

UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Chimica Biologica

SCUOLA DI DOTTORATO IN BIOCHIMICA E BIOTECNOLOGIE

INDIRIZZO IN BIOTECNOLOGIE

XXIII CICLO

**Structural studies of
Helicobacter pylori proteins
relevant for gastric colonization**

Direttore della scuola: Ch.mo Prof. Giuseppe Zanotti

Coordinatore d'indirizzo: Ch.mo Prof. Giorgio Valle

Supervisore : Ch.mo Prof. Giuseppe Zanotti

Dottoranda: Lorenza Sisinni

31 Dicembre 2010

UNIVERSITÀ DEGLI STUDI DI PADOVA



Dipartimento di Chimica Biologica

SCUOLA DI DOTTORATO IN BIOCHIMICA E BIOTECNOLOGIE

INDIRIZZO IN BIOTECNOLOGIE

XXIII CICLO

**Structural studies of
Helicobacter pylori proteins
relevant for gastric colonization**

Direttore della scuola: Ch.mo Prof. Giuseppe Zanotti

Coordinatore d'indirizzo: Ch.mo Prof. Giorgio Valle

Supervisore: Ch.mo Prof. Giuseppe Zanotti

Dottoranda: Lorenza Sisinni

Index

Chapter 1 <i>General introduction</i>	7
1.1 <i>Helicobacter pylori</i>	10
1.2 Successful colonization of the mucous layer	12
1.3 Interaction with gastric Epithelium and Adherence	15
~ 1.3.1 Adhesion it is used to cause cellular damage and inflammation	16
~ 1.3.2 Adhesion to avoid mechanical clearance and promote and persistence	22
~ 1.3.3 Use of the cell surface as a site of replication	30
1.4 Conclusion	31
Chapter 2 <i>Production and characterization of H. pylori proteins involved in virulence and colonization</i>	32
2.1 Introduction	34
2.2 Project	36
2.3 Results and discussion	37
~ 2.3.1 Cag proteins	37
~ 2.3.2 Adhesion Process	38
~ 2.3.3 Essential components of metabolic pathways	40
~ 2.3.4 Secreted factors involved in stress	44
2.4 Conclusion	46
Chapter 3 <i>CagL from Helicobacter pylori</i>	47
3.1 Introduction	49
3.2 Materials and Methods	51
3.3 Results and discussion	55
3.4 Conclusion	58

Chapter 4 <i>Helicobacter pylori</i> adhesin A (HP0797)	59
4.1 Introduction	61
4.2 Materials and methods	64
4.3 Results and discussion	70
Chapter 5 <i>Acidic stress response factor from Helicobacter pylori</i>	73
5.1 Introduction	75
5.2 Materials and Methods	77
5.3 Results and discussion	80
Conclusions	89
References	91

Chapter 1

General introduction

Contents

1.1 *Helicobacter pylori*

1.2 Successful colonization of the mucous layer

1.3 Interaction with gastric Epithelium and Adherence

↻ 1.3.1 Adhesion it is used to cause cellular damage and inflammation

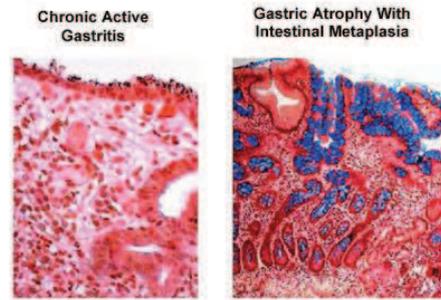
↻ 1.3.2 Adhesion to avoid mechanical clearance and promote and persistence

↻ 1.3.3 Use of the cell surface as a site of replication

1.4 Conclusion

Gastric carcinomas are the second worldwide leading cause of death due to cancer. Despite being a consequence of the increased human longevity, the rate of gastric cancer has been declining from the early 20th Century onward (Parkin and Pisani, 2005; Herrera and Parsonnet, 2009). This drop in incidence, particularly in industrialized nations, is the result of the diminishing prevalence of *Helicobacter pylori* infection, an organism that is often present in the human flora.

Gastric Mucosal Lesions During Chronic *Helicobacter pylori* Infection



Interest in *H. pylori* as a cause of cancer started after the pioneering discoveries of Marshall and Warren in the 1980s (Marshall and Warren, 1984). Prior to the identification of this organism, it was known that gastric adenocarcinomas typically arose as a consequence of chronic gastritis. After the relationship between *H. pylori* and chronic gastritis was established, investigators began to be interested in the causal role of *H. pylori* in gastric cancer. The first studies that examined the association between *H. pylori* and gastric cancer were ecological, comparing the regional prevalence of *H. pylori* with the incidence of gastric cancer (The Eurogast Study Group, 1993). Subsequently, numerous observational studies of diverse design confirmed these findings. Epidemiological and pathophysiological data led to the consensus that *H. pylori* is one of the

world's most important causes of cancer. In 1994, the International Agency for Research on Cancer declared *H. pylori* a type I carcinogen, or a definite cause of cancer in humans (IARC Working Group, 1994).

H. pylori infects more than half of the world's human population and gives rise to gastroduodenal diseases, such as peptic ulcer in about 10% and gastric adenocarcinoma in 1–2% of infected people. In 2002, an estimated 1.9 million cases, 17.8% of the worldwide incidence of cancer, were considered to be attributable to infectious diseases, with *H. pylori* infection as the leading cause (5.5% of all cancers), followed by human papilloma viruses, hepatitis B and C viruses, Epstein-Barr virus, human immunodeficiency virus, and human herpesvirus-8 (Ohkusa, 2001). *H. pylori* is responsible for about 75% of all non-cardia gastric cancers and 63.4% of all stomach cancers worldwide (Ohkusa, 2001; Herrera and Parsonnet, 2009).

1.1 *Helicobacter pylori*



Figure 1.1 Cartoon representing *Helicobacter pylori* invading epithelial cells.

H. pylori is a Gram-negative, spiral-shaped bacterium that is characterized by its many unipolar flagella, that confer to it a corkscrew-like motility. (Herrera and Parsonnet, 2009). It is named from of its spiral or helical shape (Fig 1.1). The organism is approximately 0.6 μm thick, taking the shape of a flat spiral (Westblom and Midkiff, 1991).

It prefers a microaerophilic (reduced oxygen) environment and can be cultured in gas jars with *Campylobacter* gas generating envelopes (Oxoid), or in carbon dioxide (CO₂) incubators at 37°C.

Prospective studies revealed that the risk of development of gastric carcinoma is much greater in *H. pylori*-infected populations than in uninfected populations (Uemura and Schlemper, 2001). Despite the close association of *H. pylori* infection with non-cardia gastric cancer, most infected persons do not develop the disease. The clinical outcome of *H. pylori* infection is determined by multiple factors, including

host genetic predisposition (especially certain cytokine polymorphisms), (Amieva and El-Omar, 2008), bacterial strain heterogeneity and environmental factors, such as dietary salt intake (Forman and Burley 2006).

H. pylori is a highly heterogeneous bacterial species, both genotypically and phenotypically, and is highly adapted for survival in the gastric niche. The genomic diversity of *H. pylori* parallels that of its host species, consistent with colonization of the earliest humans and co-migration out of East Africa at least 60,000 years ago (Linz and Achtman, 2007).

H. pylori has probably accompanied humans for tens of thousands of years (Ghose, 2002; Falush, 2003; Linz and Achtman, 2007) and it has been hypothesized that *H. pylori* colonization could have provided benefits to its human carriers and hence provided a selective advantage during long periods of human history (Blaser, 2006).

In the modern world, *H. pylori* infections are responsible for a heavy toll of morbidity and mortality as a consequence of ulcer disease, lymphoma of the mucosa-associated lymphoid tissue (MALT) and, the most dangerous complication of *H. pylori* infection, gastric adenocarcinoma. These selective pressures operate on a microorganism that possesses specific capabilities for diversification by mutation and recombination. The bacteria need to interact indirectly and directly with the hosts, requiring both fixed bacterial surface molecules to provide adherence and soluble molecules that are either surface-bound or secreted, and act on their respective host receptors. As such, one might assume that adaptation processes could lead to alterations in three groups of bacterial genes: first, genes in systems that affect intra-bacterial mutation, DNA uptake, repair and recombination themselves; second, genes that favor bacteria–bacteria interactions for the purpose of inter-bacterial genetic exchange, including decreasing or increasing the barrier of genetic exchange; and third, genes that influence bacterial properties that modulate host interaction (adherence and immune response) (Suerbaum and Josenhans, 2007).

H. pylori heterogeneity and the association of certain *H. pylori* products with specific diseases (virulence factors) has been intensively investigated over the past two decades since *H. pylori* was first cultured. In an attempt to bring order to a sometimes chaotic field, it has been proposed that the following criteria should be met for designation of a virulence factor:

- the *H. pylori* determinant should have a disease or other *in vivo* correlation,
- there should be epidemiologic consistence across populations and regions,
- there should be biologic plausibility,
- the biologic activity should be reduced or eliminated by gene deletion and restored by complementation (Lu and Graham, 2005; Lodato and Mazzella, 2009).

1.2 Successful colonization of the mucous layer

The successful life-lasting colonization of the human stomach by *H. pylori* is achieved through a combination of factors, which address the different challenges presented by the harsh environment. In general, the gastric mucosa is well protected against bacterial infections. Despite that, *H. pylori* is highly adapted to its ecological niche, using a bundle of polar flagella, a potent urease and additional features that allow its oriented swimming and multiplication in the mucus, attachment to epithelial cells, evasion of the immune response, and persistent colonization and transmission.

Urease

H. pylori dedicates several genes to the codification of a group of ureases. Cytosolic urease is a Ni²⁺-containing enzyme (Hu and Mobley, 1990; Labigne and Courcoux, 1991) that hydrolyses urea into NH₃ and CO₂. The biosynthesis of urease is governed by a seven-gene cluster, including the genes encoding UreA (26.5 kDa) and UreB (60.3 kDa) subunits. Urease is a hexa-dimeric cytoplasmatic protein, arranged as a double ring of 13-nm diameter, that determines the survival of the bacterium in acid condition. Urea is taken up by *H. pylori* through a proton-gated channel (Weeks and Sachs, 2000); its hydrolysis generates ammonia, necessary to create a neutral layer around the bacterial surface (Phadnis, 1996). The urea channel is positively regulated by protons. This allows the organism to survive and grow in the stomach in the presence of the usual gastric urea concentrations. The absence of urea transport at neutral pH prevents high urease activity in the absence of gastric acidity, as occurs during digestion. In fact, urea is toxic to the bacterium at neutral pH, since it generates an alkaline environment. The combination of a high level of a neutral pH-optimum urease and an acid-regulated urea channel explains why *H. pylori* is unique in its ability to inhabit the human stomach. (Weeks and Sachs, 2000).

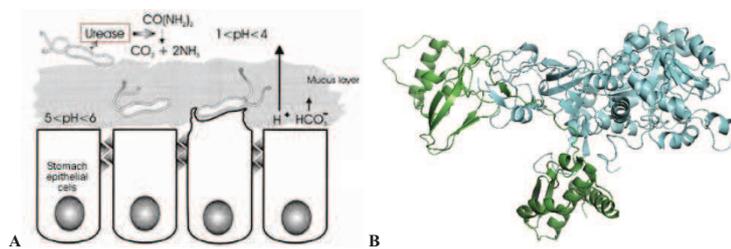


Figure 1.2. A) Mechanisms of survival in *Helicobacter pylori*. Urea, which is present in the stomach, is transported through UreI into the bacterial cytoplasm, where urease converts it into NH₃ and CO₂. From Montecucco et al, 1999. B) Ribbon representation of the monomeric unit of the *H. pylori* Urease. Each unit contains the α -subunit (light blue) and the β -subunit (green) (Ha, 2001) [PDB accession code 1E9Z and 1E9Y].

Flagella

Most *H. pylori* bacilli swim within a narrow band of the protective mucus gel that is constantly being secreted and renewed (Schreiber, 2003). To remain in the mucus layer, *H. pylori* need to utilize its polar flagella for motility. The motility is provided by two-to-six polar, sheathed flagella, the filaments of which consist of two flagellin types, the major FlaA and the minor FlaB (Kostrzynska and Trust, 199; Josenhans and Suerbaum, 1995).

Both the ability to swim with flagellar motion and the ability to control the direction of movement by chemotactic responses are essential for *H. pylori* colonization (McGee, 2005; Ottemann and Lowenthal, 2002).

A study about the spatial orientation of *H. pylori* within the gastric mucus layer have demonstrated that they actively remain within 25 μm of the surface of the epithelium and orient themselves based on the pH gradient in the mucus, avoiding the acidic distal regions (Schreiber, 2004). If these pH gradients are eliminated experimentally, *H. pylori* loses its spatial organization and, if the bacteria are placed at pH 4 or lower, they lose their motility in a matter of minutes (Schreiber, 2005). Therefore, it seems that a constant sensing and responding to pH gradients is critical for the survival of the bacterium in the stomach. A recent report identified an *H. pylori* chemoreceptor responsible for the negative chemotaxis in pH gradients. *H. pylori* increases its swimming speed when placed near an acidic gradient (Croxen, 2006). It also changes its swimming paths to favor movement away from acid. The sheathing of the flagella is believed to protect bacteria from the acidic environment in the stomach (Spohn and Scarlato, 2001).

A factor just as important as motility is the shape of *H. pylori*. Its curved S-shape enables the bacteria to bore through the mucin web (Slomiany, 1992). Its shape and motility may be the primary reasons why only *H. pylori*, and no other bacteria, is able to colonize the gastric mucosa. Most other *Epsilonproteobacteria* (*Campylobacter* spp., *Helicobacter* spp., etc.) have other shapes than *H. pylori*. For example, *Helicobacter heilmannii*, which has a cork-screw shape, does not penetrate the mucin layer but colonizes the gastric glands (Andersen, 2007).

1.3 Interaction with the gastric epithelium and adherence

Approximately 20% of *H. pylori* in the stomach is found adherent to the surface of mucus epithelial cells.

The bacterium comes into contact with the mucin layer that covers the epithelial cells either by an active or a passive process, as it moves actively toward areas with the highest concentrations of urea and bicarbonate that are found in the mucosa (chemo-attraction) (Yoshiyama, 1999). The contact with the mucin results in a strain-dependent interaction (adherence) between the mucin and *H. pylori*.

Adhesion involves specialized molecular interactions with the gastric mucosa that may lead to intimate attachment and modification of the cell surface and the underlying cytoskeleton.

Adherence to sialic acid in the mucin seems to be a common feature in most *H. pylori* strains. *H. pylori* has at least six adhesins for sialic acid, three genes of which (*hpaA*, *nap*, *sapA*) have been identified (Wadström and Borén, 1996; Testerman and Mobley, 2001). Another adhesin is BabA, which binds specifically to the Lewis B (LeB) antigen in mucin MUC5AC (Van de Bovenkamp and Van Male, 2003). In a study by Linden et al. (Linden and Hedenbro, 2004), all of the *H. pylori* strains tested adhered to MUC5AC at acid pH, whereas only BabA-positive strains adhered to MUC5AC and MUC1 at neutral pH. Not all *H. pylori* strains contain all of the sialic acid's adhesins, which contributes to the strain variation in *H. pylori*.

Individual types of genetically determined mucins of the host might facilitate the colonization of *H. pylori* compared to other types of mucin, a topic that has been better studied. The primary colonization may even take place in the oral cavity, since *H. pylori* has been shown to adhere to MG2 in the human salivary mucin (Prakobphol and Rosen, 1999).

Several recurrent themes in the biology of *H. pylori* adhesins exemplify the importance of heterogeneity in this system:

- no individual adhesin is essential for attachment to the gastric mucosa, indicating redundancy of adhesive mechanism;

✎ expression of adhesins is different between strains and variable within a single strain over time, and these mechanisms of variability and adaptation are controlled at the genetic level by on/of switching of adhesion gene expression, gene inactivation, or recombination. There is a dynamic adaptation of each adhesin to its receptor and mucosal glycosylation at different localization and over the time;

✎ adhesive interactions contribute to inflammation and are likely involved in disease progression.

At this point it is important to clarify why does *H. pylori* adhere to the cell surface. Three main hypotheses about this question have been put forward and they are not mutually exclusive.

1.3.1 Adhesion used to cause cellular damage and inflammation

H. pylori adhesion causes damage to the epithelium, induces inflammation, and eventually delivers toxins. The successful attachment to the cell receptors is guaranteed by at least five types of adhesins or adherence-associated proteins (AlpA, AlpB, HopZ, BabA and SabA), and other known outer membrane proteins, like those belonging to the OMP family (Tomb, 1997). *H. pylori* OMPs are encoded by about 4% of the entire chromosome, a percentage higher than that of any other bacterial genome (Doig, 1999).

H. pylori is very well adapted to its host, a fact that allows it to persist for years or decades in its niche. In addition, it has to be flexible enough to rapidly accommodate to a new host during infection. The high genetic variability between different *H. pylori* strains is based on microdiversity (at the gene level) as well as macrodiversity (at the genomic level) (Odenbreit and Haas, 2002). The large group of OMPs is probably of

considerable importance for optimal adaptation of *H. pylori* to its host, especially in young people. All the OMPs family members show significant sequence similarities in the N- and C- terminal domains (Tomb, 1997), whereas their central part is rather diverse. *H. pylori* uses several strategies to generate diversity in the large group of OMPs. One mechanism relies on slipped-strand mispairing (SSM), which involves the deletion or insertion of nucleotides in homopolymeric tracts located in the gene promoter region or the 5' gene sequence (coding repeats) (Stern and Meyer, 1986). This is only an example of the mechanisms used by the bacterium to get into the habit of the physiological characteristics of each human stomach, conferring high versatility to the *H. pylori* phenotype.

Adherence is believed to help protecting the bacteria from gastric acidity, as well as from displacement due to peristalsis, but its role in the infection process is not fully understood.

Two highly homologous OMPs, the adherence-associated lipoproteins A and B (AlpA and AlpB), are involved in *H. pylori* adherence to human gastric histo-tissue sections (Odenbreit, and Haas, 1996; Odenbreit and Haas, 1999), although the corresponding receptors for these proteins are not known. At present, it is being ascertained that AlpAB presents two channel-forming membrane pores, hopB and hopC, which are organized in an operon (Exner and Hancock, 1995). Through designing the mutant strain AlpA or AlpB, Odenbreit testified AlpAB-specific adherence and concluded that the adherence was independent of the composition of the lipopolysaccharide (LPS) (Odenbreit and Haas, 2002).

Another hypothetical adhesin, the HopZ protein (Peck and Knapp, 1999), was identified in the outer membrane fraction, and the relevant gene was cloned and sequenced from 15 different strains. The protein was shown to mediate binding to gastric epithelial cells, but that it does not function as a porin.

Among other adhesins expressed by the bacterium, two are dominant: the Lewis *b* blood group antigen binding adhesin, BabA, and the sialic acid-binding adhesin, SabA. These adhesins recognize specific carbohydrate moieties of the gastric epithelium, the Lewis *b* antigen, Leb, and the

sialyl-Lewis x antigen, sLex, respectively, which promote infection and inflammatory processes in the gastroduodenal tract (Petersson, 2006). When transgenic mice expressing the Lewis-B antigen were infected with *H. pylori*, the mice showed increased bacterial attachment, more severe chronic gastritis, and a parietal cell loss (Guruge,1998). The SabA adhesin is the key molecule in the activation of human neutrophils. SabA adhesin stimulates human neutrophils through selectin-mimicry.

Interestingly, a protein that is not an adhesin modulates the oxidative burst, which could tune the impact of the *H. pylori* infection for establishment of balanced and chronic inflammation of the gastric mucosa (Petersson, 2006). This protein is HP-NAP, or neutrophil-activating protein, a 150 kDa oligomeric protein composed by twelve identical subunits. It promotes the adhesion of human neutrophils to endothelial cells, the production of reactive oxygen radicals and it acts through a cascade of intracellular activation events (Yoshida,1993; Evans, 1995). It has been hypothesized that HP-NAP might induce a state of moderate inflammation, which alters the epithelial tight junction, leading to the release of nutrients from the mucosa (Blaser, 1993).

The quaternary structure of HP-NAP is similar to that of other dodecameric bacterial ferritins (Dps-like family), but it has a different surface potential charge distribution, which could well account for its unique ability in activating human leukocytes (Zanotti, 2002).

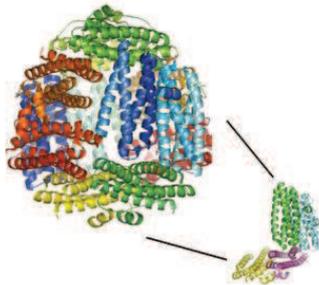


Figure1.3. Structure of *H. pylori* Neutrophil activating protein (HP-NAP). Ribbon representation of the dodecameric and quaternary assembly. (Zanotti, 2002) [PDB accession code1J14].

HP-NAP plays an important role in immunity. Vaccination of mice with HP-NAP induces protection against *H. pylori* infection. This is consistent with the finding of HP-NAP specific antibodies present in the majority of *H. pylori* infected patients (Satin and Rossi, 2000).

The contact with the bacterium is intimate and almost irreversible, and leads to a profound rearrangement of the plasma membrane below the zone of contact (Smoot, 1993; Segal and Tompkins, 1996). During the invasion of epithelial cells by *H.pylori*, the plasma membrane changes shape and extends to contact a large portion of the bacterial surface, after an *H. pylori*-induced reorganization of the underlying actin cytoskeleton (Montecucco and Rappuoli, 2001).

The delivery of the major *H. pylori* virulence factors CagA and VacA are intimately related to adhesion as well, suggesting that a major role of adhesion is the delivery of toxins. The increased inflammation and damage caused by adhesion are of adaptive value to *H. pylori*, bringing to the release of nutrients into the gastric lumen.

VacA is a pore forming cytotoxin that was identified when the supernatants of *H. pylori* broth culture were found to cause aberrant vacuolation of cultured cell, (Leunk, 1988; Del Giudice and Rappuoli, 2001). Vacuolating cytotoxin A, or VacA, gene encodes a protoxin of approximately 140 kDa mass. An amino terminal signal sequence and carboxy-terminal fragment are proteolytically cleaved to produce a ~88kDa mature toxin (Cover and Blaser, 1992). The mature VacA can be further nicked into two smaller peptides that remain non-covalently associated (Cover, 1994). Two domains of VacA, p33 and p55, have been identified based on partial proteolysis of p88 into fragments of 33 and 55 kDa, respectively (Telford, 1994; Fig.1.4). When expressed independently and then mixed, p33 and p55 can physically interact and reconstitute vacuolating toxin activity (Ye, Blanke, 1999; Torres and Cover, 2004). The N-terminal p33 domain (residues 1–311) contains a hydrophobic sequence (residues 6–27) involved in pore formation (Vinion-Dubiel and Cover, 1999; McClain and Cover, 2003), whereas the p55 domain (residues 312–821) contains one or more cell-binding domains (Reyrat and Telford, 1999; Wang and Wang, 2000). The first domain is highly conserved, whereas the second domain is genetically diverse. VacA purifies as a large oligomeric complex (>900 kDa), which under the electron microscope appears as a “flower” like structure of about 30 nm diameter (Lupetti and Telford, 1996; Lanzavecchia and Teldord, 1998). This structure is thought to be composed of one or two

rings, each comprising 6-7 VacA monomer. The oligomeric complex can be disassembled in monomers when exposed to acidic or basic pH. The monomeric forms show increased cytotoxicity with respect to oligomers (Yahiro and Hirayama, 1999). It is believed that the flower-like ring complex of VacA is the conformation it adopts when inserted into membranes (Czajkowsky, 1999). Mature toxin molecules are either secreted in the extracellular space or they may be retained at the bacterial surface.

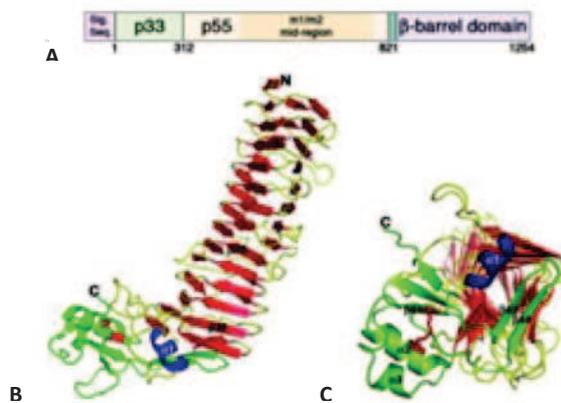


Figure 1.4. A) The *vacA* gene encodes a 140-kDa protoxin. The mature 88-kDa VacA toxin contains two domains, designated p33 and p55. The midregion sequence that defines type m1 and m2 forms of VacA is located within p55. (B) The VacA p55 fragment adopts a β -helix structure that is composed of three parallel β -sheets (red) connected by loops of varying length and structure (yellow). The α -helix in blue ($\alpha 1$) is contained within one of these loops but is highlighted in blue to show how it caps the end of the β -helix. The C-terminal domain (green) has a mixture of α/β secondary structure elements and contains a disulfide bond (red), not previously observed in an autotransporter passenger domain structure. (C) This view represents a rotation of the molecule in *b* by $\approx 90^\circ$ into the plane of the page.

