

The importance of CagA protein in *Helicobacter Pylori* infection

Azucena Arévalo, Bact.,¹ Alba Alicia Trespalacios, MSc,² William Otero, MD.³

¹ Junior Investigator, MSc Candidate, Faculty of Science, Microbiology Department, Universidad Javeriana

² Professor of Microbiology, Medical Microbiology Specialization Director, PhD Candidate, Universidad Javeriana, Faculty of Science, Department of Microbiology

³ Internist, Gastroenterologist, Epidemiologist, Professor of Medicine, Gastroenterology Unit, National University of Colombia.

Received: 25-05-09
Accepted: 14-10-09

Summary

Helicobacter pylori is a microorganism able to colonize gastric mucosa in humans where it can produce chronic gastritis and other type of complications. *H. pylori* is present approximately 20-50% in the industrialized countries but in developing countries its prevalence is the highest because approximately 80% of people are infected with the bacteria. In general this bacteria is variable in its genome but the greatest genetic plasticity is located at 40kb DNA segment, known as a pathogenicity island (PAI), inside of this DNA segment there are *cagA* gen which coding for CagA protein and genes that coding for type IV secretion system that is necessary for export CagA protein into target cell. *cagA* gen is important because it is a marker of PAI presence and because the presence of *cagA* has permitted classified *H. pylori* strains in *cagA+* and *cagA-*, which is of great importance due *cagA+* strains are more virulent than *cagA-* strains, although the principal importance of *cagA+* strains is its special association with gastric cancer. The aim of this review is study the functions of pathogenicity island genes and its association with gastro duodenal pathologies developing.

Key words

Helicobacter pylori, *cagA*, pathogenicity island.

INTRODUCTION

Helicobacter pylori (*H. pylori*) are gram negative bacteria that colonize the body and the antrum of the human stomach causing gastritis and complications such as gastric and duodenal ulcers, MALT lymphoma and gastric adenocarcinoma (1-3). The outcome of the disease depends on environmental factors, the host and the bacterium (1-3). Among the characteristics attributed to the organism are several virulence factors which allow it to colonize the gastric mucosa, survive in the acidic environment of the stomach and evade the immune response (1, 2).

Unlike other enteric pathogens such as Salmonella and Yersinia that have evolved mechanisms to invade the M cells of the intestinal epithelium, *H. pylori* remains primarily in the gastric mucus outside of the gastric epithelial cells. However, thanks to several virulence factors, it is capable of triggering signals in host cells that interfere with

basic cellular processes and which ultimately culminate in the onset of disease (1, 4). One of the most studied virulence factors encoded within the pathogenicity island is the *cagA* gene the protein it codes for, Cag A (2-4). *H. pylori* strains capable of coding for this protein are called Cag A positive strains. The gene has been found in approximately 90% of all patients infected with the microorganism whose presence has been associated statistically with duodenal ulcer, gastric mucosa atrophy and gastric cancer (5).

The objective of this review is to broaden the outlook on the importance of the pathogenicity island to *cagA* in *H. pylori* and to analyze how some of the genes in this island contribute to the development of gastroduodenal diseases.

PAI GENERAL CHARACTERISTICS

Bacterial evolution is not a continuous process. It can result from horizontal transfer by acquiring DNA segments of

unknown origin which are integrated into the bacterial chromosome by homologous recombination (4). This new portion of the integrated DNA is known as an island. Its DNA can code for multiple proteins involved in iron storage systems, metabolic enzymes, secretion systems, cell surface proteins, adherence factors, toxins and so on (4, 6). The pathogenicity islands (PAIs) are characterized by the following: they have a guanine-cytosine content (GC) different from the rest of the genome, they have a constant codon which adapts to the bacterial chromosome, they are surrounded by direct repeats (DR); they are associated with gene transfer RNA (tRNA); and they have genes that code for mobile factors such as integrase, transposase and insertion sequence elements (IS) (6). The importance of the DR regions of tRNA genes and IS elements is that they act as deletion sites for PAIs causing them to be unstable (6). However, these regions can also act as DNA binding sites. In *H. pylori*, one of the sequences which serves as a binding site is the *glr* gene encoding glutamate racemase (6). Thanks to the presence of these regions after the initial acquisition of foreign DNA, PAIs can be optimized according to the needs of the recipient cell (4).

Pathogenicity islands are regulated by genes which code for regulatory factors located within the same island, although these regulatory genes may also be involved in regulating genes outside the PAI. However, a PAI may also be regulated by genes located outside of it which at the same time regulate housekeeping genes (6).

H. PYLORI PATHOGENICITY ISLAND

The concept of PAIs was initially developed by to describe the acquisition of DNA segments by pathogenic *E. coli* strains which were absent in nonpathogenic strains (7). In 1996 Censini et. al. found the presence of PAI in *H. pylori*, these findings have since been confirmed and extended by subsequent studies. The *H. pylori* PAI was initially called *cag* (cytotoxin associated gene) because it was thought to be associated with the expression of *VacA* (vacuolating cytotoxin). However, it was subsequently observed that *VacA* and PAI are independent (4, 8).

