Supplemental Information

TABLE OF CONTENTS

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

H. pylori Cultures, Strains, Sequence Analysis and Antibodies

Culture Conditions

Bacterial plate cultures were performed on BD BBL Brucella agar (Becton, Dickinson and Company, MD, USA), enhanced with 10% citrated bovine blood (Svenska LabFab, Ljusne, Sweden) and 1% Isovitox Enrichment (Dalynn Biologicals, Calgary AB, Canada) supplemented with 4 µg/mL of amphotericin, 10 µg/mL of vancomycin, and 5 µg/mL of trimethoprim (Sigma Aldrich, St. Louis, MO, USA). *H. pylori* cultures were grown at 37°C in a mixed-gas incubator (Forma Scientific Inc., AB Ninolab, Väsby, Sweden) under micro-aerophilic conditions with 10% CO_2 and 5% O_2.

Bacterial broth cultures for the 2D-DIGE-analysis (Figure 7A and 7B; Figure S7A) were performed in 50 ml Brucella broth containing 5% heat-inactivated newborn calf serum (Invitrogen, Carlsbad, CA) supplemented with 10 µg/mL vancomycin, 2.5 units/mL polymyxin B, 5 µg/mL trimethoprim, and 2.5 µg/mL amphotericin B (Sigma Aldrich St. Louis, MO, USA) in 5% CO_2 with gentle agitation (60 rpm) to mid-log phase (16–17 hours).

Strains

Laboratory strains: 17875/Leb is a spontaneous mutant of *H. pylori* strain CCUG17875 that binds ABO/Leb antigens, but not sialylated antigens. The 17875*babA1A2* "double" mutant 17875babA1::kan babA2::cam is a deletion (null) mutant where both *babA1* and *babA2* genes were removed from *H. pylori* strain CCUG17875, and this is referred to as the *babA1A2*-mutant (Mahdavi et al., 2002).

Clinical isolates:

Strains from Sweden (22 strains), Japan (4 strains), and Spain (one strain) have been described (Aspholm-Hurtig et al., 2004). Strains from Peru (59 strains) were reported in (Aspholm-Hurtig et al., 2004; Kersulyte et al., 2010).

Swedish strains SO-1 and SO-2: The Orebro1 (SO-1) patient was 43 years old and diagnosed with non-reflux dyspepsia. The Orebro2 (SO-2) patient was 29 years old and diagnosed with gastric reflux symptoms. Ulcer disease was not present in the two patients and consequently, due to the standard treatment procedures in Sweden at the time of the first visit, neither of them were treated with antibiotics for *H. pylori* infection. All patients underwent gastroscopy in 1990 and 1999 at Orebro University Hospital, Sweden. At the first endoscopy, only one biopsy from the gastric antrum was obtained, and at the second follow-up endoscopy biopsies were collected from both antrum (A) and corpus (C). The corresponding isolates are referred to as A90, A99, and C99, respectively (Gustavsson et al., 2005).

Greek strains: Isolates were from 44 patients diagnosed with gastric or duodenal ulcers and non-ulcer dyspepsia or esophagitis. *H. pylori* clinical sweeps were collected from both antrum and corpus (Panayotopoulou et al., 2007).

Indian strains: 17 Indian clinical isolates were collected at the National Institute of Cholera and Enteric Disease, Kolkata, India.

GC-patient (Colombian strain): The Colombian isolates Y0 (NQ392) and Y3 (NQ1707) were obtained from a patient (here denoted the GC-patient) diagnosed with mixed, predominantly incomplete intestinal metaplasia. The isolate NQ4060 was obtained at the later 16-year time point when the patient was diagnosed with dysplasia and three clones (Y16-1, 2 and 3) were collected (Kennemann et al., 2011).
All clinical isolates were obtained with informed consent for molecular genetic studies in protocols approved by local human studies committees.

Sequence Analysis

Sequence analysis of *babA*, *babB* and *cysS*, *glr* genes were carried out using genomic DNAs of different *H. pylori* strains with primers from Eurofins MWG Operon (Germany). Genomic DNA was isolated using the E.Z.N.A. ® Tissue DNA Kit (Omega Bio-Tek, USA). Primers used for amplification and sequencing of *babA*, *babB* and *cysS*, *glr* gene fragments are listed in Table S4; "Sequence analysis of

babA, babB and house-keeping genes*"*. Amplification was by Dream Taq DNA polymerase (Fermentas/ Thermo Scientific, Lithuania). All *babA*, *babB* and *cysS*, *glr* gene products were sequenced by Eurofins MWG Operon (Germany). Sequence analysis including alignments using the ClustalW algorithm was performed using DNASTAR/Lasergene Software Package.

PCR to identify *babA* loci and 5'-end sequence.

Locus A with primers; HypDF1-F - BabAR1-R, expected fragment length ~2.5 kbp
Locus B with primers; S18F1-F - BabAR1-R, expected fragment length ~ 1 kbp
Locus C with primers; Hp1AS-F - BabAR1-R, expected fragment length ~ 1.5 kbp.

Amplification program:

Step	Temperature, ^{0}C	Time	Cycles
Initial denaturation	95	3 min	
Denaturation	95	30 sec	40 cycles
Annealing	55	30 sec	
Extension (1 kbp / 1 min)	72	Kbp×1min	
Final extension	72	5 min	

The PCR products were loaded on the 1% gel with 6×Loading buffer and separated by electrophoresis for 1 hour, 90 V. The identified PCR fragment was cut from gel and DNA was extracted using the E.Z.N.A® Gel Extraction Kit (Omega Bio-Tek). The *babA* gene was sequenced with the BabAR1-R (reverse) primer.

PCR amplification of the full *babA* loci and 3'-end sequence.

Locus A with primers; HypDF1-F – 1539-R, expected fragment length ~4.1 kbp

Locus B with primers; S18F1-F – 1539-R, expected fragment length ~ 2.6 kbp

Locus C with primers; Hp1AS-F – 1539-R, expected fragment length ~ 3 kbp.

Amplification program:

Step	Temperature, ^{0}C	Time	Cycles
Initial denaturation	95	3 min	
Denaturation	95	30 sec	40 cycles
Annealing	52	30 sec	
Extension (1 kbp / 1 min)	72	Kbp×1min	
Final extension	72	5 min	

A The PCR products were loaded on the 1% gel with 6×Loading buffer and separated by electrophoresis for 1.5 hour, 90 V. The selected PCR fragment was cut from gel and DNA was extracted using the E.Z.N.A® Gel Extraction Kit (Omega Bio-Tek). The *babA* gene was sequenced with the 1539-R (reverse) primer.

PCR amplification and sequence of the mid part of *babA*.

PCR of *babA* with primers; OP9647-F – A27-R, expected fragment length ~1.4 kbp.

Amplification program:

Step	Temperature, ^{0}C	Time	Cycles
Initial denaturation	95	3 min	
Denaturation	95	30 sec	40 cycles
Annealing	52	30 sec	
Extension (1 kbp / 1 min)	72	Kbp×1min	
Final extension	72	5 min	

The PCR products were loaded on the 1% gel with 6×Loading buffer and separated by electrophoresis for 1 hour, 90 V. The ~1.4 kbp PCR fragment was cut from gel and DNA was extracted using the E.Z.N.A® Gel Extraction Kit (Omega Bio-Tek). The *babB* gene was sequenced with the A27-R (reverse) primer.

PCR amplification and sequence of *babB*.

PCR of *babB* with primers; SE22-F and A72-R, expected fragment length ~1.3 kbp.

Amplification program:

Step	Temperature, ^{0}C	Time	Cycles
Initial denaturation	95	3 min	
Denaturation	95	30 sec	40 cycles
Annealing	52	30 sec	
Extension (1 kbp / 1 min)	72	Kbp×1min	
Final extension	72	5 min	

The PCR products were loaded on the 1% gel with 6×Loading buffer and separated by electrophoresis for 1 hour, 90 V. The ~1.3 kbp PCR fragment was cut from gel and DNA was extracted using the E.Z.N.A® Gel Extraction Kit (Omega Bio-Tek). The *babB* gene was sequenced with primers SE22-F and A72-R.

PCR amplification and sequence of the *cysS* and *glr* housekeeping genes.

PCR of *glr* with primers; 5733 and 6544, expected fragment length ~ 800 bp.

PCR of *cysS* with primers; cysS-F and cysS-R, expected fragment length ~ 1.3 kbp).

Amplification program:

Step	Temperature, ^{0}C	Time	Cycles
Initial denaturation	95	3 min	
Denaturation	95	30 sec	40 cycles
Annealing	53	30 sec	
Extension (1 kbp / 1 min)	72	Kbp×1min	
Final extension	72	5 min	

The PCR products were loaded on the 1% gel with 6×Loading buffer and separated by electrophoresis for 1 hour, 90 V. The ~1.3 kbp *cysS* PCR fragment and the ~0.8 kbp *glr* PCR fragment were cut from gel and DNA was extracted using the E.Z.N.A® Gel Extraction Kit (Omega Bio-Tek). The *glr* gene was sequenced with primers; 5733 and 6544, and the *cysS* gene was sequenced with primers cysS-F and cysS-R.

Antibodies

Two batches of the polyclonal anti-BabA antibodies were used in this study. The central domain of BabA for the antibody Vite was produced by the T. Borén lab according to the procedure for Antibody AK277 by Odenbreit et al., 2002. Rabbit immunization was performed by Agrisera, (Vännäs, Sweden).

THE FOLLOWING PROCEDURES ARE PRESENTED IN EXPERIMENTAL ORDER AND ACCORDING TO SECTION HEADINGS

1. Acid Inactivation and Reversal of Leb Binding

Labeling of *H. pylori* cells. Bacteria were labeled according to previously described methods **(**Aspholm et al., 2006**)**. Depending on the experiments, the following series of fluorophores were used: Fluorescein isothiocyanate (FITC), (Sigma, St. Louis, MO), Fluor® 488 (Alexa488) and Alexa Fluor® 555 (Alexa555) (both from Molecular Probes, Life Technologies Corp., Oregon, USA) or CF™ 555 succinimidyl ester (CF555) (Sigma, St. Louis, MO).

Buffers used for pH titration. The binding reaction was performed in a series of citrate-phosphate buffers of pH 2–6 in incremental steps of 0.5 pH units (final concentrations 15 mM citrate and 15 mM phosphate), supplemented with 0.1% bovine serum albumin (BSA) Cohn fraction (Saveen Werner AB, Sweden), 0.05% (v/v) Tween-20, and ionic strength adjusted to 0.15 with NaCl (I = 0.15).

***In vitro* adherence to human gastric tissue** (Figure 1A; Figure S1A). FITC labeling and *in vitro* bacterial adhesion was assayed with minor modifications (Aspholm et al., 2006). Bacteria were pre-incubated in citrate-phosphate buffers at pH 2, 4, and 6, for 10 min and then applied to histo-tissue sections for 30 min. After washing, the sections were immersed in 0.15 M NaCl before mounting with Dako Fluorescent Mounting Medium (Dako North America, Inc., CA, USA). Bacterial adherence was digitalized with a Zeiss AXIOcam MRm (Carl Zeiss AB, Stockholm, Sweden) with optical magnification of 200×. Zeiss AxioVision software v.4.5 was used to quantify bacterial adherence. Each value is the mean ± SEM of three different fields. ApoTome (Carl Zeiss AB) was used in Figure 1A to increase sharpness. A human gastric tissue section was also subjected to standard hematoxylin and eosin (H&E) histochemical staining.

pH-dependent binding of *H. pylori* to human mucin (Figure 1B; Figure S4C). A human gastric specimen was obtained after informed consent and approval of the Lund University Hospital, Lund, Sweden, ethics committee. The specimen was isolated from the macroscopically normal mucosa of a tumor-affected stomach, as evaluated by a clinical pathologist. The mucosa was scraped using a microscope slide in order to obtain the surface mucin. Isolation of mucin, preparation of mucin samples, and ELISA were performed as described (Skoog et al., 2012). Prior to the binding experiments, mucin samples were diluted in 4 M GuHCl to 6 µg/mL, and 96-well Polysorb plates were coated with 100 µL of mucin overnight at 4°C. The plates were washed three times with PBS 0.05% Tween and then blocked for 1 hour with 200 µL of blocking buffer. A total of 100 µL of bacterial suspension with an OD_{600} = 0.1 was diluted 1:20 in citrate-phosphate buffer with pH ranging from 2 to 6 in increments of 0.5 pH units and were incubated in a bacterial shaker at 37°C for 2 hours. The plates were washed three times with PBS 0.05% Tween. *H. pylori* binding to mucin was detected with anti-whole cell *H. pylori* rabbit serum diluted 1:1000 as described previously (Skoog et al., 2012). Absorbance data were normalized and analyzed with GraphPad Prism v 6.0 (GraphPad Software Inc., La Jolla, CA, USA) and presented as the mean ± SEM.

Blood group antigen conjugates. Two different types of fucosylated blood group antigen conjugates were used; Semi-synthetic Leb- and Lea-glycoconjugate with natural purified or synthesized oligosaccharides linked to human serum albumin (Leb-HSA and Lea-HSA, respectively) were provided by Isosep (Isosep AB, Tullinge, Sweden) or synthesized by S. Oscarson (co-author) and used for all binding experiments, and for Leb affinity column for native BabA protein purification. For the series of SPR/Biacore tests at acidic pH, the PolyAcrylAmide-based Leb conjugate, Leb-PAA (Lectinity Holdings, Inc., Moscow, Russia) was used instead.

Labeling of Leb-HSA by ^{125}I. The Leb-HSA conjugate (IsoSep AB, Tullinge, Sweden) was ^{125}I-labeled (^{125}I-Leb-conjugate) by the chloramine T method (Aspholm et al., 2006).

Analysis of BabA binding properties by radioimmunoassay (RIA). ^{125}I-labeled Leb-HSA conjugate ("hot conjugate") or ^{125}I-labeled Leb-HSA diluted with unlabeled conjugate ("cocktail") (Aspholm et al., 2006) were mixed with 1 mL of bacterial suspension (OD_{600} = 0.1) in PBS-Tween pH 7.4, including 1% BSA and 0.05% (v/v) Tween-20 (blocking buffer). Leb binding of *E. coli* full-length recBabA from SO-2 (as in Figure 3F) was tested at OD_{600} = 10.0). Following incubation, bacteria were pelleted by centrifugation at 20,000 × *g*, and the ^{125}I in the pellet and in the supernatant was measured using the 2470 Wizard^2 Automatic Gamma counter (PerkinElmer, Waltham, MA, USA) giving a measure of binding activity (% binding).

Acid-induced release of Leb receptors (Figure 1C). The 10 ng/mL of ^{125}I-Leb-conjugate was mixed with cells of strain 17875/Leb (OD_{600} = 1.0) in pH 6.0 citrate-phosphate buffer and incubated for 1 h at room temperature. The bacterial cells were pelleted by gentle centrifugation. The pellet was re-suspended in pH 2 citrate-phosphate buffer. After 0.5–5 min incubation, the reaction was stopped by centrifugation at 20,000 × *g* and ^{125}I-Leb-conjugate bound to the pelleted bacterial cells was analyzed with the 2470 Wizard^2 Automatic Gamma counter (PerkinElmer, Waltham, MA, USA). The experiment was repeated three times.

Leb binding analysis at equilibrium of BabA binding parameters for Leb at different pH such as in Figures 1D, 6A, 7G, S3E, S7F. The association (affinity) constant (K_a) and binding capacity (B_{max}) of the BabA-Leb binding were determined by radioligand binding experiments at equilibrium (Motulsky and

Neubig, 2001). Briefly, *H. pylori* cells were incubated in the presence of a fixed concentration of ^{125}I-Leb-HSA conjugate and an incremental series of concentrations of unlabeled "cold" Leb-HSA conjugate at room temperature. Long-term incubation (17 h) was used to approach equilibrium. Bacterial cells were diluted so that less than 10% of the ^{125}I-Leb-HSA was bound to the bacterial cells in the absence of unlabeled conjugate. Cells of the non-Leb-binding *H. pylori* 17874 strain were added as a "carrier" (OD_{600} = 0.05) to create a visible pellet (after centrifugation) and to assess the non-specific binding. Appropriate bacterial cell dilutions, concentration of ^{125}I-Leb-HSA (in the range of 0.2–1 ng/mL), and a range of concentrations of unlabeled Leb-HSA (usually within 0.1 ng/mL–10 µg/mL) were separately optimized for each strain. Binding parameters of *E. coli* full-length recBabA were determined in a saturation binding experiment (Motulsky and Neubig, 2001) measuring equilibrium binding to increasing series of concentration of ^{125}I-Leb-HSA (0.01–10 ng/mL) and assessing non-specific binding of *E. coli* cell at the same ^{125}I-Leb-HSA concentrations. Binding reactions were performed either in a series of citrate-phosphate buffers pH 2-6 (see "Buffers used for pH titration") or in a blocking buffer (PBS with 0.05% Tween-20 and 1% BSA) at pH7.4. The reaction volume was 1 mL. ^{125}I-Leb-conjugate bound to the pelleted bacterial cells was analyzed with the 2470 Wizard^2 Automatic Gamma counter (PerkinElmer, Waltham, MA, USA). Binding curves were fitted with GraphPad Prism version 5.0 or 7.0 (GraphPad Software Inc., La Jolla, CA, USA). The affinity constant (K_a) was calculated as $1/K_d$; binding capacity (B_{max}) was normalized to nM per 1mL of OD_{600} = 0.1. For the SO-2 antrum and corpus clones, affinities at pH 6 and pH 4 (Figure S3E) were analyzed as described (Aspholm et al., 2006).

Acid sensitivity in BabA-mediated Leb-binding (pH_{50}) i.e. pHgram, analyzed by RIA, such as Figures 1D-1E; 2A, 2D, S2A-B, 3A-3B, S4B, 5A, 5E-5G, S5B, 6A, 6D, 6F, 6G, S6D-E, 7G and S7F, with pH_{50} displayed by box-plots in Figure 6F. Binding of bacterial cells (OD_{600} = 0.1) to Leb-conjugate cocktail (300 ng/mL of Leb-conjugate, 0.3 ng/mL of ^{125}I-Leb-conjugate) was assessed in a series of nine citrate-phosphate buffers with pH values ranging from pH 2 to pH 6 in incremental steps of 0.5 pH units. After 1 h incubation at room temperature, the reaction was stopped by centrifugation, and ^{125}I-Leb-conjugate bound to pelleted cells was determined as described above. The results were presented as "pHgrams" where pH-dependent Leb binding is presented as relative to maximal binding. The sensitivity of different strains to pH was expressed as their midpoints, pH_{50} (interpolated pH value at a 50% decrease in binding relative to maximal binding).
Mann–Whitney or Kruskal–Wallis tests in GraphPad Prism version 5.0 (GraphPad Software Inc., La Jolla, CA, USA) were used (as specified in the results section) to compare pH_{50} values for different groups of isolates or populations (Figure 6F). The Spearman correlation coefficients were calculated in GraphPad Prism to study relationship between pH_{50} values of strains from several studied populations, the level of their binding to Leb cocktail at neutral pH (pH 6) (Figure 6G), and the affinity and maximal binding capacity (B_{max}) of the same strains (Figure S7F).

Test for reversibility of acid-induced inactivation of Leb binding (Figure 1E). Cells of strain 17875/Leb (OD_{600} = 0.1) were incubated in a series of citrate-phosphate buffers from pH 6 to pH 2 for 1 h. The bacterial suspensions were reconditioned to pH 5 by adding a second series of corresponding buffers, mixed with Leb-conjugate cocktail (300 ng/mL of Leb-conjugate and 1 ng/mL of ^{125}I-Leb-conjugate), and bound ^{125}I-Leb was measured as described above. The results are presented as 're-constituted' pHgrams. The acid stability of Leb-conjugate was tested by 1 h exposure of Leb-conjugate to pH 6 to pH 2 buffers and reconditioning to pH 5. "Acid-treated" Leb was then applied to 17875/Leb cells for 1 h to test the residual Leb-binding activity (in green). Experiments were repeated several times, and representative experiments are shown.

Kinetics of Leb binding by RIA (Figure S1B). Cells of strain 17875/Leb (OD_{600} = 0.1) were first incubated at pH 2 to acid inactivate Leb binding or at pH 5 in the appropriate citrate-phosphate buffer. After 2 min, the pH of the suspensions was reconditioned to pH 5 and ^{125}I-Leb-conjugate (1 ng/mL) was added. After 1 to 60 min, the reaction was terminated by centrifugation, and bound conjugate was measured. The experiment was repeated three times.

Tolerance of BabA-mediated Leb binding during repeated pH cycles (Figure 1F), "the yo-yo endurance experiment". Cells of 17875/Leb (OD_{600} = 0.1) were treated with one to five pH shifts between

pH 5 and pH 2.5. After 15 min incubation in citrate-phosphate buffer at pH 2.5, the suspensions were diluted into pH 5 buffer, incubated for 15 min, pelleted by gentle centrifugation, re-suspended in pH 2.5 buffer, etc. After each cycle of pH changes from pH 2.5 to pH 5, the subset of samples was analyzed by RIA with Leb cocktail (300 ng/mL of Leb-conjugate and 1 ng/mL of ^{125}I-Leb-conjugate) for 1 h at room temperature. The binding reaction was stopped by centrifugation, and the amount of bound conjugate was measured as described above. The experiment was run in duplicate. The results are expressed relative to the control in which cells were treated by similar cycles of buffer changes and centrifugations but the pH was kept constant at pH 5.

pH stability of Alexa488 fluorophore (Figure S1E). Alexa488 (10 mg/mL in DMSO) was diluted in PBS-T to 10 µg/mL, and the NHS-groups were quenched with Tris-HCl pH 8.8. The Alexa solution was diluted 10 times with pH 6 citrate-phosphate buffer and incubated for 5 min at room temperature after each pH adjustment. A total volume of 100 µL of solution was applied to a FLUOTRAC™ 200 96W black polystyrene Microplate (Greiner Bio-One GmbH, Germany). Fluorescence intensity was measured on a Victor3 V multi-label plate counter (PerkinElmer, Waltham, MA). pH ranged from 6 to 2 and was lowered in steps of 0.5 pH units by addition of ortho-phosphoric acid followed by neutralization to pH 6 by addition of 5 M NaOH. The measurements were performed in triplicate and corrected for background.

