

Supporting Information

Nešić et al. 10.1073/pnas.1320631111

SI Results

Hydrophobic Patch. The patch in Cytotoxin associated gene A (CagA) is composed of residues from disparate portions in the protein sequence, including Phe26, Ile105, Val107, Thr111, Phe114, Ile175, Pro207, Trp212, Ile215, Phe219, and Phe221. The residues in Apoptosis-stimulating Protein of p53-2 (ASPP2) that bind in this patch are as follows: the aliphatic portion of Lys751, Tyr754, Thr758, Ala761 Met762, and Ile765 (Fig. 2B).

Hydrogen Bonding Contacts. These hydrogen bond interactions include the following (ASPP2 residue listed first, followed by CagA residue): the side chain of Lys751(NZ) to the main chain Gly210(O), the side chain of Tyr754(OH) to the main chain of Gln170(O), the side chain of Gln755(NE2) to the side chain of Ser218(OG), the side chain of Gln755(OE1) to the main chain of Gly209(N), the side chain of Thr757(OG1) to the side chain of Asp119(OD1), the side chain of Thr758(OG1) to the side chain of Ser171(OG), and the main chain of Ala760(O) to the side chain of Gln123(NE2).

Experimental Procedures

Purification of the *Helicobacter pylori* 26695 CagA and ASPP2 Complex. Full-length *H. pylori* 26695 CagA was cloned by PCR from genomic DNA (obtained from the American Type Culture Collection, reference ATCC 700392D) into the pCDNA4/His-Max TOPO (Invitrogen) vector (1). N-terminal domain constructs of CagA (19–257) and (19–235), which were used in this study, were derived from that original plasmid by PCR and subcloned as an N-terminal hexahistidine fusion protein into an engineered version of the pCDFDuet-1 vector (Novagen), which contains two tandem hexahistidine tags separated by linker and followed by rhinovirus 3C protease recognition sequence (1). ASPP2 (726–782) was amplified by PCR from a human ASPP2 cDNA clone (Clone Id: 4454266, Open Biosystems; catalog no. MHS1010-99621958) and ligated as a GST-fusion protein into modified pGEX 4T3 (GE Healthcare), which also contains a 3C protease recognition sequence between the GST tag and the *aspp2* gene. Proteins were coexpressed in BL-21 *Escherichia coli* bacterial cultures after induction with 1 mM IPTG at 26 °C overnight. GST-tagged ASPP2 and hexahistidine-tagged CagA were copurified on Ni-NTA Sepharose (Qiagen) in an elution buffer consisting of 20 mM Hepes at pH 7.5, 200 mM NaCl, and 250 mM imidazole. Affinity tags were removed by site-specific proteolytic cleavage with recombinant hexahistidine-tagged 3C protease (1:100), the protease (as well as uncleaved protein) removed by an additional pass through Ni-NTA column. The CagA–ASPP2 complex was further purified by cation-exchange (Fast Flow SP Sepharose, 20 mL of bead volume; GE Health), and gel-filtration chromatography (Superdex 200 HiLoad 16/60 column; GE Healthcare) using an ÄKTA FPLC (GE Health). Before loading the ion-exchange columns, a buffer containing the protein complex was diluted to 50 mM NaCl. After equilibrating the columns with washing buffer (20 mM Hepes at pH 7.5 and 2.5 mM DTT), protein complexes were loaded and unbound material removed with 3 column volumes of washing buffer. Proteins that were bound to the columns were subjected to a gradient increase of salt concentration (from 0 to 500 mM NaCl in 20 mM Hepes at pH 7.5 and 2.5 mM DTT using at least 20 column volumes). CagA–ASPP2 complexes bind strongly to cation-exchange resins and elute at 300 mM NaCl. As a final step in the purification, fractions eluted from the SP Sepharose column containing intact CagA–ASPP2 complexes were pooled

together, concentrated by using Centricon Plus-80 (Millipore), and run on a gel filtration column (Superdex 200 HiLoad 16/60; 120 mL) in the gel filtration buffer (200 mM NaCl, 20 mM Hepes at pH 7.5, and 2.5 mM DTT). The peak fractions were examined by SDS/PAGE, and proteins were visualized by staining with Coomassie blue solution [50% (vol/vol) methanol, 0.05% Coomassie brilliant blue R-50, and 10% (vol/vol) acetic acid in water] followed by destaining solution [5% (vol/vol) methanol and 7% (vol/vol) acetic acid in water].