According to analyses of *H. pylori*'s genome, its PAI is considered to be the most important zone of variability in its entire genome (9). It consists of an open reading frame (ORF) of 40 kb surrounded by 31pb of direct repeats. These contain recombination sites corresponding to the last nucleotides of the *glr* gene coding for glutamate racemase (4). The *cag* PAI lies between a *glr* gene and an ORF, the function of which has not yet been described. It is known as ORF5 and was designated HP0519 by Tomb et. al. in 1997 (10). Both at the left and to the extreme right end the PAI has insertion sequences called IS605 elements.

They vary in number depending on the strain. Strains with many insertion elements have been designated as type II strains, which are less virulent than type I strains containing a PAI (11). In addition two of the IS605 elements are responsible for splitting the PAI into two regions which are designated *cag* I and *cag* II (4, 7), Figure 1A.

H. pylori *cag* PAI can be found as a single unbroken unit, divided into two regions by an insertion sequence, or as a partially deleted DNA fragment. C. Audibert et. al. (5) assessed the structure of *H. pylori* *cag* PAI strains and found some that had no *cag* PAI disruption, some with *cag* PAI split in two, and others without *cag* PAI. They also reported the finding strains that had only middle region (IS605) and *cag* II, as well as strains with only the middle region and *cag* I, and other strains with only *cag* II, only *cag* I or only IS 605 (5).

Some genes on the island encode proteins similar to those involved in DNA transfer (*Vir* and *Tra*) and in the export of toxins (PTL). However, these genes do not have conserved operons of *Vir*, *Tar* and PTL proteins as in other microorganisms. This suggests that PAI is not derived directly from those systems. The type IV secretion system is found within the proteins involved in transfer (12). Sequencing of strains NCTC 26695 and 11638 has found 30 proteins encoded by the PAI genes have a signal sequence. Its size is approximately 25 amino acids, suggesting that PAI has a Gram-negative origin (4). It also codes for eight internal membrane proteins with at least two membrane domains. The percentage of nucleotide identity among PAI strains is approximately 97% which means that among these strains the remaining 3% may represent significant differences in important biological functions (12).

cag II and cag. Censini et. al. (7) designated a segment of DNA from *cagA* + strains which is present in PAI as *cag* II, and the segment that contains the *cagA* gene as *cag* I. However a study by Akopyants NS, et. al. (10) found that portions of the *cag* II segment hybridizes with segments of *cagA*- strains indicating that on both sides of this segment there are sequences which are common to all strains of *H. pylori*. (10)

Studies by Akopyants NS, et. al. (10) show that *cag* II is a 19kb segment that consists of 15 open reading frames unique to *cagA* + strains. 64% of their content is A+T, while 65% of *cag* I's content is A+T and 61% of the content throughout the entire chromosome of the microorganism consists of A+T. Four ORFs found in *cag* II (Table 1) have genes which encode proteins homologous to *A. tumefaciens* and *B. pertussis*, required for the release of toxins and for transfer of DNA. In addition, to the right of *cag* II ORF 21 and ORF 22 encode transposases of IS elements homologous to those in other bacteria and corresponding to repeated sequences ("RS2"), known as IS 605 (10).