Real-time adhesion of *H. pylori* by the LigandTracer Green System (Figure S1C) The LigandTracer Green instrument (Ridgeview Instruments AB, Uppsala, Sweden) was used to quantify pH-dependent interaction of *H. pylori* cells to immobilized Leb-constituting receptors in real time.

a) Real-time adhesion to Leb and Lea conjugates (Figures S1D and S1F). The polystyrene Petri dish (Sarstedt, Numbrecht, Germany) was coated with 2 mL of 4 µg/mL Leb-conjugate (receptor) or Lea-conjugate (non-binding control) in pH 9 carbonate buffer in two opposite domains of the dish (target and reference area, respectively) for 12 h at +4°C. The dish was washed with PBS-Tween, blocked for 1 h with 5% BSA PBS-Tween, and washed again. Alexa488-labelled cells of 17875/Leb or 17875 *babA1A2* with OD_{600} = 0.02 were re-suspended in 3 mL of pH 6 citrate-phosphate buffer, placed into the dish, and incubated with the LigandTracer Green with slow rotation of the dish. After 10 h of incubation at room temperature, pH was decreased stepwise by 0.5 pH unit steps by addition of ortho-phosphoric acid to pH 3, and then the pH was reconditioned by stepwise addition of 5 M NaOH. The dish was incubated by rotation for 30 min at each pH step followed by a full hour at pH 3. The measured fluorescent signal represents dynamic changes in the amount of 17875/Leb cells bound to the immobilized Leb conjugate receptors. The signal was background corrected by subtraction of signal from the negative reference, i.e. the Lea-coated reference surface area. The signal was corrected for the pH-dependent changes in the fluorescence of the Alexa488 fluorochrome.

b) Real-time adhesion to Leb-CHO cells (Figures 1G and S1G). Stable clone 2C2 of Chinese hamster ovary cells CHO/K1 expressing human Leb antigen (Leb-CHO) was grown in Dulbecco's Modified Eagle Medium (Gibco, Invitrogen; Paisley, UK) supplemented with 10% heat-inactivated fetal bovine serum (FBS, Gibco, Invitrogen) (Löfling et al., 2008). To obtain monolayer-grown cells, 150,000 freshly split Leb-CHO cells were re-suspended in 3 mL of fresh medium and applied to the positive read-out domain in a Nunclone™ Surface Petri dish (Nunc A/S, Roskilde, Denmark). The opposite non-treated side of the dish was used as the reference. The plate was incubated at 37°C under 7% CO_2 atmosphere in a Sanyo incubator (Sanyo Electric Co., Ltd, Japan) until the confluence of monolayer reached 80%, which was confirmed by light microscopy. Liquid was carefully discarded and rinsed with PBS. Cells were fixed according to (Löfling et al., 2008). The plate was blocked with 5% BSA-PBS for 1 h at room temperature and then washed three times with PBS. The plate was pre-incubated with 3 mL of citrate-phosphate buffer pH 6 for 20 min at room temperature. For the adhesion study, 100 µL of Alexa488-labeled *H. pylori* 17875 of OD_{600} = 1.0 was applied to the buffer in the dish, placed into the LigandTracer Green system (Ridgeview Instruments AB Uppsala, Sweden) and incubated with slow rotation of the dish for 9 h at room temperature. The pH of the bacterial suspension was gradually decreased to pH 2 by ortho-phosphoric acid. pH was next reconstituted to neutral conditions by 5 M NaOH. The fluorescence intensity of bacterial cells attached to Leb-CHO cells or controls were measured in the rotating dish and registered once every minute by the fluorescence detector. At the end of the experiment, the suspension

was aspirated, the Petri dish was rinsed with PBS, and the dish was visualized with the Zeiss microscope to evaluate the specific bacterial and Leb-CHO cell interaction and, in addition, to evaluate background binding. For contrast staining, the Leb-CHO cells were exposed for 20 min to a combination of GelRed nucleic acid stain (Biotium Inc., Hayward, CA, USA) diluted 1:1000 and rhodamine phalloidin (Invitrogen Molecular Probes, Eugene, Oregon, USA) diluted 1:500 in PBS.

c) Real-time adhesion to buccal epithelial cells (BECs) (Figures 1H and S1H). The buccal mucosal surface of an individual with secretor (ABO/Leb)-positive phenotype was scraped with a Cytobrush Plus Cell collector (CooperSurgical. Inc. Berlin, Germany) (Strömberg and Borén, 1992). The brush was swirled in PBS-Tween, and BECs were washed three times in PBST under gentle vortexing. For immobilization to the Petri dish, a new method was developed. BEC cells were first counted with a Bürker counting chamber, and 350,000 cells were re-suspended in 3 mL of PBS with 50 µL of DMSO containing 50 µg of biotin-amido-hexanyoyl-6-aminohexanoic acid N-hydroxy-succinimide ester (biotin) (Sigma, Steinheim, Germany) and incubated on a rotary shaker for 45 min at room temperature. Biotinylated BECs were then washed three times with PBS and blocked twice by 15 min incubation with 5% BSA followed by three washes with PBS. The final cell pellet was re-suspended in PBS. A Sarstedt Petri dish (Sarstedt AG & Co., Numbrecht, Germany) was prepared in advance by coating a local part/domain of the dish with 3 mL of PBS mixed with 100 µg of MegaCell^{R}-Streptavidin (Cortex Biochem, San Leonardo, CA, USA) and incubated overnight at +4°C. Non-attached beads were discarded, and the dish was gently rinsed with PBS. A total of 150,000 biotin-labeled BECs were added to the plate and incubated in PBS at room temperature for 2 h. Unbound cells were discarded, and the rest were fixed for 20 min with 1% aqueous solution of paraformaldehyde (PFA). The dish with attached cells was blocked overnight at +4°C and rinsed again with PBS. The plate was pre-incubated with 3 mL of phosphate-citrate buffer pH 6 for 20 min at room temperature and then placed into the LigandTracer Green system. A total of 100 µL of OD_{600} = 1.0 Alexa488-labeled *H. pylori* 17875/Leb was added to plate and incubated for 9 hours at room temperature with slow rotation of the dish. The pH gradient from pH 6 to pH 2 was performed as described previously. For the duration of the experiment, the fluorescence intensity of bacterial cells attached to the BECs or control (opposite) areas were measured every minute during rotation of the Petri dish. Microscopic evaluation of the dish was performed with a Zeiss microscope with BECs stained with DAPI.

2. The BabA protein Confers Polymorphic Acid Sensitivity to *H. pylori* Leb Binding

SDS-PAGE and immunoblot detection of BabA. SDS-PAGE was performed with 7.5% Tris-HCl Ready Gels (Bio-Rad Laboratories, Hercules, CA, USA). Protein extracts in Laemmli sample buffer with 2% β-mercaptoethanol were boiled for 10 min before loading on gels. For immunoblots, proteins were transferred to a polyvinylidene difluoride (PVDF) membrane (Bio-Rad, USA) and blocked with 5% skim milk. Polyclonal AK 277 (Odenbreit et al., 2002) diluted 1:6000 was used as primary antibody, followed by HRP-goat α-rabbit antibody diluted 1:1000 (DakoCytomation, Denmark A/S). Signals were developed with ECL chemiluminescence (SuperSignal West Pico, Thermo Scientific/Pierce, IL, Rockford, USA). For visualization and quantification of the signal, the Kodak Image Station 2000R (Eastman Kodak Company, Rochester, NY) and Kodak 1D Image Analysis Software v.4.0.5f7 (Kodak) were used. The 78 kDa BabA protein migrates in SDS-PAGE with a molecular mass of 75 kDa. SDS-PAGE staining was performed with PageBlue™ Protein Staining Solution (Thermo Scientific, Rockford, IL, USA) and SilverXpress Silver Staining Kit (Invitrogen/ Life Technologies, Carlsbad, CA, USA).

Competitive real-time adhesion of *H. pylori* strains of different acid sensitivity of Leb-binding measured by the LigandTracer Green System (Figures 2B and S2C). Two *H. pylori* strains, SW7 (acid resistant) labeled with CF555 and SW38 (acid sensitive) labeled with Alexa488 were re-suspended in 3 ml of pH 6 citrate-phosphate buffer, each to OD_{600} = 0.02. Application of bacterial suspension to the Petri dish with immobilized Leb and Lea conjugates was performed as described above. Following 10 h of incubation with rotation of the dish in the LigandTracer at room temperature, the pH was reduced by the addition of ortho-phosphoric acid stepwise by one pH unit steps and 15 min incubation at each step until pH 2 and, after 30 min incubation, reconditioned stepwise back to pH 6 by addition of 5 M NaOH. The

dish was further incubated until 17 h from the beginning of the experiment. The Leb and Lea areas of the dish were examined with the fluorescent microscope both at pH 6 (the 10 h time point) and reconditioned pH 6 (the 17 h time point). The numbers of SW7 and SW38 cells attached to the Leb-coated area were quantified with the ImageJ v.1.47 software (http://imagej.nih.gov/ij) in 7–10 field views and were corrected by the background binding to the Lea-coated area. The data are presented as means ± SD. The numbers of attached cells (sqrt-transformed) were compared by two-way ANOVA with Bonferroni post-test comparisons in GraphPad Prism 5.0.

BabA-Leb binding by "Far-Western blot overlay" at different pH (Figure 2C). Bacterial cells of strain 17875/Leb were treated with non-reducing SDS sample buffer (2% SDS, 5% glycerol, 40 mM Tris, pH 6.8). Solubilized proteins were size-separated by SDS-PAGE and Western transferred to a PVDF membrane according to (Ilver et al., 1998; Mahdavi et al., 2002; Aspholm et al., 2006) with minor changes. To refold the immobilized BabA protein and reconstitute Leb binding, the membrane was treated overnight with 10 mM n-Octyl-ß-D-Glucopyranoside (abbreviated octyl-glucoside) (Alexis Biochemicals Corporation, Göteborg, Sweden) in citrate-phosphate buffer pH 5. The membrane was applied to the Mini-Protean II Multiscreen, and the series of channels of the chamber were filled with octyl-glucoside buffers at pHs ranging from 2 to 6 along with 2 µg/mL of biotinylated Leb conjugate (Aspholm et al., 2006), and this was incubated in the multiscreen chamber overnight and washed with TBST buffer (50 mM Tris-HCl, 150 mM NaCl, and 0.05% (v/v) Tween 20, pH 7.4). Signal from bound Leb conjugate was developed with HRP-streptavidin and ECL-chemiluminescence. Each experiment was carried out twice with equivalent results. The 78 kDa BabA protein migrates in non-reduced SDS-PAGE as a molecular mass of 65 kDa.

Stability of the BabA protein expression under conditions of pH shift (Figures S2D).
H. pylori strain SW109 grown on blood agar plate was used to inoculated Brucella broth containing 1% Isovitox and 10% fetal calf serum (Gibco) to an OD_{600} of 0.01. The culture was grown for 48 h, to OD_{600} = 0.4 at 37°C (5% O_2, 10% CO_2, and 85% N_2) before the culture was split in three parts. The pH of the cultures was adjusted by addition of a set amount of HCl to pH 4 or pH 2, and as a control, equal amount of sterile water was added to the pH 7 cultures. The cultures were left to grow for additional 24 h before samples were collected and analyzed by immunoblot analysis. As a control samples were also collected before the pH shift and the pH of the cultures was checked after 24 h. SDS PAGE analysis was made using NuPage Novex Bis-Tris protein gels (4-12% or 12%) using 1x MOPS SDS as running buffer (Life Technologies). The gels were either stained with PageBlue Protein stain (Thermo Scientific) or transferred to PVDF membrane using Trans-Blot SD Semi-Dry Transfer Cell (Bio Rad). Immunoblot analysis was performed as previously described (Bäckström et al, 2004). Antibodies against SabA (AK278) and BabA (AK253) (Odenbreit et al., 2002) were used in combination with secondary α-rabbit IgG-HRP (P0160, DAKO A/S, Denmark). Blots were developed with SuperSignal (Pierce, Rockford, IL) ECL and detected on High Performance Chemiluminescence film (GE Healthcare). Protein densities were measured by ImageJ software (NIH) and normalized to the corresponding PAGE Blue stained SDS-PAGE gels. The expression level in the pH 7 culture was set to 1 and relative protein expression was quantified and plotted.

Replacement of *babA* in strain SW38 with *babA* from strain SW7 (Figures 2D and S2E). The *babA* gene of strain SW38 was first deleted by transformation with genomic DNA from strain 17875*babA1*::kan *babA2*::cam (both *babA* genes replaced with drug-resistance determinants) and selection for chloramphenicol-resistant transformant colonies (Mahdavi et al., 2002). Six representative Cam^R colonies were purified. Loss of BabA synthesis and of Leb binding was confirmed by BabA-immunoblots and RIA analysis, respectively, in each of these six transformants. One of these SW38 derivatives was used as the recipient for a second transformation with genomic DNA from strain SW7. In this case the bacterial-DNA mixture was incubated overnight on non-selective plates and then suspended in Brucella broth. The suspension was incubated with 7 µg of biotinylated Leb conjugate for 2 h, and Leb-binding bacterial cells were obtained by three sequential cycles of biopanning to progressively enrich for the desired transformants (Aspholm et al., 2006). Each biopanning cycle entailed a 2 h incubation with 20 µL streptavidin-coated magnetic beads (Qiagen, Hilden Germany) in 2 mL Brucella broth, washing the beads with Brucella broth containing 1% BSA, spreading the beads with attached bacteria on non-selective agar

growth medium, and incubating overnight. In the final plating, the bacteria-loaded beads were spread at a density of about 5,000 colony-forming units per plate and were incubated until single colonies were visible. To identify colonies with Leb binding activity, replicas were made on nitrocellulose membranes, which were then incubated for 2 h with 2 µg/mL biotinylated Leb conjugate in 50 mM Tris, 150 mM NaCl, and 0.05% (v/v) Tween-20, pH 7.4 (TBST) buffer supplemented with 1% BSA. The membranes were washed in TBST, and Leb-binding colonies were detected by incubating the membranes with peroxidase-conjugated streptavidin in TBST with 1% BSA for 1 h followed by the addition of 4-chloro-1-naphtol tablets (Sigma Aldrich, St. Louis, MO, USA). The Leb-binding clones were examined for the presence of *babA* from strain SW7 by sequencing. Chromosomal DNA was prepared from the Leb-binding clones, and the central part of *babA* was amplified by PCR and sequenced as described (Aspholm-Hurtig et al., 2004; Backström et al., 2004; Colbeck et al., 2006) above for sequence analysis of *babA* and *babB*. One such transformant, made using SW38ΔbabA as the recipient and SW7 as the donor (SW38*babA*SW7), was denoted Trans38-7.

Native BabA protein purification (Figures 2E, S2G and Table S1). *H. pylori* 17875/Leb was grown on plates at 37°C for 35 h under microaerophilic conditions as described above. Cells were harvested in 30 mM phosphate buffer pH 6.7, frozen in liquid nitrogen, and stored as a paste at −80°C until used. A total of 3.4 grams of cell paste was suspended overnight at 4°C in 500 mL detergent buffer (30 mM Hepes (4-(2-Hydroxyethyl)-1-piperazineethanesulfonic acid) (Roche Diagnostic GmbH, Mannheim, Germany), 30 mM 3-(dodecyldimethyl-ammonio) propesulfonate (abbreviated ZW 3-12) (Sigma Aldrich, St. Louis, MO, USA), 50 mM NaCl, pH 8, and 0.05 U/µL of DNase I (Invitrogen/Life Technologies, Stockholm, Sweden)). Insoluble material was pelleted by centrifugation at 25,000 rpm for 1 h. The supernatant was filtered through 0.22 µm Durapore membrane filters (Millipore, Carrigtwohill, Co. Cork, Ireland) and then diluted with 250 mL of 30 mM ZW 3-12. The membrane protein extract was applied to a Source 30 S (GE Healthcare, Uppsala, Sweden) cation exchange column with flow rate of 1 mL/min and bed volume of 180 mL equilibrated with 30 mM Hepes and 30 mM ZW3-12, pH 8. BabA protein was eluted with a linear gradient of 0–200 mM NaCl in 30 mM Hepes and 30 mM ZW 3-12, pH 8. Fractions containing BabA were identified by immunoblotting with BabA AK277 antibody diluted 1:6,000. BabA-containing fractions were pooled, diluted 10-fold v/v with 30 mM Hepes and 30 mM octyl-glucoside, pH 8, and affinity purified using the Leb-HSA conjugate as the receptor. The Leb-affinity matrix was constructed by chemical immobilization of 35 mg of Leb-conjugate to 20 mL of NHS-activated Sepharose 4B Fast Flow matrix (GE Healthcare, Uppsala, Sweden). The Leb-column was equilibrated with 30 mM Hepes, 30 mM octyl-glucoside, pH 8.0. The pooled BabA fractions were applied to the Leb column at 1 mL/min. Bound BabA protein was eluted in a linear pH 8–3 glycine gradient buffer. BabA eluted from the column at pH 3.6–3.4, and the BabA-containing fractions were quickly neutralized (pH 6). All chromatography procedures were conducted at 4°C. Fractions containing BabA were identified by immunoblotting with BabA AK277-antibodies diluted 1:6000. One major protein band of the expected 75 kDa was seen on SDS-PAGE under reducing conditions. BabA was isolated in parallel from strains SW38 and SW7 with the same method and with the same yield. For analysis of levels of BabA protein in the different purification steps, samples were separated by SDS-PAGE, immunoblotted, and probed with Ab-AK277. BabA signals were quantified with Kodak 1D Image Analysis Software (Eastman Kodak Company, Scientific Imaging System, Rochester, NY, USA) and related to pure BabA diluted 1:100 in octyl-glucoside buffer, pH 8, as the reference sample. Protein concentrations were assessed by Nanodrop ND-100 at 280 nm (NanoDrop Technologies, In., USA) and by BCA Protein Assay Kit (Pierce/Thermo Scientific). All BabA preparations were stored in aliquots at −80 °C.

Native BabA protein analysis by MALDI-TOF mass spectrometry (Figure S2H). The mass spectrometry analysis of the purified BabA protein was performed on a Voyager DE-STR MALDI-TOF instrument (AB Sciex, the KBC Proteomic Facility, Umeå University) with sinapinic acid as the matrix.

pH-sensitive immunohistostaining by purified BabA protein (Figures 2F and S2I). Deparaffinized and rehydrated human gastric mucosal tissue sections were blocked for 3 h with 2% BSA in PBS and 0.05% (v/v) Tween-20 (PBST). Tissues were washed briefly in 0.15 M NaCl and 0.005% (v/v)) Tween-20. A 2 µg/mL solution of affinity-purified BabA from strain 17875 was added to tissues in citrate-

phosphate buffer pH 6.0 overnight at 4°C. As a negative control, 2 μg/mL BabA was mixed with 40 μg/mL Leb-HSA conjugate for 1 h in PBST prior to binding sections. For the pH 2→6 experiment, BabA was diluted to 20 μg/mL in citrate-phosphate buffer pH 2, incubated for 10 min, and then further diluted 1:10 with pH 6.6 citrate-phosphate buffer to reach pH 6 and a final BabA concentration of 2 μg/mL. Tissues were washed three times for 10 min with citrate-phosphate buffers at pH 6, pH 4, and pH 2 citrate-phosphate buffer, and then briefly with 0.15 M NaCl and 0.05% (v/v) Tween-20 to ensure that residual acidity did not interfere with antibody binding. The slides with bound BabA were incubated with the BabA-specific antibody Vite (Odenbreit et al., 2002) diluted 1:4000 in PBST and incubated for 3 h at room temperature. Biotinylated-mouse-anti-rabbit (Sigma, Stockholm, Sweden) diluted 1:2000 in PBS and 0.005% Tween-20 was added, and the incubation was continued at room temperature for 1 h. To visualize bound BabA, Alexa555-Streptavidin (Molecular Probes) diluted 1:100 in PBST was added for 30 min. Slides were then incubated with 0.5 μg/mL DAPI (Sigma, Stockholm, Sweden) in PBST for 7 min. BabA binding was digitized at 200× magnification with a Zeiss AXIOcam MRm. Digital quantification was done on three visual fields/tissue, and nine 0.3-μm thick Z-stack ApoTome grid projections were merged. The mean grey-value (intensity) of 15–40 equally sized areas along the epithelial lining of the merged images was measured. Each image was quantified three times to ensure reproducibility in measurements (Figure S2I). Background intensity from the negative (Leb treated) control was subtracted. All quantification was done using ImageJ for Macintosh v 1.39u (http://rsbweb.nih.gov/ij/).

Leb binding by native BabA protein at different pHs with enzyme-linked immunosorbent assay (ELISA) detection (Figure 2G). ELISA was used to measure binding of Leb-HSA conjugate to purified BabA. Wells of a flat bottom NUNC-Immuno MaxiSorp plate (Nunc A/S, Roskilde, Denmark) were coated with 25, 100 or 300 ng of BabA from strains 17875, SWI7 and SW38, respectively. Protein was diluted in 10 mM octyl-glucoside, 150 mM carbonate buffer, pH 9, and applied to the plate at room temperature for 2 h. Wells were washed with PBST, and blocked overnight at 4°C with PBST and 2% BSA. Leb-HSA conjugate was biotinylated according to (Aspholm et al., 2006) and diluted in citrate-phosphate buffers containing 10 mM octyl-glucoside at pH ranging from 2 to 6. Plates were incubated with 25 ng of Leb-HSA conjugate/well for 2 h at room temperature, washed five times with PBST, and incubated for 1 h with peroxidase-labeled streptavidin diluted 1:1000 (Roche Diagnostics Scandinavia AB, Bromma, Sweden). The reactions were developed with orthophenylene diamine hydrochloride (Sigma) and analyzed at A_{490} with a SpectramaxPlus microplate spectrophotometer (Molecular Devices, Sunnyvale, CA, USA). Negative control wells contained all components except BabA.

Alterations in native BabA protein folding followed by circular dichroism (CD) and fluorescent spectrophotometry (Figures S2J) Secondary and tertiary structural changes of full-length purified native BabA induced by acidic pH were analyzed by circular dichroism (CD) with a Jasco J810 spectropolarimeter (Jasco Cooperation, Tokyo, Japan). Path length of the cuvette was 0.1 cm for far-UV and 0.5 cm for near-UV light measurements. Far-UV CD was performed at 190–250 nm and the final spectrum consisted of 8 accumulated spectra, whereas near-UV CD were performed at 250–310 nm and were accumulated 20 times. The assay mixture contained purified BabA in citrate-phosphate buffer, pH 7, with 10 mM octyl-glucoside. pH was reduced from pH 7 to pH 2.5 by 28% ortophosphoric acid. The pH was measured by indicator strips (Merck KG, Darmstadt, Germany). Effects of pH on CD were followed for 15 and 30 min.

Detection of acid dependent secondary structural transitions of recBabA$_{526}$ (Figures S7B) from clones 81G antrum and 81G corpus was analyzed by Far-UV spectra circular dichroism (CD) with a Jasco J810 spectropolarimeter (Jasco Cooperation, Tokyo, Japan). A series of pH titration buffers with pH 5.5, 4.5, 3.5 and 2.5 was used described in "1. Acid Inactivation and Reversal of Leb Binding / Buffers used for pH titration".

