**Pathogenic mechanisms of the oncoprotein CagA in H. pylori-induced gastric cancer (Review)**

SHUAI-YIN CHEN, RONG-GUANG ZHANG and GUANG-CAI DUAN

Department of Epidemiology and Biostatistics, College of Public Health, Zhengzhou University, Zhengzhou, Henan 450001, P.R. China

Received May 13, 2016; Accepted September 16, 2016

DOI: 10.3892/or.2016.5145

**Abstract.** Infection with *Helicobacter pylori* is the strongest risk factor for the development of chronic gastritis, gastric ulcer and gastric carcinoma. The majority of the *H. pylori*-infected population remains asymptomatic, and only 1% of individuals may progress to gastric cancer. The clinical outcomes caused by *H. pylori* infection are considered to be associated with bacterial virulence, genetic polymorphism of hosts as well as environmental factors. Most *H. pylori* strains possess a cytotoxin-associated gene (cag) pathogenicity island (cagPAI), encoding a 120-140 kDa CagA protein, which is the most important bacterial oncoprotein. CagA is translocated into host cells via T4SS system and affects the expression of signaling proteins in a phosphorylation-dependent and independent manner. Thus, this review summarizes the results of relevant studies, discusses the pathogenesis of CagA-mediated gastric cancer.

**Contents**

1. Introduction
2. Cytotoxin associated gene pathogenicity island (cagPAI) and cytotoxin-associated gene A (*cagA*)
3. Translocation and phosphorylation of CagA protein
4. CagA phosphorylation-dependent or independent effects
5. CagA destroys the host cells via epigenetic modifications

1. Introduction

*Helicobacter pylori* is a spiral-shaped, flagellated, microaerophilic Gram-negative bacillus first described in 1982 by Marshall and Warren. *H. pylori* is thought to colonize the gastric mucosa of >50% of the world’s population, with the higher prevalence in the developing countries (1,2). However, the majority of *H. pylori*-infected population is asymptomatic, but resulting in chronic gastritis. Only 10% of infected individuals occurred symptomatic diseases. Furthermore, experimental and epidemiological studies have indicated that *H. pylori* infection indeed increase the risk of gastric cancer. Based on this evidence, the World Health Organization International Agency for Research on Cancer classified *H. pylori* as class I carcinogen in 1994. The estimated total of infection-attributable cancer is 1.9 million cases every year, which is 17.8% of the global cancer burden. The principal agent is *H. pylori*, ~63.4% of all stomach cancer or 5.5% of the global cancer burden would be attributable to *H. pylori* infection.

The interindividual differences in risk of *H. pylori*-induced gastric diseases involve significant heterogeneity of both host genetics and *H. pylori* strain virulence factors. In the *H. pylori*-associated diseases pathogenic mechanisms, several strain-specific virulence factors were reported, such as *cagA* (cytotoxin-associated gene A), *vacA* (vacuolating cytotoxin A), *hpaA* (*Helicobacter pylori* adhesin A), *babA* (blood group antigen binding adhesin), *dupA* (duodenal ulcer-promoting gene A), *iceA* (induced by contact with epithelium) genes. One of the main virulence factors is CagA, which is associated with higher risk of gastric cancer and peptic ulcer. CagA protein can interact with intercellular proteins and activate signaling pathways through both tyrosine phosphorylation-dependent or independent mechanisms. Here, we review the possible underlying pathogenic mechanisms of the oncoprotein CagA in *H. pylori*-induced gastric diseases.

2. Cytotoxin associated gene pathogenicity island (cagPAI) and cytotoxin-associated gene A (*cagA*)

The cag PAI is an ~40-kb DNA, which likely was acquired by horizontal gene transfer from another strain in the course of evolution. The cag PAI contains ~31 genes including *cagA*, *cagB*, *cagC*, *cagL*, *cagM*, *cagI*, *cagY*, which encode the CagA protein and functional components of a type IV secretion system (T4SS) (3,4). cag PAI is found in >95% East Asian strains, whereas 60% of Western strains isolated are cag PAI-positive (5). In some strains, cagPAI is split into a right segment (cagI) and a left segment (cagII) by an insertion
sequence (IS605). IS605 was associated with gastric cancer that was higher in *H. pylori* isolated from patients with gastric carcinoma than in patients with duodenal ulcer or chronic gastritis (6).