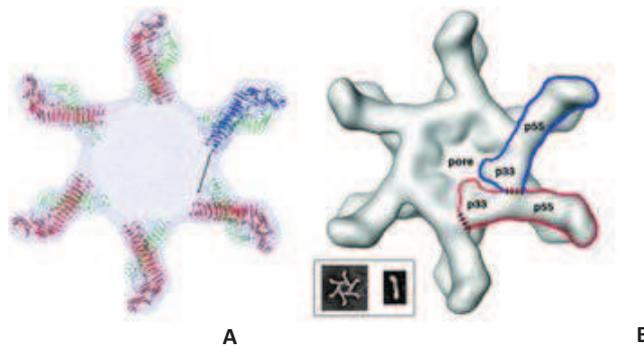


Figure 1.5. A) Docking the p55 crystal structure into a 19-Å cryo-EM map of the VacA dodecamer results in a model for oligomerization. Twelve p55 subunits are shown docked into a 19- Å cryo-EM map of a VacA dodecamer. An arrow indicates the space that the blue molecule will occupy if p33 extends the β -helix structure of p55. (B) An oligomerization model in which p33 interacts with the N-terminal portion of p55 from the neighboring subunit is proposed. Regions of contact between p33 and p55 are depicted with dashed lines. (*Inset*) EM images of a VacA hexamer and a VacA monomer. The shape of a VacA hexamer (*Inset*) is similar to the shape of a single layer within the dodecamer. The rod-like shape of the p88 VacA monomer (*Inset*) supports a model in which the β -helix observed in p55 will extend into p33.

VacA is released and placed in between the mucus layer and the apical domain of the stomach epithelial cells. VacA may not need specific receptors on the cell surface. Upon the insertion into the plasma membrane, VacA forms anion specific channels of low conductance (Szabo, 1999). The presence of vacuolar ATPase proton pump on the membrane of the endosome increases hydrogen ion concentration inside the lumen of the endosome. In the presence of ammonia generated by urease, NH_4^+ are accumulated inside the endosome: this leads to water influx and vesicle swelling, leading in turn to the vacuole formation. VacA alters the tight junction and increases the permeability of ions through the paracellular route. These nutrients are essential for *H. pylori* growth. VacA may mediate this activity by specific interaction with the recently identified cytosolic protein VIP54 (de Bernard and Montecucco, 1995).

The role of VacA *in vivo* is not well understood. It has been speculated that VacA may allow *H. pylori* to acquire nutrients by damaging the

epithelial barrier or causing paracellular leakage of small molecules. It is also possible that its main role is the suppression of the T-cell immune response, although it has not been established whether enough VacA penetrates beyond the epithelium for this purpose. It is also not known whether the adhesion-mediated delivery of VacA is more important than the role of soluble VacA, given the tendency of the monomers to oligomerize into inactive structures in solution (Ogura, 2000; Amieva and El-Omar, 2008).

1.3.2 Adhesion to avoid mechanical clearance and to promote the persistence

The second hypothesis to explain *H.pylori* adhesion is that it evolved as a way to avoid mechanical clearance. The gastric environment is dynamic and presents several mechanisms of clearance, including the constant exocytosis of mucopolysaccharides, the secretion of gastric juices and the peristaltic movement of gastric walls. In these conditions, bacterial population that colonizes mucosal surfaces have developed several adhesive strategies, like pili or fimbriae, to avoid being cleared from the mucosa (Mulvey, 1998; Amieva and El-Omar, 2008).

Many Gram-negative bacteria have elaborated alternative systems to connect at adhering cells. A first group of molecular syringe derives from duplication and re-elaboration, through evolution, of flagella (type III secretion system). The second type of molecular syringe (type IV secretion system) derives from the evolution of conjugative pili, and comprises the cag system. Despite type III and type IV secretion systems share similar structural architecture, no apparent similarity is evident between the sequences of the protein components of the two families of secretion systems, whereas homology exists among members of the same family (Montecucco and Rappuoli, 2001). Generally, type IV secretion system (TFSS) are involved in conjugative DNA transfer of prokaryotes and in the delivery of bacterial virulence factors into the eukaryotic cells. The best studied TFSS is that of the plant pathogen *Agrobacterium tumefaciens* (Christie, 2005). These bacteria inject a

DNA plasmid in the host plant cells to induce the formation of a plant tumor or gall.

H. pylori is characterized by the presence of a genomic insert, called *cag* pathogenicity island (PAI), which contains 31 open reading frames that codify for its TFSS, an apparatus of 40kDa. Some of these ORFs share significant homology with the virulence (*vir*) genes *virB4*, *virB7*, *virB8*, *virB9*, *virB10*, *virB11*, and *virD4* of the so-called VirB/D complex of type IV secretion systems known from *Agrobacterium tumefaciens* and *Borderella pertussi* (Odenbreit, 2000; Christie, 2000; Cascales and Christie, 2003; Fischer *et al.*, 2001). It was thus hypothesized that the *cag* PAI of *H. pylori* could serve as a novel transport system for secretion of virulence factors (Segal, 1999).

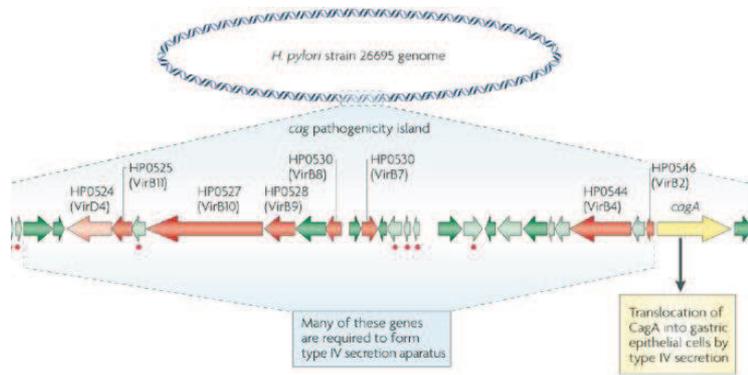


Figure 1.6. Arrangement of *cag* PAI genes in *H. pylori* strain 26695. Most of the *cag* genes are probably involved in the assembly of the type IV secretion system that translocates the protein CagA into the cytoplasm of gastric epithelial cells. Seven genes (marked in red) show similarity to components of the type IV secretion system of the plant pathogen *Agrobacterium tumefaciens*. Proteins encoded by the island are involved in two major processes, the induction of interleukin-8 (IL-8) production by gastric epithelial cells and the translocation of CagA from the bacterium into host cells. All genes depicted by arrows in dark shades of red and green are essential for IL-8 induction, whereas lighter shades of red and green indicate genes that are not involved in this process. The arrows marked with a red dot indicate genes that are not required for translocation of CagA, the non-marked genes are essential for translocation. (From the following article: "*Helicobacter pylori* evolution and phenotypic diversification in a changing host" Sebastian Suerbaum and Christine Josenhans *Nature Reviews Microbiology* 5, 441-452, June 2007)

The VirB genes can be grouped into three classes:

- ✎ the putative channel or core components (VirB6-10),
- ✎ the energetic components (the NTPases VirB4, VirB11),
- ✎ the pilus-associated components (VirB2 and, possibly, VirB3 and VirB5) (Fig.1.7)

VirB1 is a transglycosylase that acts for the localized lysis of the murein layer at the site of T4SS assembly. As mentioned above, the *Hp cag* PAI contains up to 32 genes encoding almost all VirB proteins and VirD4 as well, along with several auxiliary factors. While the role of most of the T4SS specific accessory factors is unknown, the function of CagF and CagL was only recently elucidated.

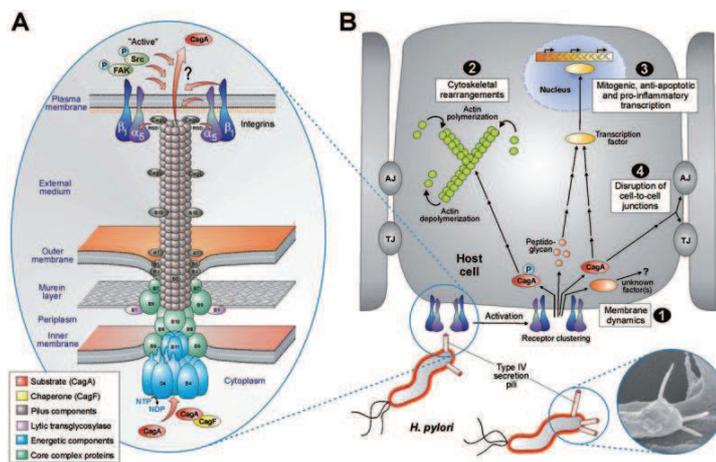


Figure 1.7. Model for the assembled T4SS machinery and its role in *Hp*-induced cell signalling. A. Hypothetical model of the T4SS machinery. The T4SS is a multi-component protein complex spanning the inner and outer membranes of *Hp*. Current knowledge of T4SS functions and cellular localization of its components is shown in a simplified manner. The coupling protein VirD4 and structural components (VirB1–11) are typically required for secretion and are positioned according to their proposed functions. This transporter enables secretion of substrates (CagA, peptidoglycan) from the bacterial cytoplasm directly into the cytoplasm of infected host cells. The CagL protein interacts with integrin receptors via its RGD motif to deliver CagA across the host cell membrane and to activate the Src tyrosine kinase for CagA phosphorylation. B. Functional activity of the T4SS pilus requires integrin receptors of the host. Scanning electron micrograph of T4SS pili is shown at the bottom (electron micrograph by Dr Manfred Rohde, HZI Braunschweig, Germany). The T4SS and CagA are involved in

numerous cellular effects, including membrane dynamics, actin cytoskeletal rearrangements, nuclear signaling and disruption of cell-to-cell junctions as indicated. AJ, adherens junction; TJ, tight junction. (From the following article: "Role of type IV secretion in *Helicobacter pylori* pathogenesis" Steffen Backert and Matthias Selbach Microreview Cellular Microbiology 10(8), 1573–1581, 2008)

CagF is a chaperone-like protein that binds close to the C-terminal secretion signal of the CagA effector protein and is crucial for the translocation of CagA (Couturier *et al.*, 2006; Pattis *et al.*, 2007). CagL is a pilus-covering protein that acts as a specialized adhesin that connects the T4SS with target cells (Kwok *et al.*, 2007).

CagA is the only protein known so far to be injected by *H. pylori* into the host cells. This is the first evidence of a functional secretion apparatus in *H. pylori*. To test the relevance of each *cagPAI* gene for CagA translocation into host cells, seventeen out of 27 genes were deleted, without causing polar effect on the expression of the downstream genes (Fischer and Haas, 2001). It is not known where the *cag* pathogenicity island arose or how *H. pylori* acquired it, and CagA presents no homology with other known proteins.

CagA was discovered in the early 1990s by independent pioneering work in the labs of Martin Blaser, Jean Crabtree and Antonello Covacci (Cover *et al.*, 1990; Crabtree *et al.*, 1991; Covacci *et al.*, 1993). It was initially found as a marker for disease, since patients with antibodies against this protein show higher rates of both peptic ulcers (Nomura, 2002) and gastric carcinoma (Blaser, 1995; Huang, 2003). The disease outcome is determined by multiple factors, including both the bacterial genotype and genetic predisposition of the host. *H. pylori* isolates are surprisingly different in both their genome sequences and their virulence (Backert and Selbach, 2008).

The reason CagA is associated with disease is still not completely understood. In addition, the benefits of these factors for *H. pylori* itself remain enigmatic. The cellular biology of Cag A and its ability to activate signaling mechanisms and to affect structure, differentiation, and behavior of epithelial cells has become a fascinating area of investigation. Currently, it is known that CagA can activate a number of signal transduction pathways that resemble signaling by growth factor receptors. Simultaneously, CagA is involved in binding and perturbing

the function of the epithelial junctions, resulting in aberrations in tight junction function, cell polarity, and cellular differentiation.

H. pylori is an excellent model system to study bacterial-induced epithelial cell signaling cascades which are of relevance to neoplasia. A key feature of the increased risk of developing gastric cancer is gastric epithelial hyperproliferation and possibly suppression of apoptosis by chronic *H. pylori* infection.

The functional activation of these T4SSs requires a signal from the host cell and the interaction with a specific receptor. Host cell integrins were recently shown to directly interact with the *H. pylori* CagL protein and are so far the only T4SS receptor known (Kwok *et al.*, 2007). CagL is a cell adhesion molecule that mediates cell–cell and cell–extracellular matrix interactions by binding through the Arg-Gly-Asp (RGD) motif. It mediates the binding of the pilus to integrin $\alpha_5\beta_1$ and is required for injection of CagA (Kwok *et al.*, 2007).

This binding to local membrane ruffling has a lot of consequences at different levels in the host cells, as in the actin cytoskeleton, transcriptional responses and cell-to-cell junctions. Clustered integrins assemble into actin-rich structures known as focal adhesions, where integrin signaling is mediated predominantly by the tyrosine kinases FAK (focal adhesion kinase) and Src (Mitra and Schlaepfer, 2006). CagL-dependent stimulation of integrin $\alpha_5\beta_1$ and FAK also activates Src, the tyrosine kinase which phosphorylates CagA. The latter is an elegant mechanism, by which the T4SS ensures phosphorylation of CagA directly at the site of injection and triggers signaling, leading to host cell motility and elongation. Immuno-electron microscopy indicated that CagL was present along the length of the T4SS pilus and mutational tests showed that changing Arg or Gly in CagL's RGD motif to Ala diminished or eliminated both CagA phosphorylation and the cell elongation phenotype that the internalized CagA protein can elicit.

Interestingly, the kinases responsible for phosphorylation of CagA are known oncogenes. First, the Src family kinases (SFKs) which control cytoskeletal processes, cell proliferation and differentiation in normal cells. They are also key players in carcinogenesis and were found to

phosphorylate CagA (Selbach *et al.*, 2002; Stein *et al.*, 2002). Second, Abl kinases (c-Abl and Arg) were also identified to directly phosphorylate the EPIYA motifs of CagA (Poppe *et al.*, 2007; Tammer *et al.*, 2007). Interestingly, *H. pylori* controls the activity of SFKs and Abl in a very specific and time dependent manner. While Src is only activated during the initial stages of infection (0.5–2 h) and then rapidly inactivated, Abl is continuously activated by the bacterium with strongly enhanced activities at late time points (2–8 h), supporting a model for the successive phosphorylation of CagA by Src and Abl (Tammer *et al.*, 2007). Cells infected by *H. pylori* elongate and scatter, a morphology that has been originally referred to as the ‘hummingbird phenotype’ (Segal *et al.*, 1999). This phenotype results from two successive events: (i) the induction of motility leading to cell scattering and (ii) host cell elongation (Moese *et al.*, 2004). It appears that CagAPY can induce cellular elongation by causing a cell retraction defect in focal adhesions, an event that is still not fully understood (Bourzac *et al.*, 2007).

One important aspect of injected CagA is the observation that it is tyrosine-phosphorylated (CagAPY) by host cell kinases (Covacci and Rappuoli, 2000). The phosphorylation sites are Glu-Pro-Ile-Tyr-Ala (EPIYA) motifs repeated up to five times in the C-terminal half of CagA (Backert *et al.*, 2001; Stein *et al.*, 2002). These EPIYA repeats are flanked by repetitive DNA sequences involved in recombination which could explain the variability in the number of these motifs in CagA variants (Aras *et al.*, 2003; Kim *et al.*, 2006) as well as differences in pathogenicity between different *H. pylori* strains (Hatakeyama, 2008).

The needs for CagL and integrins for CagA injection and phosphorylation in AGS cells suggest that these events and the successive cellular regulatory changes might be far less efficient if CagA was not delivered at the integrin-containing focal adhesion sites, where FAK and SRC are clustered, and if FAK and SRC were not activated by CagL–integrin adherence.

At this point it is necessary to clarify how the observed CagL- and integrin-dependence of CagA injection and phosphorylation in nonpolarized AGS cells does relate to infection and virulence in

unperturbed human gastric tissues *in vivo*. At least three possibilities can be considered, described below.

One hypothesis assumes that some CagA protein delivery by *H. pylori* binds intimately to the apical epithelial surface by generalized adhesins, like BabA protein, which is specific for fucose containing Lewisb antigen glycans that are abundant on gastric epithelial cells (Ilver, 1998). Any CagA delivered from such attached bacteria should have little if any exposure to integrin or activated FAK-SRC complexes and therefore might remain largely non-phosphorylated. This aspect is correlated with the onset of infection: non-phosphorylated intracellular CagA protein triggers regulatory changes that then lead to disruption of tight junctions and activation of transcription factors, which in turn increase cell motility and the levels of different factors. By helping to disrupt epithelial integrity and cell polarity and deregulating cell growth, it is also possible to increase *H. pylori*'s access to integrin. This would lead to increased CagA phosphorylation and the phosphorylation-dependent regulatory changes that also figure so importantly in *H. pylori*'s manipulation of host cell regulatory programs.

Two alternatives hypotheses invoke disruption of tight junctions to give *H. pylori* access to the normally buried integrins. The first one exploits the damage to tight junctions caused by ammonia, that *H. pylori*'s urease generates (Lytton and Beck, 2005). The other suggests that apoptosis and cell shedding, a normal activity of epithelial tissues, transiently expose basolateral surfaces of adjacent epithelial cells (Pentecost and Amieva, 2006). Each means of disrupting tight junctions should increase the access of highly motile *H. pylori* to normally buried integrins and thereby allow CagA delivery entirely via CagL–integrin interactions. In these cases, the balance of phosphorylated and non-phosphorylated forms of CagA, and thus of the regulatory circuits controlled by each form, would be shaped by the activities of kinases that CagA encounters, and that themselves seem to be regulated and to change during the course of infection (Tammer and Beckert, 2007).

On the other hand, with a transient adhesion to the epithelial cell surface *H. pylori* remains mechanically attached and also protected from the extracellular environment. This is an important characteristic that

distinguish *H. pylori* from other microbes, such as uropathogenic *Escherichia coli*, a largely extracellular organism that is unable to bladder epithelial cells, invade cells to establish a persistent intracellular niche and avoid mechanical clearance (Martinez and Schilling, 2000).

There is no evidence that *H. pylori* can replicate intracellularly and there are no specific factors known to be essential for cell invasion. Antibiotics have a poor intracellular penetration, suggesting that even small numbers of bacteria breaching the epithelial barrier could have an important role in the persistence in the organism (Dubois and Borén, 2007; Amieva and El-Omar, 2008).

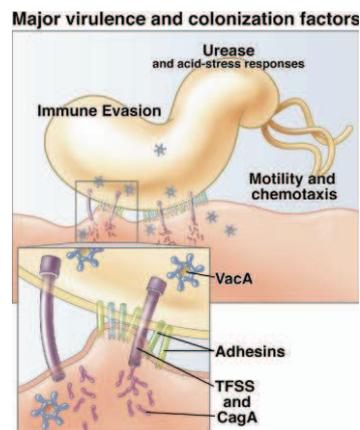


Figure 1.8. *Helicobacter pylori* major virulence and colonization factors. Schematic diagram of a single spiral *H. pylori* bacillus attached to the cell surface. *Inset* shows higher magnification of the adhesion site and delivery of VacA and CagA toxins to the host cell. As noted in the text, *H. pylori* utilizes many modalities to colonize the gastric epithelium, and some of these adaptive strategies contribute to disease progression. Urease and various other stress responses are important in surviving transient exposure to low pH environments. However, most *H. pylori* avoid the acidic lumen of the stomach by swimming toward the mucosal cell surface, using their polar flagella and chemotaxis mechanisms. Swimming and chemotaxis are important to avoid being swept into hostile microenvironments and cleared with the mucous flow. Some *H. pylori* actively adhere to the cell surface using a variety of specific OMP adhesins that recognize glycoproteins on the host cell surface. Attachment to the host cell allows *H. pylori* to deliver toxins that control various aspects of host cell function. The best studied, CagA and VacA, are illustrated. Multiple other mechanisms of immune evasion and immune modulation are utilized by *H. pylori*, including the generation of cell surface diversity through OMP allelic variation, molecular mimicry of host cell surface glycans, and modification of LPS and flagella to reduce recognition by the host innate immune system (From the following article: "Host-Bacterial Interactions in *Helicobacter pylori* Infection" Manuel R. Amieva and Emad M. El-Omar REVIEWS IN BASIC AND CLINICAL GASTROENTEROLOGY, GASTROENTEROLOGY 134:306-323, 2008).

1.3.3 Use of the cell surface as a site of replication

It has been proposed that *H. pylori* uses the cell surface as a site of replication. There are no experimental studies that directly explore the possibility for *H. pylori* to divide actively while attached to the cell surface. Nevertheless, there is a recent study showing that *H. pylori* is capable of obtaining cholesterol directly by attachment to the host cell and that incorporates and modifies this cholesterol within its own cell membrane (Wunder and Meyer, 2006; Amieva and El-Omar, 2008). Cholesterol is a physiological constituent of membranes that is crucial for their physiological properties (Simons and Vaz, 2004). By extracting cholesterol from gastric epithelial membranes, the gastric epithelial cells up regulate expression of gene involved in cholesterol synthesis during *H. pylori* infection (Guillemin and Falkow, 2002; Wunder and Meyer, 2006). This up-regulation could reflect either a compensatory response of epithelial cells to infection-associated loss of cholesterol or a form of counterattack of the host, trying to increase the local cholesterol amount to avert bacterial immune escape. *H. pylori* under acidic conditions (Slonzewski and Mobley, 2000) possibly facilitate the acquisition of cholesterol extracted from epithelial cells of the stomach.

1.4 Conclusion

In the following chapters the work performed on *H. pylori* proteins during my thesis is described. In chapter II there is an overview of all the proteins cloned and expressed; in chapter III, IV and V the structural studies on proteins CagL, Hp0797 and Hp1286 are reported, respectively. CagL is a protein on the *pilus* surface and it is a specialized adhesin implicated in the recognition and connection process between the T4SS and target cells. In this thesis we propose a structural model of the oligomerization state of CagL in solution, using the structural technique of “small angle x-ray scattering” (SAXS). The recombinant protein HpaA (Hp0797) was crystallized, but the structure of this protein has not yet been solved, since crystals diffracted only at 4Å. The optimization of crystals is in progress. The comparison between Hp0797 and Hp0410 demonstrate that proteins with the same function use different strategy to recognize the host cell. Probably Hp0797 uses the glycolipid binding domain, an hairpin structure containing a water-exposed aromatic residue.

The structure of Hp1286 was determined and refined at a resolution of 2.1Å. The electron density map showed that a ligand was present in the hydrophobic cavity. The ligand, identified by mass spectrometry, was found to be erucamide. The 3D structure of HP1286 suggests for the protein a role in the storage and transport of long-chain fatty acid(s) or amide(s). The evidence that the protein is secreted, together with the fact that the stomach mucosa, where *H. pylori* establishes persistent colonization and causes chronic inflammation, is rich in lipids, support the hypothesis that the protein sequesters fatty acids or amides present in the environment of the bacterium.

Chapter 2

Production and characterization of H. pylori proteins involved in virulence and colonization

Contents

2.1 Introduction

2.2 Project

2.3 Results and discussion

↗ 2.3.1 Cag proteins

↗ 2.3.2 Adhesion Process

↗ 2.3.3 Essential components of metabolic pathways

↗ 2.3.4 Secreted factors involved in stress

2.4 Conclusion

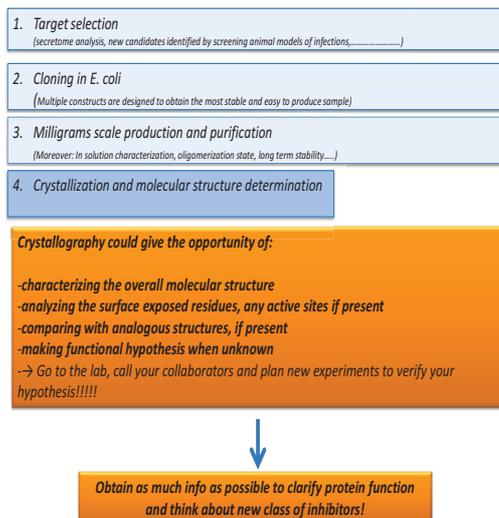
2.1 Introduction

The imbalance between aggressive and defensive factors determines the outcomes of gastric lesions under the exposure to noxious etiologies represented by either a relative increase in aggressive factors or a considerable decrease in protective factors (Robert, 1978). The gastric mucosa is continuously challenged by a variety of aggressive factors of both endogenous and exogenous nature, including excess secretion of gastric acids and pepsin, ethanol, reactive oxygen species, non-steroidal anti-inflammatory drugs (NSAIDs), and *Helicobacter pylori* infection (Hills and Lichtemberger, 1983).