Fucolectins-H2 binding by RIA at different pH (Figure S2K). H2-BSA conjugate (H2 conjugate) was obtained from IsoSep AB, Tullinge, Sweden, and was ^{125}I-labeled (^{125}I-H2-conjugate) by the chloramine T method (Aspholm et al., 2006). Purified *Anguilla anguilla* (AAA) lectin and *Ulex europaeus* (UEA-I) lectin were covalently linked to an agarose gel and obtained from EY Laboratories (EY

Laboratories Inc., San Mateo, CA, USA). AAA and UEA-I were diluted in 0.15 M NaCl, 0.05% Tween-20 to 20 µL/mL and mixed with 25 ng/mL of ^{125}I-H2-conjugate in nine different citrate-phosphate buffers for 30 min at room temperature. The pH in all vials was then re-constituted to pH 5 by a corresponding series of buffers, and after a 3 h incubation at room temperature the amount of bound/free ^{125}I-conjugates was measured. A reconstituted pHgram of 17875/Leb binding to ^{125}I-Leb-conjugate was used as shown in Figure 1E as the control.

3. Gastric Corpus/Antrum BabA Differences in Acid-Sensitivity in Leb Binding

Cloning, expression, and purification of recombinant SO-2 BabA$_{526}$ in *E. coli* (applied in Figures 3D and S3C). SO-2 antrum and corpus *babA* genes were PCR-amplified from chromosomal DNA with Phusion DNA Polymerase (Thermo Scientific) using primers NdeI_O2babA_F and NotI_O2babA_R, and the chimeric Pro199 to Leu199 (called LG) and Gly431 to Glu431 (called PE) corpus mutants were made by PCR site-directed mutagenesis using primers O2_PtoL_F and O2_PtoL_R or O2_GtoE_F and O2_GtoE_R, respectively. Fragments were run on a 1% agarose gel and purified using the E.Z.N.A Gel Extraction Kit (Omega Bio-Tek). Antrum, corpus, LG, and PE *babA* amplifications were used as templates for amplification of the *babA*$_{547}$ using the Phusion DNA Polymerase (Thermo Scientific) with primers 5'-babA-NcoI and 3'-babA547-BamHI. Final PCR products were digested with NcoI and BamHI, cloned into the pOPE-vector (Schmiedl et al. 2000; Jones et al., 1990) (GenBank Y14585), and verified by DNA sequencing at MWG Operon, Germany.
The following primers were used for this cloning:
NdeI_O2babA_F (5'GATCCATATGAAAAAAACTCTTTTACTCTCTC)
NotI_O2babA_R (5'CAAGGCGGCCGCTTAATAAGCGAACACATAGTTCAAATAC)
O2_PtoL_F (5'AGGGGAAGGGAATGACAACTGCTCGCTACGAGCCACAGGTGTAGGTAAAC)
O2_PtoL_R (5'GTTTACCTACACCTGTGGCTCGTAGCGAGCAGTTGTCATTCCCTTCCCCT)
O2_GtoE_F (5'CTTTCGCTTCCGGTTGCGCGTATGTGGAACAAACCATAACAAATCTAA)
O2_GtoE_R (5'TTAGATTTGTTATGGTTTGTTCCACATACGCGCAACCGGAAGCGAAAG)
3'-babA547-BamHI (5'TCGTATGGATCCGTTACGCCCTAATTCTTGGTTG)
The pOPE-*babA* plasmid was transformed into the *E. coli* strain XL10 Gold (Stratagene, USA), plated on selective LB plates supplemented with 100 µg/mL carbenicillin, 12.5 µg/mL tetracycline, and 0.1 M glucose and incubated overnight at 37°C. A single colony was inoculated into 50 mL pre-warmed LB supplemented with 100 µg/mL carbenicillin, 12.5 µg/mL tetracycline, and 0.1 M glucose (hereafter called LB$_{CTG}$) and grown in a baffled flask overnight at 37°C (~150 rpm). A total of four liters of LB$_{CTG}$ was divided into six to eight baffled 2 L flasks and pre-warmed to 37°C before 1 mL of the overnight culture was used to inoculate each flask. Cultures were grown at 37°C (~150 rpm) until OD$_{600}$ ~0.5 was reached. The cultures were immediately cooled down on ice water, and protein expression was induced by the addition of IPTG to a final concentration of 75 µM. Protein expression was carried at 20°C for ~16 hours. Bacteria were pelleted at 5000 × *g* and re-suspended in 200–250 mL ice-cold spheroblast buffer (50 mM Tris-HCl pH 8.0, 20% sucrose, 1 mM EDTA) and shaken for ~2 h at 4°C on a rocking table. From here on, all work was done on ice or at 4°C. The suspension was centrifuged at 30,000 × *g* for 45–60 minutes, and the supernatant, representing the periplasmatic extract, was dialyzed twice against 5 L of PBS overnight. The extract was then passed through a 0.2 µm sterile filter (Nalgene, USA), and the concentrations of NaCl and imidazole were adjusted to 500 mM and 25 mM, respectively. The extract was loaded using a peristaltic pump at ~1 mL/minute into a 1 mL HisTrap Ni-NTA column (GE Healthcare, USA) pre-equilibrated with 10 mL of binding buffer (PBS, 500 mM NaCl and 25mM imidazole). The column was washed with 10 mL binding buffer before elution of protein with elution buffer (binding buffer with 500 mM imidazole). Fractions (0.5 mL) were collected and the protein content was recorded at A$_{280}$ nm with a NanoDrop 1000 spectrophotometer. Protein-containing fractions were pooled and dialyzed twice against 2 L of PBS overnight. After dialysis, any precipitated protein was removed by Eppendorf-scale centrifugation at 18,000 × *g* for 20 minutes, and the protein concentration was determined at A$_{280}$ nm on a NanoDrop 1000 spectrophotometer. The typical yield was 1–2 mg of soluble protein per liter of culture. Protein aliquots were frozen in liquid nitrogen and stored at −80°C. A small sample from each

protein preparation was used to determine the quality of the preparations (PageBlue-stained polyacrylamide gels). If necessary, proteins were concentrated by centrifugation in ultrafiltration columns (Vivaspin 500, 10,000 MWCO PES, Sartorius Vivascience Biotech, Göttingen, Germany). Recombinant proteins were separated on a Superdex 200 16/60 (GE Healthcare) at 1 mL/min in PBS to isolate monomers.

Binding analysis of SO-2 BabA$_{526}$ by Biacore (Figure 3D: Figure S3C). Sensorgrams were recorded at 25°C using a Biacore 3000 instrument (GE Healthcare). BabA monomers were isolated by size exclusion chromatography with a Superdex200 16/60 (GE Life Science) column in PBS with a flow rate of 1 mL/min at 4°C. The protein was detected by measuring the absorbance at 280 nm. The monomeric fractions were pooled and stored at −80°C until further use. Gel filtration-separated monomeric BabA was immobilized on an NTA chip (GE Healthcare) with standard amine coupling chemistry as follows. The NTA chip was loaded with 0.1 M $NiSO_4$ at a flow rate 20 µL/min for 5 min. The NTA chip was activated with NHS/EDC at a flow rate of 5 µL/min for 7 min. Monomeric BabA in PBS separated by gel filtration was used for immobilization to a level of 7000–8000 RU and thereafter deactivated with ethanol amine for 7 min. The use of an NTA chip allowed for immobilization in PBS buffer and the use of the His-tag on BabA$_{547}$ to concentrate the protein on the surface, which prevented possible protein aggregation due to the low pH buffer that is commonly used for amine-coupling immobilization. The interaction of BabA and Leb was studied by passing 10 µg/mL Leb-PAA over the chip 10µg/mL at a flow rate of 20 µL/min and a 4 min injection. To regenerate the surface, the flow rate was increased to 100 µL/min followed by an injection of 25 µL of diethylamide (diluted 1:100). The relative binding was calculated by dividing the maximum signal at different pHs by the maximum signal at pH 6. For 81G antrum and corpus isolates, BabA$_{1\text{-}526}$ binding was analyzed at 35°C, and binding at pH 5.5 was used as the reference.

pH dependent BabA mediated binding to Leb conjugate by ProteOn (Figure S3D). Surface Plasmon Resonance (SPR) analysis was performed on ProteOn XPR36 (Bio-Rad, CA, USA). Leb-HSA was immobilized at pH 4.0 to a GLM chip (Bio-Rad) according to the manufactures instructions. As control Lea-HSA was used as reference surface. The experiment were performed in PBS with 0.05% Tween-20 at 25°C. *H.pylori* was injected at OD 0.2, 0.1 and 0.05 for ten minutes at a flow rate of 25µL/min followed by 10 minutes dissociation phase. Unspecific binding was corrected by subtracting the Lea-HSA surface. An injection with 0.5% SDS was used for regeneration.

Expression of full-length 720aa BabA in the outer membrane of *E. coli* (Figure 3E). The experiments were performed according to (Fei et al., 2011) with minor modifications. BabA was cloned from genomic DNA of SO-2 (corpus) with primers:
1539-F (5'-ATG AAAAAAACTCTTTTACTCTCTCTC-3')
1539-R3 (5'-GGTTGATTGTTCATGGCT ATTAATG-3')
1539-R (5'-TTAATAAGCGAACACATAGTTCAAATAC-3')
1539-F3 (5'-CATTAATAGCCATGAACAATCAACC-3').
Fragments were stitched together, and restriction sites for the cloning were added with primers:
5'-babA-NcoI (5'-GAATAGGGCCATGGCGGAAGACGACGGCTT TTACAC-3')
3'-babA-BamHI (5'-CCGATAGGGATCCATAAGCGAACACGTAAT TC-3').
babA was digested with NcoI and BamHI and cloned into the pOPE-vector (GenBank Y14585). The pOPE101-vector, originally designed for the expression of antigen binding variable regions (VH and VL) from hybridoma cells (Jones et al., 1990), contains a pelB-leader sequence immediately followed by a *Nco*I-*Bam*HI cloning cassette where the different variants of *babA*$_{547}$ have been inserted. An IPTG-inducible lac promoter/operator region upstream of the pelB leader sequence controls expression from the plasmid. Downstream of the *Bam*HI-site is a c-*myc* tag for detection of the recombinant protein, and this tag is followed by a $(His)_6$ tag that aids in the purification of the recombinant protein. The vector has a carbenicillin resistance marker. Sequencing confirmed an insert of the correct sequence. Mutagenesis of corpus *babA* in the pOPE-vector was performed using the QuickChange Multi Site-Directed Mutagenesis Kit from Stratagene (AH diagnostics, Stockholm, Sweden) according to manufacturer's recommendation with following primers: Pro to Leu (5'-CACCTGTGGCTCGTAGCGAGCAGTTGTCA-3')

and Gly to Glu (5'-TTTTAGATTTGTTATGGTTTGTTCCACATACGCGCAACCGGAAG-3'). The correct mutations were confirmed by sequencing. The pOPE-*babA* plasmid was transformed into the *E. coli* strain XL10 Gold (Stratagene, USA) and plated on selective LB plates supplemented with 100 µg/ml carbenicillin, 12.5 µg/ml tetracycline, and 0.1 M glucose and incubated overnight at 37°C. A single transformant from a warm transformation plate was inoculated into 50 ml pre-warmed LB supplemented with 100 µg/mL carbenicillin, 12.5 µg/mL tetracycline, and 0.1 M glucose (hereafter called LB_{CTG}) and grown in a baffled flask overnight at 37°C (~150 rpm). Four liters of LB_{CTG} were divided into six to eight baffled 2 L flasks and pre-warmed to 37°C before 1 mL of the overnight culture was used to inoculate each flask. Cultures were grown at 37°C (~150 rpm) until an optical density (A_{600} nm) of ~0.5 was reached. The cultures were immediately cooled down on ice water while the temperature of the incubator was reduced to 20°C. Protein expression was induced by the addition of IPTG to a final concentration of 75 µM for ~16 h. Bacterial cells that expressed full-length recombinant BabA were pelleted and washed three times in 0.15 M NaCl and 0.05% Tween (v/v), and the concentration was adjusted to OD_{600} = 10. This suspension was used to perform RIA.

4. Correlation Between BabA Acid Sensitivity in Binding and the BabA Structure

Prediction of BabA secondary structure (Figures 4A and S4A). The BabA protein consensus secondary structure prediction was performed with the following prediction servers:
PRISPRED v. 3.3 (http://bioinf.cs.ucl.ac.uk/psipred/) (Jones, 1999),
Predictprotein (https://www.predictprotein.org) (Yachdav et al., 2014),
PREDATOR v.2.1.2 (http://mobyle.pasteur.fr/cgi-bin/portal.py? - forms::predator) (Frishman and Argos, 1997).

Prediction of intrinsically unstructured regions (IURs) (Figures 4A) of the BabA protein was based on integrated results obtained from DisEMBL™ v. 1.5 (http://dis.embl.de) (Linding et al., 2003a) and GlobPlot 2 (http://globplot.embl.de) (Linding et al., 2003b).

5. Correlation Between BabA Acid Sensitivity and Geographic Disease Patterns.

Phylogenetic tree based on 98 full-length BabA amino acid sequences (Figure S5A). Sequences were retrieved from UniProt http://www.uniprot.org and aligned with MegAlign Pro in the Lasergene software package (DNASTAR, Inc. Madison, USA). The phylogenetic tree was prepared with Tree View and the BIONJ algorithm. Pairwise distances between and within the sequences were calculated with Kimura parameters.

Prevalence of amino acids in the Key-position 199 (Figure 5C and Table S2). These figures are based on BabA sequences from several different populations (Aspholm et al., 2004) and the Indian and Greek sequences from this paper. The amino acid prevalence in the Key-position among different countries and native populations are depicted by the relative size of the amino acid's 1-letter code.

Shuttle vector design genetic complementation by conjugation, and recombinant expression of full-length BabA in *H. pylori* (Figures 5E-5G and S5B). The *H. pylori* shuttle vector pIB6 is based on the pHel3 plasmid (Heuermann and Haas, 1998), which has been further modified to contain the *alpA* promoter upstream of a multiple cloning site were the various *babA* alleles were introduced (Jiménez-Soto et al 2013). For construction of pIB6, plasmid pSO97 was used as a template to delete the *alpA* and *alpB* structural genes by inverse PCR using primers SO157 and SO158 flanked by NdeI and NotI restriction sites, which resulted in plasmid pIB1. The complementary oligonucleotides IB13 and IB14 were used to insert a multiple cloning site between the NdeI and NotI site of pIB1, which yielded pIB6 (Jiménez-Soto et al., 2013). Indian I9 and I18 *babA* genes were PCR-amplified from chromosomal DNA with Phusion DNA Polymerase (Thermo Scientific) using primers Indian_BabAF with either I9A_BabAR for I9 or I18_BabAR for I18. Fragments were run on a 1% agarose gel and purified using the E.Z.N.A Gel Extraction Kit (Omega Bio-Tek). I9 Glu to Gly and I18 Gly to Glu mutants were made by PCR site-directed mutagenesis using primers I9A_F_Gly and I9A_R_Gly for I9 or I18_F_Glu and I18_R_Glu for I18. Fragments were cloned in pIB6 using NdeI and NotI restriction sites to make up shuttle vector plasmids I9-SV, I18-SV, I9M-SV, and I18M-SV. These shuttle plasmids as well as shuttle vector

plasmids I9A-M1, I9A-M2, I9A-M3, I18deltaKQ and I9D198A, which were obtained by site-directed mutagenesis of the I9-SV plasmid and I18-SV plasmid, respectively, at Genscript, (NJ, USA), were transformed into a *dapA*-negative *E. coli* strain, β2150. This "donor" strain strictly depends on exogenously supplied diaminopimelic acid; hence, the removal of diaminopimelic acid provides an efficient counter-selection against this donor. All "donor" strains with the various *babA* shuttle vector plasmids were conjugated into the *H. pylori* strain P1Δ*babA*. Plasmids from obtained *H. pylori* clones were re-isolated, and the full *babA* sequence of one representative clone from each mutant type was sequenced at MWG Operon, Germany, using primers ShuttleVector_F, ShuttleVector_R, and I1F. All tested clones showed the correct sequence. At least three clones of each mutant type were later used for pH-titrations of Leb-binding and showed similar pH tolerance.
The following primers were used:
SO157: (5'GATCGTCGACCATATGTTTCCTTATCAATGGGATGC)
SO158: (5'GATCGTCGACGCGGCCGCAAGCTCAAGGCCTTTTATAGG)
IB13: (5'TATGGTCGACATTGGTAACTAGTAGATCTGC)
IB14: (5'GGCCGCAGATCTACTAGTTACCAATGTCGACCA)
ShuttleVector_F: (5'CAAGCTCTGATGTAACTAAAT)
ShuttleVector_R: (5'AAGGGTTTTGGGTTTGTTTTTA)
I1F: (5'CACTTGCTCAGGGGAAGGGAA)
Indian_BabAF (5'GATCCATATGAAAAAACGCATCCTT)
I9A_BabAR (5'CAAGGCGGCCGCTTAGTAAGCGAACACATAAT)
I18_BabAR (5'CAAGGCGGCCGCTCAGTAAGCGAACACGTAAT)
I9A_F_Gly (5'TCGCTTCCGGTTGTGCCTATGTGGGGCAAACCATAACGAATCTAAAAAACA)
I9A_R_Gly (5'TGTTTTTTAGATTCGTTATGGTTTGCCCCACATAGGCACAACCGGAAGCGA)
I18_F_Glu (5'TCGCTTCCGGTTGCGCCTATGTGGAACAAACCATAACGAATTTAAACAACA)
I18_R_Glu (5'TGTTGTTTAAATTCGTTATGGTTTGTTCCACATAGGCGCAACCGGAAGCGA)

Dynamic force spectroscopy (Figures 5H and 5I). Rupture force measurements were conducted on single BabA-Leb bonds on living bacterial cells by means of dynamic force spectroscopy in a force measuring optical tweezers (FMOT). The FMOT-system was based on an inverted Olympus IX71 microscope with a high numerical aperture oil-immersion objective lens (Olympus UPlanF1 100x) and a continuous wave Nd:YVO_4 laser (Spectra-Physics, Millennia IR10) for optical trapping. A more detailed description of the measurement system is given elsewhere (Fällman et al., 2004; Andersson et al., 2006; Björnham et al., 2009). Large polystyrene beads with a diameter of 9.5 µm were immobilized on a coverslip by means of a drying procedure and thereafter coated with poly-L-lysine. A sample of 25µL buffer including small beads (diameter of 3.2 µm) and bacterial cells of strain 17875/Leb at a low concentration was placed onto the coverslip and inserted in the FMOT. For pH 7.4 the buffer was PBS with 1% Bovine Serum Albumin (BSA), for pH 3.6 – 5.5 a citrate buffer (15 mM citrate, 15 mM phosphate, 0.1% BSA and NaCl 0.15 ion strength) was used. A single bacterium cell was first mounted on the waist of a large bead where after a small Leb-coated bead was trapped by the FMOT and carefully brought into contact with the bacterium to create a single BabA-Leb bond. The actual measurement was then executed by retracting the coverslip, with the large bead and the mounted bacterium, whereby the force in the system was recorded until the bond ruptured (Björnham et al., 2009). As a relevant control for specificity of BabA-Leb binding the 17875*babA*-null mutant devoid of BabA, was used to assess non-specific binding.

6. Adaptation of BabA Acid Sensitivity During Disease Progression Into Gastric Cancer.

H. pylori macaque **output strain used for analysis of BabA-mediated Leb-binding by Alexa555-Leb conjugate.** The 17-week macaque-passaged (output) strain J166, defined as "*H. pylori* 17 week infection" (Solnick et al., 2004) was tested for binding to Alexa555-labeled Leb-conjugate with *H. pylori* CCUG17875 (positive) and 17875*babA1A2*-mutant (negative) controls.

Labeling of Leb-HSA conjugate by Alexa555 fluorochrome (used in Figure S6B). A total of 0.5 mg of Leb-HSA conjugate was dissolved in 100 µL of 0.1 M sodium bicarbonate buffer, pH 9.0. One mg of Alexa555 was dissolved in 100 µL of DMSO. A total volume of 10 µL reactive dye solution was then mixed with HSA-Leb conjugate solution and incubated in the dark for 1 h. The excess of free dye was removed by two gel-filtration passes over a PD-10 column (GE Healthcare, Buckinghamshire, UK) equilibrated with PBS and 0.05% Tween. Labeled glycoconjugates were divided into 10 µL aliquots and kept at −20°C until use.

***In vitro* BabA-mediated binding of *H. pylori* isolates with Alexa555-labeled Leb-HSA conjugates** (Figure S6B). *H. pylori* cells were collected from plates and washed twice with PBS-Tween and 0.1% BSA. Cell suspensions of all strains/clones were adjusted to OD_{600} = 0.1. In addition, the suspension of CUG17875 (WT) was gradually diluted with a suspension of 17875*babA1A2*-mutant by 1:10, 1:100, and 1:1000 times. A total volume of 1 mL of cell suspension was mixed with 1 µL of Alexa555-labeled Leb-HSA and incubated for 1 h at room temperature in the dark. A total volume of 15 µL of the suspensions was applied to microscope slides and subjected to fluorescence microscopy.

Strains of *H. pylori* used for rhesus macaque long-term infection: *H. pylori* strain USU101 was isolated from a patient with gastric adenocarcinoma and used for infection of rhesus macaques. The animal 81G was administered the nitrosating carcinogen ethyl-nitro-nitrosoguanidine (ENNG) five days a week (Liu et al., 2009). The 81G antrum and corpus output sweeps were collected at the 5-year time point by gastroscopy.

Colony membrane screening for Leb-binding monkey output clones (Figure S6C). Leb-affinity-panning for isolation of clones positive for Leb-binding was established according to (Aspholm et al., 2006) with minor modifications. *H. pylori* were applied at a density of 500 colonies per plate on blood agar plates supplemented with 0.004% Tetrazolium Red (Sigma-Aldrich) for 3–4 days until single colonies appeared. Colonies were transferred onto a nitrocellulose membrane (Bio-Rad, Hercules, CA, USA), allowed to dry for 20 min, washed three times with PBS plus Tween-20 (the bacteria were scraped off after the second wash), and blocked overnight with 1.5% gelatin/0.5% BSA. The next day, the membrane was incubated with 0.5 µg/mL biotinylated Leb-HSA glycoconjugate for 2 h at room temperature, washed three times, and then incubated with a 1:5,000 dilution of streptavidin-peroxidase (Roche, Stockholm, Sweden) for 1 h at room temperature. After three washings with PBST, the membrane was developed with 4-chloro-1-Naphtol tablets (Sigma-Aldrich) for 10 minutes. Leb-binding colonies were identified by the colony spots changing color from red to dark brown/black. As positive and negative controls, we used the 17875/Leb strain and 17875*babA1A2*-mutant.

Expression and purification of recombinant BabA from 81G monkey (Figure S6F) **for SPR analysis.** 81G *babA* was amplified from genomic DNA of antrum and corpus isolates by a two-step PCR with a the *babA*-specific primer SE25: (5'-CATGCCCTTGGTTTAATGCTAC-3') and the complementary primer SE26 (5'-GTAGCATTAAACCAAGGGCATG-3'), which also simultaneously removed an NcoI restriction enzyme site within the *babA* gene, and the two primers 3'-babA547 BamHI and 5'-babA-NcoI (described above) that also adds the restriction sites for cloning *babA* into the pOPE vector. Stitched *babA* was inserted into pOPE-vector, transformed, expressed, and purified as described in the section "**Cloning, expression, and purification of recombinant SO-2 $BabA_{1-526}$ in *E. coli***".

