RESEARCH ARTICLE

Dynamics of Lewis b Binding and Sequence Variation of the *babA* Adhesin Gene during Chronic *Helicobacter pylori* Infection in Humans

Sandra Nell,[a,b] Lynn Kennemann,[a] Sandra Schwarz,[a*] Christine Josenhans,[a,b] Sebastian Suerbaum[a,b]

Institute of Medical Microbiology and Hospital Epidemiology, Hannover Medical School, Hannover, Germany[a]; German Center for Infection Research (DZIF), Hannover, Germany[b]

* Present address: Sandra Schwarz, Institute of Medical Microbiology and Hygiene, University of Tübingen, Tübingen, Germany.

S. Nell, L. Kennemann, and S. Schwarz contributed equally to this article.

ABSTRACT ***Helicobacter pylori*** **undergoes rapid microevolution during chronic infection, but very little is known about how this affects host interaction factors. The best-studied adhesin of *H. pylori* is BabA, which mediates binding to the blood group antigen Lewis b [Le(b)]. To study the dynamics of Le(b) adherence during human infection, we analyzed paired *H. pylori* isolates obtained sequentially from chronically infected individuals. A complete loss or significant reduction of Le(b) binding was observed in strains from 5 out of 23 individuals, indicating that the Le(b) binding phenotype is quite stable during chronic human infection. Sequence comparisons of *babA* identified differences due to mutation and/or recombination in 12 out of 16 strain pairs analyzed. Most amino acid changes were found in the putative N-terminal extracellular adhesion domain. One strain pair that had changed from a Le(b) binding to a nonbinding phenotype was used to study the role of distinct sequence changes in Le(b) binding. By transformations of the nonbinding strain with a *babA* gene amplified from the binding strain, *H. pylori* strains with mosaic *babA* genes were generated. Recombinants were enriched for a gain of Le(b) binding by biopanning or for BabA expression on the bacterial surface by pulldown assay. With this approach, we identified several amino acid residues affecting the strength of Le(b) binding. Additionally, the data showed that the C terminus of BabA, which is predicted to encode an outer membrane β-barrel domain, plays an essential role in the biogenesis of this protein.**

IMPORTANCE ***Helicobacter pylori*** **causes a chronic infection of the human stomach that can lead to ulcers and cancer. The bacterium can bind to gastric epithelial cells with specialized outer membrane proteins. The best-studied protein is the BabA adhesin which binds to the Lewis b blood group antigen. Since *H. pylori* is a bacterium with very high genetic variability, we asked whether *babA* evolves during chronic infection and how mutations or recombination in *babA* affect binding. We found that BabA-mediated adherence was stable in most individuals but observed a complete loss of binding or reduced binding in 22% of individuals. One strain pair in which binding was lost was used to generate *babA* sequences that were mosaics of a functional allele and a nonfunctional allele, and the mosaic sequences were used to identify amino acids critically involved in binding of BabA to Lewis b.**

Received 11 November 2014 **Accepted** 17 November 2014 **Published** 16 December 2014

Citation Nell S, Kennemann L, Schwarz S, Josenhans C, Suerbaum S. 2014. Dynamics of Lewis b binding and sequence variation of the *babA* adhesin gene during chronic *Helicobacter pylori* infection in humans. mBio 5(6):e02281-14. doi:10.1128/mBio.02281-14.

Editor Martin J. Blaser, New York University

Address correspondence to Sebastian Suerbaum, suerbaum.sebastian@mh-hannover.de.

This article is a direct contribution from a Fellow of the American Academy of Microbiology.

The human gastric pathogen *Helicobacter pylori* infects more than half of the world's population. Infection is usually acquired during early childhood, persists lifelong, and results in chronic inflammation of the gastric epithelium, which remains asymptomatic in most cases. However, long-lasting chronic infection can lead to severe sequelae including gastric and duodenal ulcers, gastric cancer, or mucosa-associated lymphoid tissue (MALT) lymphoma (1).