Since Marshal and Warren discovered the association of *Helicobacter pylori* infection with peptic ulcer diseases (PUDs) in 1983, multiple studies have been performed on the role of *H. pylori* infection in the pathogenesis of other gastrointestinal pathologies, such as gastric adenocarcinoma and mucosa-associated lymphoid tissue lymphoma, and *H. pylori* eradication has been reported with benefits (Hunt and Lam, 1998; Mohammad and Hunt, 2010). The recognition of *H. pylori* as the key pathogen in gastroduodenal diseases with a variety of clinical manifestations, ranging from mild dyspeptic symptoms to severe complications, has posed an increasing demand for therapeutic interventions. The development of effective treatment for *H. pylori* infection has resulted in an immense change in the clinical management of upper gastrointestinal diseases, including curative antibiotic therapeutic strategies for low-grade gastric-mucosa associated lymphoid tissue lymphomas and *H. pylori* related peptic ulcers. The wide use of antibiotic therapies for *H. pylori* infection has also increased the number of therapeutic failures. Recent data show a decreasing efficacy of these therapies worldwide (Fischbach and Evans, 2007; Laine and Levine, 2000; Treiber and Klotz, 2007). This has brought to the request of improved therapeutic strategies and of the development of new drugs (Selgrad and Malfertheiner, 2008). Several factors are responsible for *H. pylori* eradication failure and there is a significant relationship between therapy failure and the bacterial strain, the host and/or the administered therapy. Recent data suggest a decreasing therapy efficacy of the seven-

day triple therapy as recommended in the European guidelines. The success rate in most European, Asian and North American countries is constantly declining (Altintas and Camdeviren, 2004; Gumurdulu and Boyacioglu, 2004). Currently, combination therapies employing one proton pump inhibitor (*e.g.* meprazole) and two or three antibiotics (*e.g.* metronidazole, amoxicillin, and clarithromycin) have been used as the preferred treatment against *H. pylori* infection. Nowadays it is necessary to found novel therapeutic strategies. In addition, there is a global challenge for the development of an effective vaccine. A possible strategy for the development of new drugs against *H. pylori* could be the structure determination of proteins essential in the infection mechanisms. This knowledge could be used to better understand the molecular mechanism in which targets are involved and to design new drugs. The strategy adopted in my research can be summarized in the scheme presented below.

In the remaining of the chapter all the proteins analyzed are briefly summarized. The most successful among them are described at length in the next chapters.



2.2 Project

The first step of this strategy is the selection of targets. Proteins involved in the following pathways have been chosen:

- Cag proteins (CagL, HP539)
- Proteins involved in the adhesion process (HP0797, HP0638 and HP0815)
- Essential components of metabolism (HP0283, HP1290 and HP1249)
- Secreted factors involved in oxidative stress (HP1286, HP0170)

Milligrams of protein are required for structural studies. For this reason, they must be produced in a heterologous system, like *E. coli*. For each target an active, stable, soluble and homogenous fraction of protein is required. Two major difficulties arise with this strategy: protein insolubility and low protein yield. Moreover, the main obstacle to three-dimensional structure determination remains the crystal growth process. Bioinformatics and literature information have been used to identify the optimal length, at the C- and/or N- terminus, of each target and used during the cloning strategies. A bioinformatics approach has also been used to determine biochemical characteristic, such as molecular weight, absorption coefficients at 280nm, isoelectric point, putative signal secretion or transmembrane domain. An exhaustive search of the amino acid sequence of each protein target against the entire database has been used to find similarities with others proteins of known function or structure. Finally, programs that allow the prediction of second or tertiary structure have been used to visualize the protein fold. For all bioinformatics analysis performed the Expasy server (www.expasy.org) was used.

2.3 Results and discussion

2.3.1 Cag proteins

CagL (HP0539) is a member of the proteins involved in the formation of a T4SS apparatus in *Helicobacter pylori*. Virulent *H. pylori* isolates harbor the *cag* (cytotoxin-associated genes) pathogenicity island, a 40 kb stretch of DNA that encodes components of a type IV secretion system (T4SS). T4SS forms a pilus for the injection of virulence factors into host target cells, such as the CagA oncoprotein. This is accomplished by a specialized adhesin of the pilus surface, the CagL protein, which binds to and activates host cell integrins for subsequent delivery of CagA across the host cell membrane.

CagL is composed of 237 amino acids. It is a protein with a theoretic isoelectric point of 8.3 and a molecular weight of 26.9 kDa. It is essential for CagA secretion into the host cell. CagL deletion mutants lose the ability to induce secretion of IL-8 by host cells (Fischer, 2001).

Integrins are transmembrane cell adhesion molecules that mediate cell–cell and cell–extracellular matrix interactions by binding to Arg-Gly-Asp (RGD) motifs. CagL contains an RGD motif that mediates the binding of the pilus to integrin $\alpha_5\beta_1$, a step required for the injection of CagA (Kwok *et al.*, 2007). Binding of CagL to integrins induces local membrane ruffling, indicative of a general effect on membrane dynamics which is the first T4SS effect on host cells.

CagL amino acids sequence analysis reveals the presence of a predicted N-terminal secretion peptide, with a predicted cleavage site between amino acids 20 and 21. For this reason the *cagL* gene was cloned and amplified from amino acid 21 to 237. The gene was ligated into the pET-28a vector to obtain His₆-CagLD, then the gene was expressed in a recombinant system using *E. coli* BL21(DE3) cells. The refolded-soluble protein was characterized with small angle X-ray scattering (SAXS), circular dichroism (CD) spectroscopy, and chemical cross-linking. Details about purification, refolding and structural characterization are described in Chapter 3.

2.3.2 Adhesion Process

H. pylori adhesin A (**HpaA**, HP0797) has been described as an adherence factor for blood cells, despite controversial evidences (O'Toole and Trust, 1995). HpaA mediates binding to sialic acid in vitro (Evans and Graham, 1988; Evans and Lee, 1993); it is a putative neuraminylactose-binding hemagglutinin, and it has also been described as a flagellar sheath protein (Jones and Penn, 1997). HpaA is a surface-localized antigen recognized by human antibodies (Lundstrom and Bolin, 2001; Blom, and Svennerholm, 2001; Evans and Graham, 1988; Lundstrom and Bolin, 2001; O'Toole and Trust, 1995; Mattsson, and Svennerholm, 1998; Yan, and Shao, 2005). Expression of HpaA protein has previously been found to be highly conserved among *H. pylori* isolates (Bolin, and Svennerholm, 1995; Yan and Shao, 2005). Furthermore, genomic studies (Alm and Trust; Tomb, 1997) have shown no significant sequence homologies of HpaA with other known proteins. Taken together, these evidences make HpaA a putative candidate as a vaccine antigen against *H. pylori* infection. It has to be clarified if HpaA itself directly mediates receptor binding or whether it is involved in facilitating the adhesin transport and folding, or if it exerts regulatory functions.

HpaA has a molecular weight of 26 kDa and a theoretical isoelectric point of 6.5. These data are relative to the protein without the predicted N-terminal secretion signal.

The gene was amplified and then cloned into pET151/D-TOPO vector using the TOPO cloning kit (Invitrogen), which permits directional cloning into expression systems. The recombinant protein has been expressed in *E. coli* (DE3) cells and purified using a metal-chelate affinity chromatography (GE Healthcare), followed by a gel filtration using a Superdex 75 10/30 (GE Healthcare). Single crystals of reasonable size diffracted to the maxima resolution of 4Å, giving data of insufficient quality for structure determination. Despite many efforts, crystals quality could not be improved. Details of the purification and crystallization process are described in Chapter 4.

Outer membrane protein 13 (omp13, HP0638). Omp13 is a protein with a molecular weight of 34.2 kDa and a basic theoretical isoelectric point of 9.78. The interest for this protein derives from the fact that this class of proteins is a potential candidate for a vaccine. Several literature studies involved HP0638, which is classified as an outer inflammatory protein (oipA). The complete genome sequence has revealed the presence of 32 outer membrane proteins (OMPs) in *H. pylori*. Knockout of the HP0638 gene in 81% of *cag*-positive strains reduced IL-8 production at approximately 50%. The three *cag*-positive strains in which IL-8 levels were unchanged by HP0638 knockout had five or seven CT dinucleotide repeats in the 5' region, resulting in a frame shift and truncation. Although *cag*-negative isolates produced a limited IL-8 response, *cag*-negative strains that contained a functional HP0638 gene produced an increase of more than 3-fold of IL-8 induction than *cag*-negative non-functional HP0638 strains. Previous studies have hypothesized that functional HP0638 gene may be an important virulence factor in relation to the risk of clinically significant outcomes of *H. pylori* infection. *hp0638* was isolated and cloned without the nucleotide sequence that codifies for its first 18 amino acids, since bioinformatics analysis predicts a secretion signal at the N-terminus. The nucleotide sequence from amino acid 18 to 300 was amplified and cloned into the pET 151/D-TOPO expression vector using the TOPO cloning kit (Invitrogen). Finally the recombinant His₆-HP0638 expression level was detected using *E.coli* (DE3) cells. Several difficulties were encountered during the purification process using a metal-chelate affinity chromatography (GE Healthcare), the major one being protein solubility. A possible strategy to obtain HP0638 in solution could be to refold it after purification.

Flagellar motor protein (MotA, HP0815). The bacterial flagellar motor is a molecular machine that rotates helical filaments and allows bacteria to swim toward nutrients, optimal temperatures or other factors that favor survival (DeRosier, 1998). The flagellar motor is embedded in the bacterial cell wall, spanning the outer and inner membranes. The MotA and MotB protein components of the motor form a stator complex that uses the gradient of protons across the cytoplasmic membrane to generate

the turning force (torque) applied to the FliG component of the rotor (Manson et al., 1977; Blair and Berg, 1988). MotA has four α -helices that span the cytoplasmic membrane, with the rest of the molecule (about two thirds) forming a globular cytoplasmic domain (Zhou et al., 1995). MotB is anchored to the cytoplasmic membrane through its N-terminal hydrophobic-helix, with the bulk of the protein in the periplasm (Chun & Parkinson, 1988).

In *H. pylori*, the MotA–MotB complex attaches to the cell wall *via* the MotB periplasmic domain. Motility by the flagellar motor is required during the initial colonization of the stomach (Eaton et al., 1992) and is needed for the bacteria to attain full infection levels (Ottemann and Lowenthal, 2002). The molecular mechanism of the torque generation by the stator complex remains obscure, owing to a lack of structural information for the stator and the rotor–stator interface.

HP0815 has a molecular weight of 27.7 kDa and an acidic isoelectric point of 5.3. It is predicted to have a secretion signal at the N-terminus. For this reason the protein was isolated and cloned without the nucleotide sequence that codified for its first 20 amino acids. The nucleotide sequence that corresponds to amino acids from 21 to 257 was amplified and cloned into the pET 151/D-TOPO expression vector using the TOPO cloning kit (Invitrogen). This construct was used to perform some HP0815 expression trials, but in all conditions tested the level of protein expression was too low to perform a purification and the protein was shown to be toxic, leading to bacterial death. To improve the expression levels of this protein some different strategies have to be adopted, for example the use of different vectors or different expression systems.

2.3.3 Essential components of metabolic pathways

AroB 3-dehydroquinate synthase (HP0283). The *aroB* gene encodes the enzyme 3-dehydroquinate synthase that catalyzes one of the early steps in the shikimate pathway. This pathway, which creates aromatic molecules from sugar precursors, is present in prokaryotes, fungi and plants, but it is absent in mammalian cells. The predicted amino acid sequence of the *H. pylori aroB* gene product showed significant

homology (30 – 40% identity and 50 – 60% similarity) to 3-dehydroquinate synthases from various other prokaryotes and eukaryotes. The single gene on a plasmid was biologically active in *E. coli*. It suppressed the specific phenotype of *aroB* mutants by restoring the shikimate pathway-dependent synthesis of aromatic amino acids and the production of the siderophore enterobactin. In bacteria, aromatic amino acids and para-amino benzoic acid (PAB) are synthesized from the common precursor chorismate, which is the product of the shikimate pathway. PAB is the precursor for folic acid synthesis, a substance that is not produced by chordates. As a consequence, mutants of pathogenic bacteria, which lack any of the enzymes of the shikimate pathway, require aromatic amino acids and PAB for growth, and are thus unable to grow in the host. The shikimate pathway of bacteria, including the function of the AroB protein, has been extensively studied in *Escherichia coli* (Frost and Knowles, 1984; Millar and Coggins, 1986; Patnaik and Liao, 1994; Pittard, 1996). The enzyme catalyzes the formation of 3-dehydroquinate by cyclization of the sugar precursor 3-deoxy-D-arabino-heptulosonate-7-phosphate. The reaction, which depends on NAD⁺ as a cofactor, represents the second step in the shikimate pathway. In *E. coli* and other enteric bacteria, the shikimate pathway is linked to iron metabolism, as the synthesis of the siderophore enterobactin depends on chorismic acid from the shikimate pathway (reviewed in Earhart, 1996). As a consequence, *aroB* mutants of *E. coli* fail to grow in minimal media and cannot survive iron deprivation. *hp0283* was isolated, amplified and cloned in the pET151/D-TOPO expression vector using the TOPO cloning kit (Invitrogen). Expression trials for this protein have not yet been performed, but it could be an interesting target to study.

AroE shikimate dehydrogenase (HP1249). Shikimate dehydrogenase (SDH, EC 1.1.1.25) catalyzes the fourth reaction in the shikimate pathway, and is responsible for the NADPH-dependent reduction of 3-dehydroshikimate to shikimate. SDH belongs to the superfamily of NAD(P)H-dependent oxidoreductase. In plants, including *Pisum sativum* and *Nicotiana tabacum*, SDH is associated with 3-dehydroquinate dehydratase to form bifunctional enzyme (Bonner and Jensen, 1994;

Deka and Coggins, 1994). In fungi and yeast, such as *Aspergillus nidulans* and *Saccharomyces cerevisiae*, SDH exists as a component of the penta-functional AROM enzyme complex that catalyzes steps 2–6 within the shikimate pathway (Charles and Hawkins, 1986; Duncan and Coggins, 1987). In most bacteria, SDH functions as a single mono-functional enzyme. There are two SDH orthologues, AroE and YdiB, in *E. coli*, *Salmonella typhimurium*, *Streptococcus pneumoniae*, and *Haemophilus influenzae*. AroE is strictly specific for shikimate, while YdiB utilizes either shikimate or quinate as substrates in the shikimate or quinate pathway. However, the complete genome sequence of *H. pylori* has revealed the presence of only AroE, which plays an essential role in its metabolism. Recently, the three-dimensional structure of AroE from several bacteria such as *E. coli*, *Methanococcus jannaschii*, and *H. influenzae*, and YdiB from *E. coli*, including structures of enzyme-cofactor complexes, have been published (Benach and Hunt, 2003; Ye S and McRee, 2003). All these structures reveal a common fold comprising two domains that are responsible for binding substrate and NADP cofactor. The detailed structural information might expedite the discovery of novel SDH inhibitors and further of antimicrobial agents, though few SDH inhibitors have yet been reported so far. *hp1249* was isolated, amplified and cloned in the pET151/D-TOPO expression vector using the TOPO cloning kit (Invitrogen). Expression trials for this protein have not yet been performed.

Nicotinamide mononucleotide transporter (pnuC, HP1290). The pyridine cofactors NAD and NADP are electron carriers essential for both catabolic and biosynthetic redox reactions. In addition, NAD is used by bacterial DNA ligase to activate single-strand ends prior to joining (Olivera and Lehman, 1968; Zimmerman and Gellert, 1967) and serves as a precursor of cofactor B₁₂ (Maggio-Hall and Escalante-Semerena, 2003.). The pathway in which pnuC is involved has been extensively studied in *Salmonella* (Starai and Escalante-Semerena, 2002; Andreoli and Grover, 1972; Foster and Moat, 1979; Zhu and Roth, 1989; Spector and Foster, 1985; Tirgari and Foster, 1986; Zhu and Roth, 1989). A role for PnuC in transport is consistent with the

multiple membrane-spanning domains of this molecule (Sauer and Reidl, 2004) and the presence of a functional signal sequence (Zhu and Roth, 1989). It was initially thought that PnuC transports intact NMN, based on very convincing double-labeling experiments that demonstrated cotransport of the nucleotide phosphate and pyridine ring (Liu and Olivera, 1982). The role of NadR in transport was initially attributed to a posited regulatory interaction between internal NadR and the PnuC transporter (Foster and Spector, 1990; Zhu and Roth, 1991). This contrasted with the finding that in *Haemophilus influenzae* both the two periplasmic phosphatases convert NMN to NmR prior to transport (Kemmer and Reidl, 2001). The *Salmonella* NadR protein is now known to have additional enzymatic activities, serving also as a transcriptional repressor (Grose and Roth, 2005). NadR(R) represses transcription of the *nadB* and *pncB* genes and the *nadA-pnuC* operon when levels of NAD are high (Cookson and Roth, 1987; Spector and Foster, 1985). When NAD levels are low, the repressor activity is lost and NadR perform two enzymatic activities, NmR kinase activity (Kurnasov and Osterman, 2002) and NMN adenylyltransferase activity (Raffaelli and Magni, 1999), both of which are inhibited by NAD through a feedback (Grose and Roth, 2005). The NmR kinase, NadR(T), contributes to the transport of pyridine by trapping NmR inside the cells as the charged pyridine compound NMN. The NMN adenylyltransferase activity is contributed by a domain that also mediates feedback regulation of all three activities by NAD (Grose and Roth, 2005). In order to try to characterize HP1290 in *Helicobacter pylori*, we have produced different *pnuC* constructs, based on literature and bioinformatics analysis. PnuC is a protein with 220 amino acids, corresponding to a molecular weight of 25.4 kDa. HP1290 is an integral membrane protein with seven trans-membrane helices and with a basic theoretical isoelectric point of 9.3. Firstly, we have cloned the protein without its predicted secretion signal sequence, plus a N-terminus hexa-histidine tag, into the pET 151/D-TOPO expression vector using the TOPO cloning kit (Invitrogen). After preliminary expression tests in which no protein expression was detected, different *E. coli* strains (as BL21(DE3), C41(DE3) and C43 (DE3), and BL21(DE3)pLysS) were tried, without success. It was decided to clone

the full-length protein using different strategies. *hp1290* gene was cloned in pTTQ18, a vector usually used to express membrane proteins; as a second attempt, the protein was cloned in pET20, a vector with the sequence that codifies for the His₆-tag at the C-terminus, and in pET28a, a vector with the His₆ at the N-terminus; finally, the protein was cloned without any tag. All these constructs were then used to perform expression trials, but without any positive result. HP1290 is an integral membrane protein that is known to be hard to produce and purify. A future possibility could be the use of a new expression strategy, like a cell-free system.

2.3.4 Secreted factors involved in stress

Acidic stress response factor (HP1286). HP1286 is a member of the YceI family and it plays a relevant role in bacterial colonization and persistence in the stomach. In addition to major virulence factors that contribute to the inflammatory response and to epithelial cell damage proteins, CagA and Vac A (Hatakeyama, 2006; Parsonnet and Hiatt, 1997; Cover and Blaser, 1990; Cover and Blaser 1993), there are several secreted proteins implicated in the damage that have been identified, but for most of them the effective role on secretion, the physiological effect and relevance of this secretion are often unclear. One of these proteins is Hp1286, which has been found present in the external medium by different independent study.

HP1286 is a protein of 182 amino acids. It was classified as belonging to the YceI-like family of proteins analyzing its primary sequence (Karow and Georgopoulos, 1991). The protein has been cloned, expressed and its structure determined, as extensively described in Chapter 5.

HP0170. HP0170 is a remote homologue of CheZ. Motile and chemotactic bacteria create an expanded colony, relative to non-motile or non-chemotactic strains. This process is more studied in archea and other model organism, such as *E. coli*. Flagellar rotation is controlled by the core chemotaxis signaling pathway. The first components of this pathway are hemoreceptors, which detect chemical cues in the environment and relay this information to the flagellar motor (reviewed in Stock and Surette, 1996). Ligands bind to a chemoreceptor and indirectly affect the

conformation of a histidine kinase, CheA. CheA phosphorylates the response-regulator CheY, and phosphorylated CheY (CheY-P) influences flagellar motor rotation. CheW couples the chemoreceptors to CheA, and this protein is required for signaling. Chemotactic microbes also contain proteins involved in removing the CheY-P signal by dephosphorylating CheY. One of these proteins is CheZ, which accelerates the intrinsic autodephosphorylation of CheY-P. CheZ has been hypothesized of having evolved recently, since orthologues are found in only a limited group of prokaryotes, the α - and β -proteobacter (Szurmant and Ordal, 2004). Another protein that may attenuate the CheY-P signal is CheV, a protein that contains a C-terminal CheW-domain linked to an N-terminal CheYlike domain that has recently been shown to be involved in the phosphate flow in *H. pylori* chemotaxis (Jimenez-Pearson *et al.*, 2005). These proteins can function both as CheW-redundant coupling proteins (Rosario *et al.*, 1994) and like phosphate-sink CheYs. The latter accept phosphates from CheA when the flagellar-motor CheY is fully phosphorylated (Sourjik and Schmitt, 1998). Phosphate sinks act in some way similarly to CheZ, interrupting the flow of phosphate to CheY and thus allowing CheY-P dephosphorylation. Unlike CheZ, CheV proteins are found in most bacteria, including *H. pylori*, but not in *E. coli* or archaea (Szurmant and Ordal, 2004).

hp0170 was isolated, amplified and cloned into the pET 151/D-TOPO expression vector using the TOPO cloning kit (Invitrogen). The recombinant His₆-HP0170 was expressed in *E. coli* BL21(DE3) cells. During the purification process, performed using a metal-chelate affinity chromatography (GE Healthcare), several problems arose, mainly due to lack of protein solubility. In future a refolding strategy on the protein purified from inclusion bodies will be tried.

2.4 Conclusion

The project presented here was aimed at the determination of the three dimensional structure of proteins and to their analysis and *in vitro* characterization. It is well established that the structure of a protein is closely related to its function and, in that sense, a structural genomics approach has a great potential. Here we have addressed several potential targets, including proteins involved in different metabolic pathways or essential for pathogen colonization or virulence. All genes presented here were cloned using different strategy; 7 over 9 of them were expressed; 3 over 7 were soluble and all these three proteins have been structurally characterized, at least partially. They are extensively described in the following chapters.

Chapter 3

CagL from Helicobacter pylori

Contents

3.1 Introduction

3.2 Materials and Methods

3.3 Results and discussion

3.4 Conclusion

3.1 Introduction

The Gram-negative bacterium *Helicobacter pylori* lives near the surface of the human gastric mucosa and is one of the most successful bacterial pathogens. The discovery of *H. pylori* by Robin Warren and Barry Marshall about 25 years ago had fundamental consequences for current understanding and treatment of stomach diseases (Warren and Marshall, 1983; Backert and Selbach, 2008). A wide range of Gram-negative bacteria pathogens, like *H. pylori*, translocate virulence factors into host target cells by multisubunit transport apparatuses known as type-IV secretion systems (Cascales and Christie, 2003; Backert, and Meyer, 2006) which injects CagA into host cells (Covacci and Rappuoli, 2000).

The gastric pathogen exploits integrin receptors for the injection of virulence factors into mammalian cells (Kwokand and Backert, 2007). This is achieved by a type IV secretion system (T4SS) consisting of 11 VirB protein orthologs (encoded by *virB1-11* genes) and the so-called coupling protein (VirD4, an NTPase). The VirB proteins of *Agrobacterium* can be grouped into three classes: (i) the putative channel or core components (VirB6-10), (ii) the energetic components (the NTPases VirB4, VirB11) and (iii) the pilus-associated components (VirB2 and, possibly, VirB3 and VirB5). These proteins are encoded by a 40-kb gene cluster known as the *cag* (cytotoxin-associated gene) pathogenicity island (*cagPAI*; Peek and Blaser, 2002; Amieva and El-Omar, 2008).

The role of most of the *cagPAI*-specific accessory factors is unknown, but the functions of CagF and CagL have been elucidated. CagF is a chaperone-like protein that binds close to the C-terminal secretion signal of the CagA effector protein and is crucial for the translocation of CagA (Couturier, 2006) (Pattis, 2007). CagL is a pilus-covering protein which acts as a specialized adhesin that connects the T4SS with target cells (Kwok, 2007).

Integrin $\alpha_5\beta_1$ binds to CagL, a small 26-kDa protein. The latter is encoded by the open reading frame HP0539 in the *cagPAI* (Kwokand and Backert, 2007). CagL is predicted to be a functional VirB5 ortholog and a structural component of the T4SS pilus (Backert, and Waksman,

2008; Aly and Baron, 2007). It has no significant sequence homology to any known eukaryotic protein and it is a protein highly conserved among pathogenic *H. pylori* strains.

Many integrin ligands, such as fibronectin and vitronectin, carry the arginine-glycine-aspartate (RGD) motif, which serves as a recognition site for integrins. CagL is the only cagPAI-encoded gene product that contains an RGD motif. CagL RGD motif has been shown to be important for the interaction with the $\alpha_5\beta_1$ integrin (Kwokand and Backert, 2007; Tegtmeyer and Backert, 2010). However, recent evidences in yeast two-hybrid screens have shown that other T4SS proteins, such as CagY (VirB10), CagN, and the effector protein CagA can also bind β_1 integrin *in vitro* (Jime'nez-Soto, and Haas, 2009), confirming that *H. pylori* targets this integrin member as a receptor for the T4SS. However, mutation of the RGD motif in CagL had no effect in T4SS functions, such as the phosphorylation of injected CagA (Jime'nez-Soto, and Haas, 2009).

CagL appears to function as a specialized adhesin that not only anchors the type-IV secretion apparatus to the host surface through binding to integrin, but also promotes signal transduction. CagL triggers the activation of FAK, a key sensor of integrin engagement. Upon activation, FAK is autophosphorylated and binds the SH2 domain of Src, resulting in an active FAK–Src signaling complex². The activation of FAK and Src coincides with the occurrence of CagA-pY in infected cells, indicating that FAK activation and subsequently the activation of Src rapidly results in tyrosine phosphorylation of CagA (Kwokand and Backert, 2007).