7. Acid Sensitive and Reversible BabA Multimerization

***H. pylori* used for rhesus macaque short-term infection:** *H. pylori* strain *J166* (Hansen and Solnick, 2001) was used for short-term infection and to construct a *babB* deletion mutant with following primers:

Upstream_F_NotI	AAC GCG GCC GCC TCA CCA CGC AGA GGA AG
Upstream_R_PstI	AAC CTGCAG GAT TGT TCA GCT TTT CAT AAT TGTC
CAT_F_SacI	AAC GAG CTC GCG GAC AAC GAG TAA AAG AG
CAT_R_PstI	AAC CTG CAG GCA GGA CGC ACT ACT CTC G
Downstream_F_SacI	AAC GAGCTC TAT GTG GAA CAA ACC ATA ACG AA
Downstream_R_XhoI	AAT CTC GAG TCA GCA TTC ACT TCC TTT TTG A

Because the *babA* and *babB* genes are highly conserved in their 5' and 3' regions, a *babB* knockout was

constructed that retained these regions of homology and replaced the variable region between them with a CAT cassette in the reverse orientation. Upstream, downstream, and CAT primer pairs were used to amplify DNA fragments of 1636, 2068, and 877 bp, respectively. Following purification and restriction digestion, the amplicons were ligated to pBluescript (Agilent Technologies, Wilmington, DE) and transformed into One Shot Top10 competent cells (Life Technologies, Grand Island, NY). The resulting plasmid was used to deliver the CAT cassette to J166 by allelic exchange. Disruption of *babB* was verified by PCR and loss of expression on Western Blot. This *babB* knockout was used to infect rhesus macaques via oral gavage. After 8 weeks, gastric biopsies were taken and single colonies were saved for analysis by PCR and sequencing. An isolate was identified that no longer expressed BabA due to gain of a CT pair near the 5' end of the gene. This output strain – the phenotypically *babAbabB* deletion mutant – was compared to the *babB* deletion input strain by 2D DIGE (described below).

Separation of *H. pylori* cell extract by 2D DIGE (Figures 7A, 7B and S7A). Liquid-cultured *H. pylori* were centrifuged for 10 min at 5700 × *g*, and pellets were washed in TE buffer (10 mM Tris, 1 mM EDTA, pH 8) and frozen at −20°C. Frozen pellets were then re-suspended in 1 mL of lysis buffer containing one dissolved "Complete mini EDTA-free protease inhibitor cocktail" tablet (Roche Diagnostics) in 7 ml TE buffer chilled on ice. Cells were sonicated, and debris was removed by centrifugation (2040 × *g* for 5 min). Protein was quantitated by Bradford assay.

Protein lysates of *H. pylori* samples were labeled with 200 pmol of Cy3 (green) or Cy5 (red) fluorescent dyes and loaded onto 24 cm, pH 7–11 nonlinear Immobiline IPG DryStrips (GE Healthcare, Piscataway, NJ) for first dimension separation. IEF separation was carried out using the Ettan IPGphor II (GE Healthcare) using the following running conditions: 30 V rehydration for 12 h, 500 V for 1 h, 1,000 V for 1 h, and 8,000 V for 85,000 Vh. The IPG strips were then reduced and alkylated and then loaded onto 26 cm × 20 cm precast 12.5% Tris-glycine polyacrylamide gels (Jule, Inc., Milford, CT) and run at 2 W/gel constant power at 22ºC using an Ettan DALT 12 (GE Healthcare) until the bromophenol blue dye-front reached the end of the gels (approximately 16 h). Gels were imaged using a Typhoon 9410 imager (GE Healthcare) at 100 µm resolution. Laser power was adjusted for the optimization of sensitivity and the prevention of oversaturation. Cy3 dye was excited at 550 nm and emission spectra obtained at 570 nm; Cy5 dye was excited at 650 nm and emission spectra obtained at 670 nm. All gel images were cropped to the same size using ImageQuant v5.2 (GE Healthcare). Gel images were analyzed using the various modules of the DeCyder v7 software package. The Differential In-gel Analysis (DIA) module was used to determine optimal spot detection settings and to determine the differential protein spots to be excised for protein identification. The two BabA protein bands were excised from the gel using an ITSI Biosciences Screen Picker after identification of BabA by immunoblots with the BabA antibody AK 277. Further protein digestion and identification was performed, and the resulting peak lists were searched against a bacterial protein database using the MASCOT search engine using the following parameters: up to two missed cleavages, peptide mass tolerance of 1.2 Da, fragment mass tolerance of 0.6 Da, fixed modification carbamidation, and variable modification oxidation (Chromy et al., 2013).

ZW extraction and *In vitro* BabA oligomerization (Figure 7D). Pellets, obtained from 250 µL of bacterial suspensions of OD_{600} =1.0, were re-suspended in 25 µL of citrate-phosphate buffers containing 10 mM ZW 3-12 with a pH range from 3.5 to 6 in increments of 0.5 pH units. After 1 min incubation at room temperature, suspensions were mixed with an equal volume of 2% SDS Laemmli buffer under non-reducing room temperature conditions. Samples were neutralized and separated by SDS-PAGE. BabA protein-specific bands were detected by immunoblot with Vite antibody.

pH-dependent *In vivo* BabA oligomerization (Figures 7E and S7D). *In vivo* oligomerization was analyzed with *H. pylori* 17875/Leb. Pellets, obtained from 1 mL of bacterial suspensions of OD_{600} =1.0, were re-suspended in 200 µL in a series of citrate-phosphate buffers with a pH range from 2.5 to 6 in increments of 0.5 pH units. The suspensions were incubated at room temperature for 30 min for pH-dependent oligomerization of BabA. Cross-linking was performed by mixing 0.05% of glutaraldehyde solution (Fisher Scientific, Sweden) with the bacterial suspensions, followed by immediate reconditioning by addition of 1 M NaOH to pH 7. Samples were then incubated for 2 min at room temperature to allow for short-term cross-linking of BabA oligomers. The glutaraldehyde cross linker was then quenched by

addition of 1M Tris-HCl, pH 8, followed by 10 min incubation at room temperature. The samples were pelleted by centrifuging for 5 min at 13,000 rpm and solubilized in 200 µL SDS-Laemmli buffer under reducing conditions for 10 min at 100°C. SDS-PAGE–separated BabA monomers and oligomers were detected by immunoblots with Vite-antibody and developed with IRDye 800CW goat anti-rabbit secondary antibody (LI-COR Bioscience, NE, USA) diluted 1:20.000. Signal was visualized by the Odyssey Sa scanner and quantified using Image Studio™ Lite Ver 4.0 (LI-Cor Bioscience).

Acid-depolymerized BabA was re-constituted into multimers upon neutralization (Figures 7F and S7E). A total of 1 ml of *H. pylori* suspension, OD_{600} = 1, was pelleted, re-suspended in 1 ml of citrate-phosphate buffer pH 2.5 or pH 6, and then incubated for 30 minutes on a rotating platform at room temperature. Bacteria were pelleted again by 5 min centrifugation at 13,000 rpm, and supernatants were discarded. The pellets obtained from the bacterial suspensions exposed to pH 2.5 buffer were re-suspended in pH 6 buffer followed by incubation and centrifugation as described above. Obtained pellets were then resuspended in 200 µl of pH 9.5 carbonate buffer with 0.0125% glutaraldehyde. After 30 sec incubation, the cross-linking reaction was stopped by 50 µl of 1M Tris-HCl followed by 10 min incubation on a rotating platform at room temperature. Suspensions were centrifuged for 5 min at 13,000 rpm, and pellets were solubilized in 200 µL of SDS-Laemmli buffer with β-mercaptoethanol for 10 min at 100°C. Detection of monomers and oligomers was performed as described above.

SUPPLEMENTAL REFERENCES

Andersson, M., Fällman, E., Uhlin, B. E. and Axner, O. (2006). Dynamic force spectroscopy of *E.coli* P pili. Biophys. J. *91*, 2717-2725.

Armstrong, K. M., and Baldwin R. L (1993). Charged histidine affects alpha-helix stability at all positions in the helix by interacting with the backbone charges. Proc. Natl. Acad. Sci. USA *90*, 11337-11340.

Aspholm, M., Kalia, A., Ruhl, S., Schedin, S., Arnqvist, A., Lindén, S., Sjöström, R., Gerhard, M., Samino-Mora, C., Dubois, A., Unemo, M., et al. (2006). *Helicobacter pylori* adhesion to carbohydrates. In Methods Enzymol. (Academic Press, New York), pp. 293-339**.**

Bernarde, C., Lehours, P., Lasserre, J. P., Castroviejo, M., Bonneu, M., Mégraud, F., Ménard, A. (2010). Complexomics study of two *Helicobacter pylori* strains of two pathological origins: potential targets for vaccine development and new insight in bacteria metabolism. Mol. Cell. Proteomics. *12*, 2796-2826.

Borén, T., Falk, P., Roth, K. A., Larson, G., and Normark, S. (1993). Attachment of *Helicobacter pylori* to human gastric epithelium mediated by blood group antigens. Science *262*, 1892-1895.

Bäckström, A., Lundberg, C., Kersulyte, D., Berg, D. E., Borén, T., and Arnqvist, A. (2004). Metastability of *Helicobacter pylori* bab adhesin genes and dynamics in Lewis b antigen binding. Proc. Natl. Acad. Sci. USA *101*, 16923-16928.

Chromy, B. A., Eldridge, A., Forsberg, J. A., Brown, T. S., Kirkup, B. C., Jaing, C., Be, N. A., Elster, E., and Luciw, P. A. (2013). Wound outcome in combat injuries is associated with a unique set of protein biomarkers. J. Transl. Med. *11***,** 281.

Colbeck, J. C., Hansen, L. M., Fong, J. M., and Solnick, J. V. (2006). Genotypic profile of the outer membrane proteins BabA and BabB in clinical isolates of *Helicobacter pylori*. Infect. Immun. *74*, 4375-4378.

Dailidiene, D., Dailide, G., Ogura, K., Zhang, M., Mukhopadhyay, A. K., Eaton, K. A., Cattoli, G., Kusters, J. G., and Berg, D. E. (2004). *Helicobacter acinonychis*: genetic and rodent infection studies of a Helicobacter pylori-like gastric pathogen of cheetahs and other big cats. J Bacteriol. *186*, 356-365.

Fei, Y. Y., Schmidt, A., Bylund, G., Johansson, D. X., Henriksson, S., Lebrilla, C., Solnick, J. V., Borén, T., and Zhu, X. D. (2011). Use of real-time, label-free analysis in revealing low-affinity binding to blood group antigens by *Helicobacter pylori*. Anal. Chem. *83*, 6336-6341.

Frishman, D. and Argos, P. (1997). Seventy-five percent accuracy in protein secondary structure prediction. Proteins *27*, 329-335.

Fällman, E., Schedin, S., Jass, J., Andersson, M., Uhlin, B. E., and Axner, O. (2004) Optical tweezers based force measurement system for quantitating binding interactions: system design and application for the study of bacterial adhesion. Biosens. Bioelectron. *15*, 1429-1437.

Gustavsson, A., Unemo, M., Blomberg, B., and Danielsson, D. (2005). Genotypic and phenotypic stability of *Helicobacter pylori* markers in a nine-year follow-up study of patients with noneradicated infection. Dig Dis Sci. *50*, 375-380.

Hansen, L. M., and Solnick, J. V. (2001). Selection for urease activity during *Helicobacter pylori* infection of rhesus macaques (*Macaca mulatta*). Infect. Immun. *69*, 3519-3522.

Hennig, E. E., Allen, J. M., and Cover T. L. (2006). Multiple Chromosomal Loci for the *babA* Gene in *Helicobacter pylori.* Infect. Immun. *74*, 3046-3051.

Heuermann, D., and Haas, R. (1998). A stable shuttle vector system for efficient genetic complementation of *Helicobacter pylori* strains by transformation and conjugation.
Mol. Gen. Genet. *257*, 519-528.

Jiménez-Soto, L. F., Clausen, S., Sprenger, A., Ertl, C. and Haas, R. (2013). Dynamics of the Cag-type IV secretion system of *Helicobacter pylori* as studied by bacterial co-infections. Cell Microbiol. *15*, 1924-1937.

Jones, D. H., Schmalsteig, F. C., Dempsey, K., Krater, S. S., Nannen, D. D., Smith, C. W., and Anderson, D. C. (1990). Subcellular distribution and mobilization of MAC-1 (CD11b/CD18) in neonatal neutrophils. Blood *75*, 488-498.

Jones, D. T. (1999). Protein secondary structure prediction based on position-specific scoring matrices. J. Mol. Biol. *292*, 195-202.

Kersulyte, D., Kalia, A., Gilman, R. H., Mendez, M., Herrera, P., Cabrera, L., Velapatiño, B., Balqui, J., Paredes Puente de la Vega, F., Rodriguez Ulloa, C. A., et al., (2010). *Helicobacter pylori* from Peruvian amerindians: traces of human migrations in strains from remote Amazon, and genome sequence of an Amerind strain. PLoS One *5*, e15076.

Linding, R., Jensen, L. J., Diella, F., Bork, P., Gibson, T. J., and Russell, R. B. (2003). Protein disorder prediction: implications for structural proteomics. Structure *11*, 1453-1459.

Linding, R., Russell, R. B., Neduva, V., and Gibson, T. J. (2003). GlobPlot: Exploring protein sequences for globularity and disorder. Nucleic Acids Res. *31*, 3701-3708.

Liu, H., Fero, J. B., Mendez, M., Carpenter, B. M., Servetas, S. L., Rahman, A., Goldman, M. D., Borén, T., Salama, N. R., Merrell, D. S., and Dubois, A. (2015). Analysis of a single *Helicobacter pylori* strain over a 10-year period in a primate model. Int. J. Med. Microbiol. *305*, 392-403.

Löfling, J., Diswall, M., Eriksson, S., Borén, T., Breimer, M. E., and Holgersson, J. (2008). Studies of Lewis antigens and H. pylori adhesion in CHO cell lines engineered to express Lewis b determinants. Glycobiology *18*, 494-501.

Mahdavi, J., Sondén, B., Hurtig, M., Olfat, F. O., Forsberg, L., Roche, N., Ångström, J., Larsson, T., Teneberg, S., Karlsson, K. A., et al. (2002). *Helicobacter pylori* SabA adhesin in persistent infection and chronic inflammation. Science *297*, 573-578.

Meng, G., Surana, N. K., St Geme, J. W. 3rd., and Waksman, G. (2006). Structure of the outer membrane translocator domain of the *Haemophilus influenzae* Hia trimeric autotransporter. EMBO J. *25,* 2297-2304.

Merrell, D. S., Goodrich, M. L., Otto, G., Tompkins, L. S., and Falkow, S. (2003). pH-regulated gene expression of the gastric pathogen *Helicobacter pylori*. Infect. Immun. *71*, 3529-3539.

Motulsky, H., and Neubig, P. (2001). Analyzing radioligand binding data. Curr Protoc Protein Sci., Appendix 3. (John Willey and sons, Inc.)

Panayotopoulou, E. G., Sgouras, D. N., Papadakos, K., Kalliaropoulos, A., Papatheodoridis, G., Mentis, A. F., and Archimandritis, A. J. (2007). Strategy to characterize the number and type of repeating EPIYA phosphorylation motifs in the carboxyl terminus of CagA protein in *Helicobacter pylori* clinical isolates. J. Clin. Microbiol. *45*, 488-495.

Odenbreit, S., Kavermann, H., Püls, J., and Haas, R. (2002). CagA tyrosine phosphorylation and interleukin-8 induction by *Helicobacter pylori* are independent from alpAB, HopZ and bab group outer membrane proteins. Int. J. Med. Microbiol. *292*, 257-266.

Schmiedl, A., Breitling, F., Winter, C., Queitsch, I., and Dübel, S. (2000) Effects of unpaired cysteines on yield, solubility and activity of different recombinant antibody constructs expressed in *E. coli.* J. Immunol. Methods *242*, 101–114.

Skoog, E. C., Sjöling, Å., Navabi, N., Holgersson, J., Lundin, S. B., and Lindén, S. K. (2012). Human gastric mucins differently regulate *Helicobacter pylori* proliferation, gene expression and interactions with host cells. PLoS One *7*, e36378.

Strömberg, N., and Borén, T. (1992). Actinomyces tissue specificity may depend on differences in receptor specificity for GalNAc beta-containing glycoconjugates. (1992). Infect. Immun. *60*, 3268-77.

Yachdav, G., Kloppmann, E., Kajan, L., Hecht, M., Goldberg, T., Hamp, T., Hönigschmid, P., Schafferhans, A., Roos, M., Bernhofer, M. et al. (2014). PredictProtein--an open resource for online prediction of protein structural and functional features. Nucleic Acid Res. *42*, (Web Server issue): W337-343.

Yamaoka, Y., Ojo, O., Fujimoto, S., Odenbreit, S., Haas, R., Gutierrez, O., El-Zimaity, H. M., Reddy, R., Arnqvist, A., and Graham, D. Y. (2006). *Helicobacter pylori* outer membrane proteins and gastroduodenal disease. Gut *55*, 775-781.

Younson, J., O'Mahony, R., Liu, H., Basset, C., Grant, S., Campion, C., Jennings, L., Vaira, D., Kelly, C. G., Roitt, I. M. and Holton, J. (2009). A human domain antibody and Lewis b glycoconjugate that inhibit binding of *Helicobacter pylori* to Lewis b receptor and adhesion to human gastric epithelium. J. Infect. Dis. *200*, 1574-1582.

Suppl. figure 1.

Figure S1. *H. pylori* Binding to Leb in Human Gastric Mucosa is Acid Sensitive, Responsive, Reversible and Robust. Related to Figure 1.

(A) Digital quantification of Figure 1A, slides a-c showed that acid-sensitive binding of FITC-fluorescent *H. pylori* 17875/Leb to human gastric mucosa *in vitro* was 46% and 5% at pH 4 and pH 2, respectively, relative to 100% at pH 6.

(B) Test for acid sensitivity in Leb binding kinetics. 17875/Leb bacterial cells were first incubated for 2 min in pH 2.5 to inactivate Leb binding (dark line) or in pH 5 buffer (grey line), then reconditioned to pH 5 followed by an incubation for 1 h with ^{125}I-Leb-conjugate. Prior acidification did not affect Leb binding kinetics after reconditioning to pH 5. The slight reduction in maximal Leb binding (plateau) for the bacteria that were acidified to pH 2.5 likely stems from bacterial disruption.

(C) Acid-responsive binding by *H. pylori* in real time. LigandTracer Green by Ridgeview, Uppsala, Sweden, was used for real-time measurements of bacterial cells interactions with immobilized cells and/or glycol conjugates. The system generates kinetics data based on repeated measurements of a ligand that is bound to an area of cells or receptors that are immobilized to the dish on a slowly rotating support. In this study, fluorescent Alexa488-labeled *H. pylori* bacterial cells (green) were applied to immobilized CHO or BEC cells or to glycan receptors such as Leb-conjugate. A fluorescence detector reads the signal from the upper, almost liquid-free area (due to the inclined dish). Signal is recorded from both the "target" area and "reference" area (the negative control), which together constitute the background-corrected level of ligand (here *H. pylori*) bound to the target area of immobilized Leb-conjugate (D and F) or Leb-expressing cells (G and H). Upon bacterial binding to the Leb conjugate or host-cells over time, a response curve is generated, representing the kinetic behavior of the interaction.
http://www.kbc.umu.se/english/bicu/ligand-tracer/

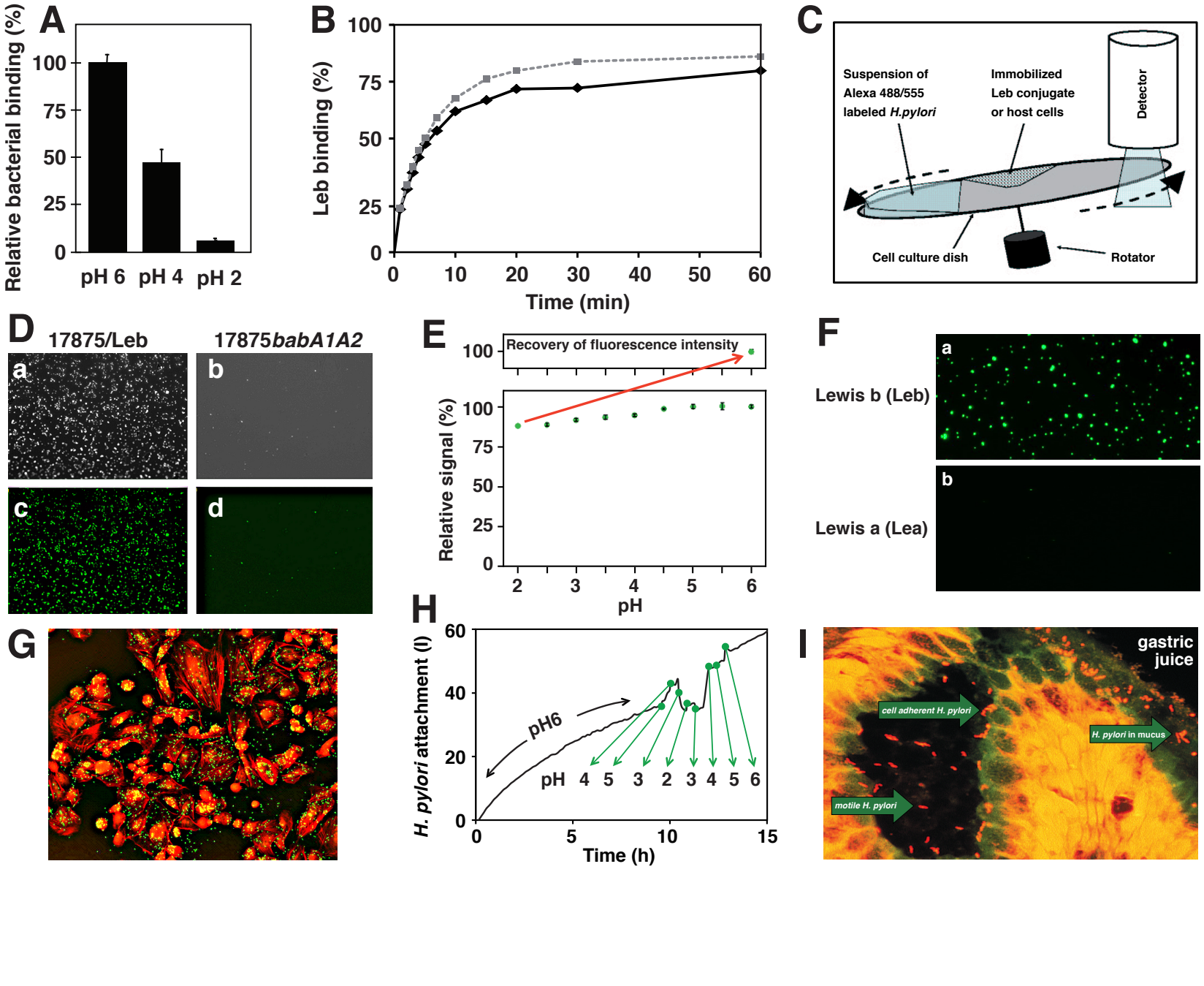
(D) Test of BabA adhesin binding in real-time. Binding of 17875/Leb bacterial cells to immobilized Leb-glycoconjugate, mimicking the natural situation in which bacteria attach to Leb on surfaces such as gastric epithelia cells or mucins, was recorded by LigandTracer. Alexa488-labeled bacteria (green) of 17875/Leb or the *babA1A2*-mutant (negative control) were added to a dish coated with Leb conjugate in the target area and incubated for 2 h. Attached bacteria were first visualized by phase contrast of 17875/Leb (a) or *babA1A2*-mutant (b) and then by fluorescence microscopy. Binding of 17875/Leb reconstituted 1.11×10^6 bacteria/cm^2 on the Leb-coated surface (c), whereas the *babA1A2* mutant showed 100-fold lower attachment, i.e. 1.20×10^4 bacteria/cm^2 (d).