Proline-rich domains of ASPP1 and iASPP were amplified from full-length cDNA clones obtained from Open Biosystems (ASPP1 Clone no. 9020517, catalog no. EHS1001-99864988; iASPP Clone no. 6068949; catalog no. MHS1010-9206150) and cloned into laboratory modified pGEX 4T3 vector. Complexes of CagA (19–235) with ASPP1 (694–755) or iASPP (511–555) were purified and analyzed in the same manner as CagA–ASPP2 complexes.

Bioinformatic and Biochemical Analysis of NTD of CagA. Using PredictProtein server (www.predictprotein.org) for the secondary structure prediction (2) and homology alignment of different CagA protein sequences (T-Coffee (3)), we identified several potential CagA subdomains that were subcloned into bacterial expression vectors as N-terminally hexahistidin fusion proteins and tested for expression, solubility, and degradation. Multiple constructs with variable N- and C-termini were examined. The construct containing residues 19–257 proved very well expressed, soluble, and stable after purification over Ni-NTA column (Qiagen). Following purification on Ni-NTA affinity column, the His tag was removed by using 3C protease during overnight dialysis into buffer consisting of 200 mM NaCl, 20 mM Hepes at pH 7.5, and 2.5 mM DTT. CagA (19–257) recombinant protein was also stable in vitro when left for a prolonged time at 4°C. Using AKTA FPLC, CagA (19–257) was tested for binding to ion exchange columns in the buffer containing 50 mM NaCl, 20 mM Hepes at pH 7.5, and 2.5 mM DTT. With calculated pI of 4.99, it was not surprising that it did not bind to a cation exchange column (SP-Sepharose), but it bound to an anion exchange column (Q Sepharose) and eluted by increasing salt concentration by gradient to 100 mM NaCl. It also eluted in a single pick from gel filtration column (Superdex 200; GE Health), which further attested for the stability of this CagA domain. To find an even more stable subdomain, we subjected the purified CagA (19–257) protein to limited proteolysis with subtilisin. This analysis produced proteolytic fragments that were separated by SDS/PAGE, visualized with PonceauS (Sigma), and sent for N-terminal sequencing. All proteolytic constructs contained the same N-terminal sequence, indicating that the cleavage occurred at C terminus. A closer examination at potential loop/helix boundaries from the predicted secondary structure of CagA protein sequence revealed two possible cleavage sites at residues 235 and 209. New constructs (19–235) and (19–209) were also well expressed and behaved similarly to CagA (19–257). To maximize the probability of detecting a protein–protein interaction, we used the largest construct for yeast two-hybrid (Y2H) screening.

Y2H Screen and Delineation of CagA-Binding Domain of ASPP2. Y2H screening was performed by Genomics and Proteomics Core Facility at German Cancer Research Center. The N-terminal domain of CagA (19–257) was cloned into the pGBT9 plasmid and used as a bait to screen Mate and Plate Universal Human

(normalized) Library (Clontech; catalog no. 630480). Clones encoding ASPP2 were isolated 15 times during the screen under moderate stringency conditions (0.4 mM 3-aminotriazole). Based on those clones, several overlapping constructs of ASPP2 proline-rich domain were designed and cloned as N-terminal GST fusion protein into modified pGEX4t3 vector. Those constructs were coexpressed with (His) 6-tagged CagA (19–257) in BL-21 *E. coli* and in vitro interaction determined by Ni-NTA pull-down assay. Minimal binding domains identified this way contained residues 684–782 of ASPP2 and residues 19–257 of CagA. Based on the limited proteolysis of this complex with subtilisin protease and N-terminal sequencing of digested products, the constructs were further trimmed to include ASPP2 (721–782) and CagA (19–235). Further deletion of 4 aa of ASPP2 (726–782) was necessary to obtain diffracting crystals of ASPP2–CagA complex.