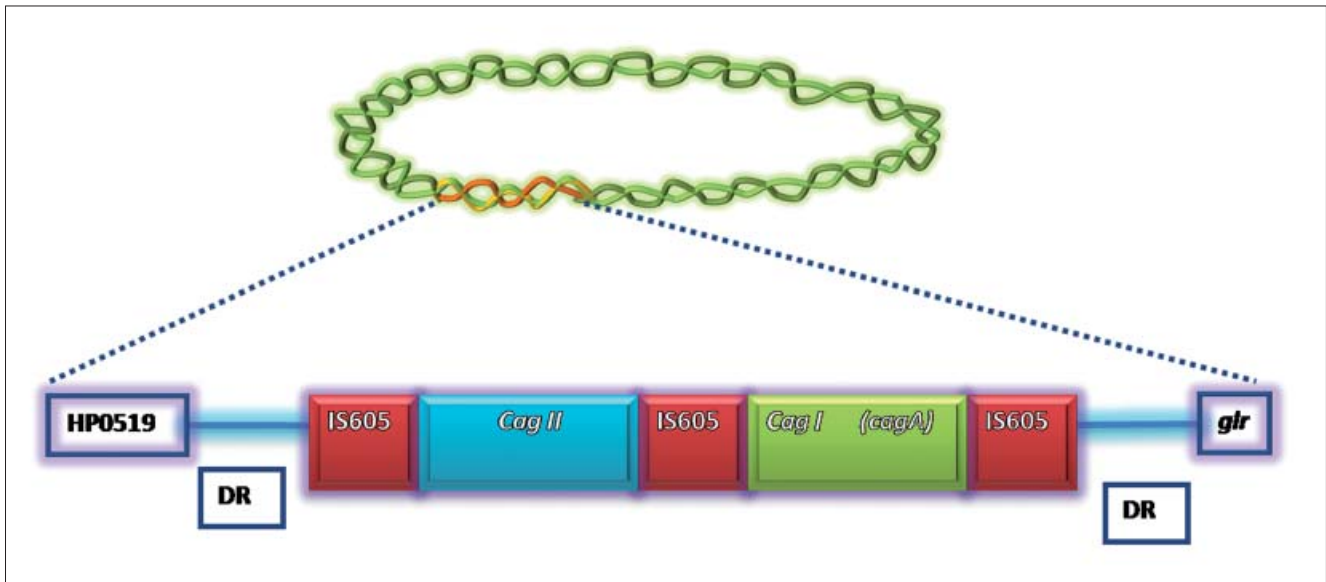


Figure 1a. Represents the pathogenicity island of *H. pylori*. This region is found between the *glr* gene and an ORF HP0519, the function of which is not known. The PAI is surrounded by direct repeat sequences (DR). It is divided into *cagI* and *cagII* by IS605 elements, the quantity of which may vary among different strains. Within the *cagI* and *cagII* regions there are genes involved in coding for T4SS. The *cagA* gene is found within *cagI*, but not *cagII*.

Table 1. *H. pylori* ORFs, the genes of which encode proteins homologous to *A. tumefaciens* and *B. pertussis*.

ORF in <i>H. pylori</i>	Homologous genes in <i>A. tumefaciens</i>	Homologous genes in <i>B. pertussis</i>
ORF 11	<i>virB 11</i>	<i>pil/H</i>
ORF 13. 14	<i>virB 10</i>	<i>pil/G</i>
ORG 15	<i>virB 9</i>	<i>pil/F</i>
ORF 10	<i>virID4</i>	_____

Type IV Secretion System Encoded by PAI

Micro-organisms have developed systems of communication between themselves and their habitat as taxis systems, quorum sensing systems and secretion systems (4). In Gram-negative pathogenic bacteria four different secretion systems have been described (which does not rule out the possibility of more). Through these systems they excrete proteins, enzymes, DNA or virulence factors (4). However, among these four secretion systems, only two have been described as capable of carrying virulence factors, the type III secretion system present in some enterobacteria and the type IV secretion system (T4SS) (4). Although these two systems are both capable of carrying virulence factors, their evolutionary origins are different. The former is related to the flagellar system, while the second probably has its ori-

gins in the bacterial conjugation apparatus. This explains why it allows transfer of DNA. *H. pylori* allows not only horizontal gene transfer, but also importation of peptidoglycan and CagA protein into the epithelial cell (4).

The components, role and position of the type IV secretion system of *H. pylori* are homologous to components of *A. tumefaciens* (identified as vir B) (13). About 18 genes identified within the PAI are responsible for encoding for the type IV secretion system in *H. pylori*. Thanks to electron microscope studies T4SS has been identified as a filamentous organelle located at one pole of the bacterial surface and induced by contact (13, 14). The model of organization for the type IV secretion system suggests that its proteins are grouped into cytoplasmic proteins (inner membrane proteins), proteins which form the central or core complex located in the periplasm, and proteins that are part of the pili or surface structure which project beyond the outer membrane (14).

Once auxiliary adhesions put the organism in contact with an epithelial cell, T4SS assembly begins. Initially, the proteins which are part of the inner membrane, such as CagE/HPO544 (VirB4) and HPO525 (VirB11), are assembled (15). They possess ATPase activity which allows translocation of the substrate. An additional protein, VirD4 which is homologous to HPO524, is responsible for delivering the substrate to the T4SS secretion machinery (15).

Next, assembly of proteins which will make up the central complex of the T4SS begins. Among these proteins is

CagT/12 (HPO532). Located at the base of the organelle, it allows assembly of the rest of the organelle. Other proteins which are assembled into the core complex include CagV/10, CagX/8 (HPO528) and CagY/7 (HPO527) which are homologous to VirB7, VirB8, VirB9, and VirB10, respectively (14 and 16). Although these proteins have been described as part of the central complex, a protein-protein interaction study done by Valerie J. Busler et al. (14) in 2006 revealed that three of these proteins, CagY / 7 CagX / 8 CagT/12 are associated with the pili formed between *H. pylori* and epithelial cell. They also propose that other proteins such as CagM/16 Cag, CagI/19, CagG/21 and CagF/22 may represent species specific core complex components of *H. pylori* (14). Because the proteins that form the core must cross the peptidoglycan, additional proteins have been identified such as HPO523 which is homologous to VirB1. It may act as a transglycosylase, smoothing the murein layer of the bacterial cell wall and facilitating T4SS assembly through the bacterial wall. This is also necessary for maturation of CagY/7 CagT/12 and for maturation of HPO539 which is a possible chaperone of these proteins (16).

