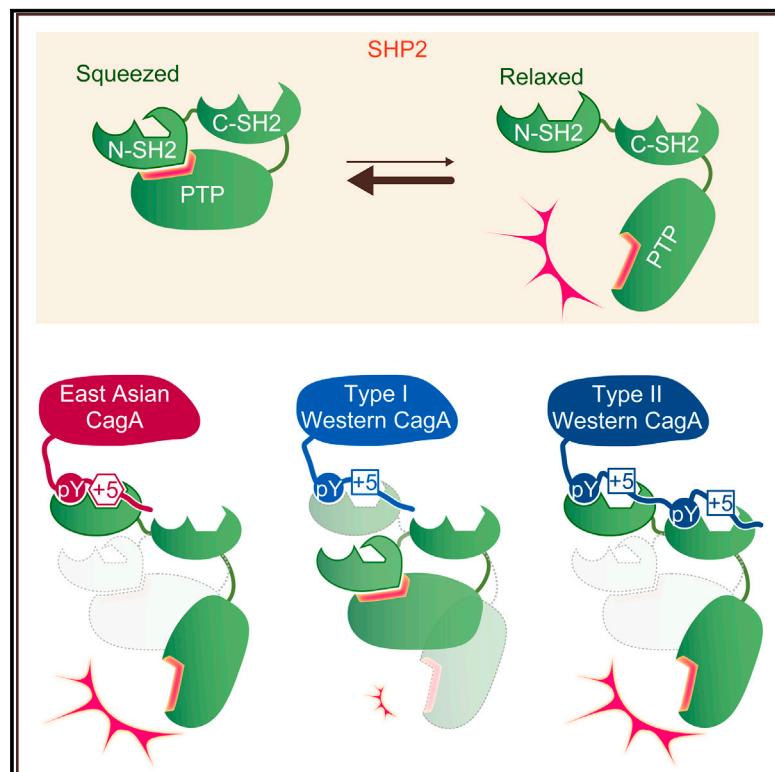


Differential Mechanisms for SHP2 Binding and Activation Are Exploited by Geographically Distinct *Helicobacter pylori* CagA Oncoproteins

Graphical Abstract



Highlights

- East Asian CagA binds to SHP2 100-fold more strongly than type I Western CagA
- East Asian CagA binds to SHP2 via a high-affinity monovalent interaction with N-SH2
- Type II Western CagA binds to SHP2 via divalent interactions with N-SH2 and C-SH2
- Oncogenic CagA binding fixes N-SH2 to the “relaxed” state that activates SHP2

Authors

Takeru Hayashi, Miki Senda, Nobuhiro Suzuki, ..., Kaori Inoue, Toshiya Senda, Masanori Hatakeyama

Correspondence

toshiya.senda@kek.jp (T.S.), mhata@m.u-tokyo.ac.jp (M.H.)

In Brief

Helicobacter pylori CagA binds and deregulates SHP2 to promote gastric carcinogenesis. Hayashi et al. examine the structural and molecular determinants underpinning the CagA-SHP2 interaction and find that totally distinct mechanisms of SHP2 binding are differentially exploited by two major oncogenic CagA isoforms, revealing highly plastic evolution in bacterial virulence acquisition.

Accession Numbers

5X7B

5X94



Differential Mechanisms for SHP2 Binding and Activation Are Exploited by Geographically Distinct *Helicobacter pylori* CagA Oncoproteins

Takeru Hayashi,¹ Miki Senda,² Nobuhiro Suzuki,² Hiroko Nishikawa,^{1,3} Chi Ben,¹ Chao Tang,¹ Lisa Nagase,² Kaori Inoue,¹ Toshiya Senda,^{2,4,*} and Masanori Hatakeyama^{1,3,5,*}

¹Department of Microbiology, Graduate School of Medicine, The University of Tokyo, Tokyo 113-0033, Japan

²Structural Biology Research Center, Photon Factory, Institute of Materials Structure Science, High Energy Accelerator Research Organization (KEK), Tsukuba 305-0801, Japan

³Max Planck–The University of Tokyo Center for Integrative Inflammology, Tokyo 113-0033, Japan

⁴Department of Materials Structure Science, School of High Energy Accelerator Science, The Graduate University of Advanced Studies, Tsukuba 305-0801, Japan

⁵Lead Contact

*Correspondence: toshiya.senda@kek.jp (T.S.), mhata@m.u-tokyo.ac.jp (M.H.)

<http://dx.doi.org/10.1016/j.celrep.2017.08.080>

SUMMARY

Helicobacter pylori East Asian CagA is more closely associated with gastric cancer than Western CagA. Here we show that, upon tyrosine phosphorylation, the East Asian CagA-specific EPIYA-D segment binds to the N-SH2 domain of pro-oncogenic SHP2 phosphatase two orders of magnitude greater than Western CagA-specific EPIYA-C. This high-affinity binding is achieved via cryptic interaction between Phe at the +5 position from phosphotyrosine in EPIYA-D and a hollow on the N-SH2 phosphopeptide-binding floor. Also, duplication of EPIYA-C in Western CagA, which increases gastric cancer risk, enables divalent high-affinity binding with SHP2 via N-SH2 and C-SH2. These strong CagA bindings enforce enzymatic activation of SHP2, which endows cells with neoplastic traits. Mechanistically, N-SH2 in SHP2 is in an equilibrium between stimulatory “relaxed” and inhibitory “squeezed” states, which is fixed upon high-affinity CagA binding to the “relaxed” state that stimulates SHP2. Accordingly, East Asian CagA and Western CagA exploit distinct mechanisms for SHP2 deregulation.

INTRODUCTION

Infection with *Helicobacter pylori* cagA-positive strains plays a major etiologic role in the development of gastric cancer (Blaser et al., 1995; Parsonnet et al., 1997; Wirth et al., 1998; Franco et al., 2005; Rieder et al., 2005). The cagA gene is localized at one end of the cag pathogenicity island (cag PAI), a 40-kb *H. pylori* genomic DNA segment acquired by a horizontal transfer (Censini et al., 1996). Genes located within the cag PAI collectively encode a syringe-like apparatus, termed the type IV secretion system (TFSS), that delivers the cagA-encoded effector protein CagA into gastric epithelial cells (Covacci and Rappuoli, 2000). Upon delivery, CagA undergoes tyrosine phosphorylation at the Glu-

Pro-Ile-Tyr-Ala (EPIYA) motif, which is present in variable numbers in its C-terminal polymorphic region, termed the EPIYA repeat region, by Src family kinases (SFKs) and c-Abl (Mueller et al., 2012). Based on the sequence flanking each of the EPIYA motifs, four distinct EPIYA segments (A, B, C, and D) have been defined (Hatakeyama, 2004; Higashi et al., 2002a). The EPIYA repeat region of CagA from *H. pylori* circulating throughout the world, except East Asia, comprises EPIYA-A, -B, and -C segments in that order. On the other hand, the EPIYA repeat region of the East Asian strain-specific CagA is in an arrangement of EPIYA-A, -B, and -D segments. Thus, EPIYA-C and EPIYA-D are the hallmarks of the world standard CagA, which has been traditionally called Western CagA, and East Asian CagA, respectively. Also, notably, the number of EPIYA-C segments can vary among distinct Western CagA species through tandem duplication, mostly two or three repeats. Approximately 60%–70% of Western CagA species contain a single EPIYA-C segment (type I Western CagA), and the rest of them contain multiple EPIYA-C segments (type II Western CagA) (Nagase et al., 2015; Xia et al., 2009).

Tyrosine-phosphorylated CagA binds to and thereby aberrantly activates SHP2 (Higashi et al., 2002b), a cytoplasmic tyrosine phosphatase composed of an N-terminal tandem pair of SH2 domains, N-SH2 and C-SH2, followed by a C-terminal protein tyrosine phosphatase (PTP) domain (Hof et al., 1998). SHP2 is required for full activation of the Ras-Erk pathway, which conveys a potent mitogenic signal, and is also involved in the regulation of cell morphology and motility (Chan et al., 2008). Catalytic activity of SHP2 is physiologically regulated by the intracellular autoinhibitory N-SH2/PTP interaction and stimulated upon binding of a tyrosine-phosphorylated ligand to the SH2 domains (Barford and Neel, 1998). SHP2 has been recognized as a pro-oncogenic phosphatase, and activating mutations of SHP2 have been found in a variety of human malignancies (Chan et al., 2008; Tartaglia et al., 2006). Transgenic animals systematically expressing CagA developed gastrointestinal and/or hematological malignancies, providing direct evidence for the role of CagA as a bacterium-derived oncoprotein (Neal et al., 2013; Ohnishi et al., 2008). The transgenic mouse study also indicated the importance of CagA-SHP2 interaction in the development of neoplastic tumors *in vivo*. Complex formation of CagA with

SHP2 is mediated via tyrosine-phosphorylated EPIYA-C or EPIYA-D and N-SH2 or C-SH2. Importantly, East Asian CagA, which contains a single EPIYA-D, binds and deregulates SHP2 much more strongly than type I Western CagA (Higashi et al., 2002a). Likewise, type II Western CagA exhibits substantially elevated SHP2-binding activity compared with that of type I Western CagA (Nagase et al., 2015). Clinico-epidemiological studies have shown that gastric cancer is more closely associated with East Asian CagA than Western CagA (Azuma et al., 2004; Vilai-chone et al., 2004; Li et al., 2017). Among Western CagA species, type II Western CagA has been reported to be a distinct risk factor of gastric cancer (Basso et al., 2008; Li et al., 2017). Thus, East Asian CagA and type II Western CagA are more pro-oncogenic than type I Western CagA.

The findings described above point to a three-way association between the EPIYA polymorphism of CagA, the relative strength of CagA-SHP2 interaction, and gastric cancer risk. Here we investigated the molecular structures underpinning CagA-SHP2 interaction. We found that East Asian CagA and Western CagA employ distinct structural mechanisms for high-affinity SHP2 binding to deregulate SHP2 phosphatase and, thereby, endow cells with neoplastic traits.

RESULTS

Binding Strength of Tyrosine-Phosphorylated CagA Peptides to SHP2 SH2 Domains

To gain insights into the mechanism underlying the differential oncogenic potential between Western CagA and East Asian CagA, we quantitatively determined the binding strengths of a tyrosine-phosphorylated EPIYA-C-specific 13-mer peptide (EPIpYA-C, VSPEPpYATIDDL; pY, phosphotyrosine) and EPIYA-D-specific 13-mer peptide (EPIpYA-D, ASPEPpYATIDFD) with the individual N-SH2 or C-SH2 domain of SHP2 by surface plasmon resonance (SPR) analysis (Figures 1A and 1B; Figure S1A). The results of the steady-state affinity analysis revealed that dissociation constant (K_D) values were $22.7 \pm 1.7 \mu\text{M}$ and $60.5 \pm 4.4 \mu\text{M}$ for the binding of EPIpYA-C to N-SH2 and C-SH2, respectively, and $0.185 \pm 0.018 \mu\text{M}$ and $1.43 \pm 0.08 \mu\text{M}$ for the binding of EPIpYA-D to N-SH2 and C-SH2, respectively. It is possible that CagA regions other than the EPIYA-C or EPIYA-D segment may also be involved in determining the binding strength of CagA to SHP2. However, SPR analysis revealed that the binding affinity between the EPIpYA-C peptide and the isolated SHP2 N-SH2 domain ($K_D = 22.7 \pm 1.7 \mu\text{M}$) (Figure 1B) and the binding affinity between full-length type I Western CagA and the SHP2 tandem SH2 domains ($K_D = 24.1 \pm 0.9 \mu\text{M}$) (Nagase et al., 2015) are virtually identical, providing a rationale for using the EPIpYA peptides to quantitatively study the binary interaction between CagA and SHP2. It should also be noted that, in contrast to full-length Western CagA, full-length East Asian CagA undergoes extensive proteolytic degradation in *E. coli*. This technical complication hampered the application of full-length East Asian CagA to an in vitro binding study.