Crystallization and Structural Determination. For crystallization, both native and selenomethionine-substituted protein (SeMet) complexes were purified in a buffer containing 20 mM Hepes at pH 7.5, 200 mM NaCl, and 2.5 mM DTT by gel filtration and subsequently concentrated to 35–50 mg/mL. Crystals were grown by vapor diffusion using hanging drops formed from mixing a 2 μ L of the protein complex with 2 μ L of an equilibration buffer [21% polyethylene glycol (PEG) molecular mass 4 kDa, 200 mM Li_2SO_4 , and 100 mM Tris at pH 8.5] and 0.6 μ L of the “Silver Bullets” additive 43 (0.25% 3,5-dinitrosalicylic acid, 0.25% 3-indolebutyric acid, 0.25% naphthalene-1,3,6-trisulfonic acid trisodium salt hydrate, 0.25% trans-1,2-cyclohexanedicarboxylic acid, and 0.02 M Hepes sodium at pH 6.8; Hampton Research HR2-996-43). Higher quality crystals were obtained from SeMet protein complexes, and these crystals were used for the final refinement. For cryoprotection crystals were transferred into a buffer with a 30% PEG at 4 kDa, 200 mM Li_2SO_4 , 100 mM Tris at pH 8.5, and 5% glycerol, and flash-cooled immediately in liquid nitrogen, where they were stored until data collection.

Data were collected from SeMet substituted protein crystals at Brookhaven National Synchrotron light source beam line X29 at the selenium absorption edge and processed by using HKL2000 (4). The crystals belonged to the space group C2, with unit-cell parameters $a = 118.64 \text{ \AA}$, $b = 120.24 \text{ \AA}$, $c = 100.66 \text{ \AA}$, and α , β , and $\gamma = 90.00^\circ$, 115.64° , 90.00° . There were four heterodimers of CagA–ASPP2 in the asymmetric unit. Phases were determined by using SHELX (5), and 90% of the final model was built by ARP/WARP (6). Cycles of manual building with COOT (7) and refinement with REFMAC5 (8) resulted in a model with an R/R_{free} of 19.1%/23.8% to 2.05- Å resolution. The crystallographic statistics are summarized in Table S1, and the coordinates and structure factors have been deposited at the Protein Data Bank with ID code 4IRV.

Mutagenesis. The amino acid substitutions were introduced by PCR using *Pfu* DNA polymerase (Stratagene) and primers containing the appropriate base changes. The template plasmid was subsequently removed by digestion with DpnI restriction enzyme before transformation. All of the mutations were verified by sequencing. Purification of the mutant proteins and protein complexes was performed as previously described for wild-type proteins.

Cells and Transfection. HEK293T cells (CRL-11268; American Type Culture Collection) and gastric adenocarcinoma cell line (AGS, CRL1739; American Type Culture Collection) were cultured in DMEM supplemented with 10% FBS and the appropriate antibiotics.

All transfections were carried out by using Lipofectamine 2000 (Invitrogen), according to the manufacturer's instructions. FLAG-mCherry, FLAG-mCherry-ASPP2 20aa (746–765), and FLAG-mCherry-ASPP2 56aa (726–782) were all cloned in the retroviral expression vector pLHCX (Clontech) and used to

generate stably expressing AGS cell lines under HygromycinB (0.4 $\mu\text{g}/\text{mL}$) selection. Full-length wild-type or mutant ASPP2 were cloned into pREV-TRE (Clontech) retroviral vector and used for transient expression in HEK 293-T cells. Full-length wild-type or mutant CagA was cloned into pcDNA5FRT/TO (Invitrogen) for transient expression in HEK 293-T cells. Mutant ASPP2 and CagA genes were generated as above described.

Bacterial Strains and Infections. *H. pylori* (Hp) strain G27 and the isogenic mutants ΔVirb10 were cultured on blood agar plates or in *Brucella* broth medium supplemented with selective antibiotics and 10% calf serum and grown in 5% CO_2 at 37 $^\circ\text{C}$. Before infection, host cells were washed and incubated in DMEM supplemented with 5% calf serum and without antibiotics. Cells were infected with the specific Hp strains at a multiplicity of infection (moi) of 1:50 for 7 h.