H. pylori resides within the mucous layer in close proximity to the gastric epithelial cells (2–4). Dynamic adhesion to the mucosal epithelium plays an important role in establishing persistent colonization and induction of gastric inflammation (5). The *H. pylori* genome encodes a large number of outer membrane proteins (OMPs) (6). While the functions of most OMPs are still unknown, some of them have been shown to be involved in bacterial adhesion to host cells (7–12). The best-characterized OMP of *H. pylori* is the blood group antigen-binding adhesin (BabA) that mediates binding to fucosylated Lewis b [Le(b)] and related histo-blood group antigens located on the surfaces of gastric epithelial cells and mucins (10, 11). BabA expression has been shown to be associated with more severe inflammation and clinical disease (13). Despite this, not all *H. pylori* strains possess a *babA* gene, and not all BabA-expressing strains are able to bind to Le(b). In addition, BabA proteins from diverse *H. pylori* strains can vary substantially in their specificity and binding strength (14–16). The molecular

basis of these different binding properties is not known, and so far no structural analysis of BabA has been published.

babA belongs to the Hop group of OMPs which show remarkable homology at their 5′ and 3′ ends (6, 12). The two most closely related *babA* paralogs are *babB* and *babC* for which no function has been identified. The extensive sequence homology between *babA*, *babB*, and *babC* has been shown to enable intrachromosomal recombination between the three genes (17). *babA* and its two closely related paralogs have been found in three different chromosomal locations, referred to as locus A, locus B, and locus C, respectively (14, 16). Experimental infection of different animal models with a Le(b) binding *H. pylori* strain resulted in the frequent loss of the Le(b) binding ability in the course of chronic infection (18–20). Based on these findings, it has been suggested that the dynamic modulation of adherence might facilitate adaptation of *H. pylori* to the changing conditions in the gastric environment (5).

In the present study, we aimed to characterize the dynamics of Le(b) binding and sequence variation of *babA* during chronic infection of humans. We analyzed a total of 47 sequential *H. pylori* isolates obtained from 23 individuals during two clinical trials in Louisiana in the United States (21) and Colombia (22). We observed a relatively stable Le(b) binding phenotype among isolates from the same individual. The BabA sequences of sequential isolates differed mostly in the putative N-terminal extracellular adhesion domain. Studies of strains harboring mosaic *babA* genes with different Le(b) binding properties demonstrated the importance of the C terminus for expression of a stable BabA protein and identified specific amino acids important for ligand binding.

RESULTS

Lewis b binding and BabA expression in sequential *H. pylori* isolates. Previous studies performed in different animal models have shown that *H. pylori* frequently loses its ability to bind Le(b) during infection (18–20). To analyze the dynamics of Le(b) adherence during chronic *H. pylori* infection in humans, we determined the Le(b) binding properties and BabA expression for a total of 47 sequential *H. pylori* isolates obtained from 23 individuals participating in two clinical trials in Louisiana in the United States (21) and Colombia (22), respectively (Table 1). Twenty-three pairs of strains were isolated at intervals ranging from 3 months to 4 years after the initial biopsy specimen; for one individual, a third isolate was obtained 16 years after the initial biopsy specimen. Altogether, 27 strains (57%) exhibited Le(b) binding, with marked differences in binding efficiencies, while 20 strains (43%) did not bind to Le(b). As expected, all Le(b) binding strains expressed the BabA adhesin, while in most of the strains that did not bind Le(b) (16/20 [80%]), no BabA protein was detectable (summarized in Table 1; also see Fig. S1 in the supplemental material). The differences in Le(b) binding were significantly associated with the origin of the strain ($P = 0.014$ by the chi-square test). Almost all strains from Colombia (NQ strains; 13/15 [87%]) displayed a Le(b) binding phenotype, whereas less than half of the strains from the United States (LSU strains; 14/32 [44%]) showed binding to Le(b).

Most sets of sequential isolates (18/23 [78%]) did not exhibit significant differences in Le(b) binding or BabA expression; this included nine pairs with binding to Le(b) and eight pairs without binding. In five strain sets, we observed a statistically significant reduction (NQ366/1790, LSU1016-1/-5) or even complete loss (NQ352/1701, NQ367/1671/4191, LSU1014-1/-6) of Le(b) binding in the course of infection. One additional strain pair (LSU1054-1/-5) showed a loss of Le(b) binding in the later strain, but this effect was not statistically significant due to the unusually variable Le(b) binding of the first strain, LSU1054-1. In two cases, the loss of Le(b) binding was accompanied by a complete loss of BabA expression (NQ4191 and LSU1014-6) as detected by Western blotting.