By simply comparing the K_D values of monovalent interactions, SHP2 association with East Asian EPIYA-D was more than 100-fold higher than that with Western EPIYA-C. Because EPIpYA-D/N-SH2 binding was almost one order of magnitude

stronger than EPIpYA-D/C-SH2 binding, East Asian CagA should primarily interact with N-SH2 for complex formation with SHP2. In the type I Western CagA-SHP2 complex, again, N-SH2 may be the primary site for interaction based on the K_D value. However, the difference in binding strengths was only about 3-fold, suggesting that a substantial fraction of type I Western CagA proteins may also bind to SHP2 C-SH2. As was previously reported using tyrosine-phosphorylated CagA (Nagase et al., 2015), SPR analysis confirmed that interaction of a bisphosphorylated EPIpYA-C peptide (EPIpYA-CC; SPEPpYATIDDLGGPFPLKRHDKVDDLSKVGRSVSPEP-pYATIDD), which mimics duplicated EPIYA-C segments, with the tandem SH2 domains of SHP2 was two orders of magnitude stronger than the interaction of the EPIpYA-C peptide with N-SH2 or C-SH2 (Figures 1C and S1B). Because the K_D values of binding of EPIpYA-CC with the C-SH2-dead SH2 domains and the N-SH2-dead SH2 domains were $29.0 \pm 6.9 \mu\text{M}$ and $25.0 \pm 5.0 \mu\text{M}$, respectively, the high-affinity binding observed between the EPIpYA-CC peptide and the tandem SH2 domains of SHP2 ($0.65 \pm 0.10 \mu\text{M}$) was due to intramolecular, but not intermolecular, avidity effects. Thus, East Asian CagA strongly associates with SHP2 through monovalent high-affinity binding with EPIpYA-D ($K_D < 1 \mu\text{M}$), whereas Western CagA/SHP2 association is mediated via monovalent low-affinity binding with EPIpYA-C ($K_D > 20 \mu\text{M}$), which can be robustly potentiated ($K_D < 1 \mu\text{M}$) via multivalent interactions through EPIYA-C duplication, a unique property of type II Western CagA (Figure 1D).

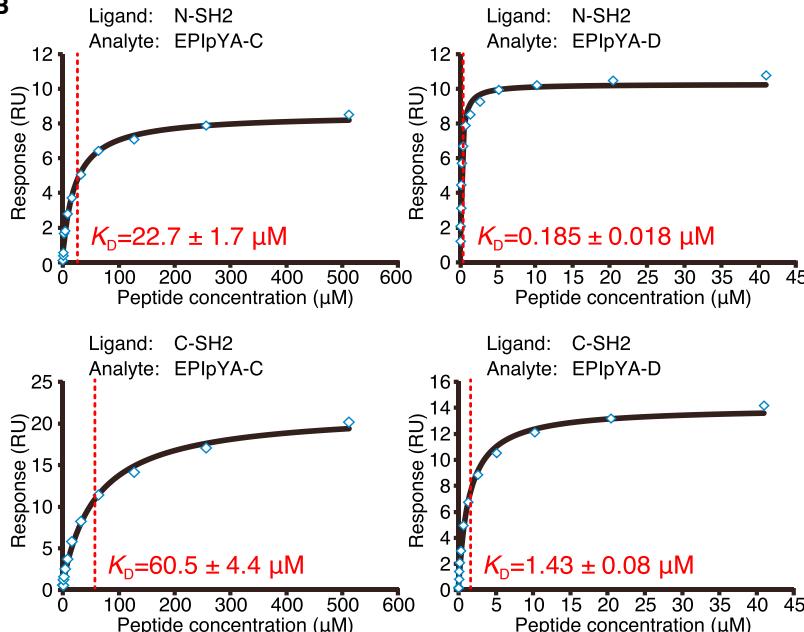
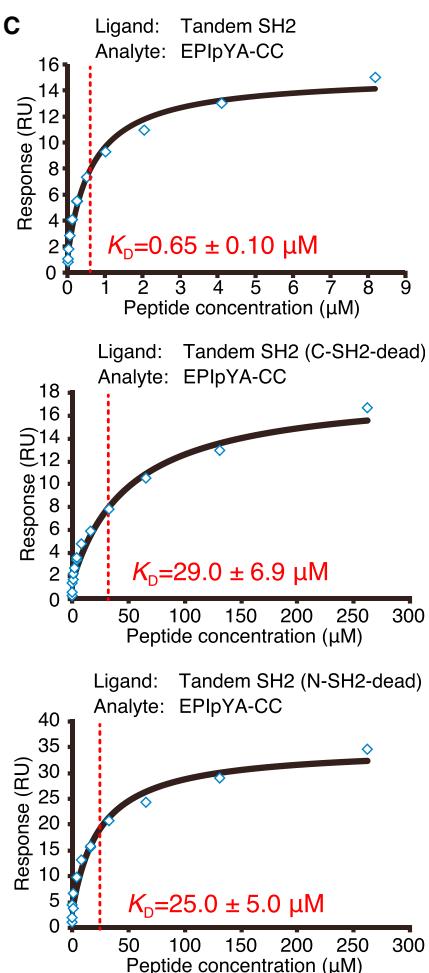
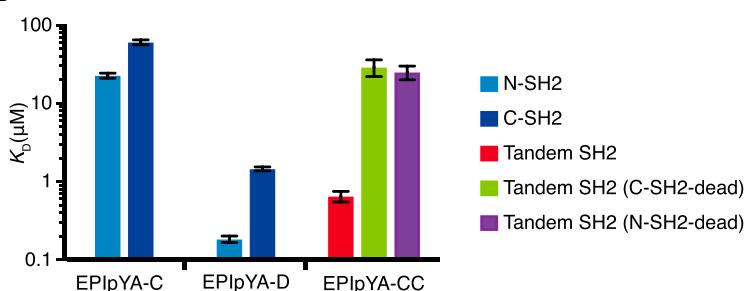
SHP2 Activation by Phosphorylated EPIYA Peptides

Next we investigated the effect of EPIpYA-C or EPIpYA-D binding on the catalytic activity of SHP2 by an in vitro phosphatase assay using *p*-nitrophenylphosphate (*p*NPP) as a substrate. The basal SHP2 phosphatase activity was extremely low with or without the non-phosphorylated peptide ligand (Figure 2A). In striking contrast, the EPIpYA-D peptide efficiently stimulated SHP2 (Figure 2A). To determine the contribution of individual SH2 domains to CagA-mediated SHP2 activation, we measured the phosphatase activity of SHP2 mutants in which one of the two SH2 domains was functionally inactivated (N-SH2-dead or C-SH2-dead) by point mutations (Figure S2) and found that C-SH2-dead SHP2 was enzymatically activated by EPIpYA-D whereas N-SH2-dead SHP2 was not (Figure 2C). Notably, N-SH2-dead SHP2 was hardly activated by EPIpYA-D, even at a concentration of $328 \mu\text{M}$ (Figure 2C). Given that the K_D value of the interaction between the SHP2 C-SH2 domain and EPIpYA-D was $1.43 \pm 0.08 \mu\text{M}$ (Figure 1B), most, if not all, of the C-SH2 domains in the N-SH2-dead SHP2 mutant should have been occupied by the EPIpYA-D peptide at $328 \mu\text{M}$ in our experimental setting. Thus, the monovalent interaction of the EPIpYA-D peptide with C-SH2 was incapable of activating SHP2. These results indicate that interaction with the N-SH2 domain is necessary and sufficient for the enzymatic activation of SHP2 by East Asian CagA.

In contrast to EPIpYA-D, the EPIpYA-C peptide only weakly activated SHP2 when it was added at extremely high concentrations ($>1 \text{ mM}$) (Figure S2C). Given this, we next examined the effect of EPIpYA-CC on SHP2 and found that EPIpYA-CC stimulated SHP2 phosphatase activity in a dose-dependent manner

A

Peptide	Amino acid sequence
EPIpYA-C	VSPEP <p>YATIDDL</p>
EPIpYA-D	ASPEP <p>YATIDFD</p>
EPIpYA-CC	SPEP <p>YATIDDLGGPFPLKRHDKVDDLSKVGRSVSPEP<p>YATIDD</p></p>

B**C****D****Figure 1. The Effect of CagA EPIYA Diversity on Binding to SHP2 SH2 Domains**

(A) The CagA peptides used in this study. Phosphotyrosine (pY) residues are indicated.

(B and C) Surface plasmon resonance (SPR) analysis of the binding reactions between monophosphorylated EPIpYA-C or EPIpYA-D peptide and isolated SHP2 single SH2 domain (B) and between bisphosphorylated EPIpYA-CC peptide and isolated SHP2 tandem SH2 domains (C). The response was plotted against peptide concentration (blue squares), and a curve was fitted (black line). Representative data from three independent experiments are shown.

(D) Summary of K_D values. Error bars represent means \pm SE.

See also Figure S1.

(Figure 2C), although the degree of activation was less than that by the EPIpYA-D peptide. Furthermore, SHP2 activation by EPIpYA-CC required both of the SH2 domains, indicating that the catalytic stimulation of SHP2 was due to stabilized CagA-SHP2 complex formation by avidity effects of multivalent interactions.