At present, the three-dimensional structure of only four cag-PAI proteins is available: CagA (HP0525), which is an ATPase located at the inner membrane (Hare and Waksman, 2007); CagZ (HP0526), a 23-kDa protein involved in the translocation of CagA (Cendron and Zanotti, 2004); CagS (HP0534), a 23-kDa protein coded by a well conserved gene in the cag-PAI but whose function remains elusive (Cendron and Zanotti, 2007), and CagD (HP545), a 25-kDa essential for CagA toxin secretion that may serve as a unique multifunctional component of the T4SS (Cendron and Zanotti, 2009).

In this study, we used biophysical methods to study the oligomeric state of CagL in solution. We report the structural model of CagL using the SAXS method and the oligomerization state in solution using cross-linking experiments and native gel. A model consistent with our observations and a survey of data is proposed.

3.2 Materials and Methods

Cloning

The nucleotide sequence corresponding to amino acids from 21 to 237 was amplified by PCR starting from *H. pylori* genomic DNA (strain CCUG 17874) using a high-fidelity thermostable DNA polymerase (Deep VentDNA polymerase, New England Biolabs) and two flanking primers: hp0539 forward and hp0539 reverse. These primers were built to introduce NdeI and XhoI restriction sites at the 5' and 3' ends, respectively. The amplified *cagL* (hp0539) was double digested with NdeI and XhoI enzymes and inserted into a pET28b expression vector (Novagen), encoding an in-frame His6 tag followed by a thrombin cleavage site at the N-terminus of the recombinant construct (His6-CagL).

Overexpression and purification

Escherichia coli BL21(DE3) cells (Invitrogen) transformed with the corresponding over expressed plasmids pET28b-*cagL* (pET28b-HP0539) were grown at 37°C in 2 liters of LB medium, with shaking, containing 25 µg/ml of kanamycin. The expression was induced at an OD600 of 0.7–0.8 by adding 1mM IPTG (isopropyl-β-D-thiogalactopyranoside) and prolonged for 4 h at 30 °C under vigorous shaking. Bacterial pellets were re-suspended in ice-cold buffer (50mM KH₂PO₄-K₂HPO₄, pH7.5, 200mM NaCl) supplemented with a protease inhibitor cocktail (complete, mini, and ethylenediaminetetraacetic acid free; Roche) and disrupted by sonication.

The over expressed CagL present in the inclusion bodies, obtained after 25 min of centrifugation at 40.000g, was solubilized in buffer (50mM KH₂PO₄-K₂HPO₄ pH7.5, 200mM NaCl, 6M guanidine hydrochloride)

and refolded in ice-cold refolding buffer (52mM Tris-HCl pH 8.2, 20mM NaCl, 834mM KCl, 1.1mM EDTA, 2.1mM reduced glutathione, 210 μ M oxidized glutathione). After refolding, CagL was further purified by metal-chelate affinity chromatography equilibrated with buffer A (52mM Tris-HCl pH 8.2, 20mM NaCl, 834 μ M KCl) at about 1–2 ml/min. After two extensive washes with 5% and 10% of buffer B (52mM Tris-HCl pH 8.2, 20mM NaCl, 834mM KCl and 500mM imidazole), the protein was eluted in a single peak at 200mM imidazole by applying a linear gradient from 10% to 100% of buffer B. The fractions containing the pure His6-CagL were pooled, concentrated by ultrafiltration (molecular weight cutoff 10 000; Millipore), diluted into buffer A, and incubated overnight at 4 °C with thrombin protease to remove the N-terminal His6 tag.

The oligomeric state of purified CagL in solution was examined by analytical size-exclusion chromatography using a Sephacryl S-200 (16/60, Amersham Biosciences) column equilibrated with buffer A. The CagL soluble fraction was purified and gel filtrated in the same way as the refolded protein.

The folded conformation of the purified CagL was tested by circular dichroism using a J-720 spectropolarimeter (Jasco Instruments) for both the refolded and soluble fractions.

Crosslinking experiment in solution

For Cross-linking reactions the General Cross-Linking Protocol (Thermo scientific, <http://www.funakoshi.co.jp/data/datasheet/PCC/22585.pdf>) was used at three different protein concentration (1mg/ml; 1,5mg/ml; 2mg/ml). The cross-linking protocol was tested with two different cross-linking reagent: DSP and 1% glutaraldehyde.

To prepare the cross-linker solution, DSP was dissolved in DMSO at a 10-25mM concentration. The protein was dissolved at increasing concentrations (1mg/ml; 1.5mg/ml; 2mg/ml) in a phosphate buffered solution (0.1M phosphate, 0.15M NaCl; pH 7.2). The cross-linker was added to the protein sample in a 20- to 50-fold molar excess. The reaction mixture was then incubated at room temperature for 30 minutes.

In the end, a solution of 1M Tris, pH 7.5, at a final concentration of 25mM was added to stop the reaction and incubated for 15 minutes.

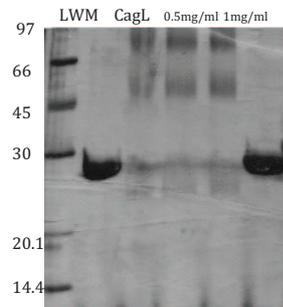


Figure 1. SDS-PAGE of the Cross linking experiment using CagL at 2mg/ml as a control (line 2), CagL incubated with 1% glutaraldehyde at three different concentrations (1mg/ml;1,5mg/ml;2mg/ml) lines 3,4,5; 2mg/ml CagL incubated with 1% glutaraldehyde and DTT to inhibit the cross linking traction as control.

Native gel

The Acid-native Gel Protocol available on line (http://wolfson.huji.ac.il/purification/Protocols/PAGE_Acidic.html) was used. A separating gel solution at 10% Acrylamide (1.5M Acetate-KOH pH 4.3; 50% Glycerol; 40% Acrylamide; H₂O; 10% APS; TEMED) and a stacking gel solution at 4% (0.25M Acetate-KOH pH 6.8; 40% Acrylamide; H₂O; 10% APS; TEMED) were prepared. Running conditions: 30mA / 250V max.

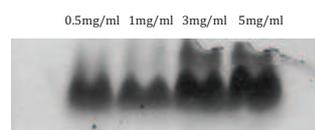


Figure2. Native gel using four different CagL concentrations: 0.5mg/ml; 1mg/ml; 3mg/ml; 5mg/ml

Secondary structure analysis of CagL

The circular dichroism (CD) spectrum of purified CagL (1.4 mg/ml) was measured at room temperature in buffer (52mM Tris-HCl, pH 8.2, 20mM NaCl, 834mM KCl), using a JASCO-J720 Spectropolarimeter and a 0.02cm path length cell. The spectra in the far-UV (190-260nm) were recorded at a scanning speed of a 50nm/min. Ten spectra were accumulated and averaged, followed by baseline correction by subtraction of the buffer. Mean residue weighted ellipticity was calculated and expressed in units of degree $\text{cm}^2 \text{dmol}^{-1}$. The spectrum deconvolution showed a content of about 85% α -helix, 2% β -sheet, 8% turns, and 5% random-coil (using the program CDNN2.1).

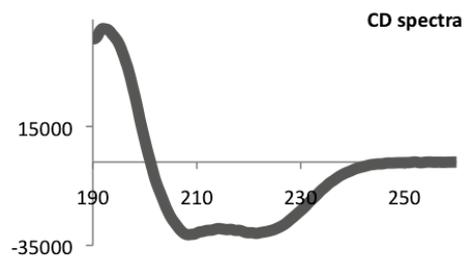


Figure3. Circular dichroism of a refolded and purified fraction of CagL.

Analytic gel filtration

The purified protein was analyzed by analytic gel filtration chromatography to characterize its oligomerization state. A small amount of sample was loaded in a Superdex 200 HR 5/150 column, equilibrated in the same CagL buffer (52mM Tris HCl pH8.2; 20mM NaCl; 834 μ M KCl). The protein eluted with a retention volume of 2,04 ml, corresponding to a molecular weight of 57,3 kDa, 2.3 times the mass of the monomer (25kDa), suggesting that the oligomerization state for CagL is in a dimeric, or possibly trimeric, form.

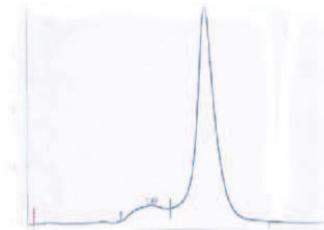


Figure 4. Analytic gel filtration (Superdex 200 HR 5/150) of refolded and purified CagL.

3.3 Results and discussion

The refolded-soluble protein was characterized with small angle X-ray scattering (SAXS), circular dichroism (CD) spectroscopy, and chemical cross-linking to evaluate the homogeneity, the oligomeric state, the radius of gyration, and the overall fold. In addition to far-UV CD spectra, SAXS and analytic gel filtration, chemical cross-linking data were obtained for CagL to confirm the correct refolding, the prevalence of the helical secondary structure content and to probe the oligomeric states of the protein. CagL exists as a multiple oligomer form in solution, as a dimer and partially as a trimer or tetramer. We cannot state if its oligomerization state is crucial for the interaction with the $\alpha_5\beta_1$ integrin. When bacterially expressed, CagL was examined on an analytic gel filtration column. It eluted as a single oligomeric complex of approximately 57.3 kDa, which is equivalent to 2.3 times the molecular weight of monomeric CagL (25 kDa).

In cross-linking experiments, three different protein concentrations (1mg/ml; 1,5mg/ml; 2mg/ml) were used, along with two different controls, i.e. CagL without cross-linker treatment and CagL incubated with cross-linker and DTT. In the presence of both cross-linkers, DPS

and glutaraldehyde, the protein formed a mixture of high molecular weight cross-linked products. It was possible to distinguish three major bands in the gel, corresponding to molecular masses of about 25kDa, 47kDa, and 70kDa. They roughly correspond to a small fraction of monomeric CagL, and to the dimer and to the trimer, respectively. A weak band of more than 90 kDa, corresponding perhaps to the tetramer, is also present.

Native gels shows two CagL predominant oligomerization state, roughly corresponding to the dimer and the trimer, but is not so easy to estimate the size of each band. As we show the major part of CagL in solution is present as a dimer and a small fraction as a trimer.

In order to obtain structural information in solution, circular dichroism (CD) spectroscopy and small-angle X-ray scattering (SAXS) were also used. Circular dichroism was used also to check the correctness of the refolding process. Recombinant CagL is in fact a protein highly expressed in *E. coli*, but a large part of it precipitates as inclusion bodies. The comparison between the secondary structure content of the soluble and refolded fractions allowed to establish the correct folding of the refolded protein. The CagL CD spectrum was also used to predict the secondary structure. It demonstrated that CagL is a protein with a very high α -helical content, about 86%.

The information on the secondary structure was used to produce by sequence homology a CagL molecular model and to use it to fit the SAXS curves.

SAXS measurements were performed at the ID14-3 beamline of ESRF. Preliminary fitting using a theoretical molecular model are reported below. It can be seen that the best fits were obtained using a tetramer or a trimer. The most likely hypothesis is that in solution more than one oligomeric species is present; fitting using different combinations of mixed molecular are in progress.

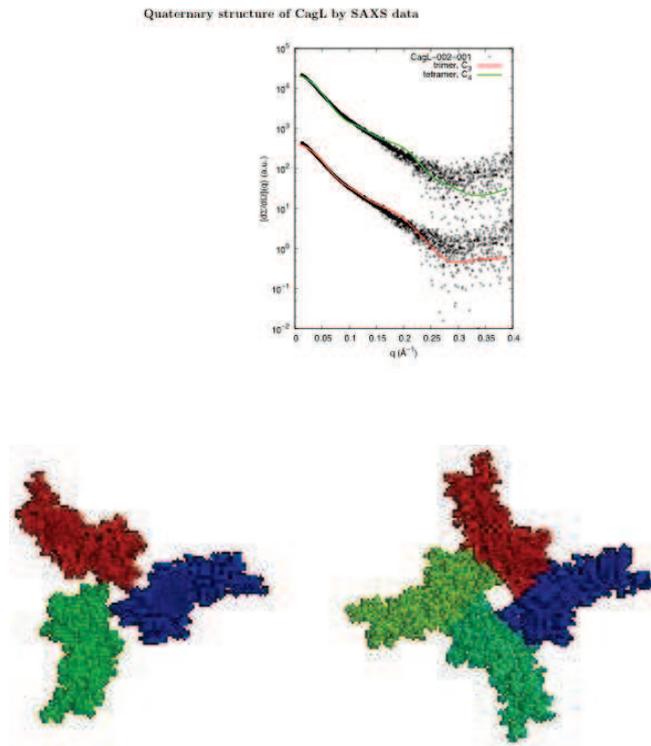


Figure 5. Oligomerization CagL model. (A) SAXS measurements performed at the ID14-3 beamline of ESRF. Preliminary fitting using theoretical molecular models are shown: trimer (red) and tetramer (green) (B) trimeric and tetrameric models of CagL used for the fitting of the curves in (A).

3.4 Conclusion

In this study we have analyzed CagL oligomerization state and structure in solution using different techniques. The protein monomer is characterized by a high α -helical content, a fact that supports the hypothesis that CagL is the orthologue of Virb5 of the *A. tumefaciens* T4SS. In solution the protein is present mainly as a dimer and as a trimer, and a model of these possible oligomerization states has been obtained through SAXS experiments. This analysis also suggests that the difficulties encountered in trying to obtain CagL crystals depend from the presence in solution of several different oligomerization forms.

Chapter 4

***Helicobacter pylori* adhesin A (HP0797)**

Contents

4.1 Introduction

4.2 Materials and methods

4.3 Results and discussion

4.1 Introduction

Among the first steps of a bacterial infection, there is the attachment to the host cell surface. It is generally admitted that many bacteria, as well as their toxin, interact specifically with discrete regions of the plasma membrane that are rich in cholesterol, sphingomyelin, and glycosphingolipids (Duncan and Abraham, 2002). These membrane microdomains, usually referred to as lipid rafts, are attractive to a wide range of pathogens, including viruses, bacteria, parasites (Fantini and Yahi, 2002). The molecular mechanism involved in raft-pathogen interactions is still poorly understood. The characterization of structural

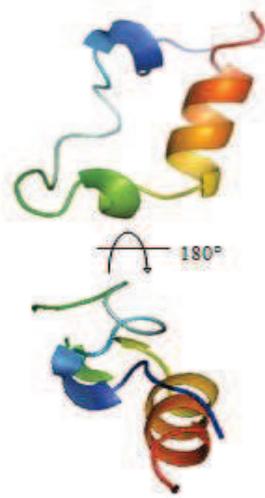


Figure 4.1 Three-dimensional structure of the glycolipid-binding domain (PDB entry 1CE4).

motifs ensuring protein binding to glycolipid receptors is of primary importance. It is possible to identify a common glycolipid-binding domain in HIV-1 surface envelope glycoprotein gp120, Alzheimer's β -amyloid peptide, and the cellular isoform of the prion protein (PrP^c) (Mahfoud and Fantini, 2002). This motif consists of a hairpin structure containing a water-exposed aromatic residue (Fig. 4.1). As many bacteria attach to host cells through binding of their adhesins to cell surface glycolipids (Karlsson, 1989), it is rational to search for glycolipid-binding domains in bacterial adhesins. The *H. pylori* adhesin A

(HpaA) has been shown to mediate the binding of the bacteria to sialic acid-containing host molecules expressed on the surface of gastrointestinal cells (Evans and Lee, 1993). *H. pylori* adhesin A is a surface-located (Blom and Smith, 2004; Evans and Graham, 1988; Lundstrom and Bolin, 2001) lipoprotein (O'Toole and Trust, 1995) that was initially described as a sialic acid binding adhesin, but supportive

evidence is still lacking. It is recognized by antibodies from *H. pylori*-infected individuals (Mattsson and Svennerholm, 1998; Yan and Shao, 2005), and expression of the HpaA protein has previously been found to be highly conserved among *H. pylori* isolates (Bolin and Svennerholm, 1995; Yan and Shao, 2005). Furthermore, genomic studies (Alm and Trust, 1999; Tomb and Venter, 1997) show no significant sequence homologies of HpaA with other known proteins. Except in specific long-passaged bacterial strains (O'Toole and Trust, 1995), HpaA is associated with the outer surface of the bacteria. As HpaA is highly conserved among *H. pylori* isolates and, as such, it is considered as a potential vaccine antigen (Lundstrom and Nystrom, 2003). Given the prominent role played by HpaA in *H. pylori* adhesion to host cells, it was sound to search for a potential glycolipid-binding domain able to recognize LacCer in the three-dimensional structure of this adhesin. It is possible to identify the putative glycolipid-binding domain (HIV-1 gp120 V3 loop) in the amino acid sequence of the protein deduced from the nucleotide sequence of this gene (Evans and Lee, 1993). As shown in figure 4.2, the motif consists of a hairpin structure containing:

- a solvent-exposed aromatic residue (Phe, Tyr, or Trp);
 - several charged residues (Asp, Glu, Arg, or Lys) with some of them oriented toward the solvent;
 - a Gly and/or a Pro residue inducing the turn in the backbone C α chain.
-

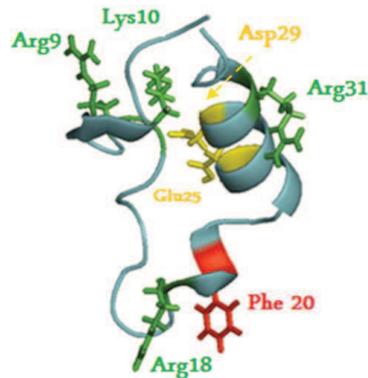


Figure 4.2. Three-dimensional structure of the glycolipid-binding domain and its interaction with GSL. HIV-1 gp120 V3 loop (PDB entry 1CE4). The presence of at least one aromatic residue, which has to be fully exposed to the solvent, is required (here Phe20, in red). Basic amino acid residues are in green and acid ones in yellow.

The charged amino acids in green and in yellow may interact with the charged polar head of membrane lipids, as sphingomyelin and gangliosides in lipid rafts. In the end, the presence of one aromatic residue, as Phe, Tyr, and/or Trp, which has to be fully exposed to the solvent, is important (Phe in red is shown in Figure 4.2). On this basis, it is possible to predict the secondary structure using the specific clusters of amino acid residues (Gly, Pro, Asp/Glu/Arg/Lys, Phe/Tyr/Trp). A similar motif can be identified in *H. pylori* HpaA.

```

      10      20      30      40      50      60
GIDPFTSPHĪ IETNEVALKĪ NYHPASEKVQ AFHEKILLLR PAFPYSHNIP NEYENKFKNQ
      70      80      90     100     110     120
TALKVEQILQ NQGYKVISVH SSDKDDFSFA QKKEGYLAVA MNGEIVLRPH PKRTIQKKSE
      130     140     150     160     170     180
PGLLFSTGLD KMEGVLPAG FIKVTILEPM TGESLDSFTM DLSELDIQEK FLKTTSSHT
      190     200     210     220     230
VGLVSTMVKG TDNSNDAIQS ALNKIFATIM QEIYKILTQT IYNLLKKPKK T

```

The previous sequence presents a peptide (yellow) containing the Phe residue (in red), a sequence PG that can induce the formation of a turn

and some charged amino acids in the vicinity of the aromatic F (KKEDK).

4.2 Materials and methods

Sequence analysis

HpaA *hp0797* (*H. pylori* strain CCUG 17874) codifies for a 260 amino acid protein. Since the full-length sequence was predicted to include a N-terminal secretion signal with high confidence (SignalP 3.0; Bendtsen *et al.*, 2004), the first 30 amino acids were excluded from the recombinant construct (His₆-Hp797). The mature protein has a molecular mass of 25,8 kDa, with a theoretical isoelectric point of 8.1. According to the CD spectrum, the content of secondary structure of Hp0797, expressed in a recombinant system and purified, is α -helix, β -turn and random coil at different levels.

Cloning

hp0797 gene was amplified by PCR from *H. pylori* CCUG 17874 genomic DNA, using as the forward primer FwHp797 (5'-CACCAGCCCGCATATTATTGAAACC-3') and as reverse the primer RwHp797 (5'-CTATTATCGGTTTCTTTTGCCTTTTAA-3'). The amplified gene fragment was cloned into pET151/D-TOPO vector using the TOPO cloning kit (Invitrogen), which permits directional cloning into expression system. The vector adds an His₆-tag upstream the Hp0797, a V5 epitope and a tobacco Etch Virus (TEV) protease cleavage site to the expressed recombinant protein. Correct cloning was verified by sequencing the DNA construct using standard primers for T7 promoters.

Gel filtration analysis

A sample of HP0797, in buffer 30mM Tris pH 8, 330 mM NaCl, coming from a preliminary purification trial (see next paragraph for details about purification) was used in order to obtain a single oligomerization state. The protein sample was divided in four parts and analyzed using different

detergents. The preliminary gel filtration fractionates in three different oligomerization state for HP0797, corresponding at trimeric, dimeric and monomeric form (Fig.4.3). Four different detergents at 0.5 CMC concentration were selected. The detergent used including C12E8, n,n-dimethyldecylamide-n-oxide (DDAO), n-Octyl- β -D-thioglucoiside and CHAPS. DDAO seemed to favor the oligomerization state in dimeric form, as shown in figure 4.4.

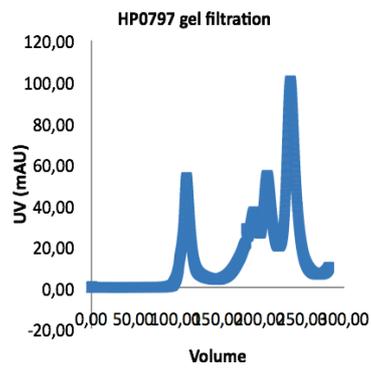


Figure 4.3 Size-exclusion chromatography profiles of HP0797 using 330mMNaCl, 30mM Tris-HCl pH8 as buffer.

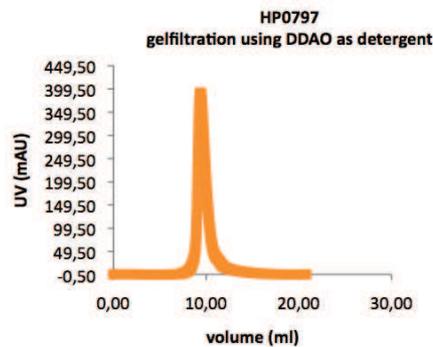


Figure 4.4. Size-exclusion chromatography profiles of HP0797 using 330mM NaCl, 30mM Tris-HCl pH8 and 1/2 CMC DDAO as buffer.

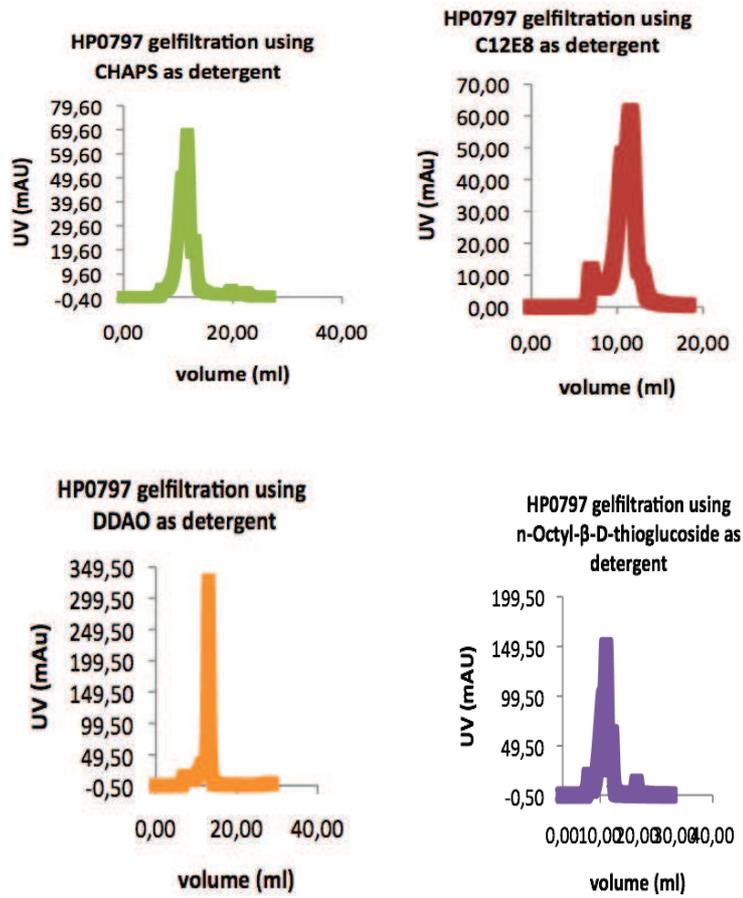


Figure 4.5. Size-exclusion chromatography profiles of HP0797 using 330mM NaCl, 30mMTris-HCl pH8 and 1/2 CMC of CHAPS, C₁₂E₈, DDAO and n-Octyl-β-D-thioglucoside, respectively, as buffer.

