**The VacA and The CagA of Helicobacter Pylori: Two Multitasking Proteins of a Multitasking Bacterium**

Santanu Chattopadhyay1*, Asish K. Mukhopadhyay2, G. Balakrish Nair1,3

1Microbiome Research Facility, Rajiv Gandhi Centre for Biotechnology, Thiruvananthapuram, Kerala, India
2National Institute of Cholera and Enteric Diseases, Kolkata, India
3Center for Human Microbial Ecology, Translational Health Science and Technology Institute, Faridabad, India

**Abstract**

*Helicobacter pylori* infects 50% of the population and 10-20% of the infected individuals develop various gastroduodenal illness, which include gastritis, duodenal ulcer, gastric ulcer, distal gastric adenocarcinoma and gastric mucosa-associated lymphoid tissue lymphoma. The bacterium was classified as type I carcinogen by WHO in 1994. Most of the pathogenic potentials of *H. pylori* have been attributed to two virulence factors, namely the vacuolating cytotoxin (*vacA*) and the cytotoxin-associated gene A (*cagA*). For some geographical regions, however, carrying the virulence specific genes and expressing them may not correlate with the *H. pylori* associated clinical manifestations. Moreover, most of the *H. pylori* infections are benign and eradication of the bacterium may increase the probability of having other diseases, like esophageal cancer. The reason for these extreme variations in clinical outcomes in relation to *H. pylori* infection is unknown, which underscores the need for studying the mechanism of developing these diseases by multidisciplinary approach.

**Introduction**

There are some qualities—some incorporate things, 
That have a double life, which thus is made
A type of that twin entity which springs
From matter and light, evinced in solid and shade.

-Edgar Allen Poe

Since the first culture of *Helicobacter pylori* from human gastric biopsy in 1983, a robust association has been shown with this bacterium and the causation of peptic ulcer, gastritis and gastric adenocarcinoma1,2. Instead of ablation of the stomach, the treatment of peptic ulcer diseases now consists of a short course of antibiotics coupled with proton pump inhibitors. *H. pylori* colonization is associated with gastric cancers and eradication of the bacterium has been shown to cure mucosa-associated lymphoid tissue lymphoma3. Later, *H. pylori* infection was also linked with extragastric diseases like cardiovascular disorder, neurodegenerative disorders and rheumatoid arthritis3. Therefore, treatment and eradication of *H. pylori* from all infected individuals have been suggested4. But on the other hand, a considerably good body of literature suggests that the colonization of *H. pylori* could actually be protective against gastro-esophageal reflux disease, Barrett’s esophagus, esophageal adenocarcinoma as well as asthma5,6. Also, we do not understand why ~80% of the *H. pylori* infected individuals remain asymptomatic throughout their lives. Thus, the true personality of this spiral gastric colonizer still appears to be dual or indefinite—lying somewhere between ‘Dr. Jekyll and Mr. Hyde’. Quite reasonably, most of the *H. pylori* research articles have been directed to reveal the ‘Mr. Hyde’ side of this bacterium. However, in spite of tremendous research leading *H. pylori* as the most cited bacterium for the past three decades, the exact mechanisms of developing various clinical manifestations still remained unresolved. This review aims to discuss some of the exciting, but paradoxical facts that are present in literature on two well known virulence factors of *H. pylori*, the vacuolating cytotoxin (*vacA*) and the cytotoxin-associated gene A (*cagA*), in relation to heterogeneous clinical outcomes.

**The vacuolating cytotoxin (VacA) Discovery.**

The first *H. pylori* toxin was discovered even before the bacterium was named *H. pylori* (that time the bacterium was named as *Campylobacter pylori*). In 1988, Leunk and colleagues found that the cell free broth-culture filtrates of 55% of the strains produced cytopathic effects in 7 of 9 cell lines that were tested for the assay7. HeLa and intestine 407 cells were highly responsive; HEp-2, WiDR, 5637, Vero and KATOIII cells were moderately responsive; Y-1 and CHO-K1 cells were unresponsive. The cellular response consisted of large membrane bound intracellular vacuoles as revealed by phase-contrast microscopy and transmission electron microscopy. The cell free substance
that possessed the vacuolating activity was sensitive to protease and was precipitable with ammonium-sulphate, indicating that it was a secreted protein toxin. Several years later, in 1992, the vacuolating cytotoxin (VacA) of *H. pylori* was purified from the cell free broth culture supernatant[9].