(E) Stability of the Alexa488 fluorochrome and fluorescence at low pH. Fluorescence spectroscopy analysis showed that the Alexa488 fluorochrome used to label *H. pylori* bacteria tolerates low pH exposure well, with merely a 12% reduction in intensity by acidification to pH 2 with ortho-phosphoric acid (green dots). The Alexa fluorochrome recovered 100% fluorescence activity when reconditioned to pH 6 with NaOH (red arrow). The binding of fluorescently labeled bacteria at each pH in Figures 1G and S1H were all corrected for the Alexa488 absorption signal at different pH values.

(F) Test of Leb *vs*. Lea blood group antigen specificity in bacterial binding. Specific BabA-mediated Leb binding was defined as bacterial binding to the immobilized Leb conjugate after subtraction of non-specific binding to the immobilized Lea antigen. Lea is similar to Leb, but Lea lacks the α1.2fucose epitope that is the characteristics of the ABO blood-group antigens. The α1.2fucose epitope is also the major determinant of the BabA adhesin binding epitope. Thus, the consequence of lack of the α1.2fucose reside is that BabA does not bind Lea (Borén et al., 1993). Leb and Lea conjugates were immobilized opposite each other in the Petri dish, and 17875/Leb bacterial binding was recorded from each position in parallel. The dish was set to slowly rotate for 10 h at pH 6. Fluorescence microscopy was used to assess bacterial binding to (a) Leb and (b) Lea. *H. pylori* bacterial BabA-mediated binding to Leb was >100-fold higher compared to Lea.

(G) Acid-responsive binding of Alexa488-labeled *H. pylori* 17875/Leb bacterial cells to recombinant Leb-expressing CHO-cells (Löfling et al., 2008) grown to confluence in the Petri dish. The pH gradient was applied to the cells at 10 h (Figure 1G). Quantification of bound bacteria to CHO cells was assessed after subtraction of signal from dish segments without CHO cells. At the end of each experiment, the integrity

of the CHO cells was verified by fluorescence microscopy in which CHO cells were stained with GelRed nucleic acid stain and rhodamine phalloidin.
(H) Acid-responsive binding of Alexa488-labeled *H. pylori* 17875/Leb bacterial cells to human buccal epithelial cells (BECs) collected by oral swabs from an individual with a positive ABO secretor phenotype (Strömberg and Borén, 1992). BECs were biotin labeled and immobilized to the dish by MegaCell^{R}-Streptavidin and conditioned to pH 6 for 10 h followed by incremental acidification to pH 2 and reconditioning to pH 6. The integrity of bound cells was verified at the end of the pH-cycle by fluorescence microscopy in which BECs were counterstained with DAPI (Figure 1H).
(I) In this acridine orange-stained biopsy of gastric mucosa from an *H. pylori*-infected individual, *H. pylori* bacteria (orange rods) are found both intimately adhering to the gastric epithelial cells (*cell adherent H. pylori)* and in close proximity to and above the epithelial cell surfaces in the mucus mucin protective layer (*H. pylori in mucus*), in addition to motile bacteria with flagellar activity (*motile H. pylori*). The viscous mucus mucin layer helps stabilize the pH gradient formed at the interface between the epithelial surfaces with buffered higher pH and the acidic luminal gastric juice. The protective mucus layer is approximately 300 μm thick in the antrum and less so in the corpus region. In this tissue section preparation, the mucus layer is seemingly rather thin, which is a result of the formaldehyde biopsy fixation procedures.


A
Relative bacterial binding (%)
100
75
50
25
0
pH 6
pH 4
pH 2
B
Leb binding (%)
100
75
50
25
0
0
10
20
30
40
50
60
Time (min)
C
Suspension of
Alexa 488/555
labeled H.pylori
Immobilized
Leb conjugate
or host cells
Detector
Cell culture dish
Rotator
D
17875/Leb
17875babA1A2
a
b
c
d
E
Recovery of fluorescence intensity
Relative signal (%)
100
75
50
25
0
2
3
4
5
6
pH
F
Lewis b (Leb)
Lewis a (Lea)
a
b
G
H
H. pylori attachment (I)
60
40
20
0
pH6
pH 4 5 3 2 3 4 5 6
0
5
10
15
Time (h)
I
gastric
juice
cell adherent H. pylori
H. pylori in mucus
motile H. pylori

Suppl. figure 2.

Figure S2. Diversity in Acid Sensitivity in Binding to Leb Among Clinical Isolates is Encoded in BabA. Related to Figure 2.

(A) pHgrams of 21 Swedish (SW) *H. pylori* strains based on bacterial binding to ^{125}I-Leb conjugate in buffers ranging from pH 2 to pH 6, including acid-resistant strain SW7 (red) and more acid-sensitive strain SW38 (green). Thus, each and every *H. pylori* strain binds to Leb with different combinations of binding affinity, binding capacity, and acid sensitivity. This complexity is illustrated here by depicting the series of pHgrams in absolute terms, whereas for easy comparison of strain-specific midpoints i.e. pH_{50} (50% of maximal Leb binding), pHgrams are preferable illustrated in relative terms (for comparison, pHgrams of the 21 Swedish strains are displayed in relative terms in Figure 2A.

(B) Test of the reproducibility of the pHgram method. The CCUG17875 strain was colony purified, and a single clone was picked (named 17875/Leb) (Mahdavi et al, 2002). This clone was then passaged on a plate and 10 isogenic clones were collected, cultured, expanded, and pH titrated for Leb binding (the mean pH_{50} of 10 replicates was 3.37, the median was pH_{50} 3.36, the minimum was pH_{50} 3.31, the maximum was pH_{50} 3.40. This gave a range of pH_{50} of 0.09 units (red lines), i.e. <0.1 pH unit difference. The pHgram of the original strain (smear) is depicted for comparison (blue). In a separate experiment, the pH_{50} of the 17875/Leb clone was 3.3 and the pH_{50} of seven isogenic single clones had a mean ± SD pH_{50} 3.39 ± 0.06, i.e. an even lower range in inter-clonal variation (data not shown).

(C) Adaptation by gain in acid sensitivity by functional pH-biopanning (recycling). Bacterial binding in real time by Alexa488 (green)-labeled bacteria of SW38 and SF555 (red)-labeled bacteria of SW7 to Leb immobilized on a Petri dish. The dish was placed in the LigandTracer instrument and set to slowly rotate. Absolute binding of bacterial cells was followed in real time by the instrument's fluorescence detector, and visualized by fluorescence microscopy. The bacterial cells were quantified by the ImageJ software at (a) 10 h and (b) 17 h, i.e. at the time point before acidification and at the time point of reconditioning from low pH (similar to Figure 1G). Through circadian pH-dependent recycling of infection, a 2-fold daily gain in reattachment would mean a >100-fold increase in a week. Such proficient clones can quickly adapt the infection to local changes in gastric acidity.

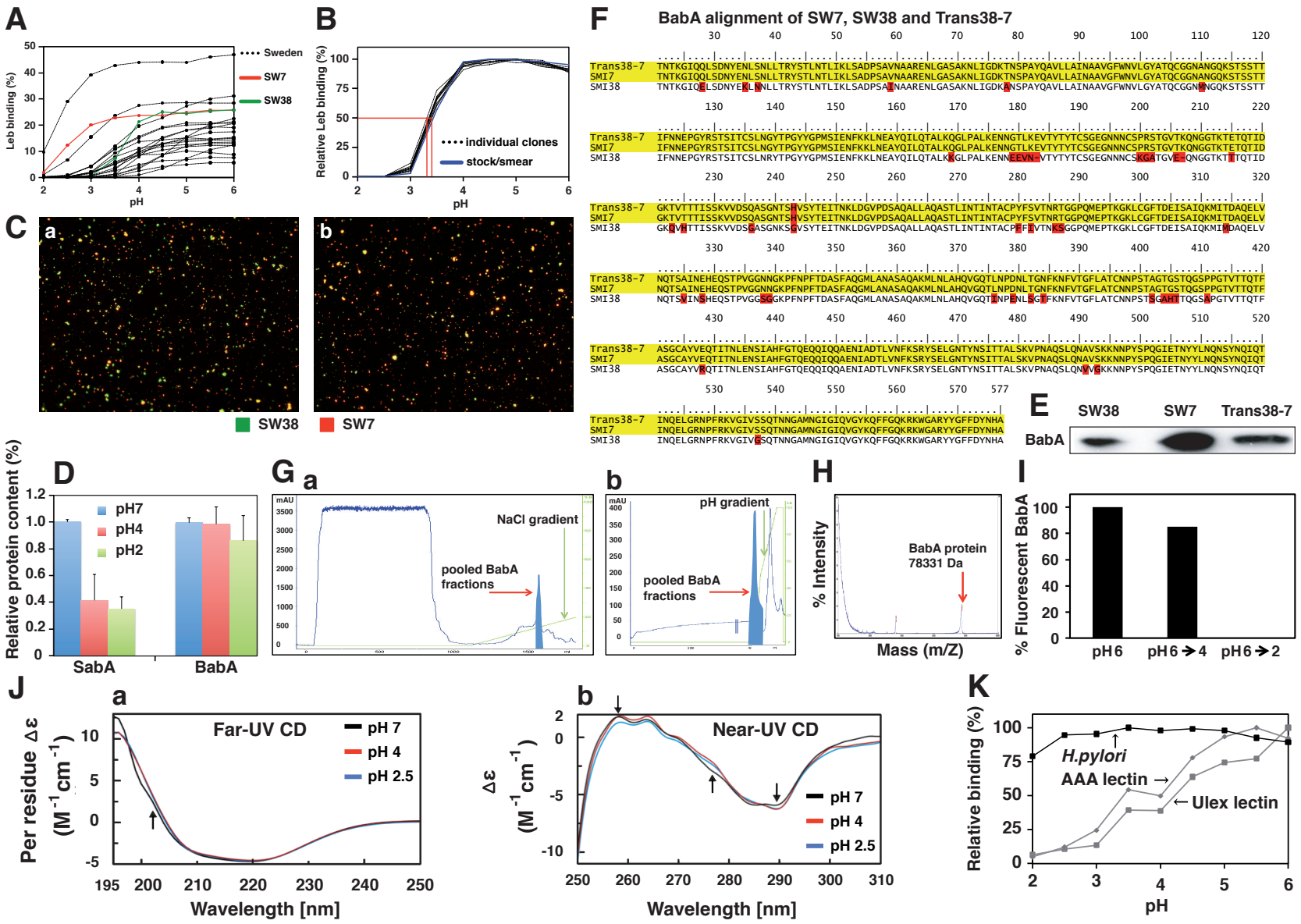
(D) To test for acid induced regulation of BabA protein, *H. pylori* strain SW109 that express both BabA and SabA was grown in broth for 48 hours at pH 7. Cells were then split into three cultures and transferred to culture media of pH 7, 4 and 2. After 24 hours of incubation, samples were collected and analyzed by immunoblot with BabA or SabA antibodies. Results showed that BabA protein levels were 99% and 86% at pH 4 and 2 respectively compare to pH 7. In contrast, SabA content were much reduced at lower pH and merely 41% and 35% SabA protein remained after 24 hours in pH 4 and pH 2 respectively (means + SD, n = 2), which is in accordance with the previous report of SabA regulated by acidic pH (Merrell et al, 2003; Yamaoka et al., 2006). The results demonstrate that BabA protein content is exceedingly acid stabile, whereas SabA tends to decline over time at lower pH.

(E) Expression of BabA in the more acid-sensitive strain SW38, the more acid-resistant strain SW7, and the SW7BabA-expressing transformant Trans38-7 with the expected donor strain SW7 *babA* allele. The BabA expression level of the Trans38-7 transformant was more similar to the parent strain SW38, as illustrated by the BabA immunoblot, and thus lower compared to the donor strain SW7.

(F) Sequence alignment of BabA SW7 and SW38 proteins shows 94% protein identity, which is typical divergence for BabA among clinical strains, whereas the BabA Trans38-7 protein sequence is identical to the donor BabA SW7 (both in yellow). The locations of all amino acid differences between Trans38-7/SW7 *vs*. SW38 are indicated in red.

(G) Purification of BabA protein to homogeneity in native (Leb binding) conformation from strain 17875/Leb. (a) Bacterial paste was solubilized in zwitterionic detergent ZW-12, and the soluble protein extract was first separated by Source 30S (GE) cation exchange chromatography (CEX). Because of the high isoelectric point (pI) of BabA (pI 8.94), this very basic protein binds to the cation column at pH 8, whereas the bulk of other proteins pass through the column (major peak to the left). The BabA protein was eluted from the column with a linear NaCl gradient (green arrow). Fractions with BabA (in blue) were identified by SDS-PAGE immunoblots with BabA antibodies. These fractions were pooled and diluted with octyl-glucoside detergent (OG) to help refold BabA into a Leb binding-proficient BabA protein

conformation. (b) The OG-solubilized and stabilized BabA preparation were applied to a Leb affinity column, and bound BabA protein was eluted with a pH gradient (green arrow). BabA fractions (in blue) were pooled and immediately pH reconditioned, and protein aliquots were stored at −80°C.
(H) The integrity of the purified native BabA protein was analyzed by MALDI-MS, and the mass was 78.331 Da, which was similar to the predicted 78.245 Da for BabA from CCUG17875. The small peak in the middle of the diagram (generated by the technique) corresponds to a 50% mass marker of the BabA protein sample.
(I) Detachment of purified native BabA protein bound to histo-tissue sections of gastric mucosa (Figure 2F), and thus reduction of Alexa555 fluorescent signal by acidification to pH 4 (15% detachment) or pH 2 (100% detachment), i.e. at pH 2 no BabA protein remained bound to the histo-tissue section epithelium), was digitally quantified from Z-stacks by the AxioVision (Zeiss) program.
(J) Circular dichroism (CD) analyses of acid-induced unfolding of purified BabA. Far-UV spectrum analysis (a). The pHs were 7 (black), 4 (red), and 2.5 (blue). Spectra are displayed in per-residue circular dichroism (Δε). Far-UV measurement of the native BabA protein suggests that the protein is 46% α-helices, 29% β-strands and turns, and 24% random structures. Changes in the Far-UV CD spectrum of the native BabA protein upon acidification showed that the combined α-helices and β-strand/turn structures of BabA at pH 7 underwent only minor changes of ellipticity at 195–206 nm upon reduction to pH 4 with no further changes caused by pH 2.5 acidification. This suggests a model in which the initial reduction in Leb-binding affinity at pH 4–3.5 (for strain 17875/Leb, Figure 1D) depends on subtle structural changes determined by protonation of amino acids with pKa values from 7.5 to 4, i.e. His (H) (pKa 6) and Glu (E) (pKa 4.25) residues (Armstrong and Baldwin, 1993). Further, acid sensitivity-dependent changes in BabA binding affinity can be due to protonation of Asp (D) (pKa 3.65), i.e. below pH 4, but no such pH 2.5-related changes in secondary structure was observed. Near-UV (CD) spectrum analysis (b). These spectra are displayed in absolute circular dichroism (Δε), and the pHs used were 7 (black), 4 (red), and 2.5 (blue), where the CD fluorescence signal reports on structural changes in domains where tryptophan (Trp) residues are present. In the BabA protein, three out of four Trp residues are located in the C-terminal β-barrel domain, and one is located in α-Helix3 (H3), which makes up the base of cysteine loop-1 (CL1) (Figure S4A). The Near-UV CD demonstrated that ellipticity at 291 nm (arrow) became more negative upon acidification from pH 7 to pH 4 and then remained stable from pH 4 to pH 2.5. This suggests that local disordering of α-helices and tertiary structure is caused by protonation of His and Glu residues. In addition, the increased ellipticity at 275 nm (arrow) corresponds to tyrosine (Tyr) residues, and protonation of these residues also takes place in the pH interval from pH 7 to pH 4 and then remains stable down to pH 2.5. The BabA Extra Cellular Domain (ECD) is rich in Tyr (Y) residues, which suggests that pH 4 induces changes in tertiary structure related to partial ECD unfolding. Finally, the decreased ellipticity at 258 nm (arrow) at pH 2.5 (in blue) refers primarily to phenylalanine and suggests that pH 2.5 acidification causes local changes in BabA tertiary structure.
(K) Tests with two blood group O antigen-binding lectins from the plant *Ulex europaeus* (UEA) and the fish/eel *Anguilla anguilla* (AAA) showed that, in contrast to BabA, they do not exhibit reversible binding activity (no robustness) after reconditioning to neutral pH from prior exposure to the acidic part of the pH gradient (reconstituted pHgram, as described in Figure 1E). Instead, the UEA and AAA lectins irreversibly lost most binding activity with only 50% recovery at pH 4, only 25% residual activity after pH 3 acidification, and with essentially no residual binding activity after prior pH 2 exposure, presumably due to denaturation. In contrast, *H. pylori* bacterial cells demonstrate almost full recovery of BabA-mediated binding activity after reconditioning to more neutral conditions after prior acid exposure to pH 2.5, and there is about 75% recovery of activity after prior acidification to pH 2.


A
Leb binding (%)
pH
Sweden
SW7
SW38
B
Relative Leb binding (%)
pH
individual clones
stock/smear
C
a
b
SW38
SW7
F
BabA alignment of SW7, SW38 and Trans38-7
Trans38-7
SMI7
SMI38
E
SW38
SW7
Trans38-7
BabA
D
Relative protein content (%)
pH7
pH4
pH2
SabA
BabA
G
a
b
mAU
NaCl gradient
pooled BabA fractions
pH gradient
pooled BabA fractions
H
% Intensity
BabA protein 78331 Da
Mass (m/Z)
I
% Fluorescent BabA
pH 6
pH 6 → 4
pH 6 → 2
J
a
Far-UV CD
pH 7
pH 4
pH 2.5
Per residue Δε (M⁻¹cm⁻¹)
Wavelength [nm]
b
Near-UV CD
pH 7
pH 4
pH 2.5
Δε (M⁻¹cm⁻¹)
Wavelength [nm]
K
Relative binding (%)
H.pylori AAA lectin
Ulex lectin
pH

Suppl. figure 3.

Figure S3. Gastric Antrum vs. Corpus Adaptation in Acid Sensitivity and Binding Affinity is Encoded in BabA. Related to Figure 3.

(A) The upper more acidic corpus and the lower less acidic antrum constitute the major regions of the human stomach, which connects to the oral cavity through the esophagus and the first part of the intestine, the duodenal region, which is the site where peptic ulcers occur.

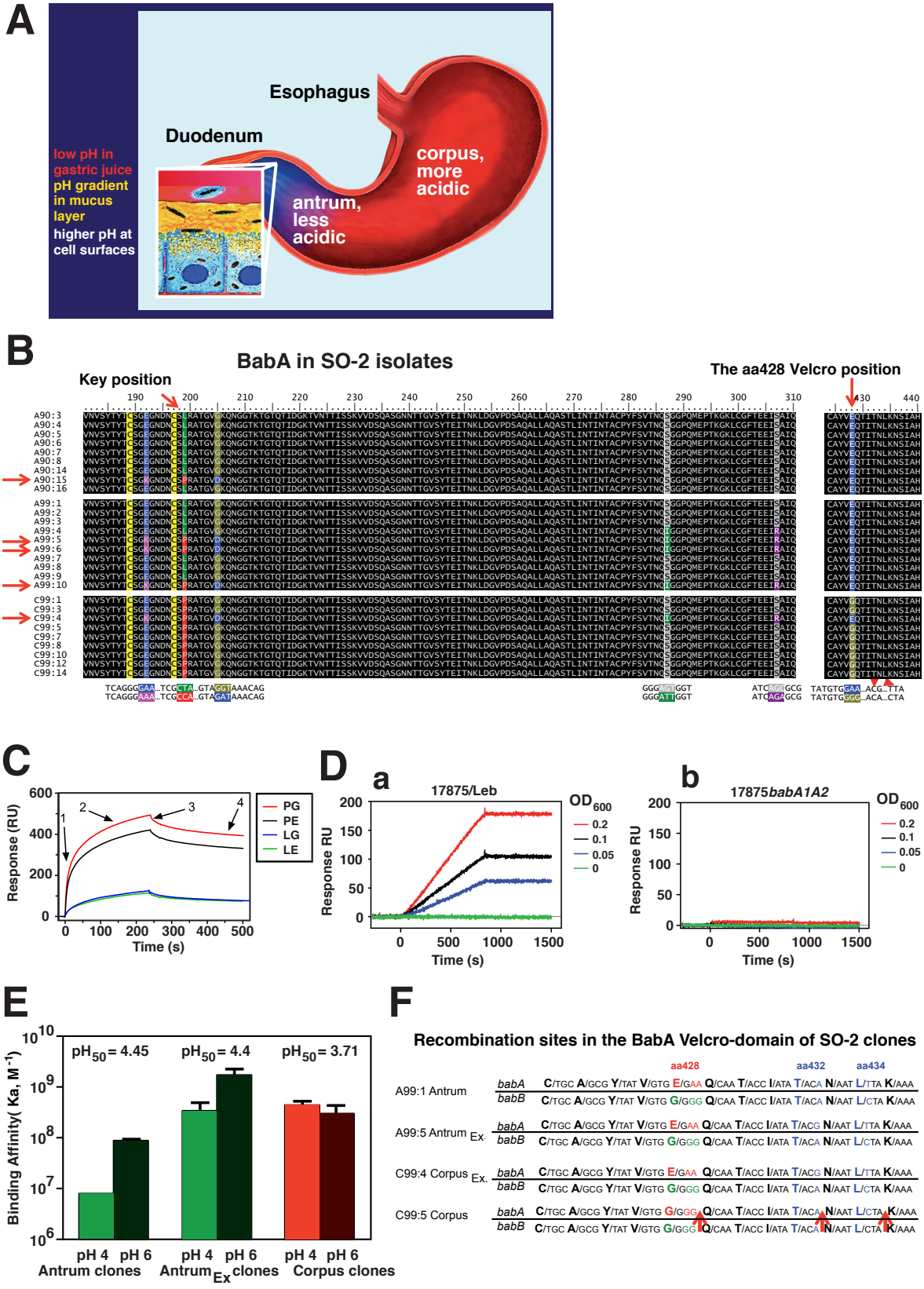
(B) The SO-2 series of BabA sequences. Nine clones were isolated from antrum in 1990 (A90:3–A90:16), ten clones were isolated from the antrum in 1999 (A99:1–A99:10), and nine clones were isolated from the corpus in 1999 (C99:1-C99:14). Residues L199/E428 were observed in the antrum clones, while P199/G428 was observed in the corpus clones. The four exceptional antrum clones (A90:15, A99:5, A99:6 and A99:10) and single exceptional corpus clone (C99:4) (Antrum-Ex and Corpus-Ex, respectively) (red horizontal arrows) carry the K192, P199, D205, and E428 signature. Nucleotide mutations are indicated with codons in colors matching the corresponding amino acid. All mutations that resulted in amino acid substitutions were derived from single nucleotide mutations in the first, second, or third codon position, except E428G, which derived from nucleotide mutations in both second and third codon position (vertical red arrows) (further described in Figure S3F). The majority of clones of exceptional phenotype (Ex) with high-affinity Leb-binding properties were isolated from antrum tissue biopsies (Antrum-Ex), and such subpopulations might have developed in relation to changes in mucosal inflammation. Such inflammation-associated shifts to more charged and sialylated glycosylation patterns followed by modulation of ABO expression might select for *H. pylori* clones with increased bacterial binding affinity (Mahdavi et al., 2002). Chronic erosive mucosal inflammation stimulates the development of atrophy with decreased cell proliferation and a reduced density of glycosylation. Preferential sialylation combined with overall reduced glycosylation, and hence a reduced number of binding sites for *H. pylori*, can stimulate adaptation of BabA binding towards increased Leb-affinity while maintaining acid sensitivity in the antrum environment. In addition to the substitution in the Key-position and Velcro Domain, there are clones that exhibit S287I and S307R mutations, both located within or in close proximity to positive selection Cluster III (which is also the third disulfide bond loop) of BabA. The SO-2 substitutions are shown by light-blue bars in Figure 4A (the numbering of amino acid residues in Figure 4A refers to BabA from the Peruvian P448 strain (Aspholm-Hurtig et al., 2004), and antrum, corpus, and exceptional clone affinities are illustrated in Figure S3E.The amino acids at positions 287 and 307 do not affect acid sensitivity or affinity in Leb binding but might have been selected for a gain in some other properties, possibly including antigenic variation.