Expression and purification

His6-HP0797 protein was overexpressed in *E. coli* BL21(DE3) strain. Cells harboring pET151-HP0797 were grown in two liters LB medium containing 100µg/ml ampicillin until 0.7OD₆₀₀ at 37°C. The protein expression was induced using 1.0 mM isopropyl thio-β-d-galactoside, and the expression was prolonged for 3 h at 28°C. Bacterial cells were harvested by centrifugation at 6000 g and stored at -80 °C. The cell pellet was resuspended in buffer containing 30 mM Tris, pH 8, 300 mM NaCl and ½ CMC of DDAO. Lysis was achieved with lysozyme (1 mg/ml) incubation, followed by multiple sonication cycles (four times, 45 min each). The resulting supernatant was isolated from the insoluble fraction by centrifugation at 40,000 g for 25 min at 4°C, and loaded onto an Ni²⁺-immobilized metal-affinity preppacked column (GE Healthcare Europe GMBH, Orsay Cedex, France) equilibrated with the buffer. After two washing steps with the buffer supplemented with 20mM imidazole and 40mM imidazole, His6-HP1286 were eluted with an imidazole gradient to 500mM. The fractions containing the protein pooled was collected and incubated overnight at 4°C with His6-TEV protease. The sample was further subjected to an immobilized metal ion affinity chromatography step to remove the His6-rTEV protease and the residual uncleaved His6-HP0797. Then the protein was concentrated by ultracentrifugation using Vivaspin 15R 10 000 MW, Sartorius. The final purification step, size exclusion chromatography (Superdex 200 HR10/ 300; GE Healthcare) after equilibration with buffer A, resulted in a single peak and a retention time roughly corresponding to a protein dimer. The monodisperse protein was separated from the aggregated fraction and concentrated to 18mg/ml for crystallization tests. The protein concentration was determined by UV/VIS spectroscopy, using at 280nm wavelength. A Cary 50 Bio UV-Visible instrument was used. Each step of the protein purification was verified by SDS-PAGE.

Circular Dichroism analysis

The circular dichroism measurement was performed using a protein solution diluted to 1mg/ml with a JASCO-J720 Spectropolarimeter and a 0.02cm length cell. The spectrum in the far-UV (190-260nm) was recorded at a scanning speed of 50nm/min. Ten spectra were accumulated and averaged, followed by baseline correction by subtraction of the buffer. Mean residue weight ellipticity was calculated and expressed in units of degree $\text{cm}^2 \text{dmol}^{-1}$. The circular dichroism spectra were deconvoluted by software program CDNN (version 2.1).

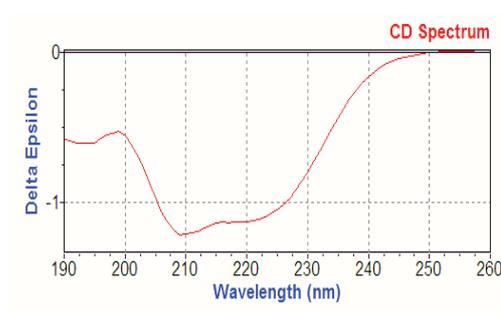


Figure 4.6. Circular dichroism of a purified fraction of HP0797

Western Blotting

Hp0797 samples were re-suspended in loading buffer and separated by SDS-page. This gel was used to transfer the resolved protein on a nitrocellulose membrane (HybondTM, GE Healthcare) with an electrophoretical system. The membrane was incubated with the primary antibody against His-tag in a blocking reagent at room temperature for 2 hours. Then the membrane was washed three times in BSA and PBS-Tween (0.15% Tween 20). The membrane was then incubated for one hour at room temperature with a secondary antibody, anti-mouse IgG, linked to alkaline phosphatase, and then diluted in the blocking reagent. The membrane was washed again with the same system described before

and developed using 5-bromo-4-chloro-3-indolyl-1-phosphate (BCIP) and nitro blue tetrazolium (NBT).

Crystallization and preliminary X-ray diffraction data

A wide range of crystallization conditions were tested with the vapor diffusion method, by using hanging and sitting drop techniques. The best crystals were obtained at 20°C using 10mg/ml of protein and the stock solution containing 0.2M Sodium citrate, 0.1M Tris HCl pH 8.5, 30% v/v PEG 400 (SS1-34 Molecular Dimensions Limited) as precipitant. Crystallization conditions were found using an automated system (Oryx drop maker, Douglas Instruments, England). A preliminary diffraction data set at 4Å resolution was collected at the ID23-2 ESRF beamline (Grenoble, France). The dataset were processed and scaled with programs MOSFLM and then SCALA as implemented in the CCP4 package (collaborative Computational Project 1994).

Molecular modeling

Using the SIM software (Local similarity program), it is possible to compare neuraminylactose-binding hemagglutinin (HP0410, PDB ID 3bgh) and HP0797 amino acid sequences. The two proteins present a high identity and sequence similarity. It is consequently possible to build a model of HP0797 using SWISS-MODEL, a fully automated protein structure homology-modeling server, accessible via the ExpASY web server, or from the program DeepView (Swiss Pdb-Viewer). This will make it possible to use the Molecular Replacement method for structure determination when better diffraction data will be available.

20.6% identity in 223 residues overlap; Score: 121.0; Gap frequency: 6.7%

```

HP0797      -----GNDPFTSPHIIIEINEVALKLNYPASEKVQALDEKI 36
3bghA      MKKGS LAIVLGSLLASGAFY TALADGMPAKQQHNTGESVELHFHYPIKGGKEPKNSHLV 60
           .. * .. *   :.* *:::*  .::  .. :

HP0797      LIIRPAFQYSDNIAKEYENKFKNQ TALKVEQILQNQGYKVISVDSSDKDDFSFAQKKEGY 96
3bghA      VLIEPKIEINKVIPESYQKEFEKSLFLQLSSFLERKGYVSQFK--DASEIPQDIKEKAL 118
           :::.* :. . . :.*:::*:::  *:::..*:::*  . . . * .:::  *:::

HP0797      LAVAMNGEIVLRPDPKRTIQKKSEFGLLFSTGLDKMEGVLPAGFIKVTILEFMSGESLD 156
3bghA      LVLRM DGNVAILED---IVEESD-----ALSEKVIDMSSGYLNLNFVEPKSEDIH 167
           *.: *:*:::  *  *  :.*:  .*.:  : : : :.*:::..*:::*  * :.

HP0797      SFMDLSELDIQEFKLTTHSSHSGGLVSTMVKGTDNS--NDAIKSALNKIFANIMQEID 214
3bghA      SFGIDVSKIKAVIERVELRRRTNSGGFVPKTFVHRIKETDHDQAIRKIMNQAYHKVMVHIT 227
           ** :.*:::  . : : : :. * :.*:*  .: : :.*:  .*: : :.* :.

HP0797      KKLTQKNLESYQKDAKELKGRNR 238
3bghA      KELS KKHMEHYEKVSSEMKKRK-- 249
           *:*:*:::*  *:*  :.*:  :.

```

4.3 Results and discussion

The recombinant HP0797 protein was expressed in *E. coli* in a soluble form, without the putative signal peptide identified using SignalP program. After gel filtration chromatography, the protein resulted to be present in at least three oligomerization forms. To obtain a single oligomerization form, different detergents, at concentration lower than CMC, were added at protein samples. One of them, DDAO, was identified as able to maintain the protein in a stable dimeric form. Monodisperse HP0797 was purified in high yield for crystallization purposes in a DDAO-containing buffer, using Ni²⁺-affinity and then size exclusion chromatography. The high purity of the protein solution was confirmed by SDS-PAGE.

Western blotting analysis was performed on samples after affinity chromatography, using an antibody against His-tag, to confirm the presence of His-tagged protein.

The study and analysis of the Hp0797 amino acid sequence show a high sequence conservation of this protein in different strains of *H. pylori*. A strong sequence homology involves also this protein with another adhesin of our bacterium, whose structure has recently been solved:

Hp0410, another member of the HpaA family. These proteins present a high sequence similarity and identity of 20.6% .

These data have made possible the realization of a theoretical structural model of Hp0797 using the SWISS-MODEL program (Fig 4.7). All data presented so far confirm the high conservation of the structural features of proteins that are member of the general HpaA family. The same considerations are not valid when considering special regions of these proteins, such as the glycolipid binding domain. The function and importance assigned to this domain in the recognition and interaction mechanism with host cells has been discussed previously. This domain is easily recognizable in the structural model of Hp0797, green in the figure 4.7B.

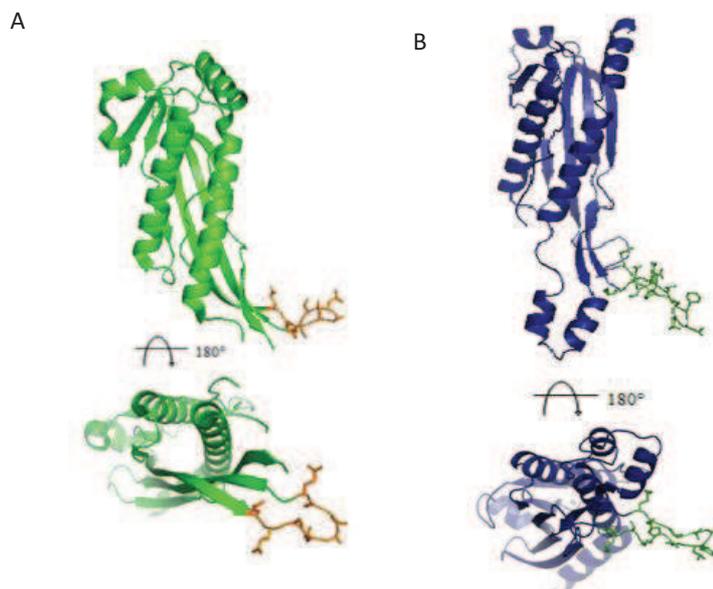


Figure 4.7. (A) In green, ribbon representation of the monomeric unit of (Bomanno and Almo, 2009) HP0410 [PDB accession code 3BGH]. In orange, the glycolipid binding portion absent in HP0410. (B) In blue, the HP0797 model, in green the glycolipid-binding portion corresponding to the amino acid sequence KKSEPGLLFSTGLDK

This motif consists in a hairpin structure containing a water-exposed aromatic residue. Probably *Helicobacter pylori* attach to host cells through binding to cell surface glycolipids of Hp0797 using this domain. For this reason it is rational to search for glycolipid-binding domains in bacterial adhesins.

The situation is totally different in the Hp0410. Although this protein is another adhesin, it differs for the total absence of the glycolipid binding domain in its sequence. The corresponding region is shorter and it has a completely different amino acid sequence in the glycolipid binding domain. This finding has two important consequences: the first is that Hp0410 certainly uses a different mechanism for the recognition and interaction with the surface of the host cells, now unknown. The second one is that the comparison between Hp0797 and Hp0410 demonstrates how proteins with the same function use different strategy to recognize the host cell.

Over the last twenty years, research has revealed the enormous complexity underlying the phenomenon of bacterial adhesion. The initial research goal was to determine the crystal structure of Hp0797 and to understand the mechanism of attachment and its effect on the bacteria as well as on the host. At this point, however, it is evident that many different forms of adhesion exist. They are the result of evolutionary pressure, and each of them may be a part of a more complex behavior strategy of the bacterium.

Chapter 5

Acidic stress response factor HP1286 from Helicobacter pylori

(The content of this chapter has been published in:

Sisinni et al., "Helicobacter pylori acidic stress response factor HP1286 is a YceI homolog with new binding specificity" FEBS J:1896-905, 2010 Apr; 277(8).)

Contents

5.1 Introduction

5.2 Materials and Methods

5.3 Results and discussion

5.1 Introduction

Helicobacter pylori is a noninvasive, gram-negative bacterium that colonizes the gastric mucosa (Chen and Diavolitsis, 1986; Hessey and Dixon, 1990; Kazi and Kazi, 1990). Gastric colonization by *H. pylori* results in a mucosal inflammatory response and is a risk factor for peptic ulcer disease and gastric malignancy (Cover and Blaser, 1996; Dunn and Blaser, 1997; Labigne and de Reuse, 1996; Mobley, 1997). Elucidation of the mechanism underlying inflammation and damage of the underlying tissue is important to understanding the development and progression of these gastrointestinal disorders. Epithelial cell damage may occur as a direct effect of bacterial habitation or as a consequence of the chronic and acute inflammatory responses induced by *H. pylori*, which may be due in part to responses to proteins released or injected by the organism (Blaser, 1992; Blaser and Parsonnet, 1994; Crabtree and Rathbone, 1991; Lee and Hazell, 1993). Several potential bacterial virulence factors that may contribute to mucosal inflammation and epithelial cell damage have been identified (Blaser, 1992; Leunk and Morgan, 1988). Major virulence factors that contribute to the inflammatory response and to epithelial cell damage are a high-molecular-mass protein encoded by *cagA* gene (cytotoxin-associated gene protein A) (Hatakeyama, 2006; Parsonnet and Hiatt, 1997), the 87-kDa vacuolating toxin A (Cover and Blaser, 1990; Cover and Blaser, 1993) encoded by *vacA* gene, and the *H. pylori* neutrophil-activating protein (Satin and Rossi, 2000; Nishioka and Montecucco, 2003). Other proteins that are secreted have been identified, but for most of them, the effective role on secretion and the physiological effect and relevance of this secretion are often unclear. One major difficulty in the correct identification of proteins secreted by *H. pylori* is its high frequency of lysis, which results in nonspecific release of the cytoplasmic contents of the bacterium (Cao and Cover, 1998; Vanet and Labigne, 1998). One protein that has been found in the external medium by many independent studies (Kim and Sachs, 2002; Toledo and Jerez, 2002) is HP1286, a polypeptide chain of 182 amino acids. The primary sequence of HP1286 suggests that it belongs to the YceI-like family of proteins (Karow and

Georgopoulos, 1991), a group of putative periplasmic proteins first described in terms of amino acid sequence, and encoded by genes located upstream of the *htrB* gene (Karow and Georgopoulos, 1991). The YceI-like family is structurally a subgroup of the lipocalin superfamily (Newcomer and Peterson, 1984). The prototype of lipocalins is retinol-binding protein (RBP), a protein of 182 amino acids present in the plasma of higher animals, and responsible for the binding and transport of retinol from the liver to the cell receptors of the tissues that need it. RBP is a monomeric protein composed of one β -barrel single domain, characterized by an internal cavity where the hydrophobic ligand is hosted (Newcomer and Jones, 1990; Zanotti and Monaco, 1993). The crystal structure of YceI has been determined for the proteins from *Thermus thermophilus* (Handa and Yokoyama, 2005) and *Escherichia coli* (Protein Data Bank ID: 1Y0G). In both cases, the protein is a homodimer, each monomer being characterized by a lipocalin fold. The *T. thermophilus* protein binds polyisoprenyl pyrophosphate, suggesting that it plays a role in isoprenoid quinone metabolism and/or transport or storage (Handa and Yokoyama, 2005). As the *T. thermophilus* protein was expressed in a heterologous system and the ligand was not added, the authors concluded that it was taken up from *E. coli*, the bacterium in which it was expressed. In the crystal structure of *E. coli* YceI protein, the compound 2-[(2E,6E,10E,14E,18E,22E,26E)-3,7,11,15,19,23,27,31-Octamethyl dotriaconta-2,6,10,14,18,22,26,30-octaenyl] phenol was found buried in the inner cavity. This is an amphipathic compound with a structure similar to that of polyisoprenyl pyrophosphate and the same number of carbon atoms.

Fatty acid amides are bioactive lipids and appear to serve a variety of functions within and outside the central nervous system in higher animals (Bialer, 1991; Jain and Wong, 1992). Erucic acid, the fatty acid precursor of erucamide, is quite common in nature. It is, for example, one of the most abundant components of different varieties of rapeseed (Hamberger and Stenhagen, 2003). Erucic acid is suitable for human consumption at low doses, but it can cause a variety of heart lesions at high doses (Charlton and Sauer, 1975). Erucamide, which was detected in pig's blood plasma, lung, kidney, liver, and brain, has been found to be

involved in the stimulation of angiogenesis, to inhibit intestinal diarrhea, and to regulate fluid volumes in other organs (Hamberger and Stenhagen, 2003). At the same time, erucamide is a contaminant of plastic materials, and is used, in particular, as a slip agent in polyethylene films (Garrido-Lopez and Tena, 2007). As neither erucamide nor any other long-chain fatty acid or amide was added during the purification and crystallization steps, the most likely hypothesis is that the ligand was taken up from *E. coli* and bound tightly enough to be conserved during all the purification steps.

At variance with the proteins from *T. thermophilus*, HP1286 presents a secretion sequence signal at the N-terminus, confirming its secretory nature. The 3D structure of mature HP1286 demonstrates that it structurally belongs to the YceI family, but that it shows an inner cavity structural adaptation for a new binding specificity.

5.2 Materials and methods

Cloning, expression, and purification

The HP1286 gene was amplified by PCR from *H. pylori* CCUG17874 genomic DNA using the following primers: forward, 5'-CACCAAACCTTATACGATTGATAAGGCAAAC-3'; and reverse, 5'-TTATTATTGGGCGTAAGCTTCTAG-3'. The construct was cloned directly into the pET151 expression vector by a Directional TOPO cloning technique (Invitrogen Ltd, Paisley, UK), which allows the introduction of a sequence coding for six Histidines upstream the HP1286 gene, spaced by a tobacco etch virus (TEV) protease for the removal of the tag in the last steps of the purification. The positive pET151-HP1286 clones were verified by sequencing. His6-HP1286 protein was overexpressed in *E. coli* (BL21 DE3 strain) using 1.0 mM isopropyl thio- β -D-galactoside, and the expression was prolonged for 3 h at 30 °C. Bacterial cells were harvested by centrifugation at 6000 g and stored at -80 °C. The cell pellet was resuspended in buffer A (30 mM Aces, pH 7.0, 200 mM NaCl), and lysis was achieved with lysozyme (1 mg/mL) incubation, followed by multiple sonication cycles (four

times, 45 min each). The resulting supernatant was isolated from the insoluble fraction by centrifugation at 40 000 g for 25 min at 4 °C, and loaded onto an Ni²⁺-immobilized metal-affinity preppacked column (GE Healthcare Europe GMBH, Orsay Cedex, France). The fractions containing His6-HP1286 were eluted with an imidazole gradient, pooled, and incubated overnight at 4°C with His6-rTEV protease. The sample was further subjected to an immobilized metal ion affinity chromatography step to remove the His6-rTEV protease and the residual uncleaved His6-HP1286. The final purification step, size exclusion chromatography (Superdex 200 HR10/ 300; GE Healthcare) with equilibration with buffer A, resulted in a single peak and a retention time roughly corresponding to a protein dimer.

Crystallization and structure determination

The purified HP1286 was concentrated to 16 mg/mL and used for crystallization trials, which were partially automated using an Oryx 8 crystallization robot (Douglas Instruments Ltd, Hungerford, UK). Several promising conditions were selected from Structure Screen I (SSI) and Structure Screen II (Molecular Dimensions Ltd, Newmarket, UK) and PACT screen (Qiagen, Hilden, Germany), but many of them gave poorly

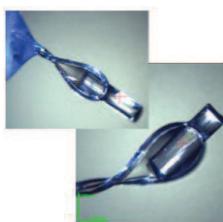


Figure 5.1 Representative crystals forms of HP1286 used in this study.

diffraction and / or disordered crystals, except for SSI no. 37 [0.2 m CH₃COONa, 0.1 m Tris/HCl, pH 8.5, 30% poly(ethylene glycol) 4000] and SSI no. 31 [0.1 m Hepes, pH 7.5, 10% isopropanol, 20% poly(ethylene glycol) 4000], which gave the best-quality diffracting crystals. In particular, these two crystallization conditions produced crystals belonging to two different space groups. Crystals of form A, grown from SSI no. 37 solution, are orthorhombic, space group P2₁2₁2₁, with a = 56.43 Å, b = 61.44 Å, and c = 94.46 Å. These data correspond to one dimer per asymmetric unit, with V_M = 2.11 Å³/Da and a solvent content of 42%. They diffract to a maximum resolution

of 2.5 Å. Form B crystals, grown from SSI no. 31, are monoclinic, space group P2₁, with a = 30.94 Å, b = 61.31 Å, c = 88.32 Å, and $\beta = 92.88^\circ$. They contain one dimer per asymmetric unit, corresponding to a V_M of 2.16 Å³ per Da and a solvent content of about 43%. Both structures were determined, but details are reported here only for form B, which provided the best diffraction pattern, at 2.1 Å resolution. The dataset used in the final refinement was measured at the microfocus beamline ID23-2 of European Synchrotron Radiation Facility, Grenoble, France. Three hundred frames of 1° oscillation each were collected with a wavelength of 0.8760 Å. Datasets were indexed and integrated with mosflm (Leslie, 2006), and merged and scaled with scala (Evans, 2006), contained in the ccp4 crystallographic package (Collaborative Computational Project Number 4, 1994). Structures were solved by molecular replacement, using phaser (McCoy and Read, 2007), starting from the model of the polyisoprenoid-binding protein from *T. thermophilus* (Protein Data Bank ID: 1WUB) (Handa and Yokoyama, 2005). Refinement was performed using the simulated annealing procedure contained in cns (Brunger et al., 1998) in the first stages of refinement and refmac5 (Murshudov and Dodson, 1997) in the subsequent steps. TLS refinement was applied in the last cycles (Painter and Merritt, 2006). Solvent molecules were added with the automated procedure of refmac5, and manually revised during the refinement. Visualization of the model and manual rebuilding were performed with coot (Emsley and Cowtan, 2004). From the first stages of the refinement, a long electron density was clearly visible in the (2|Fobs| - |Fcalc|) Fourier map inside the protein barrel of each monomer. According to the indications of the mass spectra, a molecule of erucamide was fitted inside the cavity of each monomer. Geometric parameters for the refinement of the ligand were obtained using the server <http://skuld.bmsc.washington.edu/~tmsmd> (Painter and Merritt, 2006). The final model contains 2632 protein atoms, 48 ligand atoms, and 79 solvent molecules. The final crystallographic R-factor is 0.215 (Rfree = 0.287), and the geometry of the model, checked with procheck (Laskowski and Thornton, 1993) and rampage (Lovell and Richardson, 2003), is as expected at this resolution.

The calculation of the volume of the cavity hosting the ligand was performed using voidoo (Kleywegt and Jones, 1994). The cavity was searched using a probe radius of 1.4 Å and a primary grid space of 0.75.

Mass Spectrometry

Six hundred micrograms of recombinant purified HP1286 was treated with 6 M guanidinium chloride and loaded onto a reverse-phase Jupiter C5 column (4.60 x 250 mm; Phenomenex). Elution was performed with an H₂O/acetonitrile gradient, supplemented with 0.1% trifluoroacetic acid.

The profile was monitored at 216 nm, and all of the representative peaks were collected and dried out to remove any solvent traces. The most abundant fractions were analyzed by GC-MS. GC-MS was performed with a Thermo Fisher Trace DSQ (Waltham, MA, USA). The GC operating conditions were as follows: injection port temperature of 280 °C; carrier gas He, 1.2 mL·min⁻¹; injection volume of 10 µL; column, TR-SMS Thermo Fisher (Waltham, MA, USA), 30 m x 0.25 mm internal diameter, film thickness of 0.25 µm; split mode 30 : 1; temperature program -4 min at 40 °C, raised to 150 °C at 15 °C·min⁻¹, held for 1 min, then raised to 300 °C at 10 °C·min⁻¹ and held for 2 min; and GC-MS interface temperature of 250 °C. The MS operating conditions were as follows: ion source, EI+ (70 eV); and source temperature of 250 °C. Chromatograms were recorded with total ion current monitoring. Erucamide was identified by comparing its retention time and mass spectra with those of the standard (Sigma-Aldrich).

5.3 Results and discussion

HP1286 is a protein of 182 amino acids, but as the first 17 residues are predicted to be a signal for secretion into the periplasmic space (SignalP; ExPASy website), only residues from 18 to 182 were cloned. The protein was expressed in soluble form and purified. The protein in solution is a homodimer, as demonstrated by exclusion chromatography data (not shown). Crystals were grown in two different crystal forms, both containing one protein dimer per asymmetric unit. The molecular models

of both forms are virtually identical, and the monoclinic one is described here in detail, as it diffracts to a higher resolution, 2.1 Å. Statistics on structure determination and refinement are reported in Table 1.

X-ray data		
Space group	P2 ₁	P2 ₁ ,2 ₁
Cell parameters (Å, °)	a = 30.94, b = 61.31, c = 88.32, β = 92.9	a = 56.43, b = 61.44, c = 94.46
Resolution (Å)	50.3–2.10 (2.21–2.10)	94.5–2.5 (2.64–2.5)
Independent reflections	19 383 (2823)	11 468 (1659)
Multiplicity	6.1 (6.0)	3.7 (3.8)
Completeness (%)	99.9 (99.9)	97.1 (98.4)
<I/σ(I)>	10.5 (4.9)	12.7 (2.9)
R _{merge}	0.124 (0.424)	0.090 (0.468)
σ-factor from Wilson plot	24.6	54.6
Refinement		
Total number of atoms, including solvent	2759	2706
Mean B-value (Å ²) for protein atoms, ligand, and waters	7.2–21.1–9.7	33.3–49.0–49.4
R _{int}	0.217 (23.0)	0.216 (0.240)
R _{free} (R% of reflections)	0.274 (0.278)	0.327 (0.410)
Ramachandran plot (%)		
Favored region	94.1	89.5
Allowed region	5.9	9.6
Outlier region	0	0.9
Rmsd on bond length (Å) and angles (°)	0.018, 1.9	0.022, 2.3

Table 5.1 Statistics on data collection and refinement. A wavelength of 0.8726 Å was used. Rotations of 1° were performed. The Ramachandran plot was calculated using RAMPAGE.