The *vacA* gene alleles and clinical outcomes.

All *H. pylori* strains carry a copy of *vacA*, but not all strains secrete the toxin or produce vacuolation in cells[9]. The *vacA* gene can be divided into four major parts: (a) a region near 5' end encoding the signal sequence (s); (b) the intermediate region (i); (c) the mid region (m); and (d) the 3' end encoding a type-V auto-transporter[10]. Three of these regions (s, i and m) show allelic polymorphism among *H. pylori* strains. The s region of *vacA* encodes for either s1 or s2 signal sequence of VacA precursor protein[11]. Only the toxin encoded by s1 allele carrying *vacA* gene produces vacuoles, while the toxin that is encoded by the s2 allele, due to an additional hydrophilic stretch of 12 amino acids, lacks this activity[12,13]. Consistently, the s1 allele, but not the s2 allele, of *vacA* has been shown to be associated with diseases like ulcer and cancer in the West[14]. Likewise, the mid region of the *vacA* gene can be divided into m1 and m2 (11). Any allelic combinations of signal sequences and mid-regions are possible (like s1m1 or s1m2 or s2m1 or s2m2) but s1m1 combinations are rare[11]. The secreted toxin produced from the *vacA*m1 and the *vacA*m2 show differences in cell tropism due to a 148-residue region[15,16]. The VacA encoded by m1 alleles can produce vacuoles in HeLa as well as RK13 cells, while the m2 allele express a protein that can generate vacuole in RK-13 cells, but not in HeLa cells since it fails to bind to HeLa cells[17]. The vacA*s1m1* allele is more prevalent in most geographic regions (except in South East Asia) and is reported to be associated with ulcer and cancer as compared to *vacA*s1m2[18]. Between the signal sequence and the mid-region, the *vacA* contains the intermediate region, which may have three allelic variations— i1, i2 and i3. Recent analyses suggest that the i1 is associated with more vacuolating activity and severe clinical outcome[19,21].

The VacA protein.

The 149 kDa VacA precursor protein has a ~33-residue signal peptide, a 88 kDa toxin and a 50 kDa type V auto-transporter protein[22]. Upon sequential proteolysis at both the N and the C terminal regions, only the 88 kDa middle portion is secreted as toxin with the help of the type V secretion system[10]. Cryo-electron microscopic (cryo-EM) analyses revealed that the secreted VacA may form a water soluble hexameric flower-shaped molecule (also suggested as dodecameric aggregate). However, only the VacA monomer, which is seen in acidic pH, binds to the host cell membrane and gets internalized[23,24]. The secreted toxin is significantly stable at low pH (up to 1.5) and at high temperature (56°C for 30 min)[25]. The 88 kDa VacA has two distinct domains, p33 (33 kDa; residues 1 to 311) and p55 (55 kDa; residues 312 to 821)[26]. These two domains are sometimes proteolytically cleaved and noncovalently associated, but the cleavage is not required for activity[24]. The p33 and the p55 domains can interact with each other[27]. The solved 2.4 Å crystal structure of the VacA p55 domain revealed features of autotransporter passenger domains, like predominance in right-handed parallel beta-helix[28]. Docking the crystal structure of p55 into a cryo-EM map of the VacA oligomer suggests that the p33 domain forms a ring shaped structure in the oligomer and interacts with the N-terminal portion of several p55 domains within the VacA oligomer[29,30].

Biological functions of the VacA.

The VacA is a multifunctional protein, which interferes with several cellular processes including cell death and immune response (Figure 1). Earlier studies showed that the VacA causes the appearance of acidic vacuoles, which are linked to cytoskeletal elements[29,30]. The vacuoles are rich in rab7, a marker for the late endosomes or early lysosomes[30,31]. The VacA is hypothesized to form anion selective channels in these vesicles and hence is classified as ‘pore-forming toxin’[30,31]. Finally, the anion selective channels facilitate chloride ion transport, which results in diffusion of membrane-permeable weak bases, swelling and vacuolation[31]. Several regions of the VacA (like a hydrophobic region near the N-terminal domain, residues 6–27 in p33; a region within the p55, residues 394–422) has been shown to be important, but neither the p33 nor the p55 alone can form vacuoles[10,32]. Presumably, the entire p33 and the N terminal region of the p55 (residues 1–422 or 1–672 according to another report) are required for comprehensive vacuolating activity[24,32].

