pressure on BabA to retain the ability to bind the ABO/Leb blood group antigens. The reason for this is not known but probably reflects the narrow niche in which *H. pylori* is the only permanent inhabitant. The selection for binding to the ABO/Leb structures is probably also due to the high prevalence of these receptors in the human gastric epithelium that increases the chances for *H. pylori* to find a suitable host cell for cross talk. Some diversity in receptor specificity for BabA is seen between generalist and specialist strains, but because of the high diversity it is still not known how BabA binds to the ABO/Leb structures or which amino acids promote
binding. The binding affinity is not the only selection that drives diversity of BabA. Because of the outer membrane localization of BabA, the diversity of BabA is probably also a consequence of a continuous process to evade neutralizing immune responses.

The first study that demonstrated the correlation of babA with CagA and VacA virulence factors in disease development analyzed the prevalence of the babA2, which was considered to be expressed due to an intact signal peptide sequence (Gerhard et al. 1999). However, owing to the polymorphic characteristics of BabA expression, in which not all genes express a protein and not all proteins are functional, it is interesting to speculate on actual implications of such correlation. Could merely BabA expression itself be enough to trigger a higher immune response and consequently cause more inflammation and disease? Infections of Leb transgenic mice resulted in higher levels of inflammation compared to non-transgenic mice suggesting that adherence stimulates mucosal inflammation. Since BabA is not much immunogenic during chronic infection in man, this hypothesis seems unlikely. Thus, it is reasonable to speculate that the adherence properties of BabA are correlated to disease progression. Therefore, a more complete way of correlating BabA in clinical strains to disease would be to appreciate Leb binding by the bacteria as a first indication of a BabA-positive phenotype. Consequently, in Paper III, we considered biopsies to be BabA-positive by immunofluorescence with a BabA-specific antibody. Although this does not necessarily indicate a functional BabA, it is closer to in vivo expression of BabA compared to inventory of babA genotype.

In Paper III, we found BabA to correlate with CagA and VacA1. The functional role of BabA in disease progression has not been entirely elucidated, though new results point towards a role for BabA in the delivery of VacA and CagA to the host cells (Ishijima et al. 2011). Tight adherence with immediate access to cellular junctions would not only promote more efficient delivery of VacA, but would also allow for the bacteria to swim towards the basal membrane to deliver CagA where the integrin receptors are i.e. the presumed docking station for the T4SS machinery. Many gastric cell lines, such as the AGS cells, are available for signaling studies with H. pylori, but they are variable or low in Leb expression and this makes it challenging to investigate BabA-mediated signaling. To overcome this, we constructed CHO cells with defined carbohydrate repertoires (Paper IV). Unfortunately, these hamster cells are resistant to delivery of CagA (personal observation), which might relate to differences in β1-integrins. Despite the lack of CagA delivery into the host cells, it would be interesting to study if the BabA-Leb interaction itself could trigger a signaling cascade in the cells. The functional receptor for BabA in the human gastric mucosa is not known.
However, it is known that BabA interacts with both the MUC5AC and MUC1 mucins, and might interact with non-mucin glycosylated proteins in the host cell membrane. In addition, binding mediated by glycolipid receptors could also promote signaling that could aid in promoting loosening of the cellular junctions and function as a complement to VacA during infection.

Gastric cancer is the second most common cancer worldwide and often with poor outcome because the cancer is usually detected at an advanced stage where treatment is difficult and the mortality rate is high, hence denoted the “silent killer”. Though we have gained some knowledge about the steps in cancer development, there are still many important questions to address. What factors promote a symptomatic disease? Why do some people develop cancer vs. peptic ulcer disease? Are there any benefits in H. pylori infection or should we aim to eradicate all infections? Do we need to eradicate infections to prevent disease or would it be enough to merely suppress the infectious load e.g. 10-100 fold? In this perspective, it is important to consider whether we need sterilizing vaccines or whether it would be enough if vaccines could reduce infection 100-fold. Many scientists are studying the important roles of CagA and VacA in disease development, but additionally factors must be of importance, especially since the majority of clinical isolates express CagA and VacA. If additional virulence factors could be identified that would predict those strains that are prone to develop disease, more focused treatments with antibiotics could be an option, in particular in combination with a vaccine that would prevent re-infection. Such combination would make a difference for people all around the globe.
Acknowledgements

Det har varit en underbar och spännande resa som doktorand, från början till slut. Som med mycket annat så är inte detta något man gör alldeles själv, man behöver stöd och hjälp av andra människor också!