(C) Recombinant *E. coli*-expressed $BabA_{526}$ proteins were immobilized in equal amounts with amine coupling on Ni^{2+}-chelated nitrilotriacetic (NTA) chips and analyzed for binding to Leb-PAA by SPR (Biacore) from pH 6 (here in B-figure) to pH 3. An acrylamide (poly[N-(2-hydroxyethyl)acrylamide]) (PAA)-based conjugate of Leb (Leb-PAA, Lectinity, Moscow, Russia) was injected at 10 µg/mL over the surfaces of the chips at pH 6, 5, 4, and 3. The BabA-corpus (PG) and BabA-CA chimeric recombinant proteins (with Pro199 (PE)) were more acid resistant in binding compared to the BabA-antrum (LG) and BabA-AC chimeric proteins (with Leu199 (LE)) (Figure 3D. In these experiments, the Leb-HSA albumin conjugate was substituted by the Leb-PAA polyacrylamide conjugate because of the neutral charge of the PAA conjugate and subsequent low level of non-specific binding upon acidification. It is evident from the sensograms that Leb binding by the recombinant BabA protein is complex. In the first part (1) of the association phase there is a fast transition, and this is followed by a slower phase (2) in the later part of the association process. Similarly, the dissociation is biphasic with one initial fast component (3) followed by a slower second phase (4). This is rather different from our previous results where we showed that full-length BabA exhibits a very slow dissociation phase (Younson et al., 2009), which is more similar to SPR sensograms of bacterial cells binding (C). The biphasic on-rate and off-rate might be explained by a mixture of monomeric and multimeric oligomer complexes formed by recombinant BabA, where the avidity of the multimers would make them bind faster and release slower, and vice versa for the weaker-binding monomers, which was recently verified after gel filtration separation of the recBabA protein (Moonens et al., 2016).

(D) Bacterial binding by Surface Plasmon Resonance (SPR); (a) suspension of bacteria of 17875/Leb were applied to the SPR ProteOn equipment (BioRad) in a series of dilutions over a surface of immobilized Leb conjugate. The BabA mediated bacterial On-rate in binding is fast i.e. association rate, whereas the off-rate or dissociation rate of bound bacterial cells is negligible (horizontal lines to the right); (b) In comparison, the 17875*babA1A2*-mutant that does not bind Leb, illustrates the absolute requirement of BabA adhesin for bacterial binding to ABO antigens.

(E) A subset of the SO-2 clones was analyzed for binding affinities at pH 6 and acidic pH 4. At pH 6, the antrum clones (A99:1, A99:4, and A99:7) had only 4–5-fold lower binding affinity compared to the corpus clones (C99:3, C99:5, and C99:8), whereas the exceptional antrum clones (A99:5, A99:6, and A99:10) and the single exceptional corpus clone C99:4 had >20-fold higher affinities compared to the antrum clones. At pH 4, the antrum clones (A99:1, A99:4, and A99:7) had dropped >10-fold in affinity, and because of their inherent 5-fold lower affinity at pH 6 they had in total >50-fold lower binding affinity compared to the corpus clones (C99:3, C99:5, and C99:8). In comparison, at pH 4 the exceptional antrum clones (A99:5, A99:6 and A99:10) and the single exceptional corpus clone C99:4 similarly had 50-fold higher affinity compared to the antrum clones and thus had affinities more similar to the corpus clones. The series of different affinities for the SO-2 antrum, corpus, and exceptional strains are illustrated by bars (means + SDs).

(F) Alignment of Velcro-domain of the paralogue *babB* to *babA* from the SO-2 Antrum, Antrum-Ex, Corpus-Ex, and Corpus isolates. Here *babB* is identical to corpus *babA* in the codons that correspond to Gly428, Thr432, and Leu434, but different in all three positions compared to the exceptional *babAs* and different in two of these three positions compared to antrum *babA*. The similarity in substitution patterns to USU101 and the 81G rhesus passage clones suggest that this part of the *bab* genes – the Velcro Domain – is a common target for functional recombination events.


A
low pH in gastric juice
pH gradient in mucus layer
higher pH at cell surfaces
Esophagus
Duodenum
corpus, more acidic
antrum, less acidic
B
BabA in SO-2 isolates
Key position
The aa428 Velcro position
C
Response (RU)
Time (s)
PG
PE
LG
LE
D
a
17875/Leb
b
17875babA1A2
Response RU
Time (s)
OD600
0.2
0.1
0.05
0
E
Binding Affinity(Ka, M-1)
pH50= 4.45
pH50= 4.4
pH50= 3.71
pH 4 pH 6
Antrum clones
Antrum Ex clones
Corpus clones
F
Recombination sites in the BabA Velcro-domain of SO-2 clones
aa428
aa432
aa434
A99:1 Antrum
A99:5 Antrum Ex.
C99:4 Corpus Ex.
C99:5 Corpus
babA
babB

Suppl. figure 4.

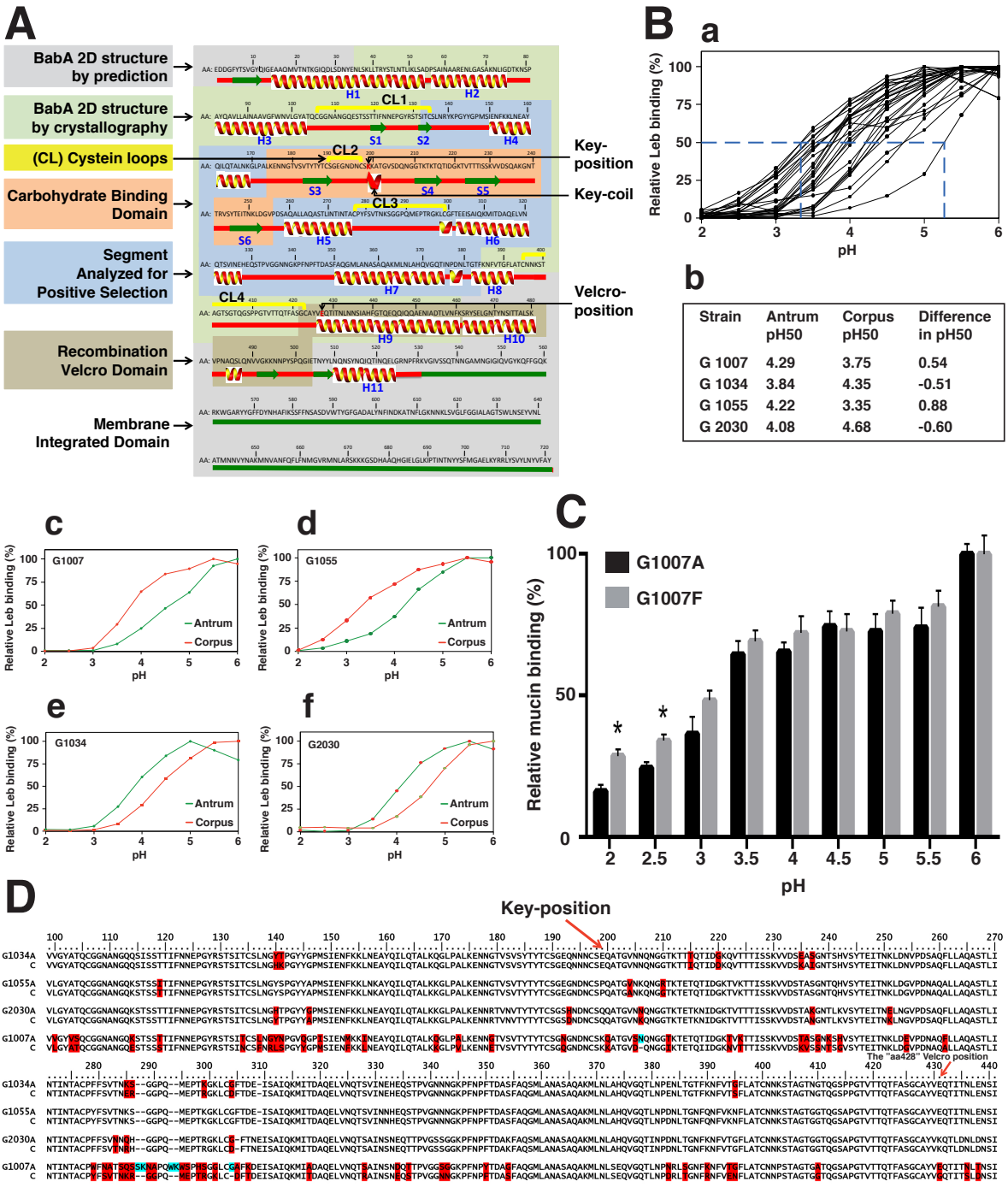
Figure S4. BabA Acid Sensitivity Determinants and Antrum vs. Corpus Adaptation in Acid Sensitivity in Binding among Clinical isolates. Related to Figure 4.

(A) BabA secondary structure and domains; The amino acid sequence is from strain 17875/Leb, and because of the polymorphism and strain variation in length of the BabA protein, the "aa428" position is here identical to position E427. The CBD domain and the secondary structure motifs (β-strands in green arrows, α-helices in red/yellow ribbons, and random coils in red strips) are derived from multiple secondary structure predictions and closely follow the recent crystal structures of BabA from strains 17875/Leb (Moonens et al., 2016) and J99 (Hage et al., 2015). For comparison, BabA from the specialist strain Peruvian P448 was similarly investigated with multiple secondary structure predictions (Figure 4A). The BabA adhesin is anchored in the outer membrane by an approximately 200 aa C-terminal predicted β-barrel domain (in green). The two major domains for Leb binding – the CBD (in orange), where the Key-position is located, and the Velcro Domain with aa428 (in brown) – are both structural parts of the Extracellular Domain (ECD). The Velcro Domain, with high rate of recombination in the corresponding gene location starts at ~aa420 (Figure 6B and 6C). The central part of the ECD (in blue) has been analyzed for adaptation and gain in Leb-binding by positive selection (Aspholm-Hurtig et al., 2004), i.e. the incidence of synonymous *vs*. non-synonymous amino acid replacements due to nucleotide mutations (Figure 4A).

(B) (a) pHgrams of the 17 Leb-binding strains from Athens, Greece (G). Out of these 17 strains, 13 strains demonstrated similar Leb binding in clones from both antrum and corpus tissue. The antrum had a median pH_{50} 3.84, and the corpus had a median pH_{50} 3.77 (mean pH_{50} 3.79), and was most similar to the median pH_{50} 3.7 of the Swedish (SW) strains (Figure 6F); (b) *H. pylori* strains G1007, G1034, G1055, and G2030 showed differences in acid sensitivity in Leb binding in clones from antrum *vs*. corpus biopsies (expressed as the difference in pH_{50} for the antrum (A) *vs*. corpus (C)). Two strains, G1034 and G2030, exhibited negative values for the antrum-corpus difference, i.e. the strains from the corpus tissue were more acid sensitive in Leb-binding compared to the clones from the antrum. (c–f) pHgrams of the four Greek strains demonstrating pH_{50} differences between antrum and corpus isolates.

(C) Purified human gastric MUC5AC mucin was probed with the Greek G1007A-antrum (pH_{50} = 4.29) and G1007F-corpus (pH_{50} = 3.75) clones. The acid-sensitive binding to human mucin is similar to the antrum and corpus clones' pH_{50} in Leb-binding (Bbc), where the corpus clones demonstrate higher binding strength at pH below 3, due to their acid-resistance in binding. The data was analyzed using a multiple t-test with Holm-Sidak's correction for multiple comparisons. Statistically significant differences between G1007A and F were observed at pH 2 ($P = 0.0005$) and at pH 2.5 ($P = 0.003$).

(D) Alignments of BabA antrum *vs.* corpus from the Greek strains G1007, G1034, G1055, and G2030. Amino acid substitutions between antrum *vs*. corpus BabA are indicated in red, and deletions are in blue. BabAs from G1034, G1055, G2030, and G1007 have E, P, Q, and K at position 199, respectively, and have E, E, K, and E/G at position 428, respectively.


A
BabA 2D structure by prediction
BabA 2D structure by crystallography
(CL) Cystein loops
Carbohydrate Binding Domain
Segment Analyzed for Positive Selection
Recombination Velcro Domain
Membrane Integrated Domain
Key-position
Key-coil
Velcro-position
H1
H2
CL1
H3
S1
S2
H4
CL2
S3
S4
S5
CL3
S6
H5
H6
H7
H8
CL4
H9
H10
H11
B
a
Relative Leb binding (%)
pH
b
Strain | Antrum pH50 | Corpus pH50 | Difference in pH50
G 1007 | 4.29 | 3.75 | 0.54
G 1034 | 3.84 | 4.35 | -0.51
G 1055 | 4.22 | 3.35 | 0.88
G 2030 | 4.08 | 4.68 | -0.60
c
G1007
d
G1055
e
G1034
f
G2030
Antrum
Corpus
C
G1007A
G1007F
Relative mucin binding (%)
D
Key-position
The "aa428" Velcro position
G1034A
G1055A
G2030A
G1007A
C

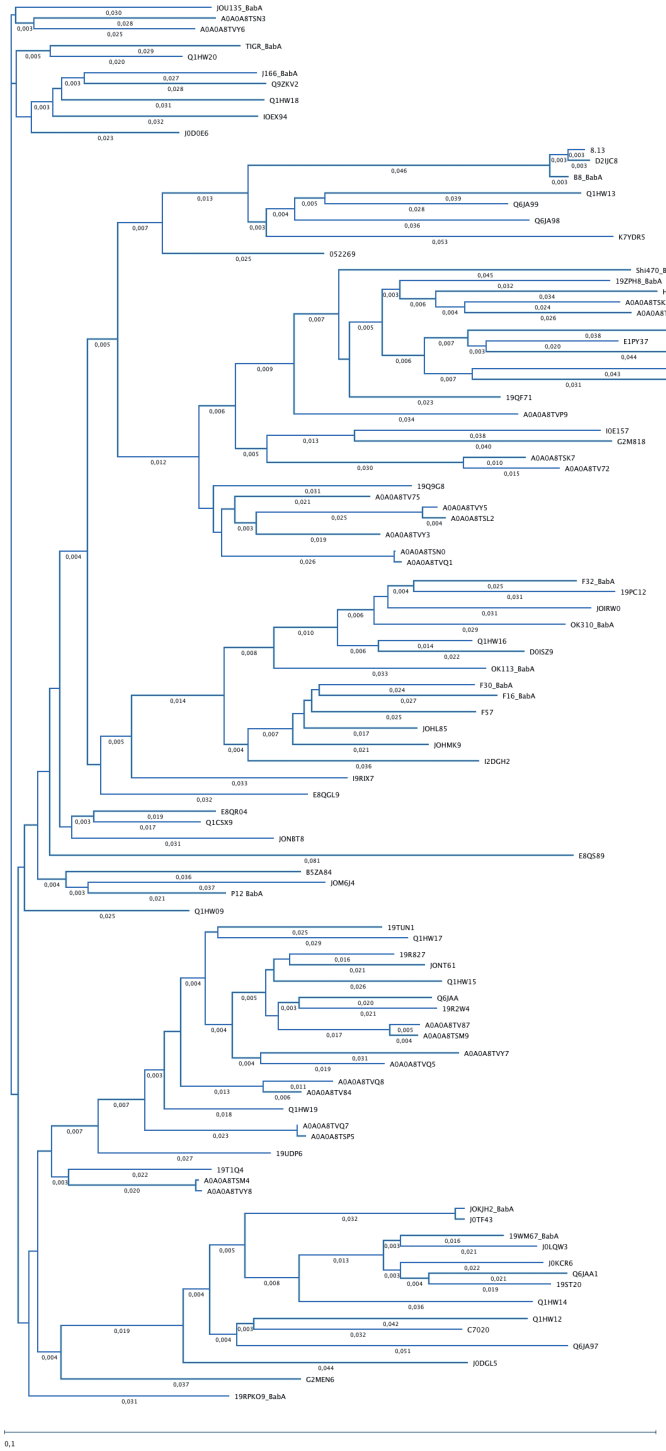
Suppl. figure 5.

Figure S5. The BabA Extraordinary Polypeptide Diversity and Acid Responsive Key-coil Among Clinical Isolates. Related to Figure 5.

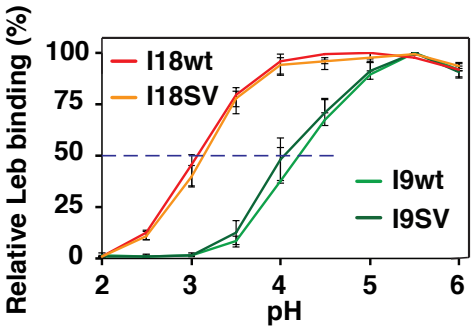
(A) The phylogenetic tree based on 98 full-length BabA amino acid sequences from strains representing different populations worldwide demonstrates the high diversity and polymorphisms in the BabA sequence. The comparison suggests that every individual has isolates with unique BabA sequences. Sequences were retrieved from UniProt http://www.uniprot.org and aligned with MegAlign Pro in the Lasergene software package (DNASTAR, Inc. Madison, USA). The phylogenetic tree was prepared with Tree View and the BIONJ algorithm. Pairwise distances between and within the sequences were calculated with Kimura parameters.

(B) Recombinant BabA from *H. pylori* I9 and I18 expressed from a plasmid shuttle vector (SV) in *H. pylori* strain P1Δ*babA* (I9-SV and I18-SV) faithfully reproduced the pH_{50} of the corresponding wild type-strains (I9wt and I18wt). For wtI9 pH_{50} = 4.21, for I9-SV pH_{50} = 4.05, for wtI18 pH_{50} = 3.06, and for I18-SV pH_{50} = 3.13.

A BabA world-wide polymorphism



B



Suppl. figure 6.

Figure S6. BabA Adaptation in Acid Sensitivity During Development of Gastric Cancer. Related to Figure 6.

(A) Tumorigenesis in rhesus macaque 81G during chronic *H. pylori* infection; (a) Normal antrum mucosa with basal gastric glands (1), and with some parietal cells (2) (horizontal orientation of tissue). (b) Normal corpus mucosa with a dominance of parietal cells (1) (vertical orientation of tissue). (c) Dysplasia of the antrum mucosa is characterized by branched structures of irregular size and shape in gastric glands and by an expanded lumen. In some glands, the lumen has expanded into several lines of glands. The glands are arranged randomly and present either as tight clusters (1) or with expanded distances between glands (2), which indicates atrophy in the gastric mucosa. (d) Dysplasia of the corpus mucosa. Here the glands are thickened and with different shapes and branching (polymorphism). They display infolding, and the cell nuclei are at different levels, i.e. cell piling (1). (e and g) Gastric adenocarcinoma with high differentiation in the antrum. The glands display an atypical structure (polymorphic and multi-layered with collapse in the individual cells) and invasive growth with infiltration into surrounding tissue (1), which results in destruction of blood vessels and bleeding (foci of hemorrhage) (2). (f and h) Gastric adenocarcinoma with high differentiation in the corpus and invasive growth and infiltration into surrounding tissue (1), including the lamina muscularis of the submucosa (2), and with atypical "giant" glands (3). Original magnification 200×.

(B) Identification of BabA-expressing Leb-binding single bacterial cells cultured from a biopsy obtained after 17 weeks of experimental infection in a rhesus macaque (Solnick et al., 2004). The *H. pylori* J166 cultures were mixed with Alexa555-labeled Leb conjugate (red), and binding was analyzed by fluorescence microscopy. A dilution series of *H. pylori* 17875/Leb was used to spike a suspension of the 17875*babA1babA2* mutant at dilutions of (a) 1:100 and (b) 1:1000. The dilution series, i.e. the mix of strains, was used as a reference to visualize the proportion of bacterial cells with intact Leb-binding properties, i.e. the number of red positive spots (Leb binding) relative to the dark background of non-binding bacterial cells. Here, the 17875wt strain stains red by Leb binding (a and b) in contrast to the 17875*babA1babA2* mutant (c), which is devoid of the *babA* gene and hence does not bind Leb. The 17-week output J166 strain from the rhesus macaque demonstrates the presence of Leb-binding bacterial cells more similar to the 1:100 dilution, i.e. ~1% of *H. pylori* bacterial cells have maintained their Leb-binding properties (d). Thus, a subpopulation of *H. pylori* bacterial cells maintains their *babA* gene and their Leb-binding properties, i.e. this subpopulation retains the expression of functional BabA during long-term experimental infection in rhesus macaques. We suggest that this might reflect gastric niche heterogeneity or the need for subpopulations that can quickly expand and return proficient adherence to infections when there a selective need for such forces, and all of this is superimposed on the corpus-antrum differences noted above.