***bab* genotyping of sequential *H. pylori* isolates.** To shed light on the genetic basis responsible for the observed differences in Le(b) binding and BabA expression (Table 1), we next aimed to determine the *bab* genotype of the sequential isolates and to identify *babA*-containing loci for the subsequent sequencing of *babA* genes. The three known chromosomal *bab* loci termed locus A, locus B, and locus C were analyzed for the presence of *babA*, *babB*, and *babC* by PCR and partial sequencing (summarized in Table 1). *bab* genotypes were highly diverse in our strain collection (Fig. 1). In total, 23% (11/47) of the isolates had a J99-like genotype (AB) with *babA* in locus A and *babB* in locus B. Only two strains (4%) showed a 26695-like genotype (BA) with the genomic location of *babA* and *babB* inverted compared to that in J99. Copies of only *babA* (AA) or only *babB* (BB) were present at both loci in 6 (13%) and 11 strains (23%), respectively. The BB genotype and the apparent lack of *babA* genes in these strains were in agreement with the lack of BabA expression and their inability to bind to Le(b). In one of the strains, we identified only a single *babA* copy in locus B but could not amplify any *bab* gene in locus A. We did not detect any *bab* gene in locus C and no *babC* gene in any of the analyzed loci. Finally, although all genomic DNAs used for *bab* genotyping were derived from low-passage-number single colonies of the respective *H. pylori* strains, we observed a mixed *bab* genotype (AB A and A AB) in 16 isolates (34%), a phenomenon that has been previously reported (18, 23). The same chromosomal location of *babA* and *babB* suggests a high frequency of recombination between the *bab* loci A and B, which may result in the rapid emergence of differing subpopulations.

Comparisons of the sequential isolates identified no change of *bab* genotype in most pairs of strains (19/23 [83%]). The remaining four strain sets differed in their *bab* genotype at one locus (LSU1014, LSU2003, LSU3001, and NQ367 triplet). Most interestingly, the third strain of the triplet (NQ4191) had replaced *babA* in locus A by a second *babB* copy, which explains both the observed absence of BabA expression and loss of Le(b) binding.

Sequence variation of *babA*. *bab* genotyping identified 36 strains in our strain collection that harbored at least one copy of *babA*. Nine of these strains did not bind to Le(b), while the other 27 strains displayed Le(b) binding to various extents. To correlate the sequence diversity of BabA proteins with the observed differential Le(b) binding capacity, we next aimed to determine *babA* nucleotide sequences of these isolates. For strains with a mixed *bab* genotype, we restricted our analysis to loci exclusively containing *babA*. Altogether, we successfully obtained complete *babA* gene sequences from 33 of the 36 strains (Table 1). For three of the 27 Le(b) binding strains, we did not succeed in sequencing a *babA* gene. However, the observation of BabA expression and Le(b) binding indicates the presence of a functional *babA* gene in these strains.

Pairwise comparisons of the *babA* nucleotide sequences of 16 pairs of sequential isolates identified four pairs with identical *babA* genes (Fig. 2 and Table 2). The remaining 12 strain pairs showed