Solution Structure Analysis of CagA-Bound SHP2

We next performed small-angle X-ray scattering (SAXS) analysis to investigate the conformational change of SHP2 upon binding with the EPIpYA-D peptide. We obtained scattering profiles of SHP2 (residues 1–527) carrying a Cys459-to-Ala mutation

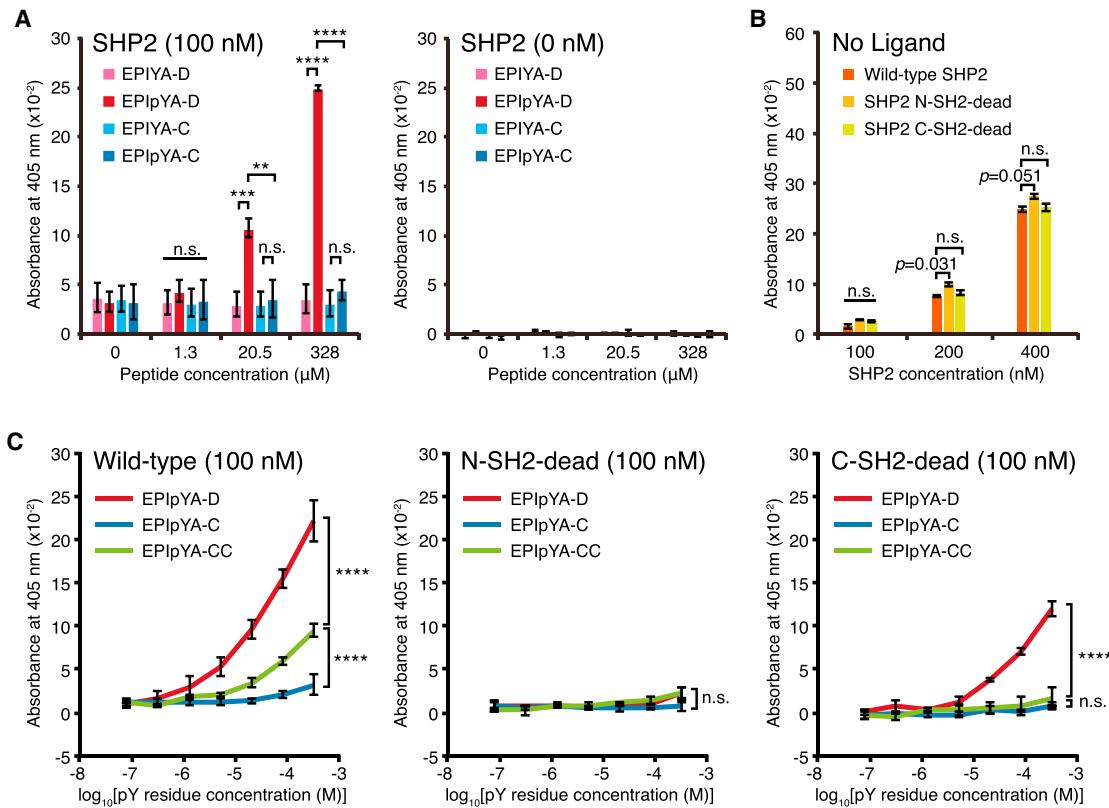
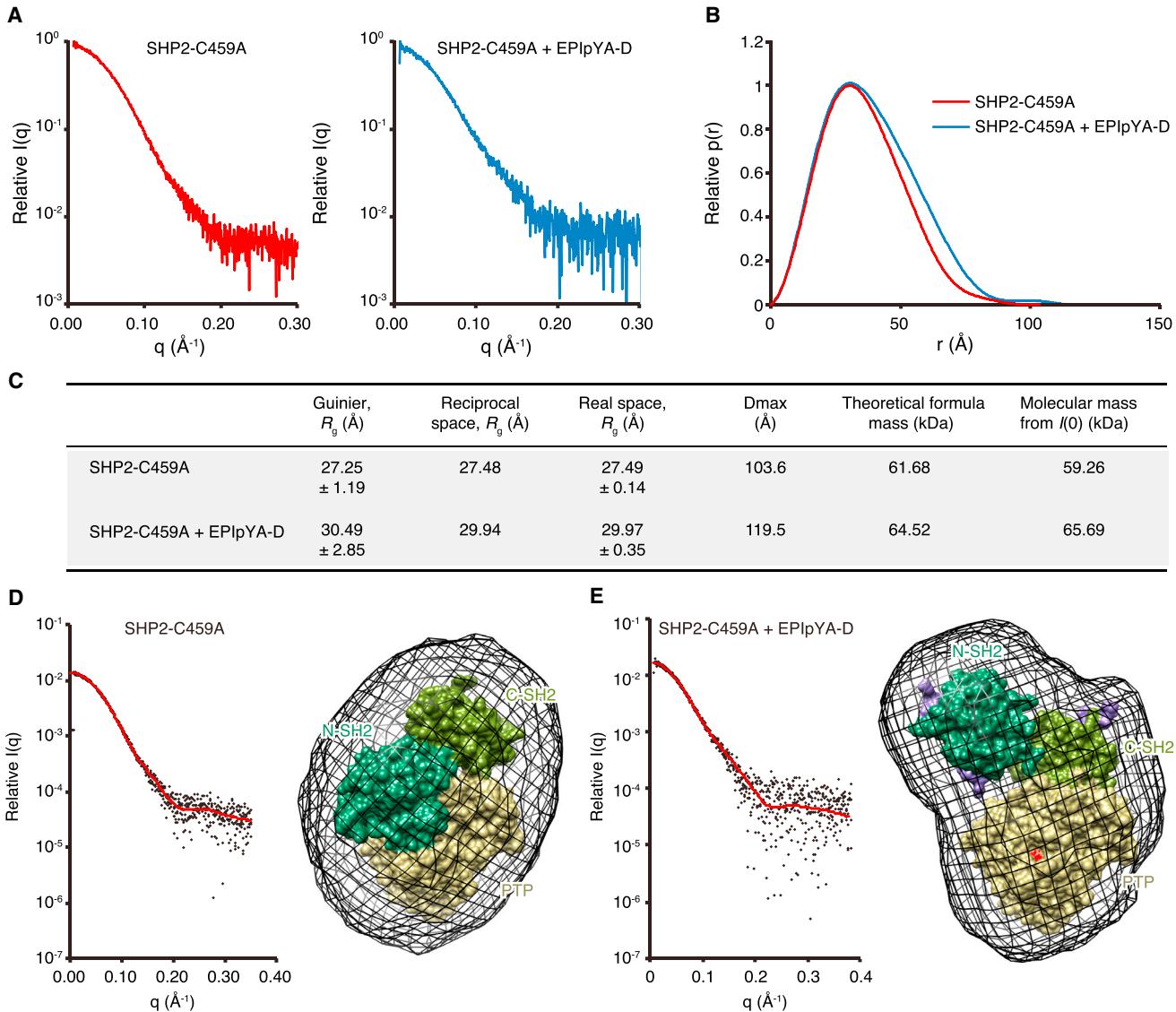


Figure 2. Enzymatic Activation of SHP2 by pY-CagA Peptides

(A) PTP activity of wild-type SHP2 at a concentration of 100 nM in the presence of phosphorylated or non-phosphorylated CagA peptides was analyzed.
(B) PTP activities of wild-type SHP2 and SH2-dead (N-SH2-dead and C-SH2-dead) mutant proteins at the indicated concentrations in the absence of the peptides were analyzed.
(C) PTP activities of SHP2 mutants at a concentration of 100 nM in the presence of the indicated CagA peptides were analyzed.
Paired one-way ANOVA with Tukey's multiple comparisons test; $n = 3$; *** $p < 0.0001$, ** $p < 0.001$, * $p < 0.01$; n.s., not significant. Error bars represent means \pm SD. See also Figure S2.

(SHP2-C459A) in the presence and absence of the EPIpYA-D peptide (Figure 3A; Table S1). The SHP2-C459A mutant has no phosphatase activity and, thus, did not dephosphorylate the phosphopeptide ligand in the reaction mixture. The radii of gyration of the peptide-free form (27.25 ± 1.19 Å from a Guinier plot and 27.49 ± 0.14 Å from the pair-distance distribution $P(r)$ function) and its maximum dimension D_{\max} of 103.6 Å indicated that the peptide-free SHP2 adopts a compact fold (Figures 3B–3D). The scattering profile of SHP2-C459A fit well with the calculated scattering profile for the crystal structure (PDB: 2SHP) (Hof et al., 1998), with a χ value of 1.107 (Figure 3D). In contrast, the peptide-bound form showed a more elongated conformation with greater structural parameters, R_g values of 30.49 ± 2.85 Å (Guinier) and 29.97 ± 0.35 Å ($P(r)$), and D_{\max} of 119.5 Å (Figure 3C). The rigid body model calculated from the scattering profile of the peptide-bound form indicated an open conformation (a χ value of 1.243; Figure 3E), in which the catalytic center of the PTP domain was exposed to the solvent and accessible to its substrate molecules (Figure 3E, right). In this regard, LaRochelle et al. (2016) reported constitutively extended structures of SHP2 mutants in which the inter-domain N-SH2/PTP interaction was disrupted by cancer-associated mutations at the interface

on these domains. Our size exclusion chromatography (SEC)-SAXS analysis demonstrated the actual conformational change of SHP2 from inactive to active states induced by binding of the EPIpYA-D peptide. We also performed SAXS experiments for the N-SH2-dead and C-SH2-dead derivatives of SHP2-C459A to evaluate the significance of peptide binding by each SH2 domain to the structural rearrangement (Figure S3). When the peptide-bound N-SH2-dead form was examined, there was only a slight difference in the SHP2 structural parameters compared with its ligand-free form (Figures S3A and S3B). In contrast, the peptide-bound form of the C-SH2-dead SHP2 mutant exhibited substantial changes in parameters, with R_g values of 28.62 ± 0.09 Å (Guinier) and 28.64 ± 0.02 Å ($P(r)$) and D_{\max} of 114.3 Å (Figure S3C). These results were consistent with the results of the SEC analysis showing that the C-SH2-dead SHP2 mutant in complex with the peptide was eluted much earlier than the N-SH2-dead SHP2 mutant with the peptide, implying extended molecular dimensions, which was followed by the N-SH2-dead SHP2 mutant with the peptide, and, finally, the peptide-free SHP2 mutants appeared (Figures S3D and S3E). Taken together, these results indicated that the monovalent binding of tyrosine-phosphorylated EPIYA-D of East Asian