In Vitro Binding Assay. Pull-down on Ni-NTA Sepharose columns was used to determine binding of wild-type or mutant hexahistidine-tagged CagA (19–235) to wild-type or mutant GST-tagged ASPP2 (726–782). The soluble fraction of bacterial cell extracts coexpressing His-CagA and GST-ASPP2 was loaded to the Ni-NTA columns in a binding buffer (20 mM Hepes at pH 7.5 and 200 mM NaCl), washed with the binding buffer plus gradual increase of imidazole (5 mM, 10 mM, 30 mM, and 50 mM imidazole), and complexes were eluted with elution buffer (binding buffer with 250 mM imidazole). The initial assessment of complex formation was performed by SDS/PAGE analysis of eluted material. If both proteins were present in the eluted fraction, the tags were removed by proteolytic digestion, and complexes further were subjected to gradient cation-exchange chromatography (Fast Flow SP Sepharose, 20 mL of bead volume, GE Health; 0–500 mM NaCl in 20 mM Hepes at pH 7.5 and 2.5 mM DTT during 150 min at a 3 mL/min flow rate). When complex stability was tested by size exclusion chromatography, 10 mg of total wild-type or mutant complexes were loaded on Superdex 200 columns [Superdex 200 HiLoad 16/60 (120 mL); GE Healthcare] in 20 mM Hepes at pH 7.5, 200 mM NaCl, and 2.5 mM DTT.

Antibodies. Anti-FLAG M2 (F1804) was purchased from Sigma. Rabbit Anti-CagA (sc-b300), mouse anti-ASPP2 (sc-10922), and mouse anti-*H. pylori* (sc-57778) were purchased from Santa Cruz Biotechnologies. Anticell Caspase-3 was purchased from Cell Signaling (9661).

Flow Cytometry. Hp-infected or uninfected cells were processed according to the Cell Signaling guidelines for flow cytometry of cells stained for cleaved Caspase-3. Data were acquired by using Becton Dickinson FACScalibur and analyzed by using FlowJo 8.6 software.

Immunoprecipitation. Cells were treated as described and harvested by scraping in PBS. Cell pellets were lysed in 0.5% Nonidet P-40-containing buffer (0.5% Nonidet P-40, 5 mM EDTA, 50 mM Tris-HCl at pH 7.5, and 100 mM NaCl). Sonication of cells in Nonidet P-40 containing buffer was carried as elsewhere described (9). All buffers were supplemented with complete protease inhibitors (Roche). Lysates were cleared of debris by centrifugation (13,500 $\times g$ for 10 min at 4 $^\circ\text{C}$). Protein concentrations were determined by BCA assay (Pierce), and an equal amount of each sample was used for immunoprecipitation. For immunoprecipitations, lysates were incubated overnight at 4 $^\circ\text{C}$ with the indicated antibodies. Twenty microliters of protein G Sepharose (Sigma) was added 1 h before completion of immunoprecipitation. Beads were washed four times in lysis buffer and eluted by boiling in sample buffer supplemented with DTT. Immunoprecipitates were resolved by SDS/PAGE and subjected to Western blotting.

Immunofluorescence. Cells were grown on coverslip and, following *H. pylori* infection, were fixed in 4% paraformaldehyde and permeabilized in 0.1% Triton X-100. Fixed and permeabilized cells were blocked in 2% BSA and incubated with the indicated primary antibodies for a 1 h at room temperature. Subsequently,

cells were incubated with Alexa Fluorochrome secondary antibodies, phalloidin-647 for actin, and mounted with Fluoromount-G (Southern Biotech) onto glass slides for the confocal image acquisition. Confocal images were taken under LSM 710 Zeiss confocal microscope.