The protein present in the asymmetric unit of both crystal forms is a dimer, formed from two identical monomers. The core of each monomer is a β-barrel formed from eight antiparallel β-strands, each strand interacting with the nearby ones through hydrogen bonds.

The topology of the barrel is illustrated in Fig.5.2, where β-strands are labeled from A to H.

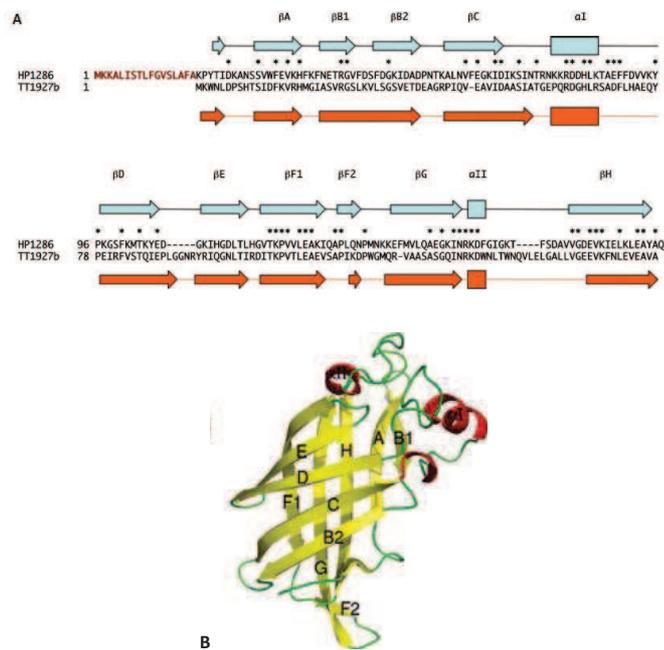


Figure. 5.2. Primary and secondary structure. (A) Amino acid sequence of HP1286 structurally aligned with that of *T. thermophilus* (Protein Data Bank ID: 1WUB). Amino acids in red represent the predicted signal of secretion to the periplasmic space, and were excluded from the expression vector. Arrows and rectangles indicate the positions of secondary structure elements, β -strands, and α -helices, respectively, for our structure (light blue) and 1WUB (orange). The assignment of secondary structures, obtained with PROCHECK, is as follows: bA, 28–35; bB1, 39–44; bB2, 48–55; bC, 60–69; bD, 97–106; bE, 109–116; bF1, 119–130; bF2, 132–135; bG, 141–152; bH, 167–180; aI, 78–85; aII, 154–156. (B) Cartoon representation of the monomer of HP1286. β -Strands, α -helices and turns are in yellow, red and green, respectively. Strands are labeled from A to H. Strands B and F, owing to some irregularities, are divided into two parts and labeled B1, B2, F1, and F2.

A α -helix (helix I), which connects strand C to strand D, and a turn of helix (helix II) at the end of strand G, complete the structure. The electron density is clearly defined for all residues from 18 to 181, with the exception of residues 57–59 of one monomer, which are part of a β -turn connecting two strands. Some of the strands present some kinks that break the continuity of the hydrogen bond patterns, and so they are formally considered to be composed of two parts. This happens for

strands β B and β F, and, in fact, they have been labeled β B1 and β B2, and β F1 and β F2, respectively.

Two hundred and two hydrogen bonds among protein atoms stabilize the 3D structure. The β -barrel forms an inner cavity that is completely closed at one end, whereas at the opposite side an opening is present next to α -helix I. Through this aperture, the internal surface of the inner cavity is in contact with the solvent.

The two monomers are spatially related by a non-crystallographic two-fold axis. The total accessible surface for the sum of the two separated monomers corresponds to 15 613 \AA^2 ; of this, 4729 \AA^2 (30% of the total surface, calculated with program Areaimol (Collaborative Computational Project Number 4, 1994), using a probe radius of 1.4 \AA) become excluded following dimer formation. The interactions between the two monomers are mainly hydrophilic, including the formation of 18 hydrogen bonds, but a few hydrophobic interactions are also present (see Table 2 for a detailed list of the interactions).

Chain A	Chain B	Hydrogen bonds
Ala25	Asn77, Arg80	AlaO-ArgNH1 AlaO-ArgND2
Asn26	His35, Arg80, Asn39	AsnOD1-ArgNH1 AsnOD1-HisNE2
Ser28	Arg76	
Trp30	Arg42, Trp30, Arg76	
His35	Glu178, Asn26	HisNE2-AsnOD1
Phe36	Phe142, Pro136, Asn135	
Phe38	Gln130, Leu133, Val144, Gln146	
Asn39	Val144, Gln146, Glu178	
Glu40	Gln146, Lys176	GluOE1-GlnOE1 GluOE2-LysNZ GluOE2-GlnNE2 GlnOE2-GlnOE1
Arg42	Glu174, Lys176	
Val44	Arg76	
Asp46	Arg76	

The structure of HP1286 is quite similar to that of polyisoprenoid-binding protein TT1927b from *T. thermophilus* (Protein Data Bank ID: 1WUB; Handa and Yokoyama, 2005): the rmsd between the two structures is 1.54 \AA for the superposition of 155 amino acids of the monomer, and 1.51 \AA for the superposition of 303 amino acids of the dimer (Fig.5.3A).

Significant differences are present in some loop regions; in particular, the long loop connecting strands G and H is longer in the *T. thermophilus* protein. A comparison of our model with YceI from *E. coli* (Protein Data Bank ID code: 1Y0G) shows that they are slightly more similar and the loop between strands G and H presents roughly the same length. Superposition with a representative member of the lipocalin family (Newcomer and Peterson, 1984), RBP (Fig.5.3B), shows that the overall motif of the core of the molecule is well preserved, but the barrel of YceI is longer, and consequently its cavity becomes much deeper. Moreover, RBP has a long C-terminal tail, about 40 amino acids, which is totally absent in the YceI family of proteins.

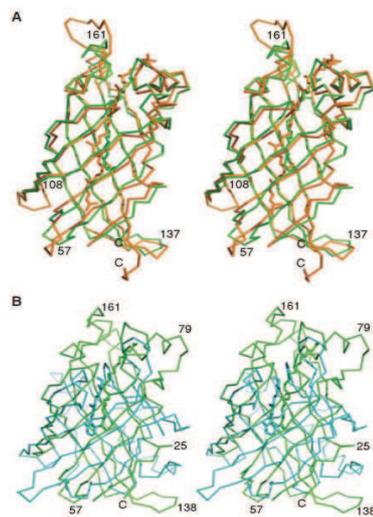


Figure 5.3. Structure superposition. (A) Superposition of the Ca chain trace of HP1286 monomer (green) superimposed on that of TT1927b from *T. thermophilus* (orange) (Protein Data Bank ID: 1WUB). Some residues of the regions that present significant differences between the two structures are labeled. The two ligands are drawn using the same colors as the corresponding proteins. (B) HP1286 chain trace (green) superimposed on a representative structure of the lipocalin family, pig RBP (cyan) (Protein Data Bank code: 1aqb [42]). The retinol bound to RBP is also shown in cyan.

Mass spectra indicated the presence, along with other contaminants, of erucamide, whose shape and length correspond to those of the electron density clearly visible inside the barrel cavity of each monomer (Fig.5.4A; see Fig.5.4B for a scheme of the labeling system of the compound).

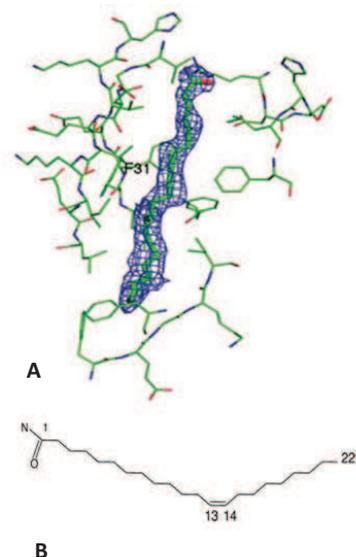


Figure 5.4. The ligand. (A) Stereo view of a detail of the HP1286 binding cavity with erucamide bound inside it. The Fourier electron density map, calculated with $(2F_{obs}-F_{calc})$ coefficients, is contoured around the ligand at 1.5σ . Portions of the protein polypeptide chain with residues in contact with the ligand (see Table 3) are shown. (B) Scheme of erucamide with the labeling system used in the text.

Other contaminants consisted of nonlinear compounds, which are incompatible with the shape of the density and the size of the protein cavity. The erucamide tail is deeply buried inside the protein cavity, which is fully hydrophobic, whereas the amidic head of the ligand is close to the open end of the cavity, which is accessible to the solvent.

The amidic group of the ligand interacts with side chain atoms of Arg80, but residues surrounding the mouth of the cavity are mostly hydrophilic or possibly positively charged: His35, His83, Lys79, Asn26, and Asn77 (see Table 3 for a list of contacts between the ligand and the protein). Another arginine, Arg153, is close to the opening of the cavity, but totally

buried inside it. Its side chain forms five hydrogen bonds with main chain carbonyl oxygen atoms, and it is possibly neutralized by Asp169, which is on the external protein surface and points towards the solvent, along with Lys154.

In each monomer, the entrance of the cavity is in contact with the external solvent, but it is partially obstructed by a loop of the other monomer. The loop connecting strands $\beta F2$ and βG protrudes from the domain core and points towards it (Fig.5.5).

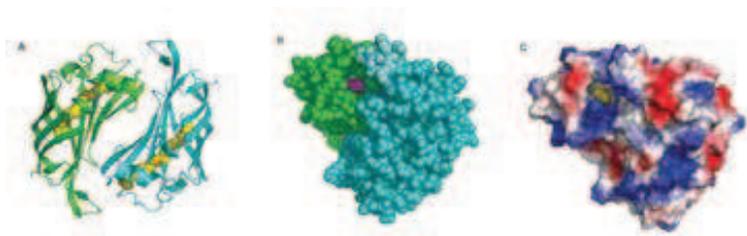


Figure 5.5. The dimer of HP1286 and the binding site. (A) Stereo view of a cartoon representation of the dimer of the protein. The two chains are in different colors, and the bound erucamide is shown as yellow spheres. (B) Space-filling representation of the HP1286 dimer. The view allows the hydrophilic terminus of erucamide (magenta) bound to subunit A (green) to be distinguished. It is possible to see how the long loop that connects strands F and G of subunit B (cyan and pale blue) partially covers the entrance of the protein central cavity. (C) Electrostatic potential surface of the protein calculated using PYMOL [41]. The view is approximately the same as in (B). The ligand has been excluded from the calculation, and is shown as yellow van der Waals spheres.

The cavity of the *H. pylori* protein is shorter with respect to that of the two homologous proteins whose structure has been determined: its volume is 151 \AA^3 , whereas that of the *T. thermophilus* protein is 233 \AA^3 . This is mainly due to the presence inside the cavity of some bulky side chain residues, namely Phe64, Leu145, Leu177, Ile22, and Ile52, which close up the cavity towards the bottom.

We cannot state that the natural ligand of the *H. pylori* protein is erucamide, but the shape and size of the cavity clearly indicate that inside the protein there is space for a roughly linear chain of about 22 carbon atoms.

The same *E. coli* could eventually have internalized some erucamide from the LB broth used to grow all of the cultures. Nevertheless, we cannot rule out the possibility that erucamide was present as a contaminant in plastic material and was taken up by the protein during some purification step. The latter event appears to be quite unlikely, as we have to assume a very high binding constant of the protein for an extraneous ligand.

The presence of a consistent number of potentially positively charged residues around the opening of the cavity supports the idea that the natural ligand(s) could be a negatively charged fatty acid, or an amide, like that tightly bound in the present structure. In contrast, both the *T.*

thermophilus and the *E. coli* proteins bind a (C40) fatty acid. Moreover, the polyisoprenyl pyrophosphate bound to the *T. thermophilus* protein is a precursor in the biosynthetic pathway of isoprenoid quinones. This indicates that, despite the fact that the three proteins belong to the YceI-like family from the point of view of the amino acid sequence and of the 3D structure, they must differ in their physiological function. This is confirmed by the presence of a secretion signal at the N-terminus of HP1286 and *E. coli* YceI protein, and the absence of anything similar in the *T. thermophilus* one.

In a study on the adaptation of *H. pylori* to acidic conditions, it was found that a UreI-negative strain, a mutant strain unable to transport urea inside the cell, induced overexpression of a relatively limited number of proteins, one of which is HP1286 (Toledo and Jerez, 2002). The method used to identify the protein was sequencing of the N-terminus, and, interestingly, the amino acid sequence found corresponds to peptide 18–29, indicating that the secretion signal had already been processed and that the protein corresponded to the mature one. Also, the other two proteins identified as being overexpressed were HP0243 and HP0485. The first, also known as *H. pylori* neutrophil-activating protein, is an iron uptake protein belonging to the class of miniferritins (Tonello and Montecucco, 1999; Zanotti and Montecucco, 2002), whereas the second is a catalase-like enzyme, and is possibly implicated in the general stress response in bacteria (Tomb et al., 1997). Moreover, it has been already observed that acid adaptations, like those described before, confer resistance to a wide range of stress conditions such as heat, salt, and H₂O₂. The 3D structure of HP1286 clearly points to a storage and transport function of some long-chain fatty acid(s) or amide(s). The evidence that the protein is secreted, coupled with the fact that the stomach mucosa, where *H. pylori* establishes persistent colonization and causes chronic inflammation, is rich in lipids, strongly supports the hypothesis that the protein sequesters fatty acids or amides present in the environment of the bacterium. This sequestering could be used to protect the external membrane from their surfactant properties and/or to supply the bacterium with the fatty acids necessary for its metabolism.

Finally, it has been shown that changes in the lipid composition of some bacteria are associated with the maintenance of a functional physiological state of the cell membrane (Guerzoni and Cocconcelli, 2001). If this holds also for *H. pylori*, HP1286 overexpression in conditions where the bacterium experiences acidic stress could be utilized to supplement the membrane with particular fatty acid chains.

Conclusions

Since the relationship between *H. pylori* and chronic gastritis was established, investigators began to take interest in the role played by *H. pylori* in this type of cancer. The project presented here aims to determine the three dimensional structure of proteins playing a role in the virulence of the bacterium and also their analysis and characterization using an approach of structural genomics. The three-dimensional structure of a protein, in addition to clarify its function, may be useful for the design of potential new drugs.

This thesis presents several potential drug targets, each consisting of proteins involved in pathogenicity or essential to the survival of the bacterium.

CagL was found to be present in different oligomerization states. This hampered the crystallization of the protein. A molecular model of the protein was nonetheless obtained using the x-ray small angle scattering technique.

The structure of Hp1286 was determined at a resolution of 2.1Å. The electron density map showed a ligand present in the hydrophobic cavity. The ligand was identified by mass spectrometry as erucamide. The three-dimensional structure of HP1286 strongly suggests for the protein a role in the storage and transport of long-chain fatty acid(s) or amide(s). The evidence that the protein is secreted, together with the fact that the stomach mucosa, where *H. pylori* establishes persistent colonization and causes chronic inflammation, is rich in lipids, support the hypothesis that the protein sequesters fatty acids or amides present in the environment of the bacterium.

Crystals of Hp0797 were grown, but they diffracted only at 4Å resolution and this hampered the possibility of solving the structure. The optimization of crystals is in progress. In parallel, the comparison between Hp0797 and another adhesin of the bacterium, Hp0410, demonstrates that proteins with the same function use different strategy to recognize the host cell.

The study of the bacterial virulence factors employed by *H. pylori* and how these interact with hosts of different genetic backgrounds is

redefining our understanding of bacterial ecology and homeostasis. Twenty-five years after its discovery, this bacterium should be regarded as perhaps the most informative of all microbial infections in biology.

References

1. Parkin DM, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. *CA Cancer J Clin* 2005; 55: 74–108.
2. V. Herrera and J. Parsonnet *Helicobacter pylori* and gastric adenocarcinoma, 2009. *Clin. Microbiol Infect.* 15, 971–976 (2009).
3. Marshall BJ and Warren JR. Unidentified curved bacilli in the stomach of patients with gastritis and peptic ulceration, 1984. *Lancet* 1984; 1:1311–5
4. The Eurogast Study Group. Epidemiology of, and risk factors for, *Helicobacter pylori* infection among 3194 asymptomatic subjects in 17 populations. The EUROGAST Study Group. *Gut* 1993; 34: 1672–1676.
5. IARC Working Group. IARC working group on the evaluation of carcinogenic risks to humans: some industrial chemicals Lyon, 15–22 February 1994. *IARC Monogr Eval Carcinog Risks Hum* 1994; 60: 1–560.
6. Ohkusa T, Fujiki K, Takashimizu I et al. Improvement in atrophic gastritis and intestinal metaplasia in patients in whom *Helicobacter pylori* was eradicated. *Ann Intern Med* 2001; 134: 380–386.
7. Westblom TU, Madan E, Midkiff BR. Egg yolk emulsion agar, a new medium for the cultivation of *Helicobacter pylori*. *J Clin Microbiol* 1991;29:819-21.
8. Uemura N, Okamoto S, Yamamoto S, Matsumura N, Yamaguchi S, Yamakido M, Taniyama K, Sasaki N, Schlemper RJ. *Helicobacter pylori* infection and the development of gastric cancer. *N Engl J Med.* 2001 Sep 13;345(11):784-9.
9. Amieva MR, El-Omar EM. Host-bacterial interactions in *Helicobacter pylori* infection. *Gastroenterology.* 2008 Jan;134(1):306-23.
10. Forman D, Burley VJ. Gastric cancer: global pattern of the disease and an overview of environmental risk factors. *Best Pract Res Clin Gastroenterol.* 2006;20(4):633-49.
11. Linz B, Balloux F, Moodley Y, Manica A, Liu H, Roumagnac P, Falush D, Stamer C, Prugnolle F, van der Merwe SW, Yamaoka Y, Graham DY, Perez-Trallero E, Wadstrom T, Suerbaum S, Achtman M. An African origin for the intimate association between humans and *Helicobacter pylori*. *Nature.* 2007 Feb 22;445(7130):915-8. Epub 2007 Feb 7.
12. Ghose, C. et al. East Asian genotypes of *Helicobacter pylori* strains in Amerindians provide evidence for its ancient human carriage. *Proc. Natl Acad. Sci. USA* 99, 15107–15111 (2002).
13. Falush, D. et al. Traces of human migrations in *Helicobacter pylori* populations. *Science* 299, 1582–1585 (2003).
14. Linz, B. et al. An African origin for the intimate association between humans and *Helicobacter pylori*. *Nature* 445, 915–918 (2007). Two analyses of the global population structure of *H. pylori* using multilocus sequence data. The first paper established that *H. pylori* can be subdivided into populations with distinct geographical distribution, and that its genes reflect human migrations. The second study compared the genetic diversity of human and *H. pylori* DNA, and provides evidence that *H. pylori* was already associated with humans at the time they migrated out of Africa 60,000 years ago.
15. Blaser, M. J. Who are we? Indigenous microbes and the ecology of human diseases. *EMBO Rep.* 7, 956–960 (2006).
16. Suerbaum S, Josenhans C. *Helicobacter pylori* evolution and phenotypic diversification in a changing host. *Nat Rev Microbiol.* 2007 Jun;5(6):441-52. Review.
17. Lodato F, Azzaroli F, Tamè MR, Di Girolamo M, Buonfiglioli F, Mazzella N, Cecinato P, Roda E, Mazzella G. G-CSF in Peg-IFN induced neutropenia in liver transplanted patients with HCV recurrence. *World J Gastroenterol.* 2009 Nov 21;15(43):5449-54.
18. Lu H, Yamaoka Y, Graham DY. *Helicobacter pylori* virulence factors: facts and fantasies. *Curr Opin Gastroenterol.* 2005;21:653–659.
19. Hu LT, Nicholson EB, Jones BD, Lynch MJ, Mobley HL. *Morganella morganii* urease: purification, characterization, and isolation of gene sequences. *J Bacteriol.* 1990 Jun;172(6):3073-80.
20. Hu LT, Mobley HL. Purification and N-terminal analysis of urease from *Helicobacter pylori*. *Infect Immun.* 1990 Apr;58(4):992-8.
21. Labigne A, Cussac V, Courcoux P. Shuttle cloning and nucleotide sequences of *Helicobacter pylori* genes responsible for urease activity. *J Bacteriol.* 1991 Mar;173(6):1920-31.
22. Weeks DL, Eskandari S, Scott DR, Sachs G. A H⁺-gated urea channel: the link between *Helicobacter pylori* urease and gastric colonization. *Science.* 2000 Jan 21;287(5452):482-5.

23. Sachs G, Scott D, Weeks D, Melchers K. Gastric habitation by *Helicobacter pylori*: insights into acid adaptation. *Trends Pharmacol Sci*. 2000 Nov;21(11):413-6. Review. No abstract available.
24. Phadnis SH, Parlow MH, Levy M, Ilver D, Caulkins CM, Connors JB, Dunn BE. Surface localization of *Helicobacter pylori* urease and a heat shock protein homolog requires bacterial autolysis. *Infect Immun*. 1996 Mar;64(3):905-12.
25. Ha NC, Oh ST, Sung JY, Cha KA, Lee MH, Oh BH. Supramolecular assembly and acid resistance of *Helicobacter pylori* urease. *Nat Struct Biol*. 2001 Jun;8(6):505-9.
26. S. Schreiber, T.H. Nguyen and M. Konradt *et al.*, Recovery from gastric mucus depletion in the intact guinea pig mucosa, *Scand J Gastroenterol* 38 (2003), pp. 1136–1143.
27. Kostrzynska M, Betts JD, Austin JW, Trust TJ. Identification, characterization, and spatial localization of two flagellin species in *Helicobacter pylori* flagella. *J Bacteriol*. 1991 Feb;173(3):937-46.
28. Josenhans C, Labigne A, Suerbaum S. Comparative ultrastructural and functional studies of *Helicobacter pylori* and *Helicobacter mustelae* flagellin mutants: both flagellin subunits, FlaA and FlaB, are necessary for full motility in *Helicobacter* species. *J Bacteriol*. 1995 Jun;177(11):3010-20.
29. McGee SE. More serious letters about obesity. *MedGenMed*. 2005;7(1):30. No abstract available.
30. Ottemann KM, Lowenthal AC. *Helicobacter pylori* uses motility for initial colonization and to attain robust infection. *Infect Immun*. 2002 Apr;70(4):1984-90.
31. S. Schreiber, M. Konradt and C. Groll *et al.*, The spatial orientation of *Helicobacter pylori* in the gastric mucus, *Proc Natl Acad Sci U S A* 101 (2004), pp. 5024–5029.
32. S. Schreiber, R. Buckner and C. Groll *et al.*, Rapid loss of motility of *Helicobacter pylori* in the gastric lumen in vivo, *Infect Immun* 73 (2005), pp. 1584–1589.
33. M.A. Croxen, G. Sisson and R. Melano *et al.*, The *Helicobacter pylori* chemotaxis receptor TlpB (HP0103) is required for pH taxis and for colonization of the gastric mucosa, *J Bacteriol* 188 (2006), pp. 2656–2665.
34. Spohn G, Scarlato V. Motility, chemotaxis and flagella. In: Mobley HLT, Mendz GL, Hazell SL, eds. *HELICOBACTER PYLORI, Physiology and Genetics*. Washington, D.C.: ASM Press, 2001;239–48.
35. Slomiany BL, Slomiany A. Role of mucus in gastric mucosal protection. *J Physiol Pharmacol*. 1991 Jun;42(2):147-61.
36. Andersen LP. Colonization and infection by *Helicobacter pylori* in humans. *Helicobacter*. 2007 Nov;12 Suppl 2:12-5. Review.
37. Yoshiyama H, Nakamura H, Kimoto M, et al. Chemotaxis and motility of *Helicobacter pylori* in a various environment. *Gastroenterology* 1999;34 (Suppl 11):18–23.
38. Wadström T, Hirno S, Borén T. Biochemical aspects of *Helicobacter pylori* colonization of the human gastric mucosa. *Aliment Pharmacol Ther* 1996;10:17–27.
39. Testerman TL, McGee DJ, Mobley HLT. Adherence and colonization. In: Mobley HLT, Mendz GL, Hazell SL, eds. *HELICOBACTER PYLORI, Physiology and Genetics*. Washington, DC: ASM Press, 2001;381–417.
40. Van de Bovenkamp JH, Mahdavi J, Korteland-Van Male AM, et al. The MUC5AC glycoprotein is the primary receptor for *Helicobacter pylori* in the human stomach. *Helicobacter* 2003;8:521–32.
41. Linden S, Mahdavi J, Hedenbro J, et al. Effect of pH on *Helicobacter pylori* binding to human gastric mucins. Identification to binding to non-MUC5AC mucins. *Biochem J* 2004;384:263–70.
42. Prakobphol A, Tangemann K, Rosen SD, et al. Separate oligosaccharide determinants mediate interactions of the low-molecular-weight salivary mucin with neutrophils and bacteria. *Biochemistry* 1999;38:6817–25.
43. Tomb JF, White O, Kerlavage AR, Clayton RA, Sutton GG, Fleischmann RD, Ketchum KA, Klenk HP, Gill S, Dougherty BA, Nelson K, Quackenbush J, Zhou L, Kirkness EF, Peterson S, Loftus B, Richardson D, Dodson R, Khalak HG, Glodek A, McKenney K, Fitzgerald LM, Lee N, Adams MD, Hickey EK, Berg DE, Gocayne JD, Utterback TR, Peterson JD, Kelley JM, Cotton MD, Weidman JM, Fujii C, Bowman C, Watthey L, Wallin E, Hayes WS, Borodovsky M, Karp PD, Smith HO, Fraser CM, Venter JC. The complete genome sequence of the gastric pathogen *Helicobacter pylori*. *Nature*. 1997 Aug 7;388(6642):539-47. Erratum in: *Nature* 1997 Sep 25;389(6649):412
44. Doig P, de Jonge BL, Alm RA, Brown ED, Uria-Nickelsen M, Noonan B, Mills SD, Tummino P, Carmel G, Guild BC, Moir DT, Vovis GF, Trust TJ. *Helicobacter pylori* physiology predicted from genomic comparison of two strains. *Microbiol Mol Biol Rev*. 1999 Sep;63(3):675-707. Review.

