Long-Awaited Structural Information of the *Helicobacter pylori* CagA Oncoprotein

The CagA protein of *Helicobacter pylori* plays an important role in gastric carcinogenesis. Upon delivery into gastric epithelial cells, CagA localizes to the plasma membrane, where it acts as an oncogenic scaffold/hub. CagA comprises a solid N-terminal region and an intrinsically disordered C-terminal tail that directs versatile protein interactions. X-ray crystallographic analysis revealed that the N-terminal CagA has a completely new structure comprising three discrete domains. Domain I constitutes a mobile CagA N-terminus, while Domain II tethers CagA to the membrane by interacting with phosphatidylserine. Domain III interacts intramolecularly with the disordered C-terminal CagA, which potentiates the oncogenic CagA action.

Chronic infection with *Helicobacter pylori cagA*-positive strains plays a critical role in the development of gastric carcinoma, the second-most common cause of cancer-related deaths worldwide. The *cagA* gene-encoded CagA protein is delivered into host gastric epithelial cells via the *H. pylori* type IV secretion system and is tethered to the inner leaflet of the plasma membrane, where it undergoes tyrosine phosphorylation at the Glu-Pro-Ile-Tyr-Ala (EPIYA) motif that is present in variable numbers in the C-terminal region by Src family kinases. Upon tyrosine phosphorylation, CagA acts as a pathogenic scaffold/hub protein that promiscuously interacts with a number of host proteins, most notably SHP2, an SH2 domain-containing tyrosine phosphatase activating mutation which is associated with a variety of human malignancies, and the polarity-regulation serine/threonine kinase PAR1/MARK. As a consequence, CagA delivery exerts unconstrained mitogenic stimulation while inducing junctional and polarity defects in polarized gastric epithelial cells. Sustained exposure to CagA also activates cell-reprogramming machinery by ectopically inducing stemness-related transcription factors in gastric cells. Such multifaceted actions of *H. pylori* CagA may cooperatively act to predispose host cells towards malignant transformation. Indeed, systemic expression of CagA in transgenic mice gives rise to spontaneous development of gastrointestinal and hematopoietic neoplasms, confirming its oncogenic potential in mammals. But why does a bacterial protein exert this kind of pathogenic action? It is well understood that protein function and structure are intimately interrelated. A research article by Hayashi *et al* [1] reported success for the first time in determining the tertiary structure of the entire CagA protein (1186 amino-acid residues) by combining X-ray crystallography and NMR spectroscopy. The crystal structure of CagA was determined by the SAD method using synchrotron radiation at PF AR-NE3A. CagA has a unique form, consisting of a solid N-terminal region (residues 1-876) and an intrinsically disordered C-terminal tail (residues 877-1186) (Fig. 1). The crystal structure of the N-terminal core predicts a square plate-like shape with approximate dimensions of 110 × 80 × 55 Å^3. The N-terminal CagA contains 23 α-helices and comprises three discrete domains, termed Domains I to III. Domain I, the most N-terminal domain, is composed of 10 α-helices, having a small interacting surface area (374 Å^2) with Domain II but has no interaction with Domain III. Because of this weak interaction, Domain I appears to be quite mobile and flexible. Domain II and Domain III comprise a protease-resistant structural core of CagA that displays an N-shaped dimodular architecture. Domains II and III are connected by a long helix α19. Domain II also contains a large anti-parallel β-sheet, which has an inserted subdomain (residues 370-446). This subdomain, located at the center of the CagA N-terminal region, tightly interacts with the β-sheet, suggesting that it comprises a rigid core of CagA. A cluster of basic residues termed the “basic patch”, which comprises a part of the Domain II surface, attaches to the acidic membrane phospholipid, especially phosphatidylserine, which is concentrated to the inner membrane surface through electrostatic interaction. This makes the CagA C-terminal tail hang freely and thereby enables Src kinases to attach a phosphate group to the EPIYA motifs in the tail. Furthermore, the flexible C-terminal tail loops back onto the solid N-terminal core to form a lariat that strengthens association of the CagA C-terminal tail with host target proteins such as PAR1 and SHP2 (Fig. 1). Elucidation of the tertiary structure significantly improves our understanding of CagA as a bacterial pathogenic scaffold/hub that is critically involved in gastric carcinogenesis. This research, which has for the first time elucidated the structural basis through which CagA promotes gastric carcinogenesis, will be of great value in developing new strategies in targeting CagA for the prevention and/or treatment of gastric cancer.

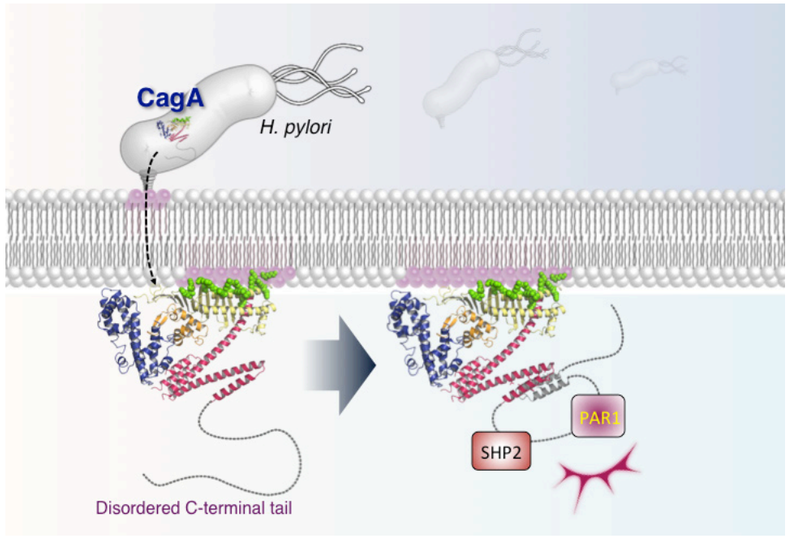

Figure 1: Tertiary structure-dependent regulation of the oncogenic potential of *H. pylori* CagA.
H. pylori CagA is delivered into host gastric epithelial cells via a bacterial micro-syringe termed the type IV secretion system, where it acts as an oncogenic scaffold/hub by perturbing intracellular signaling. The work by Hayashi *et al.* revealed that the CagA protein consists of structured N-terminal and disordered C-terminal regions [1]. Folded N-terminal CagA has a new protein structure with three distinct domains shown in blue, yellow, and red. Domain I forms the N-terminus, while Domain II tethers CagA to the inner plasma membrane via the basic patch (green)-phosphatidylserine (pink) interaction. Interaction with Domain III potentiates interaction of the disordered CagA C-terminal tail with PAR1 and SHP2, thereby enhancing the oncogenic scaffold/hub function CagA.

REFERENCE

[1] T. Hayashi, M. Senda, H. Morohashi, H. Higashi, M. Horio, L. Nagase, D. Sasaya, T. Shimizu, N. Venugopalan, H. Kumeta, N.N. Noda, F. Inagaki, T. Senda and M. Hatakeyama, *Cell Host Microbe* **12**, 20 (2012).

BEAMLINE

AR-NE3A

M. Hatakeyama[1, 2], T. Hayashi[1, 2], M. Senda[2, 3, 5] and T. Senda[4, 5] ([1]The Univ. of Tokyo, [2]JST-CREST, [3]JBIC [4]AIST, [5]KEK-PF&SBRC)