TABLE 1 Le(b) binding and *bab* genotyping of sequential *H. pylori* isolates

Stain set no.[a]	Strain[b]	Le(b) binding[c]	Stat. sign.[d]	Protein expression		*bab* genotype[e]		*babA* full-length sequencing[f]		
				BabA	BabB	Locus A	Locus B	Locus	*babA* length (bp)	BabA length (aa)
1	NQ267	4.9 ± 1.8	ns	+	−	B	A	B	2,253	750
	NQ1624	3.5 ± 0.8		+	+	B	A	B	2,241	746
2	NQ315	3.3 ± 1.0	ns	+	+	A	A	B	2,229	742
	NQ1712	4.5 ± 2.5		+	+	A	A	B	2,229	742
3	NQ331	2.3 ± 0.7	ns	+	+	AB	A	B	2,235	744
	NQ1832	4.6 ± 1.7		+	+	AB	A	B	2,235	744
4	NQ352	6.9 ± 2.6	$P \leq 0.05$	+	−	A	B	A	2,211	736
	NQ1701	1.3 ± 0.7		+	+	A	B	A	2,208	735
5	NQ366	3.7 ± 0.7	$P \leq 0.01$	+	+	AB	A	B	2,247	748
	NQ1790	2.1 ± 0.4		+	+	AB	A	B	2,247	748
6	NQ367	4.2 ± 0.8	ns/$P \leq 0.01$	+	+	A	B	A	2,250	749
	NQ1671	3.3 ± 1.3	$P \leq 0.05$	+	+	A	B	A	2,250	749
	NQ4191[g]	0.9 ± 0.1		−	+	B	B			
7	NQ372	1.8 ± 0.6	ns	+	+	A	A	B	2,232	743
	NQ1886	1.7 ± 0.7		+	+	A	A	B	2,232	743
8	LSU1010-1[g]	0.8 ± 0.3	ns	−	+	B	B			
	LSU1010-3[g]	0.9 ± 0.2		−	+	B	B			
9	LSU1013-2	4.4 ± 0.7	ns	+	+	A	AB	ND		
	LSU1013-6	5.6 ± 2.1		+	+	A	AB	A	2,235	744
10	LSU1014-1	3.8 ± 1.0	$P \leq 0.05$	+	+	A	B	A	2,235	744
	LSU1014-6	0.7 ± 0.1		−	+	A	AB	A	2,214	737
11	LSU1016-1	6.6 ± 0.5	$P \leq 0.001$	+	−	A	AB	A	2,238	745
	LSU1016-5	2.4 ± 0.5		+	+	A	AB	A	2,238	745
12	LSU1021-6	2.9 ± 1.0	ns	+	+	A	AB	A	2,220	739
	LSU1021-7	4.4 ± 2.3		+	+	A	AB	A	2,220	739
13	LSU1027-1	1.0 ± 0.1	ns	−	+	A	AB	A	759	252
	LSU1027-3	1.0 ± 0.1		−	+	A	AB	A	759	252
14	LSU1037-1[g]	0.9 ± 0.2	ns	−	+	B	B			
	LSU1037-5[g]	0.9 ± 0.1		−	+	B	B			
15	LSU1040-1[g]	0.8 ± 0.1	ns	−	+	B	B			
	LSU1040-6[g]	1.0 ± 0.1		−	+	B	B			
16	LSU1054-1	3.1 ± 1.8	ns	+	−	A	AB	A	2,226	741
	LSU1054-5	1.1 ± 0.2		+	+	A	AB	A	2,226	741
17	LSU1062-1	0.9 ± 0.2	ns	+	+	A	B	A	2,220	739
	LSU1062-3	0.9 ± 0.2		+	−	A	B	A	2,220	739
18	LSU1067-1	2.5 ± 1.1	ns	+	+	A	B	A	2,247	748
	LSU1067-5	3.2 ± 1.2		+	+	A	B	A	2,247	748
19	LSU1074-1	0.9 ± 0.2	ns	−	−	A	B	A	1,638	545
	LSU1074-4	0.7 ± 0.4		−	−	A	B	A	1,638	545
20	LSU2002-1[g]	0.8 ± 0.2	ns	−	+	B	B			
	LSU2002-8[g]	0.7 ± 0.2		−	+	B	B			
21	LSU2003-1	4.0 ± 1.4	ns	+	+	A	AB	ND		
	LSU2003-7	5.0 ± 1.8		+	−	A	A	ND		
22	LSU3001-1	1.9 ± 0.6	ns	+	−		A	B	2,214	737
	LSU3001-6	1.7 ± 0.6		+	−	A	A	B	2,217	738
23	LSU3005-1[g]	0.9 ± 0.0	ns	−	+	B	B			
	LSU3005-4[g]	1.0 ± 0.4		−	+	B	B			

[a] Sequential isolates were obtained from 23 individuals.
[b] The origin of *H. pylori* isolates is indicated as follows: NQ for the isolates from Colombia and LSU for isolates from Louisiana in the United States.
[c] Values represent mean ± SD of Le(b)/BSA binding ratios determined in at least three independent experiments. Strains are considered nonadherent if the Le(b)/BSA ratio was ≤1.5.
[d] Stat. sign., statistical significance. The values for Le(b) binding of sequential *H. pylori* isolates were compared by Welch's *t* test. The *P* value of statistically significant differences is indicated. ns, no statistically significant difference.
[e] A, *babA*; B, *babB*; AB, *babA* and *babB*.
[f] *babA* genes at the indicated locus were completely sequenced. The length of the encoded BabA proteins (in amino acids [aa]) was deduced from the corresponding gene sequences. ND, no complete *babA* sequence was determined.
[g] No *babA* gene was identified for this strain.

various levels of differences including isolated single nucleotide polymorphisms (SNPs) and clusters of nucleotide polymorphisms (CNPs) indicative of recombination. The latter could be the result of an intragenomic recombination event or due to recombination with a different *H. pylori* strain in case of coinfection with more than one strain. Altogether, the sequence similarity of *babA* ranged between 97 and 100%. Only strain pair NQ267/1624, where nucleotide polymorphisms were distributed throughout