Finally proteins that form the pili are assembled. The main component is CagC which is homologous to VirB2. It is wholly or partially coated by CagY and CagL, the latter serving as an adhesin allowing connection between T4SS and the target cell (15), Figure 1b.

Once assembled the T4SS allows transfer of CagA to the epithelial cell. This requires that the secretion system interact with a specific receptor on the host cells. This is why it has been proposed that host cell integrins might interact with CagL protein. Integrins are transmembrane cell-cell adhesion molecules and cell-extracellular matrix motifs that bind Arg-Gly-Asp (RGD). CagL contains these motifs, so it is possible that interaction of integrin $\alpha\beta 1$ and pili takes place after they inject the CagA protein. (15)

After assembly of the T4SS and adherence of the type IV secretion system to the epithelial cell, CagA protein is transferred. This requires CagF protein to act as a chaperone. It binds to the C-terminal portion of the signal sequence of CagA protein until it is translocated to the epithelial cell. (15).

cag A

cag A is a cytotoxin-associated gene which encodes the CagA protein (2). Within the cag PAI region there are 31 open reading frames (ORF). One of these ORFs encodes the immunodominant antigen CagA which is located towards the 3' end of the pathogenicity island. CagA protein was first identified in PAI and is the most important of *H. pylori*'s virulence factors. Its presence (cagA+) is a direct

marker for PAI. The protein was discovered in the early 1990s independently by Cover et al. (15) Jean Crabtree et al. (15) and Covacci et al. (15). Since then several studies have highlighted its importance in the pathogenesis of the organism (17). The molecular size of cagA varies from 120kDa to 145 kDa and depends on the number of repetitive sequences located in the 3' region of the gene. In the absence of repetitions, CagA has a molecular weight of 128 kDa. Each iteration increases its molecular weight 4kDa, up to a maximum of 4 repeats (4). In western strains two types of repeats have been found: a 57 bp repeat followed by a repeat of 102 bp. In Asian strains the same initial 57bp repetition has been found, but it is followed by a 162bp repetition (18). Protein is translocated ("shot") through the type IV secretion system into the host cell and subsequently phosphorylated at specific sites known as EPIYA. It thus interacts with various signaling pathways, triggering changes in the cytoskeleton, the morphology and the mobility of the host cell (15,19). Once CagA is translocated into the host cell it is tyrosine phosphorylated by host kinases. Phosphorylation occurs at Glu-Pro-Ile-Tyr-Ala sites (also called EPIYA motifs) which can be repeated up to five times in the middle of CagA terminal carbon. Repetitions of these motifs are surrounded by repetitive DNA sequences involved in recombination. Based on these sequences, four EPIYA sites have been termed A, B, C and D (15,20). The presence of these repeats surrounding EPIYA motifs is of great importance since they could explain both the variability in the number of motifs in CagA and differences in the pathogenicity of *H. pylori* strains. One of these differences is that between Western strains and Asian strains. Western strains are characterized by CagA protein which consists of EPIYA sites A and B followed by one to three repetitions of 34 amino acids that contain EPIYA-C. Asian *H. pylori* strains are generally characterized by expressing CagA with sites of EPIYA-D which replaces EPIYA-C (3, 15, 20). Different types of EPIYA determine Cag A proteins. Anthropologically they could identify the origin of the ancestors of a given population (18). The number of repetitions, especially in secondary protein motifs, are of particular importance in gastric cancer, the incidence of which is significantly higher in patients infected with *H. pylori* strains presenting multiple repetitions. However, the poor survival rate of these strains in acid pH limits their ability to contribute to the development of gastric cancer. For this reason it has been suggested that they develop after patients develop atrophy and hypochloridia (3, 15). Nevertheless, this suggestion remains in the realm of theory as it is not known how interactions might be carried out with the microorganism that would lead to the conditions of mucosal atrophy which allow the organism to develop alternatives that worsen the condition (21). One possibi-