The cagA gene is ≈3,500-5,000 bp located in the beginning region of cag PAI, encoding 120-145 kDa CagA protein. It is a recognized marker for the entire cag locus. The size of the cagA gene and its protein varies in different strains due to structural diversity in its C-terminal region. In Western populations, cagA-positive strain is associated with enhanced induction of gastritis, peptic ulcer, and higher risk of gastric cancer. However, in East Asia cagA gene is not associated with an increased risk of gastric diseases where almost all strains are cagA-positive (7,8). Franco et al constructed the Mongolian gerbils infection model. They indicate that loss of CagA prevents the development of cancer in this model (9). Ohnishi et al provided first direct evidence for the role of CagA as a bacterium-derived oncoprotein by transgenic mice model (10).

3. Translocation and phosphorylation of CagA protein

Translocation of CagA into host epithelial cells is the first step in the processes of CagA-induced diseases. Several different CagA proteins are involved in the translocation of CagA (11). CagL carries a RGD (arginine-glycine-aspartate) motif that is important for binding and interaction with integrin α5β1 receptor on gastric epithelial cells, and triggers CagA delivery into the target cells, as well as downstream signaling to active tyrosine kinases including FAK, Src and EGFR (12,13). CagM, along with CagX and CagT, forms an outer membrane-associated T4SS subcomplex (14). CagX and CagT interact directly, the C-terminal region of CagX is important for CagT interaction, and CagT depends on CagX for its stabilization (15). CagE is one of the energy providing components of CagA translocation. CagE is an inner membrane associated active NTPase and has multiple interacting partners including the inner membrane proteins CagV and CagB (16).

As reported, CagA also facilitates its translocation into host epithelial cells by T4SS-induced externalization of phosphatidylserine from inner leaflet of the plasma membrane. The protein binds to phosphatidylserine via Lys-Xn-Arg-X-Arg (K-Xn-R-X-R) motif present in the central region of CagA. The 2 arginine residues in K-Xn-R-X-R motif are highly conserved among CagA proteins derived from *H. pylori* strains (13,17). It has previously been reported that C-terminal CagA secretion signal and N-terminal CagA domain (D1) are crucial for efficient translocation (18). Collectively, these findings indicate that all components of this type IV secretion system, including the effector protein CagA, are encoded on the cag pathogenicity island.

Once the protein has entered these target cells, CagA localizes to the inner surface of the cellular membrane, once again by the interaction between the K-Xn-R-X-R motif with phosphatidylserine (13). Then parts of CagA proteins undergo tyrosine phosphorylation at the C-terminal Glu-Pro-Ile-Tyr-Ala (EPIYA) motifs by Src family kinases and Abl kinase, while other CagA molecules remain unphosphorylated (19-21).

CagA can be tyrosine phosphorylated at EPIYA motifs, which is found as part of repetitive region in the C-terminal of CagA. Based on the amino acid sequences surrounding each of the EPIYA motifs, four distinct EPIYA motifs (EPIYA-A, -B, -C, -D) have been classified. EPIYA-A and EPIYA-B are present throughout the world, EPIYA-C is found in strains isolated from Western countries. In contrast, East Asian strains carry East Asian CagA, which contains EPIYA-D (22,23) (Fig. 1). Through database searching and in silico analysis, Zhang et al revealed a strong non-random distribution of the EPIYA-B motif polymorphisms (including EPIYT and EPIYA) in Western *H. pylori* isolates, and provide evidence that the EPIYT are significantly less associated with gastric cancer than the EPIYA. CagA B-motif phosphorylation status is essential for its interaction with host PI3-kinase during colonization and that CagA with an EPIYT B-motif had significantly attenuated induction of interleukin-8 and the hummingbird phenotype, had higher affinity with PI3-kinase, and enhanced induction of Akt compared to the EPIYA (24). It was reported that the two SH2 domains from SHP-2 (Src homology 2-containing protein tyrosine phosphatase-2) bind to highly related sequences pY-(S/T/A/V/I)-X-(V/I/L)-X-(W/F). Intriguingly, the consensus motif perfectly matches the SHP-2-binding site of EPIYA-D (pY-A-T-I-D-F). Furthermore, EPIYA-C (pY-A-T-I-D-D) replacement of the pY + 5 position in Western CagA, reduces the binding affinity to SHP-2. This East Asian-specific sequence conferred stronger SHP-2 binding and morphologically transforming activities to Western CagA. So the East Asian CagA or larger numbers of EPIYA-C in Western CagA was associated with atrophic gastritis and increased the risk of gastric cancer. Several recent reports have demonstrated that the ABCCC type can induce the intestinal metaplasia, IL-8, perturbation of Crk adaptor proteins, anti-apoptotic effect and carcinogenic effect more significantly than ABC type (25-28) (Fig. 2).