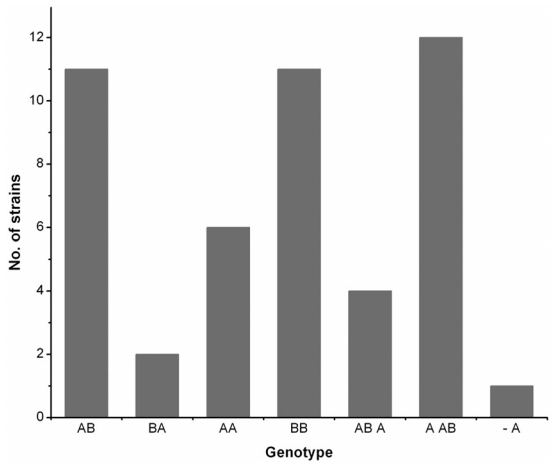

FIG 1 Distribution of *bab* genotypes. Frequency of *babA* and *babB* genes at the chromosomal loci A and B. None of the strains ($n = 47$) possessed *babC* in any of the three loci or had a *bab* gene in locus C. AB, *babA* in locus A and *babB* in locus B; BA, *babB* in locus A and *babA* in locus B; AA, *babA* in loci A and B; BB, *babB* in loci A and B; AB A, *babA* and *babB* in locus A and *babA* in locus B; A AB, *babA* in locus A and *babA* and *babB* in locus B; - A, no *bab* gene in locus A and *babA* in locus B.

the sequence, indicating that the complete gene had been exchanged by a single recombination event, displayed a markedly lower sequence similarity (91.4%). Five of the nine Le(b) nonbinding strains (LSU1027-1/-3, LSU1074-1/-4, and LSU1014-6) contained *babA* genes encoding premature stop codons at different positions (Fig. 2), which abolished BabA expression, while the other four strains harbored full-length *babA* genes encoding detectable BabA protein (Table 1; see Fig. S1 in the supplemental material). The deduced BabA amino acid sequences differed in 11 strain pairs (Fig. 2 and Table 2). Most differences were located in the predicted extracellular adhesion domain. The highest number of amino acid changes was observed in strain pairs NQ267/1624, LSU1014-1/-6, and NQ352/1701.

Next, we performed a multiple-sequence alignment (MSA) of BabA amino acid sequences from all BabA-expressing strains. After removal of duplicate sequences, the data set included 21 Le(b) binding strains and four Le(b) nonbinding strains (see Fig. S2 in the supplemental material). The highest sequence diversity was found in a region corresponding to amino acids 218 to 229 in the BabA sequence of strain J99. Based on this MSA, the sequence variability for each site was determined using the Protein Variability Server (PVS) (24). The highest diversity was observed in the putative N-terminal extracellular adhesion domain of BabA, while the C terminus, predicted to encode an outer membrane β-barrel, showed markedly less sequence variation (Fig. 3).

Characterization of strains with mosaic *babA* alleles. We next aimed to study the correlation of distinct sequence changes in *babA* with the Le(b) binding ability in an isogenic strain background. Strain pair LSU1014-1/-6, which had changed from a Le(b) binding to a nonbinding phenotype, was selected for the generation of isogenic *H. pylori* strains harboring mosaic *babA* genes. The loss of Le(b) binding in strain LSU1014-6 was accompanied by the absence of BabA expression on Western blots (see