1. Nesić D, et al. (2010) Helicobacter pylori CagA inhibits PAR1-MARK family kinases by mimicking host substrates. *Nat Struct Mol Biol* 17(1):130–132.
2. Rost B, Yachdav G, Liu J (2004) The PredictProtein server. *Nucleic Acids Res* 32(Web Server issue):W321–W326.
3. Notredame C, Higgins DG, Heringa J (2000) T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J Mol Biol* 302(1):205–217.
4. Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 276:307–326.
5. Sheldrick GM (2008) A short history of SHELX. *Acta Crystallogr A* 64(Pt 1):112–122.
6. Langer G, Cohen SX, Lamzin VS, Perrakis A (2008) Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nat Protoc* 3(7): 1171–1179.
7. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66(Pt 4):486–501.
8. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* 53(Pt 3): 240–255.
9. Buti L, et al. (2011) Helicobacter pylori cytotoxin-associated gene A (CagA) subverts the apoptosis-stimulating protein of p53 (ASPP2) tumor suppressor pathway of the host. *Proc Natl Acad Sci USA* 108(22):9238–9243.

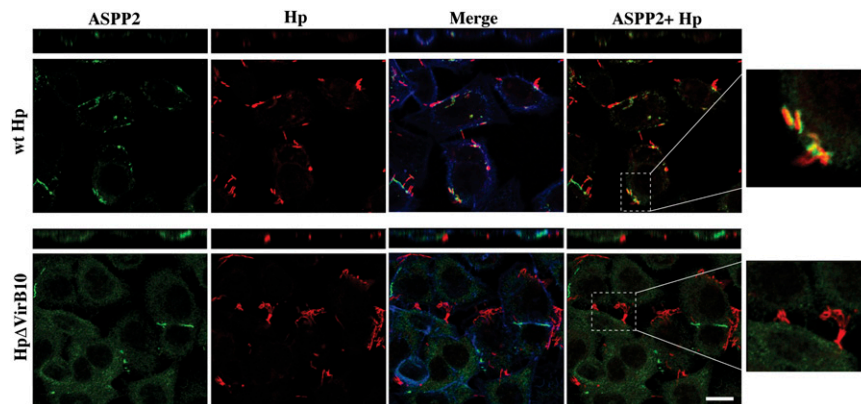


Fig. S1. Relocalization of endogenous ASPP2 during the course of *Hp* infection is CagA dependent. Endogenous ASPP2 is recruited at the sites of *Hp* infection. AGS cells were infected with the indicated *Hp* strains (moi 1:50) or left uninfected. After 7 h, cells were fixed and stained with anti-ASPP2 (green) and anti-CagA (red) antibodies and phalloidin for F-actin (blue). *Insets* show that only wild-type *H. pylori* (wt *Hp*; *Upper Inset*) strongly associates with ASPP2, whereas there is no association with *H. pylori* Δ VirB10 mutant (*Hp* Δ VirB10; *Lower Inset*). (Scale bar: 10 μ m.)