(C) Isolation of BabA-expressing Leb-binding single bacterial clones after five years of experimental infection in rhesus macaques with strain USU101 (Liu et al., 2009). *H. pylori* from the 81G animal was spread and grown on plates until single colonies appeared. The colonies were transferred onto a nitrocellulose membrane and probed with biotinylated Leb-conjugate. Three Leb-binding colonies were identified both from the antrum culture (depicted by arrows) and from the corpus culture. Though ~99% of *H. pylori* isolates recovered from this macaque did not bind Leb, i.e. 6 Leb binders out of a total of 870 clones after 5 years of infection means there are ~0.5% Leb-positive clones present as subpopulations in both the antrum and corpus mucosa.

(D) pHgrams of *H. pylori* clones isolated from the antrum (A) and corpus (C) of rhesus macaque 81G five years after challenge with strain USU101 showing BabA adaptation in acid sensitivity in Leb-binding.

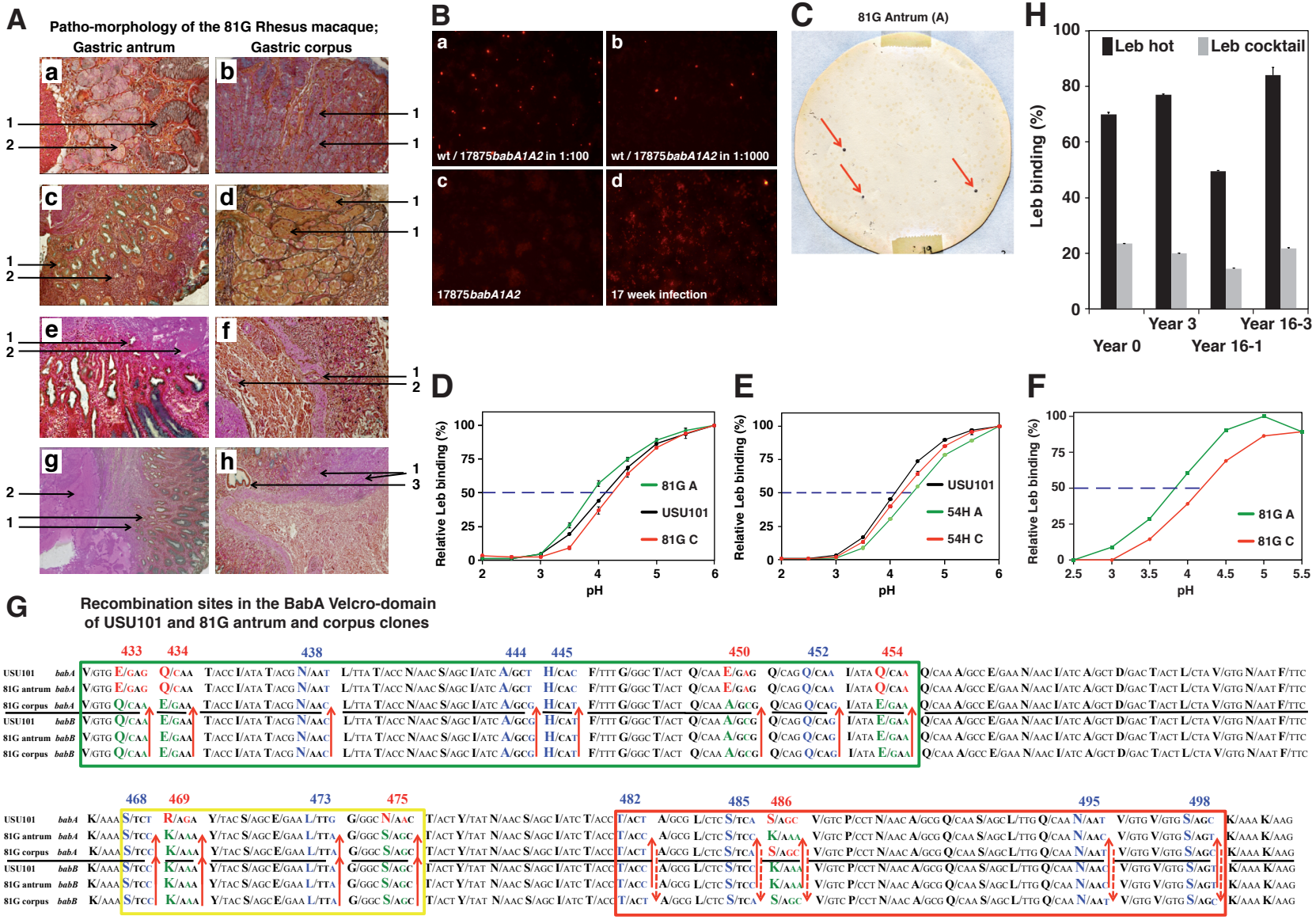
(E) The 54H antrum and 54H corpus clones had pH_{50} of 4.4 and 4.2, compared to 4.1 for the USU101 parent strain.

(F) Biacore pHgram analyses of Leb binding by recombinant $BabA_{526aa}$ protein produced in *E. coli* from the 81G antrum clone and the more acid-sensitive 81G corpus clone.

(G) Velcro Domain alignments of *babA*/BabA with the paralogue *babB*/BabB from the parental USU101 strain, the 81G antrum clones, and the 81G corpus clone. The series of amino substitutions are indicated: aa433 (identical to "aa428" in SO-2 antrum/corpus, Figure 3C), 434, 450, 454, 469, 475 and 486, and, in

addition, the synonymous mutations in codon positions 438, 444, 445, 452, 468, 473, 482, 485, 495 and 498, which corresponds to a total of 19 synonymous and non-synonymous nucleotide differences, with amino acid substitutions no. 1-4 in H9 of the corpus BabA, amino acid substitutions 5 and 6 in the common H9-H10 hinge, and substitution no. 7 in H10 of the antrum BabA (see also Figures 6B and 6C). This complexity suggests that this part of the *bab* genes, the Velcro Domain, has been the target of a series of recombination events both before and after separation of the original USU101 infection into antrum-specialized and corpus-specialized clones, i.e. recombination events between *babA* and *babB*; from *babB* to *babA* in red arrows; from *babA* to *babB* in dashed red arrows The patterns suggest that multiple recombination events in both directions have adapted and shaped BabA in the corpus region to higher acid sensitivity and reduced affinity for Leb binding and BabA in the antrum region to higher acid resistance and higher affinity for Leb binding. This is illustrated by the 0.35 pH unit difference in pH_{50} and the ~10-fold difference in affinity after five years of experimental infection (Figure 6A). Of particular interest, the recombination break points seemed to be in the Repeat-Sequence-1 (RS-1) junction (Ilver et al., 1998). The RS-1, displayed in Figures 3C, 6B, and 6E, which includes the last 900 nucleotides in the *babA/B/C* genes (from aa420 to the C-terminal aa720) and, in addition, some 500 nucleotides downstream, was originally suggested to promote homologous recombinations and shifts in genome positions of the *babA/B/C* genes (Alm et al., 2000; Colbeck et al., 2006). Here, all but one of the BabA substitutions resulted in change of charge, which can alter the electrostatic interactions between H9 and H10, and possible also with the H1 coiled-coil interaction.

(H) *H. pylori* clones from the GC-patient that had been followed by sequential collection of biopsies over 16 years were analyzed for Leb binding properties. The Radio Immuno Assay (RIA) was used both in "hot" mode and in "cocktail" mode. RIA with hot (undiluted) Leb at low nanogram levels provides limited availability of Leb and hence provides an initial assessment of Leb-binding affinity, whereas the "cocktail" mode, with an excess of Leb, provides information about the bacterial binding capacity, i.e. the numbers of bacterial binding sites or adhesin molecules. The Leb binding activity was similar both in terms of hot Leb as well as cocktail Leb over the 16-year period, i.e. in this individual neither binding affinity nor binding capacity changed dramatically over almost two decades of chronic infection.


A
Patho-morphology of the 81G Rhesus macaque;
Gastric antrum
Gastric corpus
B
wt / 17875babA1A2 in 1:100
wt / 17875babA1A2 in 1:1000
17875babA1A2
17 week infection
C
81G Antrum (A)
H
Leb hot
Leb cocktail
Leb binding (%)
Year 0
Year 16-1
Year 3
Year 16-3
D
Relative Leb binding (%)
81G A
USU101
81G C
pH
E
Relative Leb binding (%)
USU101
54H A
54H C
pH
F
Relative Leb binding (%)
81G A
81G C
pH
G
Recombination sites in the BabA Velcro-domain
of USU101 and 81G antrum and corpus clones
USU101 babA
81G antrum babA
81G corpus babA
USU101 babB
81G antrum babB
81G corpus babB

Suppl. figure 7.

Figure S7. BabA Multimerization and Reversible Acid De-Polymerization. Related to Figure 7.

(A) *H. pylori* protein expression patterns during experimental infection of rhesus macaques were analyzed by 2D-DIGE gel electrophoresis, where whole-cell protein extracts of *H. pylori* strain J166 before and after 18 weeks of experimental infection were compared. For the analysis, a genetic *babB* deletion mutant of strain *H. pylori* J166 was used as the input strain, i.e. for the experimental challenge. The absence of BabB protein would be expected to facilitate the interpretation of the migration patterns of the BabA protein because BabA and BabB have been suggested to form protein complexes (Bernarde et al., 2010). Because the BabA protein and the majority of *H. pylori* outer membrane (HOP) proteins are very basic in charge, i.e. with high isoelectric point (pI), the protein extracts were separated in the first dimension by use of pH 7–11 gradient strips to better focus BabA and HOP proteins. The whole-cell protein extract of the J166 input strain (i.e. the strain before macaque infection) was labeled green with Cy3 (a), whereas the whole-cell protein extract of the rhesus macaque 18-week passaged J166 strain was labeled red with Cy5 (b). The two protein extracts were then pooled, mixed, and separated on a gel. After protein separation in the first dimension according to pI and in the second dimension according to mass, all common proteins came out superimposed and co-localized in yellow (red + green) fluorescence, whereas proteins that are only or (highly) expressed in one or the other strain stain (more) green (such as the two bands with arrows in Figure 7A and visible in the same positions in Figure S7A) or (more) red in the 2D gels. The two green bands expressed in the J166*babB* input strain (and missing or with much lower expression in the 18 weeks macaque-passaged strain, were both identified as BabA by immunoblots with BabA antibodies (Figure 7B). The slower (upper) migrating green BabA band is presumably migrating in multimeric form (high molecular mass) in the second dimension, whereas the somewhat faster band migrates equivalent to an oligomer (Figure 7A). The BabA monomer was revealed by the immunoblot but the BabA protein is clustered together with protein of similar molecular mass and hence does not come out in green stain, but in yellow (Figure 7A and 7B).

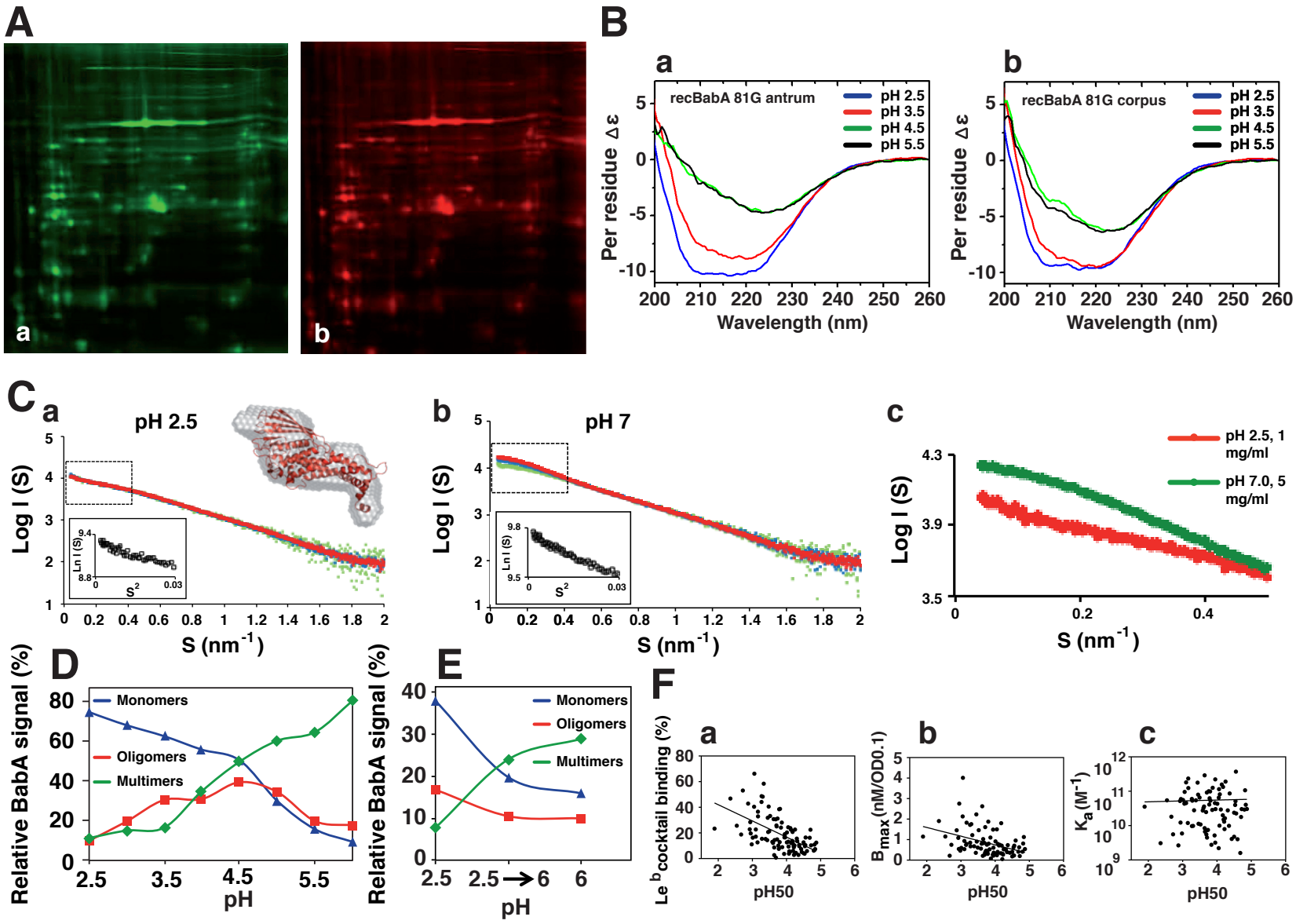
(B) Far-UV circular dichroism spectrum analysis (CD) of acid-induced unfolding of 81G antrum (a) and corpus (b) $BabA_{526}$. The pHs were 5.5 (blue), 4.5 (green), 3.5 (red), and 2.5 (black). Spectra are displayed in mean residue ellipticity ($\Delta\varepsilon$).

(C) Small angle X-ray scattering (SAXS) profiles of the BabA ectodomain at acidic and neutral pH. (a) At acidic pH the SAXS curves at different concentrations (1, 3 and 5 mg/mL colored green, blue and red, resp.) overlap, and the Guinier plot (boxed insert, 1 mg/mL) shows a linear correlation indicative of a mono-disperse solution. (a) The DAMMIF *ab initio* shape determination of the scattering curve at 1 mg/mL fits well with the monomeric structure of the BabA ectodomain (red cartoon model; Chi-square value of 1.34 for the comparison of experimental and theoretical scattering curves; PDB entry: 5F7K, Moonens *et al,.* 2016). (b) At neutral pH, the scattering curves (1, 3 and 5 mg/mL colored green, blue and red, resp.) at low angle show concentration dependence and deviate from curves for the monomeric protein as measured at low pH, indicating the formation of higher molecular weight complexes at neutral pH. (b) The Guinier plot (boxed insert; 5 mg/mL) shows the sample is aggregation free, demonstrating that the higher molecular weight species correspond to a discrete oligomer rather than non-specific protein aggregation. (c) The zoomed-in area of the dotted boxes in (a) and (b) are aligned for comparison. At the concentrations used, the ectodomain sample holds an oligomer/monomer ratio < 1 due to the absence of the transmembrane domain (see also Moonens *et al.,* 2016), precluding the calculation of an *ab initio* shape of the oligomer.

(D) Acid-responsive multimer dissociation. Application of glutaraldehyde to 17875/Leb bacteria and crosslinking of outer membrane proteins after prior exposure to a series of acidic buffers ranging from pH 2.5 to pH 6. SDS-PAGE BabA immunoblots in Figure 7E show that BabA monomers (in blue) predominated at lower pH, whereas BabA oligomers (in red) were abundant in the *H. pylori* cell membrane at pH 3.5–5, and high molecular mass multimers of BabA predominated in the *H. pylori* cell membrane at >pH 5 (in green). The BabA immunoblots of SDS-PAGE size-separated cross-linked proteins were scanned with an Odyssey Sa scanner (LI-Cor Biosciences, USA) and the IR-signal (from the secondary antibody) was processed by Image Studio Lite software.

(E) pH-responsive reconstitution of high molecular mass multimers. Reconstitution of BabA oligomerization by reconditioning to pH 6 after acid-induced multimer dissociation of BabA into monomers (Figure 7F). **pH 2.5**; Crosslinking of *H. pylori* 17875/Leb bacterial cells exposed to pH 2.5 showing significant multimer dissociation of BabA into monomers. **pH 2.5 → 6**; Crosslinking of 17875/Leb bacterial cells after prior pH 2.5 treatment followed by reconditioning with pH 6 buffer, with almost full recovery of high molecular mass BabA multimers. **pH 6**; Crosslinking 17875/Leb bacterial cells exposed to pH 6 with BabA predominant as high molecular mass multimers. The BabA IR-immunoblots were scanned and integrated as in (D). From the series of results, we conclude that acidification causes rapid dissociation of BabA multimers into monomers in the bacterial membrane and that reconstitution in binding affinity by neutralization of pH follows reassociation of BabA into multimers and the recovery of multivalency in binding. BabA multimerization is also supported by structural considerations, and the series of substitutions in BabA that affected both acid sensitivity and affinity in the rhesus macaque with gastric cancer were all located in the Velcro Domain (Figures 6B and 6C) similar to the E428G substitution in the Greek G1007 (Figure S4D) and Swedish SO-2 (Figure 3C; Figure S3B) isolates. This is particularly relevant because in BabA the CBD is carried on the Head Domain, which is supported by the coiled-coil handles of the Stalk Domain (Hage et al., 2015). Thus, Velcro Domain acts as a molecular hinge between Stalk and Head (Figure 4B), and together with the H1/10 coiled-coil structure can be presumed to increase avidity in binding through oligomerization (Meng et al., 2006).

(F) For comparison to the sets of strains in Figure 6G, the available set of strains from Sweden (n = 21)(Aspholm-Hurtig et al., 2004), Peru (n = 59 (Aspholm-Hurtig et al., 2004; Dailidiene et al., 2004), and India (n = 16 (this paper)) were tested for their correlation between acid sensitivity and binding capacity and for their correlation between acid sensitivity and binding affinity. (a) Binding capacity was first tested by RIA in "cocktail" mode with an excess of Leb-conjugate that provides information about the bacterial cell binding capacity, i.e. the numbers of bacterial binding sites or adhesin molecules. A total of 96 strains were tested, and a strong correlation was found between pH_{50} and the estimated binding capacity to Leb cocktail ($r_s = -0.60$, $p < 0.0001$, n = 96). (b) The majority of these strains (n = 87) were also analyzed for binding affinity and maximal binding capacity by Scatchard equilibrium in binding tests using a dilution series of Leb conjugate. Among the 87 strains tested, there was strong correlation between pH_{50} and maximal binding capacity (B_{max}) ($r_s = -0.42$, $p < 0.0001$, n = 87). The strong correlation between Leb "cocktail" binding and B_{max} ($r_s = 0.57$, $p < 0.0001$, n = 87) tells that binding capacity assessed with excess of Leb-conjugate (binding-saturating conditions) provides reliable data about the maximal bacterial cell binding capacity. Thus, this can serve as proxy of the total numbers of bacterial binding sites or adhesin molecules. (c) However, no correlation was found between pH_{50} and Leb-binding affinity (K_a) ($r_s = -0.04$, $p = 0.70$, n = 87). Similarly, there was no correlation between binding affinity and binding capacity ($r_s = 0.14$, $p = 0.182$, n = 87).


A
a
b
B
a
recBabA 81G antrum
pH 2.5
pH 3.5
pH 4.5
pH 5.5
Per residue Δε
Wavelength (nm)
b
recBabA 81G corpus
C
a
pH 2.5
Log I (S)
S (nm-1)
b
pH 7
c
pH 2.5, 1 mg/ml
pH 7.0, 5 mg/ml
D
Relative BabA signal (%)
Monomers
Oligomers
Multimers
pH
E
F
a
Le b cocktail binding (%)
pH50
b
Bmax (nM/OD0.1)
c
Ka (M-1)

Suppl. table 1.

Table S1. Summary of BabA Protein Purification Scheme from *H. pylori* 17875/Leb Bacterial Cells. Related to Figure 2E; Figure S2G.

The native BabA protein was isolated and purified from 24 g of *H. pylori* 17875/Leb bacterial paste harvested from 250 Brucella blood agar plates. The cell extract was enriched for outer membrane proteins by Zwitterionic detergent (ZW-12) solubilization. At pH 8.0, the majority of solubilized proteins were eliminated in the flow-through by cation exchange (CEX) (Source 30S) chromatography. Next, the eluted BabA protein fractions were pooled, and ZW-12 detergent was removed by addition of octyl-glucoside detergent for facilitated refolding. The pooled BabA material was further fractioned by Leb-affinity chromatography, where the pure BabA protein was eluted by a pH gradient and fractions were immediately neutralized. A total of 3.5 mg of homogenous BabA protein was recovered. The total yield of BabA protein from the purification process was ~17%.

Purification steps	Total protein (mg)	BabA (mg)	Enrichment	Yield %
Bacterial cell paste	3358	21.2	1	100
ZW-12 extracted protein	3010	20.3	1.1	96
Source 30 S	47.6	10.2	33.9	48
NHS-Leb	3.5	3.5	158.4	16.5

Suppl. table 2.

Table S2. Prevalence of Different Amino Acids in the BabA 199 Key-Position in Human Populations Worldwide. Related to Figure 5C.