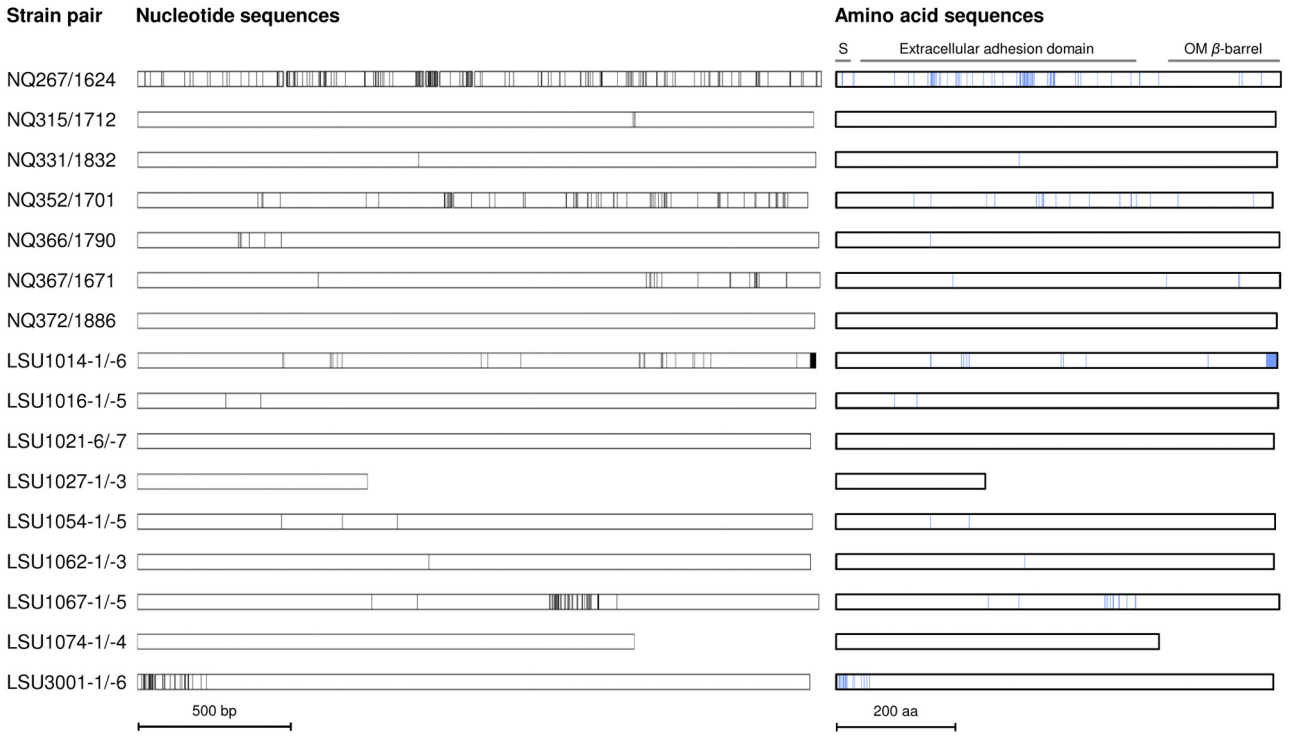

FIG 2 *babA* sequence variation between sequential isolates. Pairwise alignments of *babA* nucleotide sequences and deduced BabA amino acid sequences were performed using MUSCLE (45). The polymorphic sites in each nucleotide alignment were determined using the program PSFIND and graphically displayed with the program HAPPLOT. Polymorphic sites are displayed by vertical lines. Corresponding plots for the protein alignments were generated manually. Blue vertical lines indicate differences between sequential isolates. The signal sequence (S) as well as the predicted domain structure based on Phyre2 analysis are indicated. aa, amino acids.

TABLE 2 Comparison of *babA* sequences between pairs of sequential isolates

Strain set no.[a]	Strain	Interval between isolates (mo)	*babA* gene						BabA protein		
			Length (bp)	Sequence similarity (%)	No. of PS[b]	No. of SNPs and/or CNPs[c]	*S*[d]	*N*[d]	Length (aa)	No. of aa changes	Sequence similarity (%)
1	NQ267	36	2,253	91.41	215	1 CNP	35.5	139.5	750	70	92.94
	NQ1624		2,241						746		
2	NQ315	36	2,229	99.91	3	1 CNP	3	0	742	0	100
	NQ1712		2,229						742		
3	NQ331	36	2,235	99.95	1	1 SNP	0	1	744	1	99.86
	NQ1832		2,235						744		
4	NQ352	36	2,211	97.01	71	3 CNPs	46.75	21.25	736	19	98.09
	NQ1701		2,208						735		
5	NQ366	36	2,247	99.73	6	1 CNP	5	1	748	1	99.86
	NQ1790		2,247						748		
6	NQ367	36	2,250	99.15	19	1 SNP	14	5	749	4	99.59
	NQ1671		2,250			1 CNP			749		
7	NQ372	36	2,232	100	0		0	0	743	0	100
	NQ1886		2,232						743		
8	LSU1014-1	24	2,235	98.03	42	3 CNPs	12	11	744	23	96.63
	LSU1014-6		2,214						737		
9	LSU1016-1	12	2,238	99.91	2	1 CNP	0	2	745	2	99.86
	LSU1016-5		2,238						745		
10	LSU1021-6	12	2,220	100	0		0	0	739	0	100
	LSU1021-7		2,220						739		
11	LSU1027-1	3	759	100	0		0	0	252	0	100
	LSU1027-3		759						252		
12	LSU1054-1	12	2,226	99.86	3	1 CNP	1	2	741	2	99.73
	LSU1054-5		2,226						741		
13	LSU1062-1	3	2,220	99.95	1	1 SNP	0	1	739	1	99.86
	LSU1062-3		2,220						739		
14	LSU1067-1	12	2,247	98.26	40	2 CNPs	24	16	748	13	98.52
	LSU1067-5		2,247						748		
15	LSU1074-1	6	1,638	100	0		0	0	545	0	100
	LSU1074-4		1,638						545		
16	LSU3001-1	24	2,214	98.15	44	1 CNP	15	19	737	17	98.1
	LSU3001-6		2,217						738		