The VacA is also shown to be localized in mitochondria and due to its pore-forming activity is capable of releasing cytochrome c, a key event that precedes the apoptosis via intrinsic pathway[10]. Consistently, change in mitochondrial electrochemical membrane potential, the induction of proapoptotic proteins like Bax and Bak as well as cleavage of caspase 3 were confirmed[33–35]. Importantly, this cell death is independent of cell vacuolation[34]. On the other hand, the activation of caspase 8 and caspase 9 mediated apoptosis and caspase-independent programmed necrosis promoting action of VacA were also reported[36,37]. Another study suggested that the low-density lipoprotein receptor-related protein-1 (LRP1) can act as VacA receptor and VacA internalization through this receptor, but not through other receptors like RTPπ, RTPπ, and fibronectin, facilitates autophagyl[30]. The VacA is able to induce p38 mitogen-activated protein kinase (MAPK)/activating transcription factor 2 (ATF-2) mediated signal transduction pathway, which is also independent of cell vacuolation[29]. The activation of β-catenin signaling by the VacA depends on phosphorylation of Akt at Ser473 in a PI3K dependent pathway[40]. Finally, the *H. pylori* VacA has strong immunomodulatory actions as it is capable of inhibiting stimulation-induced proliferation of CD4(+) T cells, CD8(+) T cells, and B cells[41,42].

The cytotoxic-associated gene A (*cagA*).

Discovery.

The *cagA* protein was discovered because of its strong immunogenicity. Human serum derived from patients, who were infected with toxigenic (VacA expressing) strains of *H. pylori*, reacted strongly with a bacterial protein as discerned by Western blot analyses and ELISA[43,44]. The protein showed slower migration (120–128kDa) than VacA (88–90kDa) in SDS-PAGE and was considerably more immunoreactive than VacA[44]. It was predicted that the expression of this protein could be strongly associated with the expression of VacA and could be associated with severe clinical outcome[44]. In 1993, the gene that encodes the 120–128kDa immunodominant protein was cloned and expressed by two different groups and since its expression seemed
associated with the expression of the vacuolating cytotoxin, the gene was named cytotoxin-associated gene A (cagA)\[^{45,46}\].

### The cagA gene polymorphisms and clinical outcomes.

Analyses of the flanking region of the cagA revealed that the gene is part of a horizontally acquired pathogenicity island (PAI), which is present in ~50–70% of the Western and ~90% of the Asian H. pylori strains\[^{47}\]. The presence of cagPAI is strongly associated with the presence of vacA1m1 as well as more severe clinical outcome. The cagPAI shows characteristics that are similar to other bacterial PAIs, like different G+C content (~35%) from the rest of the genome (~39%), direct repeats and insertion sequences (eg. IS605)\[^{47}\]. The ~40kb cag-PAI contains genes that encodes a type-IV secretion system and genes that are responsible for inducing IL-8 secretion\[^{47}\].

The cagA is present at the right end of the cagPAI. The cagA has a conserved 3' end region but a highly polymorphic 3' end region. Phylogenetic analyses using 5' end conserved nucleotide sequence shows that the cagA of Western H. pylori strains and the cagA of East Asian H. pylori strains are different\[^{48}\]. However, the differences in sequences are more significant at the 3' ends of the cagA genes of Western and East Asian strains\[^{49,50}\]. The cagA 3' end encodes a stretch of five amino acids motif, EPIYA, and the tyrosine (Y) residue within this motif can become phosphorylated\[^{49}\]. The number of the EPIYA motifs (usually 3-5 for most strains, but 1 and 2 EPIYA motifs were also reported) and the amino acid sequences among the EPIYA motifs vary between the Western and East Asian H. pylori strains\[^{48,49,51}\]. The presence of East Asian specific sequence (ESS) in CagA primary structure is hypothesized to be the reason for occurrence of more gastric cancer cases in Japan due to stronger intracellular response than the Western specific (WSS) CagA sequence\[^{49,50}\]. Between any two EPIYA motifs, presence of several discrete motifs (possibly due to recombination events at the 3' end of the cagA) were also reported\[^{50,52}\]. Based on PCR, the cagA 3' end region that encodes the EPIYA motif and adjacent repeat units can be typed and different nomenclatures were proposed\[^{50,52}\].