**Figure 3. Solution Structure of SHP2 Complexed with EPIpYA-D Peptide**

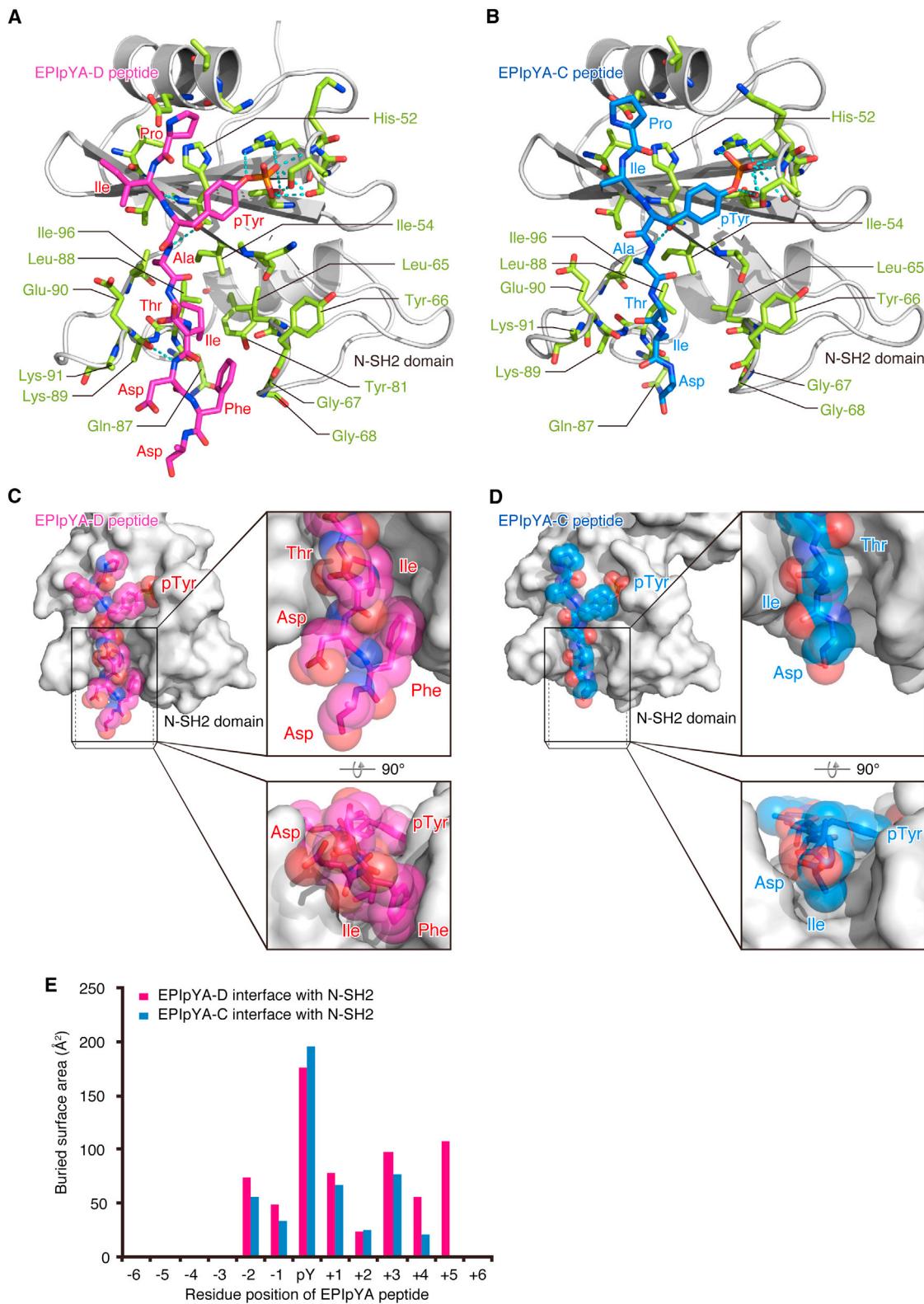
(A) X-ray scattering profiles of ligand-free (left) and EPIpYA-D peptide-bound (right) SHP2-C459A.
 (B) Normalized distance distribution functions of SHP2-C459A (red) and SHP2-C459A with bound EPIpYA-D peptide (blue) were calculated using GNOM.
 (C) Radii of gyration (R_g), maximal end-to-end distance (D_{max}) values, and molecular weights of SHP2-C459A and its complex with EPIpYA-D peptide.
 (D) SAXS data of ligand-free SHP2-C459A were evaluated by using the Crysolv program (left), and the calculated envelope and surface models are shown (right). The N-SH2, C-SH2, and PTP domains are shown in deep green, light green, and pale yellow, respectively.
 (E) SAXS data of EPIpYA-D peptide-bound SHP2-C459A were refined using the Coral program (left), and the calculated envelope and surface models are shown (right). The catalytic center in the PTP domain and EPIpYA-D peptides are shown in red and purple, respectively.
 See also Figure S3.

CagA to the N-SH2 domain is sufficient in causing drastic conformational changes in SHP2 that disrupt the autoinhibitory intramolecular interaction.

Crystal Structure Analysis of the CagA-SHP2 Interaction

To elucidate the structural basis underlying the differential SHP2-binding activity between the geographically distinct CagA species, we determined the crystal structures of the SHP2 tandem SH2 domains complexed with the EPIpYA-C

and EPIpYA-D peptides at 2.45 Å and 2.60 Å, respectively (Figures S4A and S4B; Table S2). The determined crystal structures of the tandem SH2 domains in complex with EPIpYA-C and EPIpYA-D exhibited distinct relative arrangements of N-SH2 and C-SH2 domains (Figures S4A–S4C). Because the relative dispositions of the two SH2 domains showed variety in earlier reported tandem SH2 domains, the distinct relative arrangements of the N- and C-SH2 domains could be primarily due to crystal packing effects.

**Figure 4. Crystal Structure of the Tandem SH2 Domains of SHP2 Complexed with EPIpYA Peptides**

(A and B) Residues of the N-SH2 domain in contact with EPIpYA-D (A) or EPIpYA-C (B) are shown as stick models in light green. Hydrogen bonds are shown as dashed lines.

(legend continued on next page)

The N-SH2 and C-SH2 domains of SHP2 have a conserved SH2 groove, consisting of a central three-stranded anti-parallel β sheet (β B– β D) flanked by two α helices (Figures S4A and S4B) (for secondary structure names of SHP2, see Hof et al., 1998). A phosphopeptide ligand binds to the SH2 groove in an extended conformation, with the phosphotyrosine in a deep pocket on one side of the central sheet and C-terminal residues comprising the EF and BG loops on the other side of the central sheet (Figures 4A and 4B; Figure S2A). The N-SH2 groove contains several hydrophobic residues, including Ile54, Leu65, Tyr66, Leu88, and Ile96. In contrast, there were few polar interactions between N-SH2 and each of the CagA peptides (Figures 4A and 4B). Apart from the phosphotyrosine, only one hydrogen bond and two hydrogen bonds were observed in the interaction of the N-SH2 domain with the EPIpYA-C and EPIpYA-D peptides, respectively; the main chain of His-52 in N-SH2 interacted with Ala at the +1 position from the phosphotyrosine (pY+1) in the EPIpYA-C peptide and, additionally, with Ile at the pY-1 position in EPIpYA-D. At the binding interface between each EPIpYA peptide and N-SH2, the side chain of Ile at pY+3 was oriented inwardly into the SH2 groove (Figure 4C and 4D). This Ile residue was almost buried at the interface of the SH2/EPIpYA peptide interaction, where the buried surface area of the Ile residue in the EPIpYA-D or EPIpYA-C peptide is $\sim 100 \text{ \AA}^2$ and, therefore, is considered to play a substantial role in the interaction of both East Asian CagA and Western CagA with SHP2 (Figure 4E; Figures S4D–S4F).

CagA Structures Underlying High-Affinity SHP2 Interaction

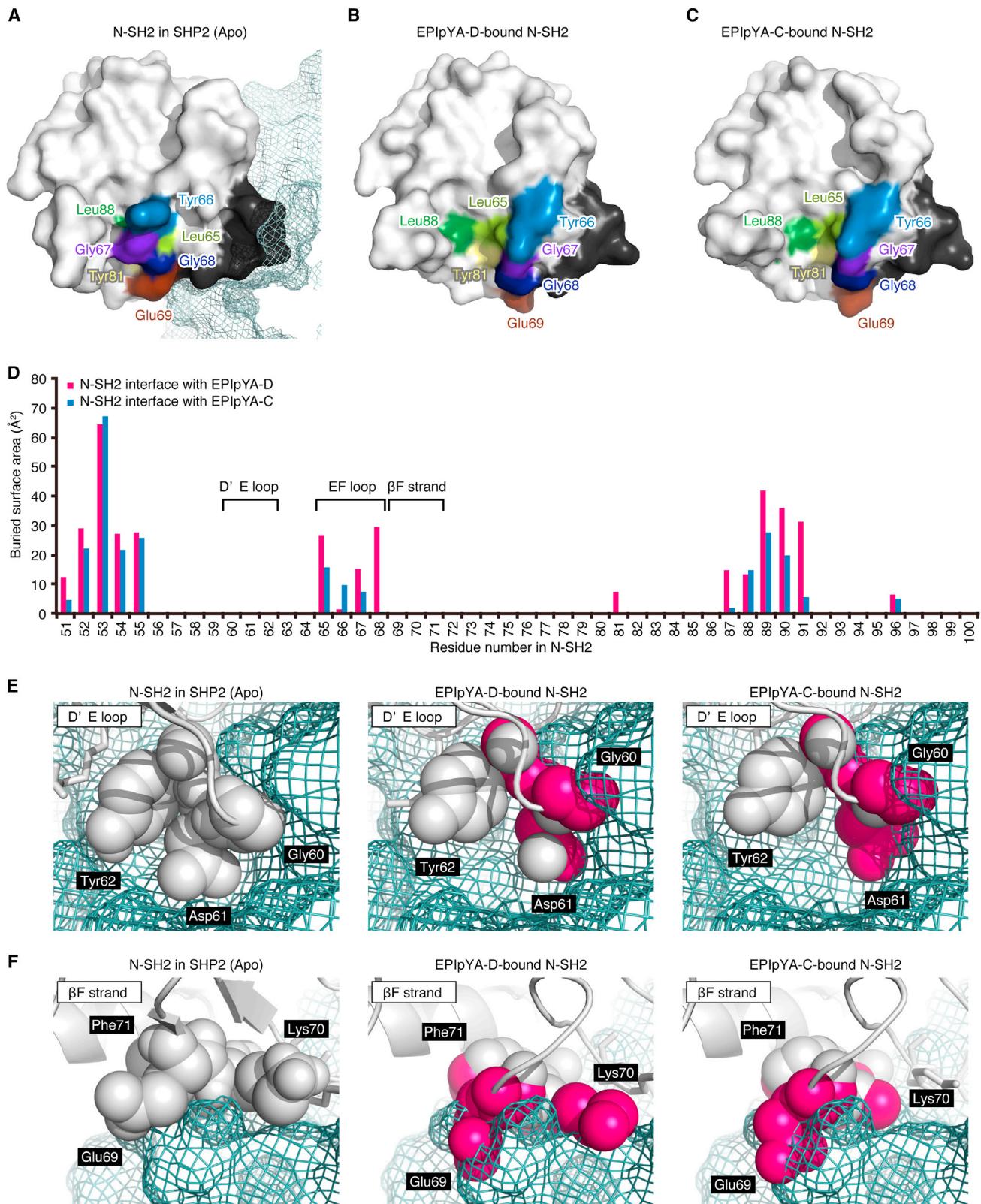
The conformations of the EPIpYA-D- and EPIpYA-C-bound SH2 domains, both N-SH2 and C-SH2, were virtually identical and, when compared with the peptide-unbound SHP2, they were primarily characterized by an extensive expansion of the distal peptide-binding groove because of outward movement of the EF loop and Glu69 in the β F strand (Figures 5A–5C). Intensive analysis of earlier and presently determined crystal structures of the tandem SH2 domains of SHP2 revealed that there are two conformers in N-SH2, relaxed and squeezed, and only the relaxed state can accommodate a tyrosine-phosphorylated peptide ligand. The relaxed state of the N-SH2 cleft created by the outward movement was due to liberation of N-SH2 from the PTP domain because it was exactly the same as the conformation of the isolated N-SH2 domain without peptide binding (Figure S7A). In other words, the relaxed state represents the default structure of the N-SH2 domain per se, and the autoinhibitory intramolecular interaction between the N-SH2 and PTP domains reduces the width of the N-SH2 cleft inwardly (squeezed state). The extended groove floor in the relaxed state of N-SH2 contained a cryptic hollow made of Leu65 (β E4, light green), Tyr81 (α B8, pale yellow), and Leu88 (BG4, green) (Figures 5A–5C), which is exposed upon outward movement of Tyr66 (EF1, blue), Gly67 (EF2, purple), Gly68 (EF3, dark blue),