[a] *babA* sequences were obtained from 16 pairs of sequential isolates.
[b] The number of polymorphic sites (PS) was determined with the program PSFind based on pairwise sequence alignments (MUSCLE) and includes substitutions, insertions, and deletions.
[c] SNP, single nucleotide polymorphism; CNP, cluster of nucleotide polymorphisms (polymorphisms are separated by <200 bp).
[d] The average number of synonymous sites (*S*) or nonsynonymous sites (*N*) was determined with the program MEGA5.

Fig. S1 in the supplemental material). However, the level of *babA* transcript expressed by LSU1014-6 was similar to that expressed by LSU1014-1 (data not shown), indicating that the difference in the Le(b) binding phenotype was indeed related to distinct protein properties. The predicted BabA protein sequences of both strains differed in 11 amino acids scattered along the sequence, plus the C terminus (Fig. 4A). The deletion of a G nucleotide at position 2172 in a stretch of 5 G's caused a frameshift in *babA* of strain LSU1014-6, resulting in 12 exchanged amino acids and a truncation of seven residues by a premature stop codon. To characterize the specific influence of these sequence differences on ligand binding and/or protein expression, we generated mosaic *babA* genes by transformation of strain LSU1014-6 with *babA* PCR products amplified from either the adherent strain LSU1014-1 or the nonadherent strain LSU1014-6 (for details, see Materials and Methods). *H. pylori* transformants were enriched for recombinants with restored Le(b) binding using a biopanning protocol or for expression of surface-located BabA via a pulldown assay. A total of 11 clones with mosaic *babA* alleles were single colony purified, subjected to *babA* gene sequencing (Fig. 4B), and characterized for BabA protein expression (Fig. S3) and Le(b) binding (Fig. 4C).

Clone 1, which possessed a *babA* gene identical to the gene in strain LSU1014-1, expressed BabA and showed restored binding to Le(b) compared to strain LSU1014-6. Compared to LSU1014-1, binding was even significantly increased (1.6-fold). Le(b) binding of all other recombinant clones was quantitated and expressed in relation to that of clone 1. These clones contained mosaic *babA* genes with combinations of LSU1014-1 and LSU1014-6 specific residues, and with the exception of clone 135, they displayed both BabA expression and distinctly different binding to Le(b).

Clone 135 lacked both BabA protein expression and Le(b) binding, while reverse transcription-PCR analysis detected *babA*-specific mRNA as in strain LSU1014-6 (data not shown). BabA of clone 135 was identical to BabA of LSU1014-1 except for the distinct and truncated C terminus. In contrast, clone 522, whose *babA* allele was identical to LSU1014-6 except for the C terminus, expressed BabA and bound to Le(b), albeit to a reduced extent. This indicated that the C terminus is important for expression of stable BabA protein, whereas the 11 amino acid differences along the BabA sequence contribute to differences in Le(b) binding. Analysis of the remaining eight clones, which all contained the