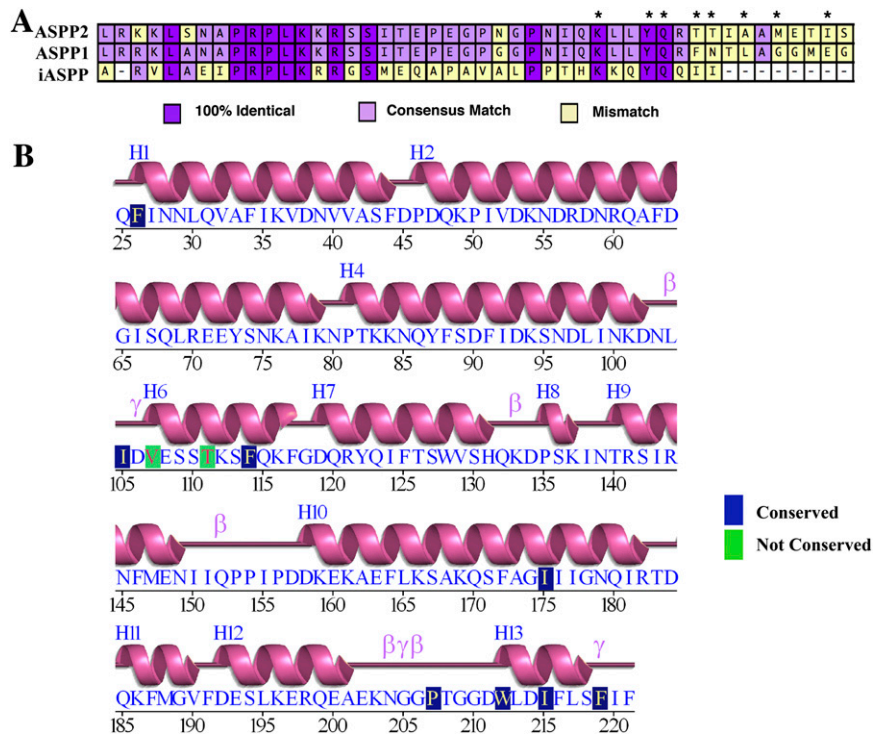


Fig. S2. Conservation of residues of ASPP2 and CagA that are involved in the complex formation. (A) ClustalW alignment of proline-rich domain of ASPP protein family: ASPP2 (721–782) and corresponding sequences of ASPP1 (694–755) and iASPP (511–555); residues involved in CagA binding are marked with asterisk. (B) Secondary structure diagram of CagA (25–221). Amino acids that are contacting ASPP2 are highlighted in boxes (blue are conserved and green are not conserved).

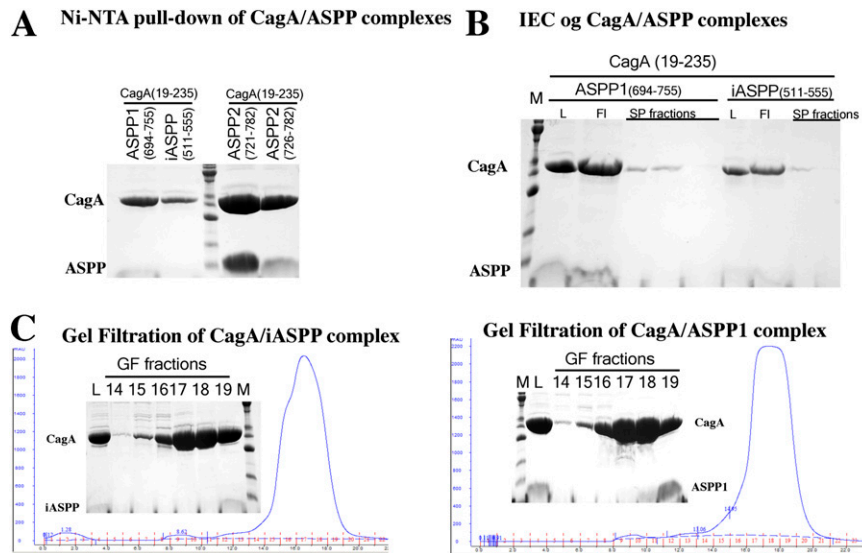


Fig. S3. Interaction of CagA(19–235) with ASPP family proteins. (A) Ni-NTA pull-down of CagA/ASPP complexes. His-tagged CagA was coexpressed with indicated GST-tagged ASPP protein in *E. coli*. The cell extracts were applied to Ni-NTA gravity columns and bound proteins eluted with elution buffer (20 mM Hepes at pH 7.5, 200 mM NaCl, and 250 mM imidazole). Following removal of affinity tags by an overnight cleavage with 3C protease, the complexes were examined by SDS/PAGE. (B) Ion-exchange chromatography (IEC) of CagA/ASPP complexes. Half of the material that was eluted from affinity columns was subjected to cation exchange chromatography (Fast Flow SP Sepharose) by using ÄKTA FPLC. Both ASPP1 (pI = 10.78) and iASPP (pI = 12.21) are expected to bind to SP Sepharose, but CagA is not (pI = 5.1). Small aliquots of loaded material (L), flow-through (FI), and eluted fractions (SP fractions) were run on SDS/PAGE. Because there is only a small amount of either ASPP1 or iASPP eluted from Ni-NTA columns, and the complex falls apart during IEC, the majority of material is present in the flowthrough, but not in the elution fractions. (C) Gel filtration profile of complex of CagA/ASPP complexes. The other half of the material that was eluted from affinity columns was subjected to size exclusion column (Superdex 200 HR 10/30 column; GE Healthcare) by using an ÄKTA FPLC (GE Healthcare). Graph represents elution profile from Superdex 200 column with SDS/PAGE of peak fractions presented inside. The proteins were visualized by Coomassie stain. M, molecular weight marker.