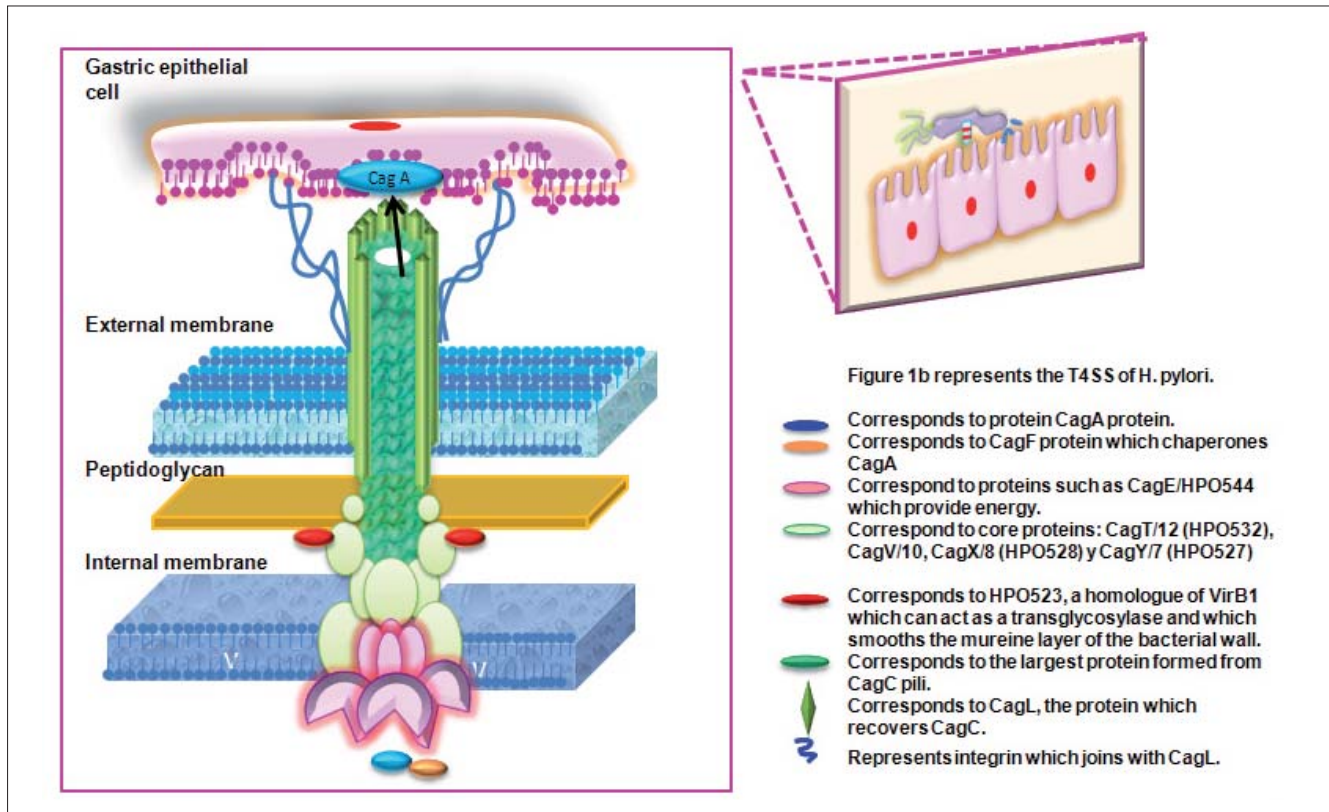


Figure 1b. Secretion system of Type IV *H. pylori* (T4SS). The figure shows the organization of proteins which form the T4SS of the microorganism. Proteins which function as ATPases organize are located in the internal bacterial membrane. Other proteins form the central complex and are available in the bacteria's periplasmic space. Finally, the proteins which form the pili, and which allow communication with the cell's through the integrin required for later injection of CagA, cross part of the peptidoglycan and the external membrane.

lity would be co-infection with *H. pylori* strains with different CagA occurring at the start, or during the course, of the disease.

In normal cells, phosphorylation is mediated by kinases of the Src family known as SFKs. Their function is to control processes in the cytoskeleton, cell proliferation and cell differentiation, however they have also been found in processes of carcinogenesis. Other kinases that phosphorylate CagA, such as c-Abl and Arg, belong to the Abl family (15). Src is activated during the initial stages of *H. pylori* infection, 0.5 hours to 2 hours, whereas Abl is continuously activated by the organism (15). Once CagA is tyrosine phosphorylated by these kinases, it specifically binds to and activates tyrosine phosphatase SHP-2, an oncoprotein, the mutations of which are associated with human malignancies. CagA deregulates SHP-2, disrupting the Erk MAP kinase and the dephosphorylated FAK (focal adhesion kinase) inducing changes in cellular morphology which adopts a phenotype known as a hummingbird shaped cell (15). Cag A also impairs cell-cell interactions independently of

phosphorylation. Cag A destroys tight junctions and causes loss of polarity in epithelial cells. It also destabilizes complex E-cadherin / β -catenin (3,22). There is a possibility that the interaction of CagA with cellular derangement may include key steps in the development of gastric cancer, although this is still being studied (22). This mechanism implies that *H. pylori* may not only have an indirect oncogenic effect through causing persistent inflammation and hyperproliferation with the risk of free radicals that injure the DNA of cells with rapid growth, but may also have a direct oncogenic effect on gastric epithelial cells through the oncoprotein Cag A, (22).