and Glu69 (β F1, brown) (Figures 5A–5C). Previous studies showed that the pY+5 residue in the EPIYA-C or EPIYA-D segment of CagA plays a pivotal role in determining the binding strength of CagA with SHP2 (Higashi et al., 2002a). Totally consistent with this, the aromatic side chain (phenyl group) of Phe at pY+5 on EPIpYA-D was sticking inwardly to fill the cryptic hollow at the binding interface (Figure 4C; Figure S4D), having significant contact with Leu65, Gly67, Gly68, and Tyr81 (Figure 5D). Of note, the aromatic side chain of the Phe residue at pY+5 was sandwiched by the $\text{C}\gamma$ atom of Ile at pY+3 and the peptide bond connecting Gly67 and Gly68. Hence, there might be a π - π interaction between the side chain of Phe at pY+5 and the peptide bond. In contrast, Asp at pY+5 in the CagA EPIpYA-C peptide did not give significant electron density, and it was thus missing from the crystal structure and, possibly, outwardly projected from the groove (Figure 4D; Figure S4E), having less contact with Leu65/Gly67 and no contact with Gly68/Tyr81 (Figure 5D). As a result, the binding interface of N-SH2 with EPIpYA-D was more extensive (487.6 \AA^2) than that with EPIpYA-C (367.1 \AA^2). The Leu65/Tyr66/Gly67/Gly68 residues in β E and EF loop residues, which were most affected by the EPIpYA-peptide binding, were sandwiched between residues that non-covalently interact with the PTP domain (residues 58–63 and residues 69–77) (Figure S2A), and, indeed, binding with EPIpYA-D or EPIpYA-C led to changes in the surface geometry of the PTP-interacting interface of N-SH2 (Figures 5A–5C, shown in dark gray), causing expanded positioning of the D'E loop (Gly60/Asp61/Tyr62) and the β F strand (Glu69/Lys70/Phe71), as shown in Figures 5E and 5F. The altered conformation causes a steric hindrance at the interface of the N-SH2/PTP inter-domain interaction when N-SH2 in the peptide-bound form was aligned with that in the SHP2 apo structure (Figures 5E and 5F). This may induce catalytic activation of SHP2 by crowding N-SH2 out of the interface of the PTP domain. In contrast to N-SH2, an indentation made of Val181 (β E4, light green), Gly182 (EF1, blue), Gly183 (EF2, purple), Gly184 (EF3, dark blue), Tyr197 (α B8, pale yellow), and Met202 (BG2, green) was already present in ligand-free C-SH2, and EPIpYA-D or EPIpYA-C binding to C-SH2 did not provoke significant conformational changes at the indentation (Figures S5A–S5D). In this case, again, Phe at pY+5 in EPIpYA-D was caught by the C-SH2 indentation whereas Asp at pY+5 in EPIpYA-C was not (Figure S4F), explaining the higher affinity in the interaction of EPIpYA-D with C-SH2 compared with that in the interaction of EPIpYA-C with C-SH2. Phe at pY+5 in EPIpYA-D may therefore act as a barb that stabilizes the complex, whereas barbless EPIpYA-C easily dissociates from SH2. These observations provided a structural basis for the strong monovalent interaction of East Asian CagA-specific EPIYA-D with SHP2 N-SH2. Given that C-SH2 has little contact with N-SH2 and PTP domains in SHP2, the lack of conformational changes in C-SH2 explains why CagA-bound C-SH2 did not catalytically stimulate SHP2. The result also

(C and D) Surface representation for the crystal structure of SHP2 tandem SH2 domains bound to EPIpYA-D (C) or EPIpYA-C (D) peptides with atomic sphere models. Enlarged areas show N-SH2/peptide interfaces.

(E) Buried surface areas of individual residues in the CagA peptides at the interface with the N-SH2 domain.

See also Figure S4.



(legend on next page)

consolidated the special role of N-SH2 as a molecular rheostat that on/off regulates the enzymatic action of SHP2 (Hof et al., 1998).

Differential Mechanisms of SHP2 Activation by Geographic Isoforms of CagA

To consolidate the importance of the Phe residue at pY+5 in EPIYA-D for SHP2 binding and activation of East Asian CagA, we conducted a binding study using mutant EPIYA peptides (Figure 6A). Ala substitution has been widely used to identify residues that are important for protein functions, including protein-protein interaction, because it enables examination of the contribution of an individual amino acid side chain to the functionality of the protein. SPR analysis revealed that the Phe-to-Ala (F/A) substitution at the pY+5 position diminished the N-SH2-binding affinities of the EPIYA-D peptide to the K_D value of $7.14 \pm 0.34 \mu\text{M}$, indicating the importance of the side chain of Phe in high-affinity SHP2 binding (Figure 6B; Figure S6A). Furthermore, the Phe-to-Ala substitution abolished the ability of EPIYA-D to enzymatically activate SHP2 (Figure 6C). To extend these observations to geographically distinct CagA isoforms, we next generated a mutant EPIYA-D peptide, EPIYA-D F/D (pY+5), in which Phe at pY+5 was replaced by Asp, the amino acid corresponding to the pY+5 position of the EPIYA-C segment, and found that the mutant EPIYA-D peptide had a dramatically reduced SHP2 binding affinity ($K_D = 60.1 \pm 5.8 \mu\text{M}$) (Figure 6B; Figure S6A) with a markedly impaired ability in SHP2 activation (Figure 6C). Interestingly, the Phe-to-Asp (F/D) substitution in EPIYA-D led to a greater reduction in SHP2 binding than the Phe-to-Ala substitution, indicating that Asp at pY+5 in EPIYA-C is negatively acting on SHP2 binding by Western CagA. Conversely, the Asp-to-Phe (D/F) substitution at pY+5 in the EPIYA-C peptide that mimics pY+5 of EPIYA-D robustly potentiated SHP2-binding affinity ($K_D = 0.51 \pm 0.04 \mu\text{M}$), which was concomitantly associated with a marked increase in the ability to stimulate SHP2 (Figures 6B and 6C). To test whether the presence of a bulky/aromatic/hydrophobic amino acid such as Phe at the pY+5 position is crucial for high-affinity SHP2 binding by the EPIYA peptide, we also generated a mutant EPIYA-C, EPIYA-C D/W (pY+5), in which Phe at pY+5 was replaced by Trp, another bulky/aromatic/hydrophobic amino acid, and found that the Asp-to-Trp (D/W) substitution also converted EPIYA-C from low affinity to high affinity in terms of SHP2 binding, which was concomitantly associated with greater activation of SHP2 (Figures 6B and 6C). These observations revealed that the pY+5 amino acid in the EPIYA-C or EPIYA-D segment is a key residue determining SHP2-binding affinity of individual CagA,

and the observations indicate the importance of a bulky/aromatic/hydrophobic residue at pY+5 of EPIYA-D, which is buried into the hollow on the N-SH2 surface, for strong SHP2 binding and activation by East Asian CagA.

Binding with EPIYA-C or EPIYA-D gave rise to virtually identical structural changes in N-SH2, including the PTP-interacting interface. Controversy exists regarding why EPIYA-D was capable of efficiently activating SHP2 whereas EPIYA-C was not. Given the two orders of magnitude difference in their binding strengths together with the interaction surface of N-SH2 with EPIYA-C being substantially smaller than that of N-SH2 with EPIYA-D (Figure 1B), we hypothesized that the differential biological activity was due to a difference in stability of the CagA-SHP2 complex. Although it was not possible to calculate the absolute values of kinetic rate constants in the interaction of EPIYA peptides with SH2 domains because of the rapid reaction, the sensorgrams at the dissociation phase presented in Figure S1A showed much faster dissociation of EPIYA-C than of EPIYA-D from N-SH2, indicating that the N-SH2/EPIYA-C complex was far less stable than the N-SH2/EPIYA-D complex. Mutational analysis also revealed that Phe at pY+5 is required for prolonged retention of EPIYA-D on the N-SH2 domain (Figure S6A). A binding competition study further demonstrated that the N-SH2/PTP inter-domain interaction was impeded by addition of the EPIYA-D peptide in a dose-dependent manner, whereas the interaction was not impeded by EPIYA-C (Figure S6B). We concluded from these observations that the CagA-SHP2 complex for which formation was mediated via N-SH2/EPIYA-C interaction was much less stable than the complex for which formation was mediated through the N-SH2/EPIYA-D interaction, which underlies the impaired enzymatic activation of SHP2 by EPIYA-C.

CagA Undergoing High-Affinity SHP2 Binding Endows Cells with Neoplastic Traits

CagA induces a morphological transformation, known as the hummingbird phenotype, in AGS gastric epithelial cells (Segal et al., 1999). The morphological change, characterized by an elongated cell shape that resembles the neoplastic cell trait known as epithelial-to-mesenchymal transition (EMT), requires CagA-mediated SHP2 activation (Bagnoli et al., 2005; Higashi et al., 2002b). We previously reported that type II Western CagA induces the hummingbird phenotype and stimulates cell invasion much more strongly than type I Western CagA, suggesting that the high-affinity CagA-SHP2 interaction ($K_D < 1 \mu\text{M}$) and subsequent SHP2 deregulation are crucial for the pathobiological CagA action that contributes to gastric carcinogenesis

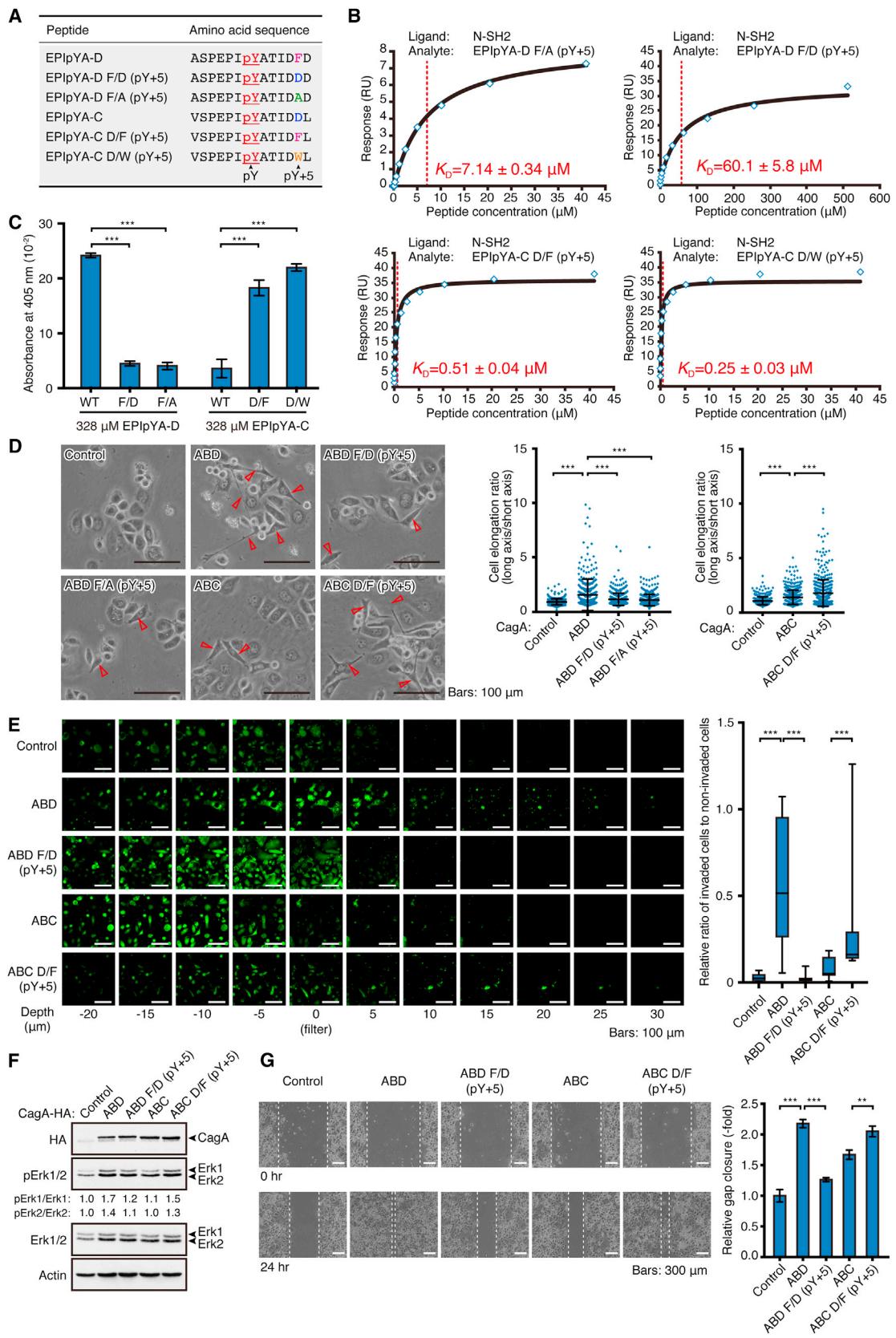
Figure 5. The Molecular Structure Underlying High-Affinity CagA-SHP2 Interaction