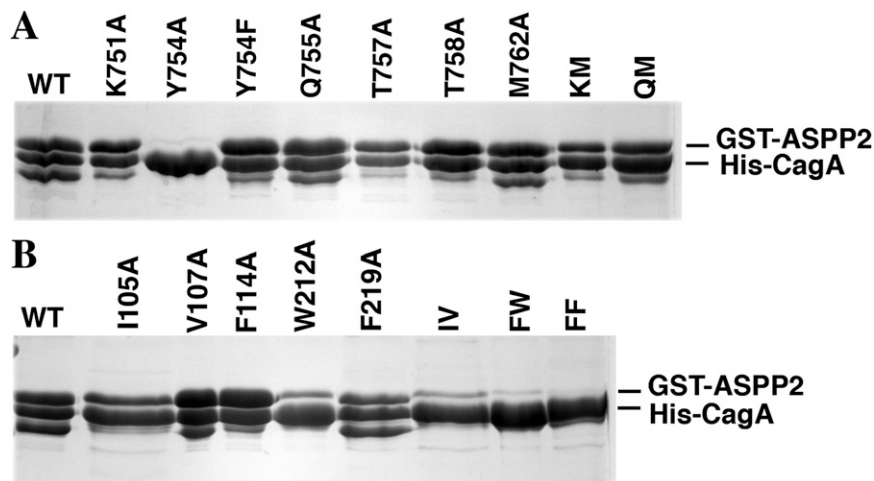


Fig. 54. Ni-NTA pull-down of mutant CagA/ASPP2 complexes. (A) Hexahistidine-tagged wild-type CagA (19–235) was coexpressed with indicated mutant GST-tagged ASPP2 (726–782) in *E. coli*. The cell extract was applied to Ni-NTA gravity columns, and bound proteins were eluted with buffer containing 250 mM imidazole (20 mM HEPES at pH 7.5, 200 mM NaCl, and 250 mM imidazole). The complex formation was determined by running 10 μ L of the eluted material on SDS/PAGE and staining gels with Coomassie dye. (B) Different mutants of hexahistidine-tagged CagA (19–235) were coexpressed with wild-type GST-tagged ASPP2 (726–782) in *E. coli*, and complexes eluted and visualized as in A. All indicated residues were changed into alanines. FF, F114A/F219A double mutant of CagA; FW, F114A/W212A double mutant of CagA; IV, I105A/V107A double mutant of CagA; KM, K751A/M762A double mutant of ASPP2; QM, Q755A/M762A double mutant of ASPP2; WT, wild type.