### The CagA protein

Purified CagA, upon partial digestion with V8 protease of *Staphylococcus aureus* results in generation of two fragments—a 100-kD N-terminal region and a ~35-kD C-terminal region\[^{54}\]. It is due to the variation in sizes of the C-terminal region, the molecular weight of the CagA varies (120–140kDa). For the N-terminal region of CagA, two crystal structures, one at 3.6Å (for residues 1-884) and another at 3.19Å (for residues 1-876), have been solved\[^{54-55}\]. According to one report, the N terminal region comprised of four domains or conserved sequence (CSS4) that are similar to other bacterial PAIs, like different G+C content\[^{56}\]. The overall N terminal region contains three domains (domain I, residues 24-221; domain II, residues 303-644 with a subdomain composed of residues 370-446; domain III, residues 645-824) and these domains form a square plate-like shape\[^{54}\]. Residues 782-820 within the N-terminal region and residues 998-1038 within the C-terminal region mediate the intramolecular interactions of CagA\[^{54}\]. The EPIYA motifs containing C-terminal region of the CagA is intrinsically disordered and carries features (like PPII conformation) for scaffolding proteins\[^{54,55}\]. A 16-residue motif (FPLXRRXXVXLDKSVG) within this region is involved in CagA multimerization within the host cell but this multimerization may occurs through the interaction with cellular proteins\[^{56,57}\].

### Biological functions of the CagA.

The *H. pylori* CagA is a master regulator that can hijack host signaling through protein-protein interactions by phosphorylation dependent or phosphorylation independent manner (Figure 1). The CagA is injected to the host cell by the cagPAI encoded type IV secretion system with a process that requires host transmembrane protein integrin. Integrins can interact with several proteins that are encoded by cagPAI including the N-terminal region of the CagA\[^{55,58}\]. The N-terminal region can interact with several other intracellular partners, such as ASPP2, RUNX3, TAK1, and TRAF6 and modulates cellular pathways.

However, the protein-protein interactions that are mediated through the CagA C terminal region are most well studied. The CagA, after being localized to the inner surface of the plasma membrane, undergoes tyrosine phosphorylation within the EPIYA motifs by the Src family protein tyrosine kinase\[^{59}\]. Phosphorylated CagA interacts with the SH2-domain-containing protein tyrosine phosphatase (SHP2) leading to deregulation of the *bona fide* oncoprotein\[^{60}\]. This event is necessary and sufficient for hummingbird phenotype and possibly for developing cancer, particularly in East Asia\[^{50,60,61}\]. Transgenic expression of the CagA induces gastric and hematopoietic neoplasms in mouse, but phospho-resistant CagA failed to develop these tumors\[^{62}\]. This phosphorylation dependant interaction occurs through EPIYA-C and EPIYA-D segments, while the EPIYA-A and EPIYA-B segments can be involved in interacting with the C-terminal Src kinase (Csk)\[^{59}\]. CagA-Csk interaction inhibits the Src dependent phosphorylation of CagA by a negative feedback loop\[^{63}\]. The CagA EPIYA motifs can also regulate cellular pathways by interacting with SHP1, Grb2, Grb7, PI-3 kinase, and Ras-GAP1 in a tyrosine phosphorylation-dependent manner\[^{59,64}\]. Among these, Grb2 can also interact to the CagA in phosphorylation-independent manner and block intrinsic apoptosis by a mechanism that involves host elements like CRK, MEK, SRE, and SRF\[^{65,66}\]. The CagA multimerization, which is independent of phosphorylation, is a prerequisite for the interaction of SHP2\[^{57}\]. The CagA disrupts TJs and induces invasiveness of AGS gastric carcinoma cells via Cdx2-dependent targeting of Claudin-2 and associates with the epithelial tight-junction scaffolding protein ZO-1, which leads to disruption of the epithelial apical-junctional complex\[^{67,68}\]. Moreover, a fraction of the internalized CagA goes to mitochondria and induces the production of reactive oxygen species, which leads to a series of events including mitochondrial DNA damage, oxidative stress to the gastric mucosa, nuclear DNA damage and all these events ultimately facilitate carcinogenesis\[^{69}\].