(A–C) Surface representation of the N-SH2 domain with a mesh model of the intramolecularly interacting PTP domain isolated from full-length SHP2 (PDB: 2SHP) (A) and N-SH2 domains from tandem SH2 complexed with EPIYA-D (B) and EPIYA-C (C). The bound peptides are hidden to highlight the cryptic hollows made by Leu65, Tyr66, Gly67, Gly68, Glu69, Tyr81, and Leu88. Residues 58–63 and 67–77, the PTP-interacting region of the N-SH2 domain, are shown in black, except Glu69, which is highlighted in brown.

(D) Buried surface areas of individual residues in the N-SH2 domain at the interface with CagA peptides.

(E and F) Conformations of the D'E loop (residues 60–62) (E) and the β F strand (residues 69–71) (F) in the ligand-free N-SH2 domain (left; PDB: 2SHP), EPIYA-D-bound N-SH2 (center), and EPIYA-C-bound N-SH2 (right). The peptide-bound N-SH2 domains were aligned with the ligand-free N-SH2 domain within full-length SHP2 (PDB: 2SHP), where the PTP domain is shown in a mesh surface model. Atoms in N-SH2 showing steric hindrance to the PTP domain are colored pink.

See also Figure S5.



(legend on next page)

(Nagase et al., 2015). To further pursue this idea, we generated an East Asian CagA mutant in which Phe at pY+5 of EPIYA-D was replaced by Asp, the pY+5 residue of EPIYA-C. We also generated a type I Western CagA mutant in which Asp at pY+5 of EPIYA-C was replaced by Phe, the pY+5 residue in EPIYA-D. We then examined the effect of these CagA mutants on the dynamics of invasion of cells into the extracellular matrix, another neoplastic trait induced by activated SHP2 (Nagase et al., 2015). Phe-to-Asp substitution of East Asian CagA, besides reducing the ability to induce the hummingbird phenotype (Figure 6D), impaired stimulation of cell invasion (Figure 6E). Conversely, the Asp-to-Phe substitution potentiated the ability of type I Western CagA to induce the hummingbird phenotype and to stimulate cell invasion into the extracellular matrix (Figures 6D and 6E). Hence, Phe at pY+5 in EPIYA-D is not only critical for high-affinity SHP2 binding but also indispensable for the pathobiological actions of East Asian CagA. In parallel, we assessed the activation status of the Ras-Erk pathway, a potent pro-mitogenic/pro-oncogenic signaling pathway that is enforced by activated SHP2 (Chan et al., 2008), in AGS cells expressing these CagA mutants. Immunoblotting analysis using an anti-phospho-Erk1/2 antibody, which specifically recognizes the activated forms of Erk1/2, showed that the magnitude of Ras-Erk activation by CagA was positively associated with its SHP2 binding strength (Figure 6F). To consolidate the relationship between SHP2 binding strength and the motogenetic activity of CagA, we conducted a wound healing assay, which reflects directional migration that is potentiated by SHP2 (Hartman et al., 2013), by expressing a series of CagA variants with differential SHP2 binding activities in AGS cells. The rate of gap closure in a wound healing assay was positively correlated with the SHP2 binding strength of CagA (Figure 6G). These observations indicated that the CagA structures underlying high-affinity SHP2 binding are also critical for the pathobiological CagA actions that endow cells with neoplastic traits.

From a viewpoint of SHP2, the N-SH2 domain and the tandem SH2 domains are involved in high-affinity binding with East Asian CagA and type II Western CagA, respectively. We thus examined the contribution of SHP2 N-SH2 and/or C-SH2 to induction of the hummingbird phenotype by CagA. To do so, we inhibited endogenous SHP2 in AGS-derived Iacc9 cells by inducible expression

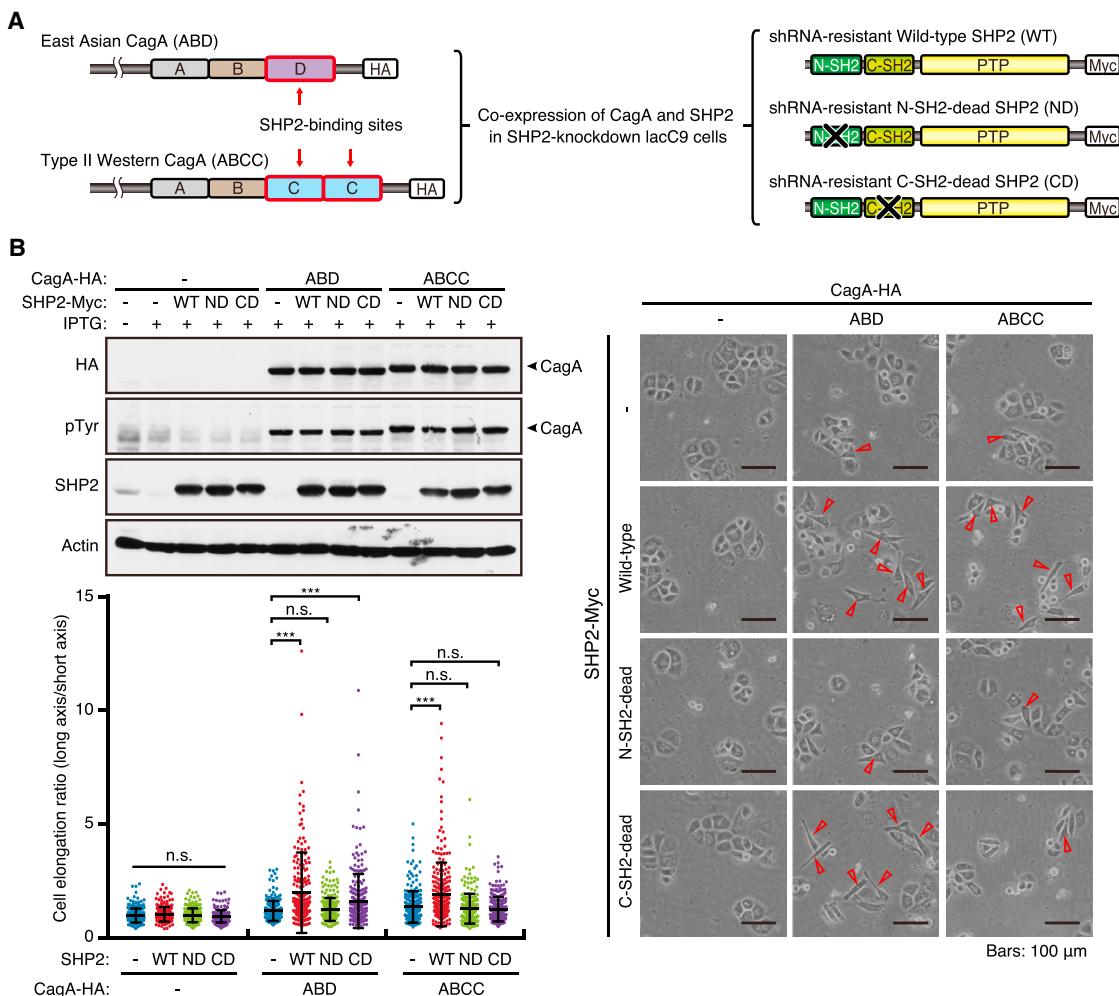
of an SHP2-specific shRNA, as previously reported, and then ectopically co-expressed CagA (East Asian CagA or type II Western CagA) and shRNA-resistant SHP2 (wild-type, N-SH2-dead, or C-SH2-dead) in the SHP2 knockdown cells (Higuchi et al., 2004). As a consequence, East Asian CagA was capable of inducing the hummingbird phenotype when it was co-expressed with shRNA-resistant wild-type SHP2 or C-SH2-dead SHP2 but not when it was co-expressed with N-SH2-dead SHP2 (Figure 7). On the other hand, type II Western CagA was capable of inducing the hummingbird phenotype when co-expressed with shRNA-resistant wild-type SHP2 but not when it was co-expressed with N-SH2-dead SHP2 or C-SH2-dead SHP2. Thus, the SHP2 structures underpinning high-affinity CagA binding are necessary for the pathobiological action of CagA.

DISCUSSION

H. pylori CagA is geographically subclassified into East Asian CagA and Western CagA, which are characterized by the presence of the EPIYA-D and EPIYA-C segments, respectively. Western CagA is further divided into type I Western CagA, which contains a single EPIYA-C segment, and type II Western CagA, which contains duplicated EPIYA-C segments. Epidemiological studies have shown that East Asian CagA and type II Western CagA are distinct risk factors of gastric cancer whereas type I Western CagA is not (Azuma et al., 2004; Vilai-chone et al., 2004; Basso et al., 2008; Li et al., 2017), raising the idea that the structural polymorphism determines the oncogenic potential of CagA. The present study revealed that East Asian CagA and type II Western CagA utilize totally distinct mechanisms for high-affinity binding with the pro-oncogenic SHP2 phosphatase, which leads to catalytic activation of SHP2. More specifically, East Asian CagA binds to SHP2 via monovalent high-affinity interaction of tyrosine-phosphorylated EPIYA-D (EPIpYA-D) with the N-SH2 domain. Whereas type I Western CagA binds to SHP2 via monovalent low-affinity interaction of tyrosine-phosphorylated EPIYA-C (EPIpYA-C) with N-SH2 or C-SH2, type II Western CagA acquires the ability to bind to SHP2 with high affinity through divalent interaction of the duplicated and tyrosine-phosphorylated EPIYA-C segments with N-SH2 and C-SH2.