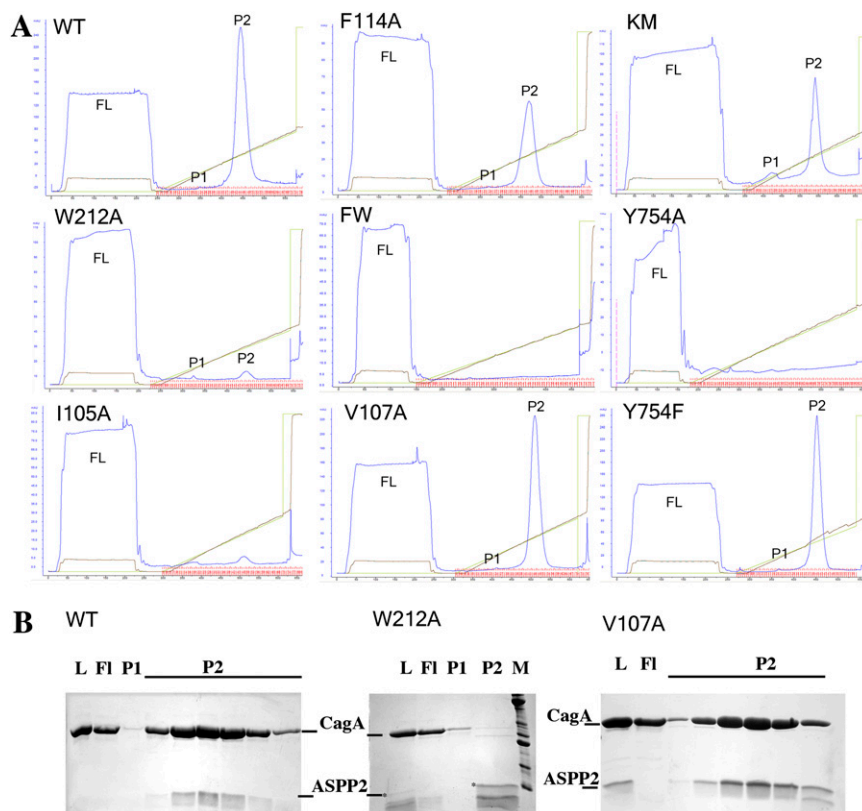


Fig. S5. Cation exchange chromatography of mutant CagA/ASPP2 complexes. (A) Different combination of wild-type and mutant hexahistidine-tagged CagA (19–235) and GST-tagged ASPP2 (726–782) were coexpressed in *E. coli* and coeluted on Ni-NTA gravity columns. Affinity tags were removed by cleavage with 3C protease, and complexes were run on Fast Flow SP Sepharose column by using ÄKTA FPLC. The complexes were loaded at low salt concentration (60 mM NaCl) and examined in a salt gradient as described in *Experimental Procedures*. CagA-binding domain of ASPP2 (726–782) is very basic with pI of 10.4, which enables CagA–ASPP2 complex to bind to SP Sepharose. If not bound to CagA, ASPP2 (726–782) is unstable and degrades rapidly, and almost cannot be detected in the elution. Thus, when the mutant complex is stable on SP Sepharose, a nice peak during elution could be seen on the chromatogram (P2), similar to the peak of wild-type complex. If the complex is disassembled, there is no elution peak. CagA has pI of 5.1, and it is not expected to bind to cation exchange columns and, therefore, it is mainly present in flowthrough (FL) when not in complex with ASPP2. Very small peak that contains CagA and GST, which are the same size, is detected sometimes (P1). (B) Individual fractions or pooled material from flowthrough (FL) and elution peaks (P1, P2) were examined by SDS/PAGE. The material from P1 fractions was concentrated to be visible on the gel. Proteins were stained with Coomassie blue stain. FI, flowthrough; FW, F114A/W212A double mutant of CagA; KM, K751A/M762A double mutant of ASPP2; L, loaded material; M, molecular weight marker; P1 and P2, elution peak; WT, wild type.

Table S1. Crystallographic statistics

Crystallographic variables	Crystallographic values
Space group	C2
Cell dimensions	
<i>a</i> , Å	118.64
<i>b</i> , Å	120.24
<i>c</i> , Å	100.66
α, β, γ , °	90, 115.64, 90
Wavelength, Å	0.97869
Resolution, Å	50.00–2.04
No. of reflections	1,517,629
No. of unique reflections	80,050
R_{merge} ,* %	7.6 (42.5)
$I/\sigma(I)$	17.3 (3.1)
Completeness, %	99.9 (99.9)
Redundancy	3.8 (3.7)
Refinement	
Resolution, Å	50.00–2.04
No. of reflections	73,785
$R_{\text{work}}/R_{\text{free}}$ †	18.9/23.4
All atoms	7,530
Protein	7,017
Water	513
Average B factor	34.0
rms deviations	
Bond lengths, Å	0.019
Bond angles, °	1.909
Ramachandran plot	
Favored regions, %	94.8
Allowed region, %	5.2
Outliers, %	0

*As defined and calculated by HKL2000 (1).

†As defined and calculated by Refmac5 (2). Values in parentheses are for the highest resolution shell.

- Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods in Enzymology: Macromolecular Crystallography* eds Carter CW, Jr., Sweet RM (Academic, New York), Vol 276, Part A, pp 307–326.
- Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Cryst D* 53:240–255.