Naomi Ohnishi, et. al. (22) demonstrated in vivo that CagA is a bacterial oncoprotein whose expression is sufficient to cause tumors. For this important research they used transgenic mice capable of expressing the Cag A protein, which may or may not phosphorylate. The formation of tumors in these mice was associated with the CagA tyrosine phosphorylation essential for CagA interaction with deregulated SHP-2 oncoprotein. This allowed abnormal

proliferation of gastric epithelial cells and altered the leukogenesis supporting the association of *H. pylori* infection and the development of B cell lymphomas (22).

Nevertheless, translocation of CagA, its subsequent phosphorylation, and subsequent effects these cause in signaling pathways and cellular morphology may be influenced by factors in addition to kinases. Ho Chin Lai, et al. (23) demonstrated that cholesterol is an important factor in the action of Cag A and that when there is a decrease in cellular cholesterol the translocation of CagA by TSS4 and phosphorylation reduces, and there is a blockade of cellular responses induced by the protein, including the form of a hummingbird phenotype and induction of IL-8 (23).

cagA PAI in patients infected with *H. pylori*

H. pylori colonize the gastric mucosa in approximately 70% to 90% of individuals in underdeveloped countries. They acquire the organism at early ages, but the majority (80%) remain asymptomatic. This which may be due in part to the strain that they have been infected with, however in susceptible individual's gastric mucosal damage ranging from gastritis and ulcers to gastric carcinomas and MALT lymphomas may develop (24).

Using the *cagA* gene as a marker to study the presence of the PAI several studies from different parts of the world have found associations between the presence of this virulence factor and the severity of patients' diseases (11, 17, 24). Hasan Umit, et. al. (24) found an association between the presence of *cagA* and the degree of inflammation in the body and antrum of the stomach, and they found that antral glandular atrophy is associated more with *cagA* + strains than with *cagA*- strains. Aime T. Franco and other authors (25) have shown that *cagA* + strains are required for induction of dysplasia and adenocarcinoma in the gastric mucosa of Mongolian gerbils. They also found the presence of other virulence factors including OipA, which has also been found in precancerous lesions in humans (25). Nilsson, et. al. (26) reported that *cag* PAI without deletions has been associated with an increased risk of the development of severe diseases, while strains with internal deletions are less virulent. A. Mohaboob et. al. (11) had results consistent with the findings of Nilsson, finding that there are associations between the histological damage caused by *H. pylori* strains with complete *cag*PAI and that partial deletions correlate with disease status. They found *cag*PAI partial deletions in patients with duodenal ulcer and functional dyspepsia, while in patients with gastric ulcers and gastric cancer complete *cag*PAI strains predominated. They also found that strains with complete *cag* PAI dominated in this population (11). Other researchers have found no differences between complete and partially dele-

ted *cag* PAI in relation to the severity of inflammation and neutrophil activity in a study of infected Korean children. That study used the presence of *cagA*, *cagE* and *cagT* to assess the structure of *cag*PAI (27).

Besides the relationship mentioned above, it has been postulated that *cagA* positive strains have the ability to induce production of interleukin 8 (IL-8) through its ability to activate MAP kinases, which regulate differentiation, proliferation, apoptosis, stress and inflammatory responses (5 and 28). However, recent studies have reconsidered the possibility of encountering a correlation between the structure of *cag* PAI and induction of IL-8. These studies found that deletions of various genes of the *cag* PAI- with the exception of *cagA* – did not generally diminish induction of production of interleukin. For this reason they postulate that different genes within *cag*PAI may be involved in the induction of IL-8, for example proteins external to the membrane such as OipA (5, 30, 31). OipA is an outer membrane inflammatory protein, its expression is regulated by a repair mechanism based on the number of cytosine-thymine (CT) repetitions in the 5' end region of the gene *oipA*. Depending on "on" or "off" status of the gene it may affect the virulence characteristics of the organism. Its expression has been associated with duodenal ulcers, gastric cancer, high densities of *H. pylori*, the presence of *cagA* + strains, severe infiltration of neutrophils, and induction of IL-8 (21, 29-31). A study by Yamaoka et. al. (30) examined the ability of four genes encoding outer membrane proteins (HPO638 (*oipA*), HPO796, HP1501 and *babA2*) to produce IL-8. They found that mutant strains HPO796, HP1501 *babA2* and even *CagE* have no significant effect on the production of IL-8. This differs from *oipA* mutant strains of which 81% were *cagA* + and in which they found a 50% reduction in the production of IL – 8. Additionally they found that *cagA*- with the active *oipA* gene produces three times more IL-8 than *cagA*- with *oipA* inactive. (30) However, the mechanisms IL-8 induction in *H. pylori* infections are still not fully understood.

CONCLUSIONS

Since *Helicobacter pylori* has the ability to induce chronic gastritis, gastric MALT and adenocarcinoma, it has been classified as a type I carcinogen. Although the factors determining the development of infection are not understood, it has been suggested that several of its virulence factors may be involved in the outcome of the disease. It has been posited that *CagA* is a marker for *cag* PAI, and if *cag* PAI is present the organism has the ability to encode for the entire machinery involved in horizontal transfer and export of proteins. In addition the presence of *cag* PAI gives a panorama of extensive genetic diversity to the microorganism.