4. CagA phosphorylation-dependent or independent effects

**Perturb host cell functions by CagA phosphorylation-dependent effects.** Once within the host cells, CagA undergoes tyrosine phosphorylation at EPIYA motifs by Src family kinases (SFKs) and Abl kinase. Along with the report that c-Src only phosphorylated EPIYA-C or EPIYA-D, whereas c-Abl phosphorylated all EPIYA motifs. CagA proteins were phosphorylated on 1 or 2 EPIYA motifs, but never simultaneously or on 3 motifs. Furthermore, none of the phosphorylated EPIYA motifs alone was sufficient for deregulation of cell growth and motility. Western CagA EPIYA-A and EPIYA-C were preferred combination phosphorylation, either across two CagA molecules or simultaneously on one (21,29).

CagA proteins were probably from a dimer in host cell by a 16-amino-acid CagA multimerization (CM) sequence, which is located downstream of the EPIYA-C motif in C-terminal region (30). After phosphorylated CagA (pCagA) from a homodimer through the CM sequence, the pCagA dimer can bind with a single SHP-2 molecule via its two homologous SH2 domains (31). The pCagA-SHP-2 complex triggers the phosphatase activity of SYP-2, which in return causes the dephosphorylation of focal adhesion kinase (FAK) and activation of Ras/MAPK/ERK signaling pathway (32-34). In addition to SYP-2, the phosphorylated EPIYA-A, -B can specifically bind to the C-terminal Src kinase (Csk) and acti-
Csk, and then the inhibitory tyrosine residues of SFKs are phosphorylated by Csk. So Csk is characterized as a negative regulator of SFKs, which results in reduced EPIYA phosphorylation of CagA. Therefore, the CagA-Csk interaction may establish a negative feedback mechanism that prevents damage to the host cells from *H. pylori*. This could be harmful to the long-term colonization of *H. pylori* in stomach (35).

CagA can associate with Crk adaptor proteins (Crk-I, Crk-II, Crk-L) EPIYA-phosphorylation-dependently. CagA-Crk interaction plays a critical role in promoting cell scattering by inducing several downstream signaling pathways, such as Sox1/His-Ras-Raf-MEK and C3G-Rap1/B-Raf-MEK (36). The human *β*-defensins (hβDs) are antimicrobial peptides that are highly active against *H. pylori* during early infection via EGFR-dependent activation of MAP kinase and JAK/STAT signaling pathways. However, during prolonged infection, hβD1 and hβD3 is subsequently downregulated by phosphorylated CagA (37,38) (Fig. 3).

**Disruption of epithelial cells by CagA phosphorylation-independent effects.** CagA also exerts numerous effects within host cells in a tyrosine phosphorylation-independent manner. A CRPIA (conserved repeat responsible for phosphorylation-independent activity) motif, FPLKRRHDVKDDLKVGLV, in the C-terminal region of CagA, which is distinct from the EPIYA motifs used for phosphorylation. The CRPIA motif in non-phosphorylated CagA was involved in interacting with the hepatocyte growth factor scatter factor receptor c-Met, which is involved in invasive growth of tumor cells. CagA binds c-Met and could represent an adaptor protein, which associates with phospholipase Cγ (PLCγ) and activates phosphatidylinositol 3-kinase/Akt signaling. This in turn led to the activation of β-catenin and NF-κB signaling, which promotes proliferation and inflammation (39,40).

CagA can interact with Grb2, which results in the activation of the Ras/MEK/ERK pathway and leads to cell scattering as well as proliferation. This ability of CagA is independent from the tyrosine phosphorylation. However, the EPIYA sequences are indispensable for the Grb2 binding and induction of the cellular responses (41).

Moreover, CagA associates with the epithelial tight-junction scaffolding protein ZO-1 and the transmembrane protein JAM (junctional adhesion molecule), causing an ectopic assembly of tight-junction components at sites of bacterial attachment, and altering the composition and function of the apical-junctional complex (42).

CagA physically interacts with E-cadherin independently of EPIYA motif phosphorylation. The CagA/E-cadherin interaction impairs the E-cadherin/β-catenin complex, causing cytoplasmic and nuclear accumulation of β-catenin (43). Then it leads to nuclear translocation of free β-catenin, where it binds to the transcriptional cofactors of the Wnt pathway and upregulates the transcription of targeted genes such as axin, cyclin and myc (44). *H. pylori* alters the E-cadherin/β-catenin complex, leading to formation of a multiprotein complex composed of CagA, c-Met, E-cadherin, and p120-catenin. This complex abrogates c-Met and p120-catenin tyrosine phosphorylation and suppresses the cell-invasive phenotype induced by *H. pylori* (45). CagA deregulation of β-catenin requires residues 1,009-1,086 and residues 908-1,012 of Western CagA and East Asian CagA, respectively, and is mediated by the CM motif of CagA (46).