45. Odenbreit S, Faller G, Haas R. Role of the alpAB proteins and lipopolysaccharide in adhesion of *Helicobacter pylori* to human gastric tissue. *Int J Med Microbiol*. 2002 Sep;292(3-4):247-56.
46. Stern A, Brown M, Nickel P, Meyer TF. Opacity genes in *Neisseria gonorrhoeae*: control of phase and antigenic variation. *Cell*. 1986 Oct 10;47(1):61-71.
47. Exner MM, Doig P, Trust TJ, Hancock RE. Isolation and characterization of a family of porin proteins from *Helicobacter pylori*. *Infect Immun*. 1995 Apr;63(4):1567-72.
48. Peck B., Ortkamp M., Diehl K. D., Hundt E., Knapp B. Conservation, localization and expression of HopZ, a protein involved in adhesion of *Helicobacter pylori*. *Nucleic Acids Res*. 1999; 27: 3325–3333.
49. C. Petersson, M. Forsberg and M. Aspholm *et al.*, *Helicobacter pylori* SabA adhesin evokes a strong inflammatory response in human neutrophils which is down-regulated by the neutrophil-activating protein, *Med Microbiol Immunol* 195 (2006), pp. 195–206.
50. J.L. Guruge, P.G. Falk and R.G. Lorenz *et al.*, Epithelial attachment alters the outcome of *Helicobacter pylori* infection, *Proc Natl Acad Sci U S A* 95 (1998), pp. 3925–3930.
51. Yoshida, N., D.N. Granger, D.J. Evans, D.G. Evans, D.Y. Graham, D.C. Anderson, R.E. Wolf, P.R. Kvietys (1993) Mechanism involved in *Helicobacter pylori*-produced inflammation. *Gastroenterology*. 105:1431–1440
52. Cover, T.L., M.K.R. Tumurru, P. Cao, S.A. Thompson, J. Blaser (1994) Divergence of genetic sequences for the vacuolating cytotoxin among *Helicobacter pylori* strains. *J. Biol. Chem* 269:10566–10573
53. Zanotti G, Papinutto E, Dundon W, Battistutta R, Seveso M, Giudice G, Rappuoli R, Montecucco C. Structure of the neutrophil-activating protein from *Helicobacter pylori*. *J Mol Biol*. 2002 Oct 11;323(1):125-30.
54. M.J. Blaser, G.I. Perez-Perez and H. Kleanthous *et al.*, Infection with *Helicobacter pylori* strains possessing cagA is associated with an increased risk of developing adenocarcinoma of the stomach, *Cancer Res* 55 (1995), pp. 2111–2115
55. Satin B, Del Giudice G, Della Bianca V, Dusi S, Laudanna C, Tonello F, Kelleher D, Rappuoli R, Montecucco C, Rossi F. The neutrophil-activating protein (HP-NAP) of *Helicobacter pylori* is a protective antigen and a major virulence factor. *J Exp Med*. 2000 May 1;191(9):1467-76.
56. Smoot, D. T. *et al.* Adherence of *Helicobacter pylori* to cultured human gastric epithelial cells. *Infect. Immun*. 61, 350–355 (1993).
57. Segal, E. D., Falkow, S. & Tompkins, L. S. *Helicobacter pylori* attachment to gastric cells induces cytoskeletal rearrangements and tyrosine phosphorylation of host cell proteins. *Proc. Natl Acad. Sci. USA* 93, 1259–1264
58. Montecucco C, Rappuoli R Living dangerously: how *Helicobacter pylori* survives in the human stomach. *Nat Rev Mol Cell Biol*. 2001 Jun;2(6):457-66. Review.
59. R.D. Leunk, P.T. Johnson and B.C. David *et al.*, Cytotoxic activity in broth-culture filtrates of *Campylobacter pylori*, *J Med Microbiol* 26 (1988), pp. 93–99. (1996)
60. Del Giudice G, Covacci A, Telford JL, Montecucco C, Rappuoli R. The design of vaccines against *Helicobacter pylori* and their development. *Annu Rev Immunol*. 2001;19:523-63. Review.
61. Cover TL, Blaser MJ. *Helicobacter pylori* and gastroduodenal disease. *Annu Rev Med*. 1992;43:135-45.
62. T.L. Cover, M.K. Tumuru and P. Cao *et al.*, Divergence of genetic sequences for the vacuolating cytotoxin among *Helicobacter pylori* strains, *J Biol Chem* 269 (1994), pp. 10566–10573.
63. Telford JL, Covacci A, Ghiara P, Montecucco C, Rappuoli R. Unravelling the pathogenic role of *Helicobacter pylori* in peptic ulcer: potential new therapies and vaccines. *Trends Biotechnol*. 1994 Oct;12(10):420-6. Review.
64. Ye D, Willhite DC, Blanke SR. Identification of the minimal intracellular vacuolating domain of the *Helicobacter pylori* vacuolating toxin. *J Biol Chem*. 1999 Apr 2;274(14):9277-82.
65. Torres VJ, McClain MS, Cover TL. Interactions between p-33 and p-55 domains of the *Helicobacter pylori* vacuolating cytotoxin (VacA). *J Biol Chem*. 2004 Jan 16;279(3):2324-31. Epub 2003 Oct 30.
66. Vinion-Dubiel AD, McClain MS, Czajkowsky DM, Iwamoto H, Ye D, Cao P, Schraw W, Szabo G, Blanke SR, Shao Z, Cover TL. A dominant negative mutant of *Helicobacter pylori* vacuolating toxin (VacA) inhibits VacA-induced cell vacuolation. *J Biol Chem*. 1999 Dec 31;274(53):37736-42.
67. McClain MS, Iwamoto H, Cao P, Vinion-Dubiel AD, Li Y, Szabo G, Shao Z, Cover TL. Essential role of a GXXXG motif for membrane channel formation by *Helicobacter pylori* vacuolating toxin. *J Biol Chem*. 2003 Apr 4;278(14):12101-8. Epub 2003 Jan 30.

68. Reytrat JM, Lanzavecchia S, Lupetti P, de Bernard M, Pagliaccia C, Pelicic V, Charrel M, Ulivieri C, Norais N, Ji X, Cabiaux V, Papini E, Rappuoli R, Telford JL. 3D imaging of the 58 kDa cell binding subunit of the *Helicobacter pylori* cytotoxin. *J Mol Biol.* 1999 Jul 9;290(2):459-70.
69. Wang HJ, Wang WC. Expression and binding analysis of GST-VacA fusions reveals that the C-terminal approximately 100-residue segment of exotoxin is crucial for binding in HeLa cells. *Biochem Biophys Res Commun.* 2000 Nov 19;278(2):449-54.
70. Lupetti P, Heuser JE, Manetti R, Massari P, Lanzavecchia S, Bellon PL, Dallai R, Rappuoli R, Telford JL. Oligomeric and subunit structure of the *Helicobacter pylori* vacuolating cytotoxin. *J Cell Biol.* 1996 May;133(4):801-7.
71. Lanzavecchia S, Bellon PL, Lupetti P, Dallai R, Rappuoli R, Telford JL. Three-dimensional reconstruction of metal replicas of the *Helicobacter pylori* vacuolating cytotoxin. *J Struct Biol.* 1998 Jan;121(1):9-18.
72. Yahiro K, Niidome T, Kimura M, Hatakeyama T, Aoyagi H, Kurazono H, Imagawa K, Wada A, Moss J, Hirayama T. Activation of *Helicobacter pylori* VacA toxin by alkaline or acid conditions increases its binding to a 250-kDa receptor protein-tyrosine phosphatase beta. *J Biol Chem.* 1999 Dec 17;274(51):36693-9.
73. Czajkowsky DM, Iwamoto H, Cover TL, Shao Z. The vacuolating toxin from *Helicobacter pylori* forms hexameric pores in lipid bilayers at low pH. *Proc Natl Acad Sci U S A.* 1999 Mar 2;96(5):2001-6.
74. Szabo, I. *et al.* Formation of anion-selective channels in the cell plasma membrane by the toxin VacA of *Helicobacter pylori* is required for its biological activity. *EMBO J.* 18, 5517–5527 (1999).
75. de Bernard, M. *et al.* Low pH activates the vacuolating toxin of *Helicobacter pylori*, which becomes acid and pepsin resistant. *J. Biol. Chem.* 270, 23937–23940 (1995).
76. Ogura K, Maeda S, Nakao M, Watanabe T, Tada M, Kyutoku T, Yoshida H, Shiratori Y, Omata M. Virulence factors of *Helicobacter pylori* responsible for gastric diseases in Mongolian gerbil. *J Exp Med.* 2000 Dec 4;192(11):1601-10.
77. M.A. Mulvey and Y.S. Lopez-Boado *et al.*, Induction and evasion of host defenses by type 1-piliated uropathogenic *Escherichia coli*, *Science* 282 (1998), pp. 1494–1497.
78. Christie PJ, Cascales E. Structural and dynamic properties of bacterial type IV secretion systems (review). *Mol Membr Biol.* 2005 Jan-Apr;22(1-2):51-61. Review.
79. E.D. Segal, J. Cha and J. Lo *et al.*, Altered states: involvement of phosphorylated CagA in the induction of host cellular growth changes by *Helicobacter pylori*, *Proc Natl Acad Sci U S A* 96 (1999), pp. 14559–14564.
80. S. Odenbreit, J. Puls and B. Sedlmaier *et al.*, Translocation of *Helicobacter pylori* CagA into gastric epithelial cells by type IV secretion, *Science* 287(2000), pp. 1497–1500.
81. Christie PJ, Vogel JP. Bacterial type IV secretion: conjugation systems adapted to deliver effector molecules to host cells. *Trends Microbiol.* 2000 Aug;8(8):354-60
82. Fischer W, Püls J, Buhrdorf R, Gebert B, Odenbreit S, Haas R Systematic mutagenesis of the *Helicobacter pylori* cag pathogenicity island: essential genes for CagA translocation in host cells and induction of interleukin-8. *Mol Microbiol.* 2001 Dec;42(5):1337-48. Erratum in: *Mol Microbiol.* 2003 Mar;47(6):1759.
83. Couturier MR, Tasca E, Montecuccio C, Stein M Interaction with CagF is required for translocation of CagA into the host via the *Helicobacter pylori* type IV secretion system. *Infect Immun.* 2006 Jan;74(1):273-81.
84. Kwok T, Zabler D, Urman S, Rohde M, Hartig R, Wessler S, Misselwitz R, Berger J, Sewald N, König W, Backert S. *Helicobacter* exploits integrin for type IV secretion and kinase activation. *Nature.* 2007 Oct 18;449(7164):862-6.
85. Fischer W, Püls J, Buhrdorf R, Gebert B, Odenbreit S, Haas R. Systematic mutagenesis of the *Helicobacter pylori* cag pathogenicity island: essential genes for CagA translocation in host cells and induction of interleukin-8. *Mol Microbiol.* 2001 Dec;42(5):1337-48. Erratum in: *Mol Microbiol.* 2003 Mar;47(6):1759.
86. Cover TL, Dooley CP, Blaser MJ.Characterization of and human serologic response to proteins in *Helicobacter pylori* broth culture supernatants with vacuolizing cytotoxin activity. *Infect Immun.* 1990 Mar;58(3):603-10.
87. Crabtree JE, Shallcross TM, Heatley RV, Wyatt JL. Mucosal tumour necrosis factor alpha and interleukin-6 in patients with *Helicobacter pylori* associated gastritis. *Gut.* 1991 Dec;32(12):1473-7
88. Covacci A, Censini S, Bugnoli M, Petracca R, Burroni D, Macchia G, Massone A, Papini E, Xiang Z, Figura N, *et al.* Molecular characterization of the 128-kDa immunodominant antigen of *Helicobacter pylori* associated with cytotoxicity and duodenal ulcer. *Proc Natl Acad Sci U S A.* 1993 Jun 15;90(12):5791-5.

89. Nomura AM, Lee J, Stemmermann GN, Nomura RY, Perez-Perez GI, Blaser MJ. Helicobacter pylori CagA seropositivity and gastric carcinoma risk in a Japanese American population. *J Infect Dis.* 2002 Oct 15;186(8):1138-44. Epub 2002 Sep 25.
90. Nomura AM, Pérez-Pérez GI, Lee J, Stemmermann G, Blaser MJ. Relation between Helicobacter pylori cagA status and risk of peptic ulcer disease. *Am J Epidemiol.* 2002 Jun 1;155(11):1054-9.
91. Kirschner DE, Blaser MJ. The dynamics of Helicobacter pylori infection of the human stomach. *J Theor Biol.* 1995 Sep 21;176(2):281-90.
92. Backert S, Selbach M. Role of type IV secretion in Helicobacter pylori pathogenesis. *Cell Microbiol.* 2008 Aug;10(8):1573-81. Epub 2008 Apr 8. Review.
93. Backert S, Selbach M. Tyrosine-phosphorylated bacterial effector proteins: the enemies within. *Trends Microbiol.* 2005 Oct;13(10):476-84. Review.
94. Kwok T, Zabler D, Urman S, Rohde M, Hartig R, Wessler S, Misselwitz R, Berger J, Sewald N, König W, Backert S. Helicobacter exploits integrin for type IV secretion and kinase activation. *Nature.* 2007 Oct 18;449(7164):862-6.
95. Mitra SK, Schlaepfer DD. Integrin-regulated FAK-Src signaling in normal and cancer cells. *Curr Opin Cell Biol.* 2006 Oct;18(5):516-23. Epub 2006 Aug 17. Review.
96. Stein M, Bagnoli F, Halenbeck R, Rappuoli R, Fantl WJ, Covacci A. c-Src/Lyn kinases activate Helicobacter pylori CagA through tyrosine phosphorylation of the EPIYA motifs. *Mol Microbiol.* 2002 Feb;43(4):971-80.
97. Poppe M, Feller SM, Römer G, Wessler S. Phosphorylation of Helicobacter pylori CagA by c-Abl leads to cell motility. *Oncogene.* 2007 May 24;26(24):3462-72. Epub 2006 Dec 11.
98. Tammer I, Brandt S, Hartig R, König W, Backert S. Activation of Abl by Helicobacter pylori: a novel kinase for CagA and crucial mediator of host cell scattering *Gastroenterology.* 2007 Apr;132(4):1309-19. Epub 2007 Feb 1.
99. Moese S, Selbach M, Kwok T, Brinkmann V, König W, Meyer TF, Backert S. Helicobacter pylori induces AGS cell motility and elongation via independent signaling pathways. *Infect Immun.* 2004 Jun;72(6):3646-9.
100. Bourzac KM, Botham CM, Guillemin K. Helicobacter pylori CagA induces AGS cell elongation through a cell retraction defect that is independent of Cdc42, Rac1, and Arp2/3. *Infect Immun.* 2007 Mar;75(3):1203-13. Epub 2006 Dec 28.
101. Backert S, Moese S, Selbach M, Brinkmann V, Meyer TF. Phosphorylation of tyrosine 972 of the Helicobacter pylori CagA protein is essential for induction of a scattering phenotype in gastric epithelial cells. *Mol Microbiol.* 2001 Nov;42(3):631-44.
102. Aras RA, Lee Y, Kim SK, Israel D, Peek RM Jr, Blaser MJ. Natural variation in populations of persistently colonizing bacteria affect human host cell phenotype. *J Infect Dis.* 2003 Aug 15;188(4):486-96. Epub 2003 Jul 31.
103. Kim SY, Lee YC, Kim HK, Blaser MJ. Helicobacter pylori CagA transfection of gastric epithelial cells induces interleukin-8. *Cell Microbiol.* 2006 Jan;8(1):97-106.
104. Lu HS, Saito Y, Umeda M, Murata-Kamiya N, Zhang HM, Higashi H, Hatakeyama M. Structural and functional diversity in the PAR1b/MARK2-binding region of Helicobacter pylori CagA. *Cancer Sci.* 2008 Oct;99(10):2004-11.
105. Ilver D, Arnqvist A, Ogren J, Frick IM, Kersulyte D, Incecik ET, Berg DE, Covacci A, Engstrand L, Borén T. Helicobacter pylori adhesin binding fucosylated histo-blood group antigens revealed by retagging. *Science.* 1998 Jan 16;279(5349):373-7.
106. Lytton SD, Fischer W, Nagel W, Haas R, Beck FX. Production of ammonium by Helicobacter pylori mediates occludin processing and disruption of tight junctions in Caco-2 cells. *Microbiology.* 2005 Oct;151(Pt 10):3267-76.
107. Pentecost M, Otto G, Theriot JA, Amieva MR. Listeria monocytogenes invades the epithelial junctions at sites of cell extrusion. *PLoS Pathog.* 2006 Jan;2(1):e3. Epub 2006 Jan 27.
108. Mulvey MA, Schilling JD, Martinez JJ, Hultgren SJ. Bad bugs and beleaguered bladders: interplay between uropathogenic Escherichia coli and innate host defenses. *Proc Natl Acad Sci U S A.* 2000 Aug 1;97(16):8829-35. Review.
109. Martinez JJ, Mulvey MA, Schilling JD, Pinkner JS, Hultgren SJ. Type 1 pilus-mediated bacterial invasion of bladder epithelial cells. *EMBO J.* 2000 Jun 15;19(12):2803-12.
110. Sauer FG, Mulvey MA, Schilling JD, Martinez JJ, Hultgren SJ. Bacterial pili: molecular mechanisms of pathogenesis. *Curr Opin Microbiol.* 2000 Feb;3(1):65-72. Review.
111. Dubois A, Borén T. Helicobacter pylori is invasive and it may be a facultative intracellular organism. *Cell Microbiol.* 2007 May;9(5):1108-16. Epub 2007 Mar 26. Review.
112. Wunder C, Churin Y, Winau F, Warnecke D, Vieth M, Lindner B, Zähringer U, Mollenkopf HJ, Heinz E, Meyer TF. Cholesterol glucosylation promotes immune evasion by Helicobacter pylori. *Nat Med.* 2006 Sep;12(9):1030-8. Epub 2006 Sep 3.

113. Simons K, Vaz WL. Model systems, lipid rafts, and cell membranes. *Annu Rev Biophys Biomol Struct.* 2004;33:269-95.
 114. Guillemin K, Salama NR, Tompkins LS, Falkow S. Cag pathogenicity island-specific responses of gastric epithelial cells to *Helicobacter pylori* infection. *Proc Natl Acad Sci U S A.* 2002 Nov 12;99(23):15136-41. Epub 2002 Oct 31.
 115. Slonczewski JL, McGee DJ, Phillips J, Kirkpatrick C, Mobley HL. pH-dependent protein profiles of *Helicobacter pylori* analyzed by two-dimensional gels. *Helicobacter.* 2000 Dec;5(4):240-7.
 116. Lancaster C, Robert A. Intestinal lesions produced by prednisolone: prevention (cytoprotection) by 16,16-dimethyl prostaglandin E2. *Am J Physiol.* 1978 Dec;235(6):E703-8.
 117. Hills BA, Butler BD, Lichtemberger LM. Gastric mucosal barrier: hydrophobic lining to the lumen of the stomach. *Am J Physiol* 1983; 244: G561-G568.
 118. Hunt RH, Lam SK. *Helicobacter pylori*: from art to a science. *J Gastroenterol Hepatol.* 1998 Jan;13(1):21-8.
 119. CRASH-2 trial collaborators, Shakur H, Roberts I, Bautista R, Caballero J, Coats T, Dewan Y, El-Sayed H, Gogichaishvili T, Gupta S, Herrera J, Hunt B, Iribhogbe P, Izurieta M, Khamis H, Komolafe E, Marrero MA, Mejía-Mantilla J, Miranda J, Morales C, Olaomi O, Oлдashi F, Perel P, Peto R, Ramana PV, Ravi RR, Yutthakasemsunt S. Effects of tranexamic acid on death, vascular occlusive events, and blood transfusion in trauma patients with significant haemorrhage (CRASH-2): a randomised, placebo-controlled trial. *Lancet.* 2010 Jul 3;376(9734):23-32. Epub 2010 Jun 14.
 120. Fischbach L, Evans EL. Meta-analysis: the effect of antibiotic resistance status on the efficacy of triple and quadruple firstline therapies for *Helicobacter pylori*. *Aliment Pharmacol Ther* 2007, 26:343-357.
 121. Laine L, Fennerty MB, Osato M, Sugg J, Suchower L, Probst P, Levine JG. Esomeprazole-based *Helicobacter pylori* eradication therapy and the effect of antibiotic resistance: results of three US multicenter, double-blind trials. *Am J Gastroenterol* 2000, 95:3393-3398.
 122. Treiber G, Malfertheiner P, Klotz U: Treatment and dosing of *Helicobacter pylori* infection: when pharmacology meets clinic. *Expert Opin Pharmacother* 2007, 8:329-350.
 123. Altintas E, Sezgin O, Ulu O, Aydin O, Camdeviren H: Maastricht II treatment scheme and efficacy of different proton pump inhibitors in eradicating *Helicobacter pylori*. *World J Gastroenterol* 2004, 10:1656-1658.
 124. Gumurdulu Y, Serin E, Ozer B, Kayaselcuk F, Ozsahin K, Cosar AM, Gursoy M, Gur G, Yilmaz U, Boyacioglu S: Low eradication rate of *Helicobacter pylori* with triple 7–14 days and quadruple therapy in Turkey. *World J Gastroenterol* 2004, 10:668-671.
 125. O'Toole PW, Janzon L, Doig P, Huang J, Kostrzynska M, Trust TJ. The putative neuraminylactose-binding hemagglutinin HpaA of *Helicobacter pylori* CCUG 17874 is a lipoprotein. *J Bacteriol.* 1995 Nov;177(21):6049-57.
 126. Evans DG, Evans DJ, Jr, Graham DY. Receptor-mediated adherence of *Campylobacter pylori* to mouse Y-1 adrenal cell monolayers. *Infect Immun.* 1989 Aug;57(8):2272–2278.
 127. Evans DG, Evans DJ, Jr, Moulds JJ, Graham DY. N-acetylneuraminylactose-binding fibrillar hemagglutinin of *Campylobacter pylori*: a putative colonization factor antigen. *Infect Immun.* 1988 Nov;56(11):2896–2906.
 128. Evans DG, Evans DJ, Jr, Moulds JJ, Graham DY. N-acetylneuraminylactose-binding fibrillar hemagglutinin of *Campylobacter pylori*: a putative colonization factor antigen. *Infect Immun.* 1988 Nov;56(11):2896–2906.
 129. Evans DG, Karjalainen TK, Evans DJ, Jr, Graham DY, Lee CH. Cloning, nucleotide sequence, and expression of a gene encoding an adhesin subunit protein of *Helicobacter pylori*. *J Bacteriol.* 1993 Feb;175(3):674–683.
 130. Evans DG, Karjalainen TK, Evans DJ, Jr, Graham DY, Lee CH. Cloning, nucleotide sequence, and expression of a gene encoding an adhesin subunit protein of *Helicobacter pylori*. *J Bacteriol.* 1993 Feb;175(3):674–683.
 131. Evans DJ, Jr, Evans DG, Engstrand L, Graham DY. Urease-associated heat shock protein of *Helicobacter pylori*. *Infect Immun.* 1992 May;60(5):2125–2127.
 132. Jones AC, Logan RP, Foyne S, Cockayne A, Wren BW, Penn CW. A flagellar sheath protein of *Helicobacter pylori* is identical to HpaA, a putative N-acetylneuraminylactose-binding hemagglutinin, but is not an adhesin for AGS cells. *J Bacteriol.* 1997 Sep;179(17):5643-7.
 133. Lundström AM, Blom K, Sundaeus V, Bölin I. HpaA shows variable surface localization but the gene expression is similar in different *Helicobacter pylori* strains. *Microb Pathog.* 2001 Nov;31(5):243-53.
 134. Mattsson A, Lönroth H, Quiding-Järbrink M, Svennerholm AM. Induction of B cell responses in the stomach of *Helicobacter pylori*-infected subjects after oral cholera vaccination. *J Clin Invest.* 1998 Jul 1;102(1):51-6.
-