### Perspective and concluding remarks

The VacA and the CagA, both of these virulence factors of *H. pylori*, are champions in multitasking. These multi-domain proteins interact with many intracellular proteins and regulate signaling pathways that participate in crucial decision making, like cell death or cell survival and immune response or immune suppression (Figure 1). Interestingly, the effects of these two toxins on several cellular pathways are apparently counteracting. Antagonistic effect between the CagA the and VacA have
been shown in cellular processes like NFAT pathway, apoptotic activity and VacA traffic to intracellular compartments[70]. Why then the expression of the CagA is highly correlated with the expression of VacA? Numerous studies have highlighted the association of vacA s1m1 and cagA with diseases. This association, however, seem to vary with geographic regions[18]. For example, strains carrying the vacAs1m1 and any variant of the cagA gene or cagPAI were not associated with diseases in India[51,71-73]. However, colonization by multiple H. pylori strains in single gastric mucosa is fairly common and single gastric biopsy based culture do not represent all the strains that colonize in a stomach. Infection with strains with vacAs1m1 and the cagA has been found in individual, who is also infected with strain carrying vacAs2m2 and lacks cagPAI[74]. At present we do not understand the significance of this association among multiple H. pylori strains (Figure 1). But, since H. pylori is a genetically diverse species that is naturally competent and capable of horizontal gene transfer, the presence of multiple H. pylori strains with different genotypes may confer some adaptive advantages, like overcoming immune responses. How the immune surveillance varies with host genetic background and how it affects clinical outcomes in relation to H. pylori colonization? Besides, the VacA and the CagA proteins, although are the most studied, are not the only virulence markers for H. pylori (eg. NapA, OipA, BabA, IceA, DupA)[18]. How the interactions among these virulence markers actually contribute in the progression of gastroduodenal diseases as well extragastric diseases like neurodegenerative disorders? H. pylori can alter the function of tight junctions but can it actually modulate the gut-brain axis[73]? If an individual is infected with H. pylori but does not have ulcer or cancer, should the individual be treated with antibiotics in order to prevent these diseases? Intriguingly, the eradication of H. pylori is also associated with esophageal cancer[8]. Moreover, antibiotic treatment changes the dynamics of the commensal gut flora and may lead to increase in number of the opportunistic pathogens, like Clostridium difficile[3]. Several lytic phages for H. pylori have been reported[56,77]. Is it possible to specifically cure H. pylori using lytic phages without affecting the microbiota? While several arguments suggested that H. pylori is indigenous microbiota of humans, the roles of VacA and CagA in the progression of gastroduodenal illness have been confirmed by numerous studies[76,78]. Therefore, the most interesting question is, how ~80% of the infected population can make a feasible relationship with H. pylori (with all strains?), while others suffer from similar infection? Likewise, why everybody among the uninfected population (~50%) is not suffering from esophageal cancer and asthma?

Figure 1. Cartoon showing the functions of the VacA and the CagA of H. pylori.

In relation to H. pylori infections, we still have more questions than answers (Table 1). Recent literature show the promise of unraveling many of these issues, but our understanding of H. pylori infected gastric mucosa is reasonably inadequate at present. A multidisciplinary interactive approach involving metagenomics, proteomics, immunology and genetics is needed to understand the mechanisms that regulate clinical outcomes in relation to H. pylori infections.

Table 1. Some important questions in relation to H. pylori infection

<table>
<thead>
<tr>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>Why the expression of CagA is highly correlated with the expression of VacA?</td>
</tr>
<tr>
<td>What are the combined roles of H. pylori virulence markers?</td>
</tr>
<tr>
<td>Does having multiple genotypes in same niche help H. pylori to evade immune surveillance?</td>
</tr>
<tr>
<td>How the immune surveillance varies with host genetic background?</td>
</tr>
<tr>
<td>Can H. pylori modulate the gut-brain axis?</td>
</tr>
<tr>
<td>Should every H. pylori infected person be treated with antibiotics?</td>
</tr>
<tr>
<td>Can lytic phages be used for the treatment of H. pylori infections?</td>
</tr>
<tr>
<td>How ~80% of the H. pylori infected individuals tolerate the infection?</td>
</tr>
</tbody>
</table>

Acknowledgement: This work was supported by Rajiv Gandhi Centre for Biotechnology (an autonomous institute, sponsored by Department of Biotechnology, Govt. of India).

Conflict of interest: The authors declare no conflict of interest.

References

3. Malferttheiner, P., Link, A., Selgrad, M. Helicobacter pylori: perspec-