CagA also disrupts the tight junction and causes loss of apical-basolateral polarity by interaction with PAR1/MARK kinase, which is a central regulator of cell polarity. Association of CagA inhibits PAR1 kinase activity and prevents atypical protein kinase C (aPKC)-mediated PAR1 phosphorylation. The PAR-aPKC system is the molecular machinery that
converts initial polarity cues in the establishment of complementary membrane domains along the polarity axis. CagA-PAR1 complex dissociates PAR1 from the membrane, collectively causing junctional and polarity defects. The PAR1 also promotes CagA multimerization, which stabilizes the CagA-SHP2 interaction (47-49). This interaction is also dependent on the CM motif of CagA. The CM sequence of CagA isolated from East Asian \textit{H. pylori} binds PAR1b more strongly than that of CagA isolated from Western \textit{H. pylori}. Within Western CagA species, the ability to bind PAR1b is proportional to the number of CM sequences (50).

Lu \textit{et al} found that CagA binds not only PAR1b but also other PAR1 isoforms, with order of strength as follows: PAR1b $>$ PAR1d $=$ PAR1a $>$ PAR1c. They also indicate that malfunctioning of microtubules and myosin II by CagA-mediated PAR1 inhibition cooperates with deregulated SHP-2 in the morphogenetic activity of CagA (51). In MDCK tissue culture model, association of CagA with MARK2 not only causes disruption of apical junctions, but also inhibition of tubulogenesis and cell differentiation (52). Furthermore inhibition of PAR1b kinase activity contributed to an increased hummingbird phenotype. CagA-mediated inhibition of PAR1b and then prevented PAR1b mediated phosphorylation, which inactivates a RhoA-specific GEF, GEF-H1 and thereby strengthens the hummingbird phenotype induced by CagA-stimulated SHP2 (53). Moreover, the interaction of

---

Figure 2. Amino acid arrangements of SHP2, East Asian CagA and Western CagA. The consensus motif perfectly matches the SHP-2-binding site of EPIYA-D ($pY$-A-T-I-D-F). In Western CagA, EPIYA-C ($pY$-A-T-I-D-D) replacement of the $pY$ + 5 position reduces the binding affinity to SHP-2.

Figure 3. The roles of the CagA in pathogenesis of \textit{H. pylori} infection in a phosphorylation-dependent and independent manner.
CagA with PAR1 is associated with reduction of an inhibitor NF-κB, called IkB kinase, which regulates microtubule stability by phosphorylating microtubule-associated proteins (MAPs). Since microtubule destabilization leads to the activation of NF-κB by promoting IkBα degradation, impairment of the microtubule system by CagA-PAR1 interaction may give a cytoskeletal cue that stimulates IkBα degradation (54).

Additionally, CagA was described to affect activity of protein kinase C-related kinase 2 (PRK2), which acts downstream of Rho GTPases and is known to affect cytoskeletal rearrangements and cell polarity (55) (Fig. 3).

5. CagA destroys the host cells via epigenetic modifications

Epigenetics may be defined as the mechanisms that initiate and maintain heritable patterns of gene function and regulation in a heritable manner without affecting the sequence of the genome. Epigenetic modifications include DNA methylation, post-translational modifications of histone proteins, microRNA expression, genomic imprinting and chromatin remodeling (56,57) (Fig. 4).

**DNA methylation.** DNA methylation is an important epigenetic modification involved in the regulation of numerous biological processes. In mammals, DNA methylation mainly occurs in the context of cytosine-phosphate-guanine (CpG) dinucleotides (58). *H. pylori* infection potently induces methylation of CpG islands (CGIs) to various degrees. Methylation levels of specific CGIs seemed to reflect gastric cancer risk in *H. pylori*-negative individuals (OR 95% CI 2.2-32). The promoter CpG islands of FLnc, HAND1, THBD, p4IARC, HRASLS and LOX gene were reportedly altered by *H. pylori* infection (59-62). *In vitro* experiments showed significant CagA aberrant epigenetic silencing of let-7 expression leading to Ras upregulation by histone and DNA methylation (63). A significantly increased risk of RUNX3 methylation (OR, 4.28; 95% CI, 1.19-15.49) was observed with a high consumption of nuts in patients with CagA-positive *H. pylori* infection (64).