Proline (P) or leucine (L) as found in the aa199 Key position of the corpus *vs.* antrum, respectively, in SO-2 strains are not ubiquitous. In Europe, 20% of the strains carry P199, which increases in prevalence to ~50% in Japan and Alaska, and to almost 100% in South American Amerindians. In contrast, in Europe, proline is often substituted with a charged amino acid, either the positively charged lysine (K), e.g. in CCUG17875 (17875/Leb) (Figures 5B and S4A), or the negatively charged glutamic acid (E) e.g. strain G1034 from Greece (Figure S4D), with a trend for higher levels of polar (including charged) residues in southern *vs*. northern Europe. In addition to the almost universal P199 among Asian and Amerindian BabA, there is isoleucine (I), the second most common Key-position in Japan and Alaska (isoleucine is similar to leucine in geometrical structure and characteristics). A special feature of the BabA Key-position in Japan and Greece is the high prevalence of the polar amino acids serine (S) and glutamine (Q), respectively. The contrasting scenario is observed in India, where the vast majority of BabA sequences are missing both the Key position and the most proximate residues, which results in with high acid sensitivity in binding. In the Table, the Key-position amino acids are grouped according to non-polar *vs.* polar characteristics, with the latter including both neutral and charged amino acids.

Population (number of strains)		Amino acid in position aa 199 (%)														
		Non-polar						Polar								
								Neutral				Basic			Acidic	Total, %
		P	L	I	A	V	Total,%	S	T	Y	Q	K	R	H	E	
Europe	Sweden (20)	25	10	-	5	-	**40**	-	10	-	5	35	-	-	10	**60**
	Spain (13)	27	-	6	-	-	**33**	-	-	-	-	40	-	-	27	**67**
	Greece (5)	20	-	-	-	-	**20**	-	-	-	40	20	-	-	20	**80**
Asia	India (16)	6	-	-	-	6	**12**	-	-	-	6	-	6	-	-	**12**
	Japan (12)	50	-	17	-	-	**67**	25	-	8	-	-	-	-	-	**33**
America	Alaska (13)	54	-	30	-	-	**84**	-	8	-	-	8	-	-	-	**16**
	Peru (23)	88	4	-	-	-	**92**	-	-	-	-	4	-	4	-	**8**

Suppl. table 3A and 3B.

Table S3A. Histo-Pathological Analysis of Gastric Mucosa from Rhesus macaque 81G. Related to Figure 6A; Figure S6A and S6D.

A rhesus macaque (81G) had received both *H. pylori* USU101 from a gastric cancer patient and the ENNG carcinogen (Liu et al., 2009; Liu et al., 2015). Biopsies were collected from both antrum and corpus tissue before the experimental infection and at month 1, 12, 22, 35, 42, 48, and 66 and at month 72 at autopsy when the animal died from internal hemorrhage of the stomach and small bowel. Necropsy showed ulceration, nodularity, and blood clots in the stomach. The histopathology analyses of the biopsies showed gastric inflammation from first month of experimental infection with atrophic gastric and intestinal metaplasia in antrum tissue starting as early as the month 12. The tumorigenic cascade continued with development of dysplasia at month 35, followed by cancer in situ (CIS) at month 48, i.e. at the beginning of the 5th year, followed by invasive cancer 18 months later at month 66. Cancer development in the gastric corpus was slower compared to the antrum with the appearance of intestinal metaplasia at month 48 and dysplasia/ cancer at month 66, albeit with invasive cancer only at 72 months, i.e. after 6 years.

Table S3B. Histo-Pathological Analysis of Gastric Mucosa from Rhesus macaque 54H. Related to Figure 6A; Figure S6E.

The *H. pylori* strain USU101 from a gastric cancer patient had been used to infect also rhesus macaque (54H) but with no supplementation of ENNG (Liu et al., 2009; Liu et al., 2015). Biopsies were collected from both antrum and corpus tissue at month 62. The histopathology analyses of the biopsies showed gastric inflammation with intestinal metaplasia (in antrum), but no dysplasia, nor cancer.

Table S3A

	Chronic gastritis	Oxyntic atrophy	Fibrosis	Atrophic gastritis	Foveolar hyperplasia	Intestinal metaplasia	Dysplasia	CIS	Invasion (cancer)	Time course
Histologic scoring of *H. pylori* infection in 81G animal	Inflammation with epithelial defects, gland dilatation	Loss of oxyntic parietal cells	Focal fibrosis	Complete loss of oxyntic parietal and chief cells, focal fibrosis	Areas of expanded surface mucous cells in fundic epithelium	Columnar elongation, mucous droplets occasionally forming goblet cells	Irregular size and shape glands, infolding, branching and cell piling;	Cellular and nuclear atypia	Penetration of basal membrane	
Morphological changes in gastric corpus mucosa										
3405	-	-	+	-	-	-	-	-	-	Pre-inoculation
3421	-	-	-	-	+	-	-	-	-	Pre-inoculation
3479	+	-	-	-	-	-	-	-	-	1 Month
3485	-	-	+	-	-	-	-	-	-	1 Month
3795	+	-	+	-	-	-	-	-	-	12 M
4051	+	-	-	-	-	-	-	-	-	22 M
4363	+	-	-	-	-	-	-	-	-	35 M
4457*	+	-	+	-	+	-	-	-	-	42 M
4583	+	+	+	-	+	+	-	-	-	48 M
4891	+	+	+	-	-	-	+	+	-	66 M
body	-	-	-	+	+	+	+	+	+	72 M Autopsy
Morphological changes in gastric antrum mucosa										
3406	-	-	-	±	+	-	-	-	-	Pre-inoculation
3422	-	-	-	-	-	-	-	-	-	Pre-inoculation
3480	+	-	+	-	-	-	-	-	-	1 Month
3486	+	-	+	-	+	-	-	-	-	1 Month
3796	+	+	+	+	-	+	-	-	-	12 M
4052	-	-	+	+	+	+	-	-	-	22 M
4364	+	-	-	+	+	+	+	-	-	35 M
4458	+	-	-	+	+	+	+	-	-	42 M
4584	+	-	-	+	+	+	+	+	-	48 M
4892	-	-	-	+	+	+	+	+	+	66 M
antrum	-	-	-	+	+	+	+	+	+	72 M Autopsy

XX M –without proliferative changes XX-M –dysplasia of glandular epithelium XX-M – Carcinoma or CIS

Table S3B

	Acute gastritis	Chronic gastritis	Oxyntic atrophy	Fibrosis	Atrophic gastritis	Foveolar hyperplasia	Intestinal metaplasia	Dysplasia	CIS	Invasion (cancer)
Histologic scoring of *H. pylori* infection in 54H animal	Infiltration of mucosal and submucosal lymphocytes with pockets of neutrophils, edema, mild mucosal defects	Inflammation with epithelial defects, gland dilatation	Loss of oxyntic parietal cells	Focal fibrosis	Complete loss of oxyntic parietal and chief cells, focal fibrosis	Areas of expanded surface mucous cells in fundic epithelium	Columnar elongation, mucous droplets occasionally forming goblet cells	Irregular size and shape glands, infolding, branching and cell piling;	Cellular and nuclear atypia	Penetration of basal membrane
The 54H animal at 62 months										
Morphological changes in gastric corpus mucosa										
4793A C	-	+	-	+	-	-	-	-	-	-
4793B C	-	+	-	+	-	-	-	-	-	-
Morphological changes in gastric antrum mucosa										
4794A A	-	+	-	+	-	-	+	-	-	-
4794B A	-	+	-	+	-	-	+	-	-	-

Suppl. table 4.

Table 4. Nucleotide sequence analysis of *babA, babB, cysS* and *glr*.
Related to Figures 2-6 and Figures S2-S6 and Suppl. Exp. Proc./ Sequence Analysis.

Gene	Population/ Isolate	Primer Name	Sequence	Reference
babA	Colombia/ GC-patient	Hp1AS-F	5′-ACCCTAATGGGCATGTGGTA-3′	Hennig *et al*., 2006
		BabAR1-R	5′-TTTGCCGTCTATGGTTTGG-3′	Colbeck *et al.,* 2006
		1539-R	5′-TTAATAAGCGAACACATAGTTCAAATAC-3′	This study
		A26-R	5′-TTGCTCCACATAGGCGCAAC-3′	Aspholm-Hurtig *et al.*, 2004
		I1-F	5′-CACTTGCTCAGGGGAAGGGAA-3′	This study
	Greece	HypDF1-F	5′-TTTTGAGCCGGTGGATATATTAG-3′	Colbeck *et al.,* 2006
		S18F1-F	5′-CTTTAATCCCCTACATTGTGGA`-3′	Colbeck *et al.,* 2006
		Hp1AS-F	5′-ACCCTAATGGGCATGTGGTA-3′	Hennig *et al*., 2006
		babA2-leader-F	5′-GCTTTTAGTTTCCACTTTGAG-3′	This study
		BabAR1-R	5′-TTTGCCGTCTATGGTTTGG-3′	Colbeck *et al.,* 2006
		OP9647	5′-GGTATCGATCCACTTCCATCAC-3′	Hennig *et al*., 2006
		SE05-R	5′-GCTGTTGAGGGGTTGTTGCATGTGGCTAAAAAG-3′	This study
		SE06-F	5′-CGAGCACGCTCATTAACACCATCAACACGG-3′	This study
		1539-R	5′-TTAATAAGCGAACACATAGTTCAAATAC-3′	This study
		A27-R	5′-TAAGTCCACACGTCAGAAGC-3′	This study
	India	J11-R	5′-TGTGTGCCACTAGTGCCAGC-3′	Aspholm-Hurtig *et al.*, 2004
		A29-R	5′-GTGCCAGCGTTGATTTGTTGTTGCATGT-3′	This study
		A10-F	5′-GCGTAGGCTATCAAATCGGT-3′	This study
		I1-F	5′-CACTTGCTCAGGGGAAGGGAA-3′	This study
		A27-R	5′-TAAGTCCACACGTCAGAAGC-3′	This study
	Sweden/ SO-2	1539II2-F	5′-TACGCAATGTGGGGGTAATGCT-3′	This study
		A37-R	5′-AATGCTCATAGGGCCGTAGTA-3′	This study
		A26-R	5′-TTGCTCCACATAGGCGCAAC-3′	Aspholm-Hurtig *et al.*, 2004
		A27-R	5′-TAAGTCCACACGTCAGAAGC-3′	This study

	Sweden/ SW7, SW38	J10-F	5′-AAAGCACCTCTTCAACCACC-3′	Aspholm-Hurtig *et al.*, 2004
		J11-R	5′-TGTGTGCCACTAGTGCCAGC-3′	Aspholm-Hurtig *et al.*, 2004
		A10-F	5′-GCGTAGGCTATCAAATCGGT-3′	This study
		A27-R	5′-TAAGTCCACACGTCAGAAGC-3′	This study
	81G from macaque	HypDF1-F	5′-TTTTGAGCCGGTGGATATATTAG-3′	Colbeck *et al.,* 2006
		S18F1-F	5′-CTTTAATCCCCTACATTGTGGA-3′	Colbeck *et al.,* 2006
		Hp1AS-F	5′-ACCCTAATGGGCATGTGGTA-3′	Hennig *et al*., 2006
		BabAR1-R	5′-TTTGCCGTCTATGGTTTGG-3′	Colbeck *et al.,* 2006
		OP9647	5′-GGTATCGATCCACTTCCATCAC-3′	Hennig *et al*., 2006
		1539-R	5′-TTAATAAGCGAACACATAGTTCAAATAC-3′	This study
		A27-R	5′-TAAGTCCACACGTCAGAAGC-3′	This study
babB	Sweden/ SO-2 (Locus A)	HypDF1-F	5′-TTTTGAGCCGGTGGATATATTAG-3′	Colbeck *et al.,* 2006
		BabBA1-R	5′-TCGCTTGTTTTAAAAGCTCTTGA-3′	Colbeck *et al.,* 2006
		1539-R	5′- TTAATAAGCGAACACATAGTTCAAATAC-3′	This study
	Sweden/ SO-2 (Locus B)	S18F1-F	5′-CTTTAATCCCCTACATTGTGGA-3′	Colbeck *et al.,* 2006
		BabBR1-R	5′-TCGCTTGTTTTAAAAGCTCTTGA-3′	Colbeck *et al.,* 2006
		1539-R	5′-TTAATAAGCGAACACATAGTTCAAATAC-3′	This study
	81G from macaque	SE22	5′-GTGGTAGCAGTTATTGGCAAGGG-3′	This study
		A72	5′-GGGTTACGCCCTAATTCTTGGTTGATGGTTTGGTTG-3′	This study
cysS	Colombia, Greece, Sweden/ SO-2	cysS-F	5’-CTACGGTGTATGATGACGCTCA	Dailidiene *et al.,* 2004
		cysS-R	5’-CCTTGTGGGGTGTCCATCAAAG	Dailidiene *et al.,* 2004
glr		5733	5’-CACATCGCCGCTCGCATGA	Dailidiene *et al.,* 2004
		6544	5’-AAGCTTTGTGTATTCTAAAATGCAAC	Dailidiene *et al.,* 2004

Suppl. table 5.

Table S5. GenBank accession numbers.

The *babA*, *babB*, *cysS* and *glr* sequences have been deposited in GenBank under the accession numbers Y613381-KY613496. Related to Figures 2-6 and Figures S2-S6.

***babA*; Sweden/ SO-2**					
Bylund.sqn	SO-2(A90-7)	KY613381	Bylund.sqn	SO-2(A90-3)	KY613382
Bylund.sqn	SO-2(A90-4)	KY613383	Bylund.sqn	SO-2(A90-5)	KY613384
Bylund.sqn	SO-2(A90-6)	KY613385	Bylund.sqn	SO-2(A90-8)	KY613386
Bylund.sqn	SO-2(A90-14)	KY613387	Bylund.sqn	SO-2(A90-15)	KY613388
Bylund.sqn	SO-2(A90-16)	KY613389	Bylund.sqn	SO-2(A99-1)	KY613390
Bylund.sqn	SO-2(A99-2)	KY613391	Bylund.sqn	SO-2(A99-3)	KY613392
Bylund.sqn	SO-2(A99-4)	KY613393	Bylund.sqn	SO-2(A99-5)	KY613394
Bylund.sqn	SO-2(A99-6)	KY613395	Bylund.sqn	SO-2(A99-7)	KY613396
Bylund.sqn	SO-2(A99-8)	KY613397	Bylund.sqn	SO-2(A99-9)	KY613398
Bylund.sqn	SO-2(A99-10)	KY613399	Bylund.sqn	SO-2(C99-1)	KY613400
Bylund.sqn	SO-2(C99-3)	KY613401	Bylund.sqn	SO-2(C99-4)	KY613402
Bylund.sqn	SO-2(C99-5)	KY613403	Bylund.sqn	SO-2(C99-7)	KY613404
Bylund.sqn	SO-2(C99-8)	KY613405	Bylund.sqn	SO-2(C99-10)	KY613406
Bylund.sqn	SO-2(C99-12)	KY613407	Bylund.sqn	SO-2(C99-14)	KY613408
***babA*; Sweden/ SW38-7, SW-7 and SW-38**					
Bylund.sqn	SW38-7	KY613409	Bylund.sqn	SW7	KY613410
Bylund.sqn	SW38	KY613411			
***babA*; Greece**					
Chernov_BabA.sqn	G1007C	KY613413	Chernov_BabA.sqn	G1007A	KY613412
Chernov_BabA.sqn	G1034C	KY613415	Chernov_BabA.sqn	G1034A	KY613414
Chernov_BabA.sqn	G1055C	KY613417	Chernov_BabA.sqn	G1055A	KY613416
Chernov_BabA.sqn	G2030C	KY613419	Chernov_BabA.sqn	G2030A	KY613418
***babA;* 81G from rhesus macaque and USU101 cancer patient**					
Chernov_BabA.sqn	81GA	KY613420			
Chernov_BabA.sqn	USU101	KY613422	Chernov_BabA.sqn	81GC	KY613421
***babB*; Sweden/ SO-2**					
ChernovBabB.sqn	SO-2(A99-5)	KY613424	ChernovBabB.sqn	SO-2(A99-1)	KY613423
ChernovBabB.sqn	SO-2(C99-5)	KY613426	ChernovBabB.sqn	SO-2(C99-4)	KY613425
***babB;* 81G from rhesus macaque and USU101 cancer patient**					
ChernovBabB.sqn	81GA	KY613427			
ChernovBabB.sqn	USU101	KY613429	ChernovBabB.sqn	81GC	KY613428
***cysS;* Greece**					
Chernovcys.sqn	1007C	KY613431	Chernovcys.sqn	1007A	KY613430
Chernovcys.sqn	1034C	KY613433	Chernovcys.sqn	1034A	KY613432
Chernovcys.sqn	1055C	KY613435	Chernovcys.sqn	1055A	KY613434

Chernovcys.sqn	2030C	KY613437	Chernovcys.sqn	2030A	KY613436
***cysS;* Colombian GC patient**					
Chernovcys.sqn	GC-year3	KY613439	Chernovcys.sqn	GC-year0	KY613438
Chernovcys.sqn	GC-year16-3	KY613441	Chernovcys.sqn	GC-year16-2	KY613440
			Chernovcys.sqn	GC-year16-1	KY613442
***glr;* Greece**					
Chernovglr.sqn	1007A	KY613443	Chernovglr.sqn	1007C	KY613444
Chernovglr.sqn	1034A	KY613445	Chernovglr.sqn	1034C	KY613446
Chernovglr.sqn	1055A	KY613447	Chernovglr.sqn	1055C	KY613448
Chernovglr.sqn	2030A	KY613449	Chernovglr.sqn	2030C	KY613450
***glr;* Colombian GC patient**					
Chernovglr.sqn	GC-year0	KY613451	Chernovglr.sqn	GC-year3	KY613452
Chernovglr.sqn	GC-year16-1	KY613453	Chernovglr.sqn	GC-year16-2	KY613454
Chernovglr.sqn	GC-year16-3	KY613455			
***babA*; India**					
MendezBabA.sqn	I17	KY613457	MendezBabA.sqn	I9	KY613456
MendezBabA.sqn	I21	KY613459	MendezBabA.sqn	I18	KY613458
MendezBabA.sqn	I60	KY613461	MendezBabA.sqn	I46	KY613460
MendezBabA.sqn	I90	KY613463	MendezBabA.sqn	I78	KY613462
MendezBabA.sqn	I93	KY613465	MendezBabA.sqn	I92	KY613464
MendezBabA.sqn	I102	KY613467	MendezBabA.sqn	I99	KY613466
MendezBabA.sqn	I110	KY613469	MendezBabA.sqn	I109	KY613468
MendezBabA.sqn	I120	KY613471	MendezBabA.sqn	I119	KY613470
***babA;* Colombian GC patient**					
MendezBabA.sqn	GC-year3	KY613473	MendezBabA.sqn	GC-year0	KY613472
MendezBabA.sqn	GC-year16-2	KY613475	MendezBabA.sqn	GC-year16-1	KY613474
			MendezBabA.sqn	GC-year16-3	KY613476
***cysS*; Sweden SO-2**					
MendezcysS.sqn	SO-2(A90-4)	KY613477	MendezcysS.sqn	SO-2(A90-6)	KY613478
MendezcysS.sqn	SO-2(A99-4)	KY613479	MendezcysS.sqn	SO-2(A99-5)	KY613480
MendezcysS.sqn	SO-2(A99-7)	KY613481	MendezcysS.sqn	SO-2(A99-9)	KY613482
MendezcysS.sqn	SO-2(A99-10)	KY613483	MendezcysS.sqn	SO-2(C99-1)	KY613484
MendezcysS.sqn	SO-2(C99-8)	KY613485	MendezcysS.sqn	SO-2(C99-12)	KY613486
***glr;* Sweden SO-2**					
Mendezglr.sqn	SO-2(A90-4	KY613487	Mendezglr.sqn	SO-2(A90-6	KY613488
Mendezglr.sqn	SO-2(A99-4)	KY613489	Mendezglr.sqn	SO-2(A99-5)	KY613490
Mendezglr.sqn	SO-2(A99-7)	KY613491	Mendezglr.sqn	SO-2(A99-9)	KY613492
Mendezglr.sqn	SO-2(A99-10)	KY613493	Mendezglr.sqn	SO-2(C99-1)	KY613494
Mendezglr.sqn	SO-2(C99-8)	KY613495	Mendezglr.sqn	SO-2(C99-12)	KY613496

Suppl. movie 1.

Movie S1 with Movie-Figure. The oncopathogen Helicobacter pylori benefits from the gastric mucus layer pH gradient for timely escape from shedding cells and mucus. By chemotaxis-driven return to the epithelium, H. pylori successfully establishes life-long infection through daily selection of the best-adapted and most prosperous bacteria. Related to Figure 1.

Helicobacter pylori (*H. pylori*) inhabit the gastric epithelium and overlying viscous mucus only hundreds of microns from the gastric lumen, which is too acidic for most microbes, *H. pylori* included. *H. pylori* is able to survive in the mucus niche because it can make use of the pH gradient formed in the mucus layer that ranges from near neutral in close proximity to the bicarbonate-buffered epithelium to the exceedingly low pH of the gastric juice. Cells and mucus are constantly being exfoliated, where the protective mucus mucin layer is renewed twice a day and the surface epithelial cells are turned over every 2–3 days. Thus, if *H. pylori* with its characteristic high binding strength would irreversibly adhere to the gastric cells and mucus by BabA-mediated binding, the *H. pylori* infection would be rapidly and spontaneously eradicated by the acidic gastric juices. Unfortunately for us humans, this is not the case; instead, *H. pylori* establish life-long persistent infections with elevated risk for gastric cancer development. This is best explained by the fact that its BabA adhesin looses its binding affinity as pH decreases and thus the bacteria are released from exfoliated gastric cells that are drifting towards the gastric lumen (red atmosphere). Chemotaxis-directed motility away from high acidity propels *H. pylori*'s return to the buffered epithelial surface (higher pH in darker blue). Here *H. pylori* will again have full access to the nutrients leaching from the cells through its manipulation of epithelial morphology, immune responses, and signaling pathways by transfer of its main virulence factors, the onco-protein CagA and the vacuolating cytotoxin (VacA) to the host cells. The continuous desquamation of mucosal debris, rapid but reversible BabA inactivation at acidic pH, chemotactic swimming to the more neutral epithelial surface, and restitution of binding activity is postulated to allow *H. pylori* to ride a continuous loop that both ensures persistence at initial infection sites and facilitates spreading throughout the gastric mucosa, including the corpus region with associated pangastritis and risk for atrophy. Such a mechanism provides for daily recycling of infection and functional selection for *H. pylori* strains with gain of fitness for the current local conditions, strains that will form the next generation infection for life-long persistence and risk for overt disease. Our hypothesis is that *H. pylori* adheres tightly to the gastric epithelium, but the bacterial cells are dislodged by the rapid turnover of epithelial cells and the mucus layer. However, by acid responsive inactivation of Leb binding, *H. pylori* detaches before it reaches the bactericidal gastric juice. Guided by chemotaxis, *H. pylori* instead successfully returns and reattaches to its niche in the gastric epithelium.