135. Mattsson A, Tinnert A, Hamlet A, Lönroth H, Bölin I, Svennerholm AM. Specific antibodies in sera and gastric aspirates of symptomatic and asymptomatic *Helicobacter pylori*-infected subjects. *Clin Diagn Lab Immunol*. 1998 May;5(3):288-93.
136. Mattsson A, Quiding-Järbrink M, Lönroth H, Hamlet A, Ahlstedt I, Svennerholm A. Antibody-secreting cells in the stomachs of symptomatic and asymptomatic *Helicobacter pylori*-infected subjects. *Infect Immun*. 1998 Jun;66(6):2705-12.
137. Bölin I, Lönroth H, Svennerholm AM. Identification of *Helicobacter pylori* by immunological dot blot method based on reaction of a species-specific monoclonal antibody with a surface-exposed protein. *J Clin Microbiol*. 1995 Feb;33(2):381-4.
138. Yan J, Mao YF, Shao ZX. Frequencies of the expression of main protein antigens from *Helicobacter pylori* isolates and production of specific serum antibodies in infected patients. *World J Gastroenterol*. 2005 Jan 21;11(3):421-5.
139. DeRosier DJ. The turn of the screw: the bacterial flagellar motor. *Cell*. 1998 Apr 3;93(1):17-20. Review. No abstract available.
140. Manson MD, Tedesco P, Berg HC, Harold FM, Van der Drift C. A protonmotive force drives bacterial flagella. *Proc Natl Acad Sci U S A*. 1977 Jul;74(7):3060-4.
141. Blair DF, Berg HC. Restoration of torque in defective flagellar motors. *Science*. 1988 Dec 23;242(4886):1678-81.
142. Zhou J, Fazio RT, Blair DF. Membrane topology of the MotA protein of *Escherichia coli*. *J Mol Biol*. 1995 Aug 11;251(2):237-42.
143. Chun SY, Parkinson JS. Bacterial motility: membrane topology of the *Escherichia coli* MotB protein. *Science*. 1988 Jan 15;239(4837):276-8.
144. Eaton KA, Morgan DR, Krakowka S. Motility as a factor in the colonisation of gnotobiotic piglets by *Helicobacter pylori*. *J Med Microbiol*. 1992 Aug;37(2):123-7.
145. Ottemann KM, Lowenthal AC. *Helicobacter pylori* uses motility for initial colonization and to attain robust infection. *Infect Immun*. 2002 Apr;70(4):1984-90.
146. Frost JW, Bender JL, Kadonaga JT, Knowles JR. Dehydroquinase synthase from *Escherichia coli*: purification, cloning, and construction of overproducers of the enzyme. *Biochemistry*. 1984 Sep 11;23(19):4470-5.
147. Millar G, Coggins JR. The complete amino acid sequence of 3-dehydroquinase synthase of *Escherichia coli* K12. *FEBS Lett*. 1986 May 5;200(1):11-7.
148. Patnaik R, Liao JC. Engineering of *Escherichia coli* central metabolism for aromatic metabolite production with near theoretical yield. *Appl Environ Microbiol*. 1994 Nov;60(11):3903-8.
149. Pittard J. The various strategies within the TyrR regulation of *Escherichia coli* to modulate gene expression. *Genes Cells*. 1996 Aug;1(8):717-25. Review.
150. Bonner CA, Jensen RA. Cloning of cDNA encoding the bifunctional dehydroquinase-shikimate dehydrogenase of aromatic-amino-acid biosynthesis in *Nicotiana tabacum*. *Biochem J*. 1994 Aug 15;302 (Pt 1):11-4.
151. Deka RK, Anton IA, Dunbar B, Coggins JR. The characterisation of the shikimate pathway enzyme dehydroquinase from *Pisum sativum*. *FEBS Lett*. 1994 Aug 8;349(3):397-402.
152. Charles IG, Keyte JW, Brammar WJ, Smith M, Hawkins AR. The isolation and nucleotide sequence of the complex AROM locus of *Aspergillus nidulans*. *Nucleic Acids Res*. 1986 Mar 11;14(5):2201-13.
153. Duncan K, Chaudhuri S, Campbell MS, Coggins JR. The overexpression and complete amino acid sequence of *Escherichia coli* 3-dehydroquinase. *Biochem J*. 1986 Sep 1;238(2):475-83.
154. Olivera BM, Hall ZW, Lehman IR. Enzymatic joining of polynucleotides. V. A DNA-adenylate intermediate in the polynucleotide-joining reaction. *Proc Natl Acad Sci U S A*. 1968 Sep;61(1):237-44.
155. Zimmerman SB, Little JW, Oshinsky CK, Gellert M. Enzymatic joining of DNA strands: a novel reaction of diphosphopyridine nucleotide.
156. Maggio-Hall LA, Escalante-Semerena JC. Alpha-5,6-dimethylbenzimidazole adenine dinucleotide (alpha-DAD), a putative new intermediate of coenzyme B12 biosynthesis in *Salmonella typhimurium*. *Proc Natl Acad Sci U S A*. 1967 Jun;57(6):1841-8. *Microbiology*. 2003 Apr;149(Pt 4):983-90.
157. Starai VJ, Celic I, Cole RN, Boeke JD, Escalante-Semerena JC. Sir2-dependent activation of acetyl-CoA synthetase by deacetylation of active lysine. *Science*. 2002 Dec 20;298(5602):2390-2.
158. Andreoli AJ, Okita TW, Bloom R, Grover TA. The pyridine nucleotide cycle: presence of a nicotinamide mononucleotide-specific glycohydrolase in *Escherichia coli*. *Biochem Biophys Res Commun*. 1972 Oct 6;49(1):264-9.

159. Zhu N, Olivera BM, Roth JR. Genetic characterization of the *pnuC* gene, which encodes a component of the nicotinamide mononucleotide transport system in *Salmonella typhimurium*. *J Bacteriol*. 1989 Aug;171(8):4402-9.
160. Spector MP, Hill JM, Holley EA, Foster JW. Genetic characterization of pyridine nucleotide uptake mutants of *Salmonella typhimurium*. *J Gen Microbiol*. 1985 Jun;131(6):1313-22.
161. Tirgari S, Spector MP, Foster JW. Genetics of NAD metabolism in *Salmonella typhimurium* and cloning of the *nadA* and *pnuC* loci. *J Bacteriol*. 1986 Sep;167(3):1086-8.
162. Sauer E, Merdanovic M, Mortimer AP, Bringmann G, Reidl J. PnuC and the utilization of the nicotinamide riboside analog 3-aminopyridine in *Haemophilus influenzae*. *Antimicrob Agents Chemother*. 2004 Dec;48(12):4532-41.
163. Liu G, Foster J, Manlapaz-Ramos P, Olivera BM. Nucleoside salvage pathway for NAD biosynthesis in *Salmonella typhimurium*. *J Bacteriol*. 1982 Dec;152(3):1111-6.
164. Foster JW, Park YK, Penfound T, Fenger T, Spector MP. Regulation of NAD metabolism in *Salmonella typhimurium*: molecular sequence analysis of the bifunctional *nadR* regulator and the *nadA-pnuC* operon. *J Bacteriol*. 1990 Aug;172(8):4187-96.
165. Kemmer G, Reilly TJ, Schmidt-Brauns J, Zlotnik GW, Green BA, Fiske MJ, Herbert M, Kraiss A, Schlör S, Smith A, Reidl J. NadN and e (P4) are essential for utilization of NAD and nicotinamide mononucleotide but not nicotinamide riboside in *Haemophilus influenzae*. *J Bacteriol*. 2001 Jul;183(13):3974-81.
166. Grose JH, Bergthorsson U, Xu Y, Sternecker J, Khodaverdian B, Roth JR. Assimilation of nicotinamide mononucleotide requires periplasmic AphA phosphatase in *Salmonella enterica*. *J Bacteriol*. 2005 Jul;187(13):4521-30.
167. Cookson BT, Olivera BM, Roth JR. Genetic characterization and regulation of the *nadB* locus of *Salmonella typhimurium*. *J Bacteriol*. 1987 Sep;169(9):4285-93.
168. Kurnasov OV, Polanuyer BM, Ananta S, Sloutsky R, Tam A, Gerdes SY, Osterman AL. Ribosylnicotinamide kinase domain of NadR protein: identification and implications in NAD biosynthesis. *J Bacteriol*. 2002 Dec;184(24):6906-17. Erratum in: *J Bacteriol* 2003 Jan;185(2):698.
169. Raffaelli N, Lorenzi T, Mariani PL, Emanuelli M, Amici A, Ruggieri S, Magni G. The *Escherichia coli* NadR regulator is endowed with nicotinamide mononucleotide adenyltransferase activity. *J Bacteriol*. 1999 Sep;181(17):5509-11.
170. Karow M & Georgopoulos C (1991) Sequencing, mutational analysis, and transcriptional regulation of the *Escherichia coli* *htrB* gene. *Mol Microbiol* 5, 2285–2292.
171. Szurmant H, Muff TJ, Ordal GW. *Bacillus subtilis* CheC and FliY are members of a novel class of CheY-P-hydrolyzing proteins in the chemotactic signal transduction cascade. *J Biol Chem*. 2004 May 21;279(21):21787-92. Epub 2004 Jan 27.
172. Jiménez-Pearson MA, Delany I, Scarlato V, Beier D. Phosphate flow in the chemotactic response system of *Helicobacter pylori*. *Microbiology*. 2005 Oct;151(Pt 10):3299-311.
173. Rosario MM, Fredrick KL, Ordal GW, Helmann JD. Chemotaxis in *Bacillus subtilis* requires either of two functionally redundant CheW homologs. *J Bacteriol*. 1994 May;176(9):2736-9.
174. Sourjik V, Schmitt R. Phosphotransfer between CheA, CheY1, and CheY2 in the chemotaxis signal transduction chain of *Rhizobium meliloti*. *Biochemistry*. 1998 Feb 24;37(8):2327-35.
175. Szurmant H, Muff TJ, Ordal GW. *Bacillus subtilis* CheC and FliY are members of a novel class of CheY-P-hydrolyzing proteins in the chemotactic signal transduction cascade. *J Biol Chem*. 2004 May 21;279(21):21787-92. Epub 2004 Jan 27.
176. Backert S, Tegtmeyer N, Selbach M. The versatility of *Helicobacter pylori* CagA effector protein functions: The master key hypothesis. *Helicobacter*. 2010 Jun;15(3):163-76. Review
177. Tegtmeyer N, Hartig R, Delahay RM, Rohde M, Brandt S, Conradi J, Takahashi S, Smolka AJ, Sewald N, Backert S. A small fibronectin-mimicking protein from bacteria induces cell spreading and focal adhesion formation. *J Biol Chem*. 2010 Jul 23;285(30):23515-26. Epub 2010 May 27.
178. Warren JR (2006) *Helicobacter*: the ease and difficulty of a new discovery (Nobel lecture). *ChemMedChem* 1, 672–685.
179. Hatakeyama M (2006) The role of *Helicobacter pylori* CagA in gastric carcinogenesis. *Int J Hematol* 84, 301–308.
180. Parsonnet J, Replogle M, Yang S & Hiatt R (1997) Seroprevalence of CagA-positive strains among *Helicobacter pylori*-infected, healthy young adults. *J Infect Dis* 175, 1240–1242.
181. Cover TL, Cao P, Lind CD, Tham KT and Blaser MJ (1993) Correlation between vacuolating cytotoxin production by *Helicobacter pylori* isolates in vitro and in vivo. *Infect Immun* 61, 5008–5012.

182. Cover TL, Cao P, Lind CD, Tham KT and Blaser MJ (1993) Correlation between vacuolating cytotoxin production by *Helicobacter pylori* isolates in vitro and in vivo. *Infect Immun* 61, 5008–5012.
 183. Nishioka H, Baesso I, Semenzato G, Trentin L, Rappuoli R, Del Giudice G & Montecucco C (2003) The neutrophil-activating protein of *Helicobacter pylori* (HP-NAP) activates the MAPK pathway in human neutrophils. *Eur J Immunol* 33, 840–849.
 184. Cao P, McClain MS, Forsyth MH and Cover TL (1998) Extracellular release of antigenic proteins by *Helicobacter pylori*. *Infect Immun* 66, 2984–2986.
 185. Vanet A and Labigne A (1998) Evidence for specific secretion rather than autolysis in the release of some *Helicobacter pylori* proteins. *Infect Immun* 66, 1023–1027.
 186. Newcomer ME, Jones TA, Aqvist J, Sundelin J, Eriksson U, Rask L and Peterson PA (1984) The three-dimensional structure of retinol-binding protein. *EMBO J* 3, 1451–1454.
 187. Newcomer M & Jones TA (1990) X-ray crystallographic studies on retinol-binding proteins. *Methods Enzymol* 189, 281–286.
 188. Zanotti G, Ottonello S, Berni R and Monaco HL (1993) Crystal-structure of the trigonal form of human plasma retinol-binding protein at 2.5-angstrom resolution. *J Mol Biol* 230, 613–624.
 189. Handa N, Terada T, Doi-Katayama Y, Hirota H, Tame JR, Park SY, Kuramitsu S, Shirouzu M and Yokoyama S (2005) Crystal structure of a novel polyisoprenoid-binding protein from *Thermus thermophilus* HB8. *Protein Sci* 14, 1004–1010.
 190. Collaborative Computational Project Number 4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* 50, 760–763.
 191. Newcomer ME, Liljas A, Sundelin J, Rask L and Peterson PA (1984) Crystallization of and preliminary X-ray data for the plasma retinol-binding protein. *J Biol Chem* 259, 5230–5231.
 192. Bialer M (1991) Clinical pharmacology of valpromide. *Clin Pharmacokinet* 20, 114–122.
 193. Jain MK, Ghomashchi F, Yu BZ, Bayburt T, Murphy D, Houck D, Brownell J, Reid JC, Solowiej JE and Wong SM (1992) Fatty acid amides: scouting mode-based discovery of tight-binding competitive inhibitors of secreted phospholipases A2. *J Med Chem* 35, 3584–3586.
 194. Hamberger A & Stenhagen G (2003) Erucamide as a modulator of water balance: new function of a fatty acid amide. *Neurochem Res* 28, 177–185.
 195. Charlton KM, Corner AH, Davey K, Kramer JK, Mahadevan S & Sauer FD (1975) Cardiac lesions in rats fed rapeseed oils. *Can J Comp Med* 39, 261–269.
 196. Garrido-Lopez A, Esquiú V & Tena MT (2007) Comparison of three gas chromatography methods for the determination of slip agents in polyethylene films. *J Chromatogr A* 1150, 178–182.
 197. Tonello F, Dundon WG, Satin B, Molinari M, Tognon G, Grandi G, Del Giudice G, Rappuoli R & Montecucco C (1999) The *Helicobacter pylori* neutrophil-activating protein is an iron-binding protein with dodecameric structure. *Mol Microbiol* 34, 238–246.
 198. Zanotti G, Papinutto E, Dundon WG, Battistutta R, Seveso M, Del Giudice G, Rappuoli R & Montecucco C (2002) Structure of the neutrophil-activating protein from *Helicobacter pylori*. *J Mol Biol* 323, 125–130.
 199. Tomb JF, White O, Kerlavage AR, Clayton RA, Sutton GG, Fleischmann RD, Ketchum KA, Klenk HP, Gill S, Dougherty BA et al. (1997) The complete genome sequence of the gastric pathogen *Helicobacter pylori*. *Nature* 388, 539–547.
 200. Guerzoni ME, Lanciotti R & Cocconcelli PS (2001) Alteration in cellular fatty acid composition as a response to salt, acid, oxidative and thermal stresses in *Lactobacillus helveticus*. *Microbiology* 147, 2255–2264.
 201. Leslie AGW (2006) The integration of macromolecular diffraction data. *Acta Crystallogr D Biol Crystallogr* 62, 48–57.
 202. Evans P (2006) Scaling and assessment of data quality. *Acta Crystallogr D Biol Crystallogr* 62, 72–82.
 203. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC & Read RJ (2007) Phaser crystallographic software. *J Appl Crystallogr* 40, 658–674.
 204. Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS et al. (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* 54, 905–921.
 205. Murshudov GN, Vagin AA & Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* 53, 240–255
-

-
206. Painter J & Merritt EA (2006) Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr D Biol Crystallogr* 62, 439–450.
 207. Emsley P & Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60, 2126–2132.
 208. Painter J & Merritt EA. (2006) Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr D62*, 439–450.
 209. Laskowski RA, Macarthur MW, Moss DS & Thornton JM (1993) Procheck – a program to check the stereochemical quality of protein structures. *J Appl Crystallogr* 26, 283–291.
 210. Lovell SC, Davis IW, Arendall WB 3rd, de Bakker PI, Word JM, Prisant MG, Richardson JS & Richardson DC (2003) Structure validation by Calpha geometry: phi, psi and Cbeta deviation. *Proteins* 50, 437–450.
 211. Kleywegt GJ & Jones TA (1994) Detection, delineation, measurement and display of cavities in macromolecular structures. *Acta Crystallogr D Biol Crystallogr* 50, 178–185.
 212. Delano WL (2002). The PyMol molecular graphics system. DeLano Scientific, Palo Alto, CA.
 213. Zanotti G, Panzalorto M, Marcato A, Malpeli G, Folli C & Berni R (1998) Structure of pig plasma retinol-binding protein at 1.65 angstrom resolution. *Acta Crystallogr D Biol Crystallogr* 54, 1049–1052. L
-

Summary

Gastric cancer is the second leading cause of cancer death worldwide. This is in part a consequence of the increased human longevity, although the rate of gastric cancer has actually been declining from the early 20th Century. Since the relationship between *H. pylori* and chronic gastritis was established, investigators began to take interest in the role played by *H. pylori* in this type of cancer. From the first studies to examine this association, several epidemiological data have confirmed the relationship between the regional prevalence of *H. pylori* and the incidence of stomach cancer.

H. pylori is a Gram-negative spiral-shaped bacterium that is characterized by two aspects: the use of several unipolar flagella that confer motility and its ureasic activity. *H. pylori* is a highly variable bacterial species, both genotypically and phenotypically, and is highly adapted for survival in the gastric niche.

The development of effective treatment options to eradicate the *H. pylori* infection has resulted in an immense change in the clinical management of upper gastrointestinal diseases. The wide use of antibiotic therapies against *H. pylori* infection, however, is related to the number of therapeutic failures. Recent data show a decreasing efficacy of these therapies worldwide. The research for new therapeutic strategies and new drugs is necessary.

The project presented here aims to determine not only the three dimensional structure of proteins playing a role in the virulence of the bacterium, but also their analysis and characterization using an approach of structural genomics. The three-dimensional structure of a protein, in addition to clarify its function, may be useful for the design of potential new drugs.

This thesis presents several potential drug targets, each consisting of proteins involved in pathogenicity or essential to the survival of the bacterium. All genes were cloned as described by using different strategies; 7 out of 9 proteins were expressed, 3 out of 7 were soluble in an heterologous enviromen; these three proteins were structurally characterized, two with a crystallographic approach and one with SAXS. The results of the studies on the latter three proteins are presented in detail: CagL, a protein belonging to cagPAI; HP1286, a periplasmic protein belonging to the family YceI that is induced by osmotic stress created by high NaCl or high acidity; and finally HP0797, an adhesin.

CagL is a protein localized on the surface of the *pilus* and acts as a specialized adhesin that connects the T4SS with target cells. The integrin complex $\alpha_5\beta_1$ binds to a small protein of 26 kDa, CagL precisely, which is encoded by the gene *hp0539* of the cagPAI. CagL seems to be a functional ortholog of VirB5, a structural component of the T4SS of *Agrobacterium tumefaciens*. In this thesis we propose a structural model of the oligomerization state of CagL in solution, using the technique of

small angle x-ray scattering (SAXS). This model is supported by other experiments, such as crosslinking and gel filtration.

Regarding the HP1286 enzyme, crystals were grown and diffraction data measured. Crystals belong to space group $P2_12_12_1$ and contain a dimer per asymmetric unit. Each monomer has a hydrophobic pocket in which there is an electron density indicating the presence of a ligand. The latter was identified by mass spectrometry as erucamide, the amide of a fatty acid of 22 carbon atoms, erucic acid ($Z\text{-CH}_3(\text{CH}_2)_7\text{CH}=\text{CH}(\text{CH}_2)_{11}\text{CONH}_2$).

The recombinant protein HpaA (HP0797), an adhesin, was expressed in *E. coli* in soluble form without the signal peptide. We have obtained crystals that diffract only at a resolution of 4 Å and attempts to improve them are in progress. A molecular model was built by homology modelling and the putative domain of glycolipids of HpaA was predicted based on its amino acid sequence.

Riassunto

Il tumore gastrico è la seconda causa di morte per cancro nel mondo. Questo fatto è in parte una conseguenza della maggiore longevità umana, anche se il tasso di cancro gastrico è effettivamente in calo rispetto ai primi anni del 20° secolo. Da quando è stata stabilita una relazione tra *H. pylori* e gastrite cronica, i ricercatori hanno cominciato ad interessarsi al ruolo svolto da *H. pylori* in questi tipi di tumore. Dai primi studi volti ad esaminare tale associazione, numerosi dati epidemiologici hanno confermato la relazione tra la prevalenza regionale di *H. pylori* e l'incidenza dei tumori allo stomaco.

H. pylori è un batterio Gram-negativo a forma di spirale che si caratterizza per due aspetti: l'utilizzo di numerosi flagelli unipolari che gli conferiscono motilità e l'attività ureasica. *H. pylori* è una specie batterica altamente eterogenea, sia genotipicamente che fenotipicamente, che si è adattata alla sopravvivenza nella nicchia gastrica. Diversi aspetti della biologia delle adesine di *H. pylori* esemplificano l'importanza dell'eterogeneità in questo sistema: (i) nessuna adesina in particolare è essenziale per il fissaggio alla mucosa gastrica, il che indica la ridondanza del meccanismo adesivo; (ii) l'espressione delle adesine è diversificata tra ceppi ed è variabile nel tempo all'interno di un unico ceppo; (iii) le interazioni adesive contribuiscono all'infiammazione ed è probabile che siano coinvolte nella progressione della malattia. Lo sviluppo di diversi trattamenti per sradicare l'infezione da *H. pylori* si è

tradotto in un enorme cambiamento nella gestione clinica delle patologie del tratto gastrointestinale superiore. L'ampio uso di terapie antibiotiche contro l'infezione da *H. pylori* è, però, correlato al numero di fallimenti terapeutici. Dati recenti mostrano una diminuzione dell'efficacia di queste terapie in tutto il mondo. Ciò ha reso necessaria la ricerca di nuove strategie terapeutiche e di nuovi farmaci.

Il progetto qui presentato mira a determinare non solo la struttura tridimensionale delle proteine descritte, ma anche la loro analisi e caratterizzazione attraverso un approccio di genomica strutturale. La struttura tridimensionale di una proteina, infatti, oltre a chiarire le funzioni della proteina target, può essere utile per la progettazione di potenziali nuovi farmaci.

In questa tesi vengono presentati diversi potenziali bersagli farmacologici, tutti consistenti in proteine coinvolte nella patogenicità o comunque essenziali per la sopravvivenza del batterio. Tutti i geni descritti sono stati clonati utilizzando strategie diverse; 7 su 9 proteine sono state espresse, 3 su 7 sono risultate solubili in ambiente eterologo; queste tre proteine sono state infine caratterizzate strutturalmente: due con approccio cristallografico ed una con SAXS. Sono presentati in dettaglio i risultati degli studi su queste ultime tre proteine: CAGL, una proteina appartenente alla cagPAI; HP1286, una proteina periplasmatica appartenente alla famiglia YceI che è indotta da stress osmotico creato da elevato NaCl o da elevata acidità; infine HP0797, un'altra adesina.

CAGL è una proteina presente sulla superficie del *pilus* e funge da adesina specializzata che collega il T4SS con le cellule bersaglio. Il complesso di integrine $\alpha_5\beta_1$ si lega ad una piccola proteina di 26 kDa, CAGL appunto, che è codificata dal gene *hp0539* della cagPAI. CAGL sembra essere un ortologo funzionale di VirB5, un componente strutturale del T4SS di *Agrobacterium tumefaciens*. In questa tesi viene proposto un modello strutturale dello stato di oligomerizzazione di CAGL in soluzione, utilizzando la tecnica di “small angle x-ray scattering” (SAXS). Questo modello viene supportato da numerosi esperimenti, quali crosslinking e gel filtrazione.

Per l'enzima HP1286 sono stati ottenuti cristalli ed i dati di diffrazione sono stati misurati. I cristalli appartengono al gruppo spaziale $P2_12_12_1$; essi contengono un dimero per unità asimmetrica. Ogni monomero presenta una tasca idrofobica nella quale c'è una densità elettronica che indica la presenza di un ligando. Quest'ultimo è stato identificato dalla spettrometria di massa, come erucamide, l'ammide di un acido grasso di 22 atomi carbonio, acido erucico ($Z-CH_3 (CH_2)_7CH = CH (CH_2)_{11}CONH_2$).

La proteina ricombinante HpaA (HP0797), un'adesina, è stata espressa in *E. coli* in forma solubile, senza il peptide segnale. Si sono ottenuti cristalli che per ora diffrangono solo alla risoluzione di 4 Å e si sta cercando di migliorarli. Il putativo dominio glicolipidico della HpaA è stato predetto sulla base della sequenza aminoacidica. La struttura

atomica HpaA potrebbe essere utilizzata anche per caratterizzare i domini glicolipidici in proteine che sono racchiuse in un guscio di lipidi.