It possesses recombination and deletion points for foreign DNA which confer advantages on the bacteria, enabling them to acquire virulence factors such as dissemination of antibiotic resistance.

Although several studies have posited a possible relationship between *cagA* + strains and the outcomes of diseases, this is still a matter for further research because the outcome does not depend solely on the organism, but also upon several additional factors. These include the environment and the immune, genetic, social and cultural characteristics of the host. It is also important to note that when considering the microorganism alone there are other factors in addition to determination of the type of *cagA* in a given population that must be considered. The features present in the motifs of the protein as well as the characteristics of *cagPAI* must be identified for the specific study population. In addition, the presence of other virulence factors of the microorganism independent of *cagPAI* that might be influence the pathogenesis of the bacterium must be verified or discarded. In other words, *H. pylori*, with all its virulence factors, is a necessary but not sufficient cause for the majority of gastric adenocarcinomas in genetically predisposed individuals.

REFERENCES

1. Amieva Manuel R, El Omar E. Host Bacterial Interactions in *Helicobacter pylori* Infection. *Gastroenterology* 2008; 134: 306-329.
2. Michael Höcker, Peter Hohenberger. *Helicobacter pylori* virulence factors—one part of a big picture. *Lancet* 2003; 362: 1231-33.
3. Osamu Handa, Yuji Naito, Toshikazu Yoshikawa. Cag A protein of *Helicobacter pylori*: A hijacker of epithelial cell signaling. *Biochemical pharmacology* 2007; 73: 1697-1702.
4. Mobley Harry LT, Mendz George L, Hazell Stuart L. *Helicobacter pylori*. Physiology and Genetics. Chapter 31. The *cag* Pathogenicity Island 2001. p. 608.
5. C Audibert, C Burucoa, B Janvier, JL Fauchère. Implication of the Structure of the *Helicobacter pylori cag* Pathogenicity Island in Induction of Interleukin-8 Secretion. *Infection and Immunity* 2000; 1625-1629.
6. Jörg Harcker, James B. Kaper. Pathogenicity Islands and the Evolution of Microbes. *Annual Review of Microbiology* 2000; 54: 641-679.
7. Stefano Censini, Christina Lange, Zhaoying Xiang, Jean E Crabtree, Paolo Ghiara, Mark Borodovsky. *cag*, a pathogenicity island of *Helicobacter pylori*, encodes type I-specific and disease-associated virulence factors. *Proc Natl Acad Sci USA* 1996; 93: 14648-14653.
8. Tummuru MK, Cover TL, Blaser MJ. Cloning and expression of a high-molecular-mass major antigen of *helicobacter pylori*: evidence of linkage to cytotoxin production. *Infection and Immunity* 1994; 61: 1799-1809.
9. Armelle Ménard, Antoine Danchin, Sandrine Dupouy, Francis Mégraud, Philippe Lehours. A Variable Gene in a Conserved Region of the *Helicobacter pylori* Genome: Isotopic Gene Replacement or Rapid Evolution? *Dna Research* 2008; 15: 163-8.
10. NS Akopyants, SW Clifton, DKersulyte, JECrabtree, BEYouree, CA Reece, et al. Analysis of the *cag* pathogenicity island of *Helicobacter pylori*. *Molecular Microbiology* 1998; 28: 37-53.
11. Mohaboob A, Aleem A Khan, Santosh K Tiwari, Niyaz Ahmed, L Venkateswar Rao, CM Habibullah. Association between *cag*-pathogenicity island in *Helicobacter pylori* isolates from peptic ulcer, gastric carcinoma, and non-ulcer dyspepsia subjects with histological changes. *Gastroenterol* 2005; 11: 6815-6822.
12. Tomb, J.F, Owen White, Anthony R. Kerlavage, Rebecca A. Clayton, Granger G. Sutton, et al. The complete genome sequence of the gastric pathogen *Helicobacter pylori*. *Nature*. 1997; 388: 539-547.
13. Stefan Kutter, Renate Buhrdorf, Jürgen Haas, Wulf Schneider-Brachert, Rainer Haas, Wolfgang Fischer. Protein Subassemblies of the *Helicobacter pylori* Cag Type IV Secretion System Revealed by Localization and Interaction Studies. *Journal of Bacteriology* 2008; 190: 2161-2171.
14. Valerie J Busler, Victor J Torres, Mark S McClain, Oscar Tirado, David B Friedman, Timothy L. Cover. Protein-Protein Interactions among *Helicobacter pylori* Cag Proteins. *Journal of Bacteriology* 2006; 188: 4787-4800.
15. Steffen Backert, Matthias Selbach. Role of type IV secretion in *Helicobacter pylori* pathogenesis. *Cellular Microbiology* 2008; 10: 1573-1581.
16. Zhong Qiao, Shao Shi-He, Cui Lei-Lei, Mu Run-Hong, Ju Xiao-Li and Dong Su-Rong. Type IV secretion system in *Helicobacter pylori*: a new insight into pathogenicity. *Clin Med J* 2007; 120: 2138-2142.
17. Leen-Jan Van Doorn, Céu Figueiredo, Ricardo Sanna, Anton Plaisier, Peter Schneeberger, Wink de Boer, et al. Clinical Relevance of the *cagA*, *vacA*, and *iceA* Status of *Helicobacter pylori*. *Gastroenterology* 1998; 115: 58-56.
18. Yoshico Yamaoka, Mototsugu Kato, Masahiro Asaka. Geographic Differences in Gastric Cancer Incidence Can be explained by differences between *Helicobacter pylori* Strains. *Internal Medicine* 2008; 47: 1077-1083.
19. Yuri Churin, Laila Al-Ghoul, Oliver Kepp, Thomas F. Meyer, Walter Birchmeier, and Michael Naumann. *Helicobacter pylori* CagA protein targets the c.Met receptor and enhances the motogenic response. *The Journal of Cell Biology* 2003; 161: 249-255.
20. Masanori Hatakeyama. Oncogenic Mechanisms of the *Helicobacter pylori* CagA protein. *Nature Reviews Cancer* 2004; 4: 688-694.
21. Y Yamaoka, O Ojo, S Fujimoto, S Odenbreit, R Haas, O Gutierrez, H M T El-Zimaity. *Helicobacter pylori* outer membrane proteins and gastroduodenal disease. *Gut* 2006; 55: 775-781.