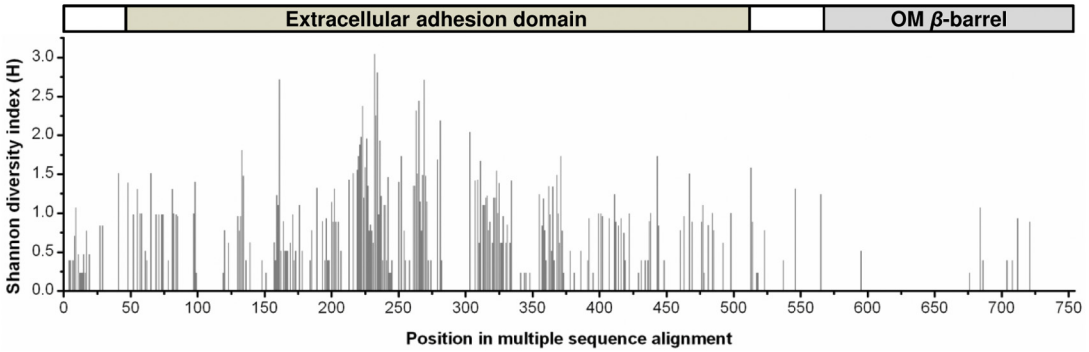

FIG 3 Sequence variability of BabA from BabA-expressing NQ and LSU strains. The Shannon diversity index (*H*) for each site in the multiple-sequence alignment of BabA (see Fig. S2 in the supplemental material) was determined using the Protein Variability Server (PVS) (24). *H* can range from 0 (the same amino acid is present at that position in all sequences) to 4.322 (all 20 amino acids are equally represented in that position). The predicted domain structure based on Phyre2 analysis is indicated on top of the figure.

intact C terminus of strain LSU1014-1 allowed a closer inspection of the influences of these 11 amino acid differences on Le(b) binding. For example, clones 12 and 743 showed no increase in Le(b) binding compared to clone 522, indicating that the specific residues at positions 421 and 630 in strain LSU1014-1 were probably less important for ligand binding. In contrast, the presence of further LSU1014-1-specific residues in clone 362 (residues 378 and 385) and clone 5 (residues 212, 213, 215, 220, 225, 378, and 385) restored Le(b) binding back to the level of clone 1, suggesting that these residues play an important role in Le(b) binding. This was further confirmed in the remaining clones, which all showed diminished Le(b) binding.

To further analyze the importance of the C terminus for protein biogenesis, we performed structure prediction analyses of the BabA sequences using Phyre2 (Fig. 5). The structural model of the N-terminal adhesion domain of BabA was based on the previously published extracellular adhesion domain of *H. pylori* SabA (25) and was predicted with high confidence for all strains (100% confidence; >21% identity). In contrast, the confidence of the structure prediction of the C-terminal β-barrel membrane domain differed considerably between Le(b) binding and nonbinding strains (Fig. 5B). While the outer membrane domain of BabA from strain LSU1014-1 was predicted with high confidence (95.3%), confidence values for the membrane domains of strain LSU1014-6 (74.6%) and clone 135 (76.8%) were markedly lower, suggesting that the missing C terminus affects the correct folding of the β-barrel and may thereby influence stability.

DISCUSSION

In this study, we analyzed the dynamics of Le(b) binding and sequence variation of *babA* during chronic *H. pylori* infection in humans. In contrast to reports from experimental infections in animal models (18–20), the Le(b) binding phenotype was relatively stable during human infection, as Le(b) binding ability changed in only 5 out of 23 sets of sequential isolates. This stable BabA phenotype might be due to the already long-standing infection and a stable physiological situation in the infected individuals when the first isolate was collected.

The reported loss of Le(b) binding in animal infection experiments has been the result of either gene conversion, frameshifts due to variation in length of repeat sequences, or mutations leading to amino acid substitutions or truncated proteins (18–20). We observed a loss of Le(b) binding due to replacement of *babA* with a second copy of *babB* in one case. In the other isolates, changes in Le(b) binding ability were caused by mutations leading to amino acid changes in BabA, which were mainly located in the predicted N-terminal adhesion domain of BabA. We did not observe loss of Le(b) binding due to phase variation in any of these isolates.

Since we analyzed only one purified *H. pylori* clone per biopsy, we cannot assess potential intrastrain variation of Le(b) binding within a stomach or between different parts of the stomach niche. Two recent studies of our group have analyzed multiple single colony isolates from two stomach biopsies for *hopZ* gene status and detected only minimal intrastrain variation (26, 27). Another study analyzing the *bab* genotype of more than one single colony isolate from each of three patients observed intrastrain genotype variation in two of the three cases (14).

Our experimental data and protein structure predictions suggest that the C terminus of BabA may be important for protein folding and possibly stability. The C terminus is predicted to encode an outer membrane β-barrel (28), and correct assembly into the outer membrane is essential for protein function. Folding and membrane insertion of OMPs are evolutionarily highly conserved and mediated by the BAM (β-barrel assembly machinery) complex (29, 30). The central and essential component of this multiprotein complex, BamA, is itself an OMP and recognizes substrate OMPs by a C-terminal signature sequence (31). This sequence is found in the majority of bacterial OMPs and is highly conserved among *Proteobacteria* (32