**MicroRNAs.** miRNAs are small, non-coding RNAs, which regulate gene expression in a sequence-specific manner. miRNAs have been implicated in the etiology, progression and prognosis of diseases, and many studies have shown that profiles of miRNA expression differ between infection and non-infection with *H. pylori* (65,66). miRNAs are involved in *H. pylori*-related pathology via the regulation of the transcription and expression of various genes, playing an important role in inflammation, cell proliferation, apoptosis and differentiation (67,68).

CagA may be involved in cellular regulation of certain miRNAs in the gastric epithelium. miRNA expression patterns in *H. pylori*-infected gastric mucosa are determined by microarray. The results found that expression levels of let-7 family miRNAs are significantly altered following infection with CagA positive strain (69). CagA enhanced c-myc, DNA methyltransferase 3B (DNMT3B) and enhancer of zeste homologue 2 (EZH2) expression and attenuated miR-26a and miR-101 expression, which resulted in the attenuation of let-7 expression (63). Using mammalian miRNA profile microarrays, miR-1290 and miR-584 expressions are upregulated in CagA-transformed cells. miR-1290 is upregulated in an Erk1/2-dependent manner, and miR-584 is activated by NF-κB. Foxa1 is an important target of miR-584 and miR-1290, which promote the epithelial-mesenchymal transition significantly (70). *In vitro*, CagA inhibited miR-370 expression, which led to overexpression of FoxM1. The upregulated FoxM1 expression altered the expression of p27 (Kip1), and promoted proliferation in gastric cells (71). CagA may function as an initiator in the process of carcinogenesis by upregulating miR-222, which further participates in the progression of cancer by promoting proliferation and inhibiting RECK (reversion-inducing cysteine-rich protein with Kazal motifs) (72). Shortly after

![Figure 4. The diagram of CagA-mediated epigenetic modifications.](image-url)
H. pylori infection, miR-372 and miR-373 synthesis is highly inhibited, leading to the post-transcriptional release of LAT52 expression and thus, to a cell cycle arrest at the G1/S transition. This downregulation of a specific cell cycle-regulating microRNA is dependent on the translocation of CagA into the host cells (73). The integrative analysis and immunohistochemistry staining validation indicated that miR-155 and miR-146b are upregulated in H. pylori-positive gastroduodenal ulcer. Further experiments in gastric epithelial cells revealed that CagA mediated upregulation of miR-155 and miR-146b, which decreases IL6 overexpression (74).

**Histone modifications.** Nucleosomes are represented by DNA wrapped around eight histone proteins, H2A, H2B, H3, and H4. Histones are primary protein components of eukaryotic chromatin and play a role in gene regulation. A histone modification is a covalent post-translational modification (PTM) to histone proteins which include methylation, phosphorylation, acetylation, ubiquitylation, and sumoylation (75,76).

Substantial research has investigated the effects of H. pylori infection on histone modification. Chromatin immunoprecipitation analysis of NCI-N87 and primary gastric cells revealed that H. pylori induce cell cycle control factor p21^{WAF1} overexpression. H. pylori is associated with hyperacetylation of histone H4 which can release HDAC-1 from the p21^{WAF1} promoter (77). cagPAI-dependent decreases of H3 phosphorylation levels at serine 10 (pH3Ser10) and threonine 3 (pH3Thr3) are observed. H. pylori causes a strong decrease of the cell division cycle 25 (CDC25C) phosphatase (78). Liang et al demonstrated that RBP2, a newly identified H3K4 demethylase, can be induced by CagA via PI3K/AKT-Spl pathway depending on AKT phosphorylation. Furthermore, the novel CagA-Pi3K/AKT-Spl-RBP2-Cyclin D1 pathway links chronic inflammation to tumor during GC development (79).

In conclusion, the CagA protein is encoded by cag PAI and delivered into the host cells during the T4SS system. Following translocation, Src and Abl kinases phosphorylate CagA on EPIYA motifs. Phosphorylated CagA can interact with SHP2, Csk, Crk and hJDP, which trigger several intracellular signaling pathways resulting in epithelial cell gene expression. The unphosphorylated CagA directly interacts with certain intracellular proteins such as PAR1b, E-cadherin/β-catenin, c-Met, Grb2 and ZO-1, then disrupts the cell-to-cell junctions and gastric epithelial cell polarity.

**References**