22. Naomi Ohnishi, Hitomi Yuasa, Shinya Tanaka, Hirofumi Sawa, Motohiro Miura, Atsushi Matsu, et al. Transgenic expression of *Helicobacter pylori* CagA induces gastrointestinal and hematopoietic neoplasms in mouse. *PNAS* 2008; 105: 1003-1008.
23. Chin Ho Lai, Yun-Chieh Chang, Shin-Yi Du, Hung-Jung Wang, Chun-Hsien Kuo, Shih-Hua Fang, Hua et al. Cholesterol Depletion Reduces *Helicobacter pylori* CagA translocation and CagA-Induced Responses in AGS Cells. *Infection and Immunity* 2008; 76: 3293-3303.
24. Hasan Umit, Ahmet Tezel, Sebnem Bukavaz, Gulbin Unsal, Muserref Otkun, Ali Riza Soyulu, et al. The Relationship between Virulence Factors of *Helicobacter pylori* and Severity of Gastritis in Infected Patients. *Dig Dis Sci* 2009; 54: 103-110.
25. Aime T Franco, Elizabeth Johnston, Uma Krishna, Yoshio Yamaoka, Dawn A. Israel, Toni A. Nagy, et al. Regulation of Gastric Carcinogenesis by *Helicobacter pylori* Virulence Factors. *Cancer Res* 2008; 68: 379-387.
26. Nilsson Anna Sillén, Lena Eriksson, Mona-Lisa Strand, Helena Enroth, Staffan Normark, et al. Correlation between *cag* Pathogenicity Island Composition and *Helicobacter pylori*-Associated Gastroduodenal Disease. *Infection and Immunity* 2003; 71: 6573-6581.
27. Jae Sung Ko, Jeong Kee Seo. *Cag* Pathogenicity Island of *Helicobacter pylori* in Korean Children. *Helicobacter* 2002; 7: 232-236.
28. Sarah Keates, Andrew C Keates, Michel Warny, Richard M Peek, Jr., Paul G. Murray, Ciarán P. Differential Activation og mitogen-Activated Protein Kinases in AGS Gastric Epithelial Cells by *cag+* and *cag-* *Helicobacter pylori*. *The Journal of Immunology* 1999; 163: 5552-5559.
29. Takahiko Kudo, Zhannat Z. Nurgalieva, Margaret E. Conner, Sue Crawford, Stefan Odenbreit. Correlation between *Helicobacter pylori* OipA Protein Expression and *oipA* Gene Switch Status. *Journal of Clinical Microbiology* 2004; 42: 2279-2281.
30. Yoshio Yamaoka, Dong H. Kwon, and David Y. Graham. A Mr 34,000 proinflammatory outer membrane protein (*oipA*) of *Helicobacter pylori*. *PNAS* 2000; 97: 7533-7538.
31. Sicheng Wen, Dominique Velin, Christian P. Felley, Likun Du, Pierre Michetti, and Qiang Pan-Hammarström. Expression of *Helicobacter pylori* Virulence Factors and Associated Expression Profiles of Inflammatory Genes in the Human Gastric Mucosa. *Infection and Immunity* 2007; 75: 5118-5126.