Figure 6. EPIYA-Dependent Endowment of Neoplastic Traits by CagA

- (A) CagA peptides used in this study. Phosphotyrosines are underlined. F/D, Phe-to-Asp substitution; F/A, Phe-to-Ala substitution; D/F, Asp-to-Phe substitution; D/W, Asp-to-Trp substitution; pY+5, the +5 position from the pY.
- (B) SPR analysis of the binding reactions between mutated CagA peptides and the N-SH2 domain. The response was plotted against peptide concentration (blue squares), and a curve was fitted (black line). Representative data from three independent experiments are shown. Actual sensorgrams are shown in Figure S6A.
- (C) Tyrosine phosphatase activity of SHP2 determined by a pNPP assay in the presence of the EPIpYA-D or EPIpYA-C peptide at a final concentration of 328 μ M. The EPIpYA-D peptide used was wild-type (WT), the Phe-to-Asp mutant at pY+5 (F/D), or the Phe-to-Ala mutant at pY+5 (F/A). The EPIpYA-C peptide used was wild-type (WT), the Asp-to-Phe mutant at pY+5 (D/F), or the Asp-to-Trp mutant at pY+5 (D/W). Paired one-way ANOVA with Tukey's multiple comparisons test, n = 3, ***p < 0.001. Error bars represent means \pm SD.
- (D) Induction of the hummingbird phenotype in AGS cells transfected with the indicated vectors (left) was quantitatively analyzed (right). Brunner-Munzel test with Bonferroni correction, n = 500, ***p < 0.001. Error bars represent means \pm SD. ABD, East Asian CagA; ABC, type I Western CagA.
- (E) Calcein-stained cells migrating into collagen gel matrices were visualized by confocal imaging at each section from the filter membrane (left). Matrix invasion activity was quantified (right). Brunner-Munzel test with Bonferroni correction, n = 10, ***p < 0.001. Whiskers represent minimum to maximum.
- (F) The effects of CagA mutants on Ras-Erk signaling were determined by phosphorylation level of Erk1/2. The protein bands were quantified by using a lumino-image analyzer, and the phospho Erk/total Erk ratios are shown.
- (G) The migratory activity of AGS cells transfected with the indicated vector was evaluated by the degree of gap closure (left) with quantitation (right). Paired one-way ANOVA with Tukey's multiple comparisons test, n = 3, ***p < 0.001, **p < 0.01. Error bars represent means \pm SD.

**Figure 7. The Role of SHP2 SH2 Domains in CagA Action**

(A) Experimental setting of the hummingbird assay performed in (B). ABD or type II Western CagA (ABCC) was co-expressed with shRNA-resistant wild-type SHP2 or SH2-dead mutant in lacC9 cells in which endogenous SHP2 expression had been knocked down by isopropyl-β-D-1-thiogalactopyranoside (IPTG)-induced shRNA targeting SHP2.

(B) Protein expression (top left) and cell morphology (right) were observed 20 hr after transfection. Arrowheads indicate cells with the hummingbird phenotype. Ratios of cell elongation were statistically analyzed (left bottom). Brunner-Munzel test with Bonferroni correction, $n = 250$, *** $p < 0.001$, Error bars represent means \pm SD.

See also Figure S7.

A comparison of K_D values revealed that the SHP2 binding strength of East Asian CagA is two orders of magnitude higher than that of type I Western CagA. We previously showed that an increase in the number of EPIYA-C segments in Western CagA from one (type I Western CagA) to two or more (type II Western CagA) elevates SHP2 binding affinity by approximately 100-fold (Nagase et al., 2015). Thus, the SHP2 binding strength of East Asian CagA is roughly comparable with that of type II Western CagA. The N-SH2 and C-SH2 domains of SHP2 recognize a similar sequence motif, pY-I/V-X-V/I/L/P (Songyang et al., 1994). The presence of an aromatic amino acid residue at pY+5 has also been reported to be involved strong SH2 binding (Lee et al., 1994), making pY-I/V-X-V/I/V/L-X-W/F a preferred recognition motif for the SHP2 SH2 domains.

The EPIYA-D segment perfectly fits the sequence, whereas the EPIYA-C segment mismatches at pY+5, and, indeed, we previously showed in a semiquantitative co-immunoprecipitation experiment that Phe at pY+5 in EPIYA-D confers strong SHP2 binding to East Asian CagA (Higashi et al., 2002a). Phosphopeptide binding to the SHP2 SH2 domain(s) has been considered to perturb the autoinhibitory N-SH2/PTP interaction and, thus, catalytically activate the PTP domain. Consistently, SAXS analyses performed in this study successfully demonstrated that SHP2 complexed with EPIYA-D adopts an extended conformation in solution. The conformational change was due to dissociation of N-SH2 from the juxtaposed PTP domain, making substrates accessible to the PTP catalytic center, as proposed previously (Hof et al., 1998). Crystal structure

analysis revealed the molecular mechanisms underpinning the CagA-SHP2 interaction. The conformation of SHP2 N-SH2 was found to be in a dynamic equilibrium between relaxed and squeezed states in terms of the phosphopeptide-binding cleft (Figure S7A). The relaxed state, characterized by an enlarged binding cleft, is the default conformation of the isolated N-SH2 domain independent of peptide binding. The squeezed state is induced by autoinhibitory interaction of N-SH2 with the PTP domain. The peptide-accessible surface area created by Leu65/Tyr81/Leu88 in N-SH2 is only 6.6 \AA^2 in the squeezed state, making it difficult to accommodate a peptide. However, the same area is dramatically expanded to 67.63 \AA^2 or 64.90 \AA^2 in the relaxed state in complex with EPIpYA-D or EPIpYA-C, respectively. In the relaxed state, the widened cleft floor creates a cryptic hollow that accommodates an aromatic amino acid such as Phe or Trp at the pY+5 position. This secondary interaction should greatly strengthen the binding of East Asian CagA to the SHP2 N-SH2 domain. The results of the present study thus argue against the currently well-accepted idea that binding with a phosphopeptide ligand induces allosteric changes in N-SH2 that, in turn, disable the inhibitory N-SH2/PTP interaction. Instead, our results indicate that binding with a phosphopeptide ligand fixes N-SH2 to the relaxed state from a dynamic equilibrium between the relaxed and squeezed states, which then enforces the dissociation of PTP from N-SH2 that gives rise to catalytic activation of SHP2 (Figure S7B).

The abovementioned mechanism for the regulation of SHP2 phosphatase activity by N-SH2 conformation, however, raises the question as to why EPIpYA-D stimulated SHP2 and EPIpYA-C did not despite the fact that both peptides bind to the relaxed N-SH2 domain. We consider that the difference is due to differential stability of the peptide/N-SH2 complex formed. Strong and stable binding of EPIpYA-D may sustain the N-SH2 conformation in the relaxed state, giving sufficient time for N-SH2 to induce allosteric changes that enforce its dissociation from the PTP domain (Figure S7B). In contrast, interaction of EPIpYA-C with N-SH2 is fragile in the absence of an aromatic residue at pY+5, causing rapid dissociation of type I Western CagA that is insufficient for N-SH2 to dissociate from PTP. Duplication of the EPIYA-C segments in type II Western CagA, however, increases the SHP2-binding activity through avidity effects and thereby stabilizes the binding of EPIpYA-C to N-SH2 sufficiently to activate SHP2. Despite comparable K_D values, however, the EPIpYA-C/N-SH2 interaction in type II Western CagA may be less stable than the EPIpYA-D/N-SH2 interaction in East Asian CagA because it is stabilized by the avidity effects through divalent binding that also include the EPIpYA-C/C-SH2 interaction. This possibility is supported by the observation that the EPIpYA-CC peptide was substantially weaker than the EPIpYA-D peptide in enzymatic activation of SHP2. In the present study, SPR analysis was used to measure the binding affinity between EPIpYA-D and the isolated N-SH2 domain, which should exist as the relaxed state that permits EPIpYA-D binding. On the other hand, an SHP2 phosphatase assay was conducted, using full-length SHP2, in which N-SH2 is in a dynamic equilibrium between the relaxed state, which allows EPIpYA-D binding, and the squeezed state, which forbids EPIpYA-D binding. Accordingly,

enzymatic activation of full-length SHP2 requires a concentration of the EPIpYA-D peptide that is higher than the K_D value of the EPIpYA-D/N-SH2 interaction.

The key CagA structures underpinning high-affinity SHP2 binding ($K_D < 1 \mu\text{M}$) are the Phe residue at pY+5 of the EPIYA-D segment in East Asian CagA and the duplicated EPIYA-C segments in Western CagA. Importantly, those CagA structures are also indispensable for induction of the EMT-like morphological transformation, stimulation of cell invasion/cell migration, and activation of the mitogenic/pro-oncogenic Ras-Erk signaling pathway. Hence, in both the Western and East Asian CagA isoforms, the mechanisms underpinning high-affinity SHP2 binding are crucial for both SHP2 activation and induction of the pathobiological CagA actions that endow cells with neoplastic traits. Because East Asian CagA and type II Western CagA are distinct risk factors of gastric cancer whereas type I Western CagA is not (Azuma et al., 2004; Vilaichone et al., 2004; Basso et al., 2008; Li et al., 2017), the oncogenic contribution of CagA may require SHP2 binding with a submicromolar K_D value that warrants enzymatic activation of SHP2 sufficient for oncogenic CagA action. In other words, the EPIYA polymorphism determines the relative SHP2-binding strength of individual CagA and, thereby, influences the magnitude of SHP2 deregulation that confers neoplastic traits on cells, providing the mechanistic link between structural diversity of CagA and gastric cancer risk.

A number of tyrosine-phosphorylated proteins interact physiologically with the SH2 domains of SHP2. Among them, some possess Phe at pY+5, and others possess two tyrosine residues in proximity, spaced by an approximately 30- to 50-amino acid stretch (Imhof et al., 2006). These SHP2-binding proteins may therefore utilize two distinct mechanisms, monovalent high-affinity interaction with N-SH2 and divalent high-affinity interaction with N-SH2 and C-SH2, for the catalytic activation of SHP2. After many years of interaction between *H. pylori* and human hosts, these mechanisms might have been exploited differentially by the two major geographic CagA isoforms, Western CagA and East Asian CagA, to elicit SHP2 deregulation that contributes to gastric carcinogenesis.

EXPERIMENTAL PROCEDURES

Details regarding plasmid construction, protein purification, cell biological assays, SAXS analysis, and X-ray crystallography are provided in the [Supplemental Experimental Procedures](#).

Synthetic Peptides

Synthetic peptides with >95% purity on high-performance/pressure liquid chromatography (HPLC) were purchased from KNC Laboratories (Japan) and Eurofins Genomics (Japan).

SPR Analysis

Purified SHP2-SH2 (tandem SH2, N-SH2, or C-SH2) protein was immobilized on a Sensor Chip CM5 (GE Healthcare) by the amine-coupling method using a Biacore X100 system (GE Healthcare) according to the manufacturer's protocol. Each CagA peptide at the indicated concentration in HBS-EP⁺ (10 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid [HEPES], pH7.4, 150 mM NaCl, 3 mM EDTA, 0.05% [v/v] surfactant P20) buffer was reacted with the immobilized SH2 protein. Three independent injections were performed, and K_D s were calculated by steady-state affinity analysis using Biacore X100 evaluation software.

In Vitro Phosphatase Assay

Purified SHP2 was incubated with CagA peptide in 100 µL phosphatase assay buffer (30 mM HEPES, pH 7.5, 120 mM NaCl, and 5 mM DTT) containing 12.5 mM pNPP at 30°C for 60 min on a 96-well plate. Dephosphorylation activity was monitored as production of *p*-nitrophenolate because of pNPP hydrolysis by measuring its absorbance at 405 nm.