Kristoffer, vår biacoreguru! Tack för trevligt sällskap i labbet under den perioden som biacorens underbara ljud förgyllde tillvaron och för all hjälp med både mina försök till pH-manuskriptet och 812-projektet. Ola och Susann, tack för ett bra samarbete och hjälp med att klura ut vad min stam egentligen binder till! Maria Azevedo, Leonor David, Jonas Löfling and Jan Holgersson, thanks for nice collaborations!

The time I have spent at the department of Medical Biochemistry and Biophysics has really been a great, and I would like to thank everyone that contributes to the nice, open atmosphere. Ett speciellt tack till Ingrid, Clas, Anna, Lola och Elisabet! Om ni inte vore så fantastiskt bra på det ni gör så skulle jag aldrig få någon forskning gjord!

Melissa, thank you so much for your help with the last experiments of My manuscript and for our scientific, and sometimes not so scientific 😊, discussions that we have had in our small office. They have really enlightened this last year! Jeanna, thanks for all the help with protein purification, discussions about microscopy and for bringing some culture (not the bacterial one) into My life. Anna S, thank you for help with RIA, for your amazing cakes and for talks about life. Rolf! jag har saknad dig sedan du gick i pension. Lena, thanks for our nice conversations in the office, sorry I’m not as in to politics as you are. Alexej och Göran, tack för all hjälp med att lära mig PCR, kloning och uttryck. Åsa, Jenny, Marie and Jalal, My best wishes for you!

Annelie, Var ska jag börja? Att få dela min doktorandtid med dig har varit fantastiskt roligt. Utan dig hade jag varit tvungen att själv åka på alla
konferenser, lära Maria och Melissa att äka längdskidor, prova vin i Napa Valley och mycket mycket annat. Tack för allt och lycka till med din disputation! Pär, jag beundrar din ambition och humor, du kan vara en doktor! Anna V, tack för att du lärde mig vävnadssnitten och mikroskopets underbara värld och för allt annat kul vi haft både i och utanför jobbet. Anna Å, Carina och Crister, tack för trevliga stunder på labbet!

Mina vänner utanför "Helicobacter-världen", Rima, tack för ditt alltid så glada och smittsamma humor, och för alla block du snittat åt mig. Emma, min underbara och ordningsamma vän! det har varit kul att äntligen få ha dig på samma institution igen så vi hunnit träffas varje dag (nästan iaf 😊). Lisa, tack för all optimism! Alla borde ha en vän som du, som hellre ser saker som spännande än jobbiga. Nu hoppas jag att vi alla tre hinner träffas ofta igen för fika och promenader. Anna och Mattias, tack för härliga middagar (och äta ute) mitt i veckan som en paus från vardagen, vet inte hur jag skulle överlevt utan!


Ett stort tack till min hela min familj, mamma, pappa, Anna & Johan, Daniel och Maria som stödjer mig i det mesta jag gör, även om det innebär att vi inte hunnit träffas så mycket senaste tiden, men det blir nog bättre efter det här. Ett extra stort tack till Maria som många gånger fått rycka in som barnvakt när tiden kniper.

David, min underbara älskade make som alltid hjälpt och stöttat mig i många år, men som gjort det lite extra den senaste tiden. Ida, min älskade dotter. Den livglädjen du sprider omkring dig har gjort detta skrivande mycket lättare och om jag bara hade haft hälften av din energi hade denna det här varit gjort på ett nafs. Jag älskar er båda otroligt mycket!

/Sara
References


Azevedo, M. et al., 2008. Infection by Helicobacter pylori expressing the BabA adhesin is influenced by the secretor phenotype. The Journal of


Carvalho, F. et al., 1997. MUC1 gene polymorphism and gastric cancer--an


Jiménez-Soto, L.F. et al., 2009. Helicobacter pylori type IV secretion