Statistical Analysis

The results of the phosphatase assays were statistically analyzed by two-tailed Welch's t test and ANOVA with Tukey's multiple comparisons test for single comparison and multiple comparisons of normally-assumed samples, respectively, using GraphPad Prism 7 (GraphPad). Non-normally distributed samples were analyzed by Brunner-Munzel (BM) test using the R package (R Project) with Bonferroni correction for multiple comparisons. The Bonferroni-corrected p values were obtained by multiplying the BM-calculated p values by the number of comparisons.

ACCESSION NUMBERS

The accession numbers for the atomic coordinates and structure factors of SHP2 tandem SH2 domains complexed with EPIYA-C and EPIYA-D peptides reported in this paper are PDB: 5X7B and 5X94.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, seven figures, and two tables and can be found with this article online at <http://dx.doi.org/10.1016/j.celrep.2017.08.080>.

AUTHOR CONTRIBUTIONS

T.H., H.N., C.B., C.T., and K.I. performed biochemical and cell biological studies. T.H., M.S., and T.S. performed X-ray crystallography. N.S. and L.N. performed SAXS analyses. T.H., N.S., T.S., and M.H. analyzed the data. T.H., T.S., and M.H. designed the study and wrote the paper.

ACKNOWLEDGMENTS

We thank Drs. N. Shimizu, S. Saito, and N. Kuwabara (KEK/SBRC) for SAXS experiments. This work was supported by CREST (120200000396); Grants-in-Aid for Young Scientists 16K19121 and 15K18495, and Grants-in-Aid for Scientific Research on Innovative Area 16H06373 and 16K15273 from MEXT, Japan; and by the Project for Cancer Research and Therapeutic Evolution (P-CREATE) (160200000291) and PDIS and Basis for Supporting Innovative Drug Discovery and Life Science Research (BINDS) from AMED, Japan.

Received: March 15, 2017

Revised: June 26, 2017

Accepted: August 23, 2017

Published: September 19, 2017

REFERENCES

- Azuma, T., Yamazaki, S., Yamakawa, A., Ohtani, M., Muramatsu, A., Suto, H., Ito, Y., Dojo, M., Yamazaki, Y., Kuriyama, M., et al. (2004). Association between diversity in the Src homology 2 domain-containing tyrosine phosphatase binding site of *Helicobacter pylori* CagA protein and gastric atrophy and cancer. *J. Infect. Dis.* 189, 820–827.
- Bagnoli, F., Buti, L., Tompkins, L., Covacci, A., and Amieva, M.R. (2005). *Helicobacter pylori* CagA induces a transition from polarized to invasive phenotypes in MDCK cells. *Proc. Natl. Acad. Sci. USA* 102, 16339–16344.
- Barford, D., and Neel, B.G. (1998). Revealing mechanisms for SH2 domain mediated regulation of the protein tyrosine phosphatase SHP-2. *Structure* 6, 249–254.
- Basso, D., Zambon, C.F., Letley, D.P., Stranges, A., Marchet, A., Rhead, J.L., Schiavon, S., Guariso, G., Ceroti, M., Nitti, D., et al. (2008). Clinical relevance of *Helicobacter pylori* cagA and vacA gene polymorphisms. *Gastroenterology* 135, 91–99.
- Blaser, M.J., Perez-Perez, G.I., Kleanthous, H., Cover, T.L., Peek, R.M., Chyou, P.H., Stemmermann, G.N., and Nomura, A. (1995). Infection with *Helicobacter pylori* strains possessing cagA is associated with an increased risk of developing adenocarcinoma of the stomach. *Cancer Res.* 55, 2111–2115.
- Censini, S., Lange, C., Xiang, Z., Crabtree, J.E., Ghia, P., Borodovsky, M., Rappuoli, R., and Covacci, A. (1996). cag, a pathogenicity island of *Helicobacter pylori*, encodes type I-specific and disease-associated virulence factors. *Proc. Natl. Acad. Sci. USA* 93, 14648–14653.
- Chan, G., Kalaitzidis, D., and Neel, B.G. (2008). The tyrosine phosphatase Shp2 (PTPN11) in cancer. *Cancer Metastasis Rev.* 27, 179–192.
- Covacci, A., and Rappuoli, R. (2000). Tyrosine-phosphorylated bacterial proteins: Trojan horses for the host cell. *J. Exp. Med.* 191, 587–592.
- Franco, A.T., Israel, D.A., Washington, M.K., Krishna, U., Fox, J.G., Rogers, A.B., Neish, A.S., Collier-Hyams, L., Perez-Perez, G.I., Hatakeyama, M., et al. (2005). Activation of β-catenin by carcinogenic *Helicobacter pylori*. *Proc. Natl. Acad. Sci. USA* 102, 10646–10651.
- Hartman, Z.R., Schaller, M.D., and Agazie, Y.M. (2013). The tyrosine phosphatase SHP2 regulates focal adhesion kinase to promote EGF-induced lamellipodia persistence and cell migration. *Mol. Cancer Res.* 11, 651–664.
- Hatakeyama, M. (2004). Oncogenic mechanisms of the *Helicobacter pylori* CagA protein. *Nat. Rev. Cancer* 4, 688–694.
- Higashi, H., Tsutsumi, R., Fujita, A., Yamazaki, S., Asaka, M., Azuma, T., and Hatakeyama, M. (2002a). Biological activity of the *Helicobacter pylori* virulence factor CagA is determined by variation in the tyrosine phosphorylation sites. *Proc. Natl. Acad. Sci. USA* 99, 14428–14433.
- Higashi, H., Tsutsumi, R., Muto, S., Sugiyama, T., Azuma, T., Asaka, M., and Hatakeyama, M. (2002b). SHP-2 tyrosine phosphatase as an intracellular target of *Helicobacter pylori* CagA protein. *Science* 295, 683–686.
- Higuchi, M., Tsutsumi, R., Higashi, H., and Hatakeyama, M. (2004). Conditional gene silencing utilizing the lac repressor reveals a role of SHP-2 in cagA-positive *Helicobacter pylori* pathogenicity. *Cancer Sci.* 95, 442–447.
- Hof, P., Pluskey, S., Dhe-Paganon, S., Eck, M.J., and Shoelson, S.E. (1998). Crystal structure of the tyrosine phosphatase SHP-2. *Cell* 92, 441–450.
- Imhof, D., Wavreille, A.S., May, A., Zacharias, M., Tridandapani, S., and Pei, D. (2006). Sequence specificity of SHP-1 and SHP-2 Src homology 2 domains. Critical roles of residues beyond the pY+3 position. *J. Biol. Chem.* 281, 20271–20282.
- LaRochelle, J.R., Fodor, M., Xu, X., Durzynska, I., Fan, L., Stams, T., Chan, H.M., LaMarche, M.J., Chopra, R., Wang, P., et al. (2016). Structural and functional consequences of three cancer-associated mutations of the oncogenic phosphatase SHP2. *Biochemistry* 55, 2269–2277.
- Lee, C.-H., Kominos, D., Jacques, S., Margolis, B., Schlessinger, J., Shoelson, S.E., and Kuriyan, J. (1994). Crystal structures of peptide complexes of the amino-terminal SH2 domain of the Syp tyrosine phosphatase. *Structure* 2, 423–438.
- Li, Q., Liu, J., Gong, Y., and Yuan, Y. (2017). Association of CagA EPIYA-D or EPIYA-C phosphorylation sites with peptic ulcer and gastric cancer risks: A meta-analysis. *Medicine (Baltimore)* 96, e6620.
- Mueller, D., Tegtmeier, N., Brandt, S., Yamaoka, Y., De Poire, E., Sgouras, D., Wessler, S., Torres, J., Smolka, A., and Backert, S. (2012). c-Src and c-Abl kinases control hierachic phosphorylation and function of the CagA effector protein in Western and East Asian *Helicobacter pylori* strains. *J. Clin. Invest.* 122, 1553–1566.
- Nagase, L., Hayashi, T., Senda, T., and Hatakeyama, M. (2015). Dramatic increase in SHP2 binding activity of *Helicobacter pylori* Western CagA by EPIYA-C duplication: its implications in gastric carcinogenesis. *Sci. Rep.* 5, 15749.

- Neal, J.T., Peterson, T.S., Kent, M.L., and Guillemin, K. (2013). *H. pylori* virulence factor CagA increases intestinal cell proliferation by Wnt pathway activation in a transgenic zebrafish model. *Dis. Model. Mech.* 6, 802–810.
- Ohnishi, N., Yuasa, H., Tanaka, S., Sawa, H., Miura, M., Matsui, A., Higashi, H., Musashi, M., Iwabuchi, K., Suzuki, M., et al. (2008). Transgenic expression of *Helicobacter pylori* CagA induces gastrointestinal and hematopoietic neoplasms in mouse. *Proc. Natl. Acad. Sci. USA* 105, 1003–1008.
- Parsonnet, J., Friedman, G.D., Orentreich, N., and Vogelman, H. (1997). Risk for gastric cancer in people with CagA positive or CagA negative *Helicobacter pylori* infection. *Gut* 40, 297–301.
- Rieder, G., Merchant, J.L., and Haas, R. (2005). *Helicobacter pylori* cag-type IV secretion system facilitates corpus colonization to induce precancerous conditions in Mongolian gerbils. *Gastroenterology* 128, 1229–1242.
- Segal, E.D., Cha, J., Lo, J., Falkow, S., and Tompkins, L.S. (1999). Altered states: involvement of phosphorylated CagA in the induction of host cellular growth changes by *Helicobacter pylori*. *Proc. Natl. Acad. Sci. USA* 96, 14559–14564.
- Songyang, Z., Shoelson, S.E., McGlade, J., Olivier, P., Pawson, T., Bustelo, X.R., Barbacid, M., Sabe, H., Hanafusa, H., Yi, T., et al. (1994). Specific motifs recognized by the SH2 domains of Csk, 3BP2, fps/fes, GRB-2, HCP, SHC, Syk, and Vav. *Mol. Cell. Biol.* 14, 2777–2785.
- Tartaglia, M., Martinelli, S., Stella, L., Bocchinfuso, G., Flex, E., Cordeddu, V., Zampino, G., Burgt, I.v., Palleschi, A., Petrucci, T.C., et al. (2006). Diversity and functional consequences of germline and somatic *PTPN11* mutations in human disease. *Am. J. Hum. Genet.* 78, 279–290.
- Vilaichone, R.K., Mahachai, V., Turwasorn, S., Wu, J.Y., Graham, D.Y., and Yamaoka, Y. (2004). Molecular epidemiology and outcome of *Helicobacter pylori* infection in Thailand: a cultural cross roads. *Helicobacter* 9, 453–459.
- Wirth, H.P., Beins, M.H., Yang, M., Tham, K.T., and Blaser, M.J. (1998). Experimental infection of Mongolian gerbils with wild-type and mutant *Helicobacter pylori* strains. *Infect. Immun.* 66, 4856–4866.
- Xia, Y., Yamaoka, Y., Zhu, Q., Matha, I., and Gao, X. (2009). A comprehensive sequence and disease correlation analyses for the C-terminal region of CagA protein of *Helicobacter pylori*. *PLoS ONE* 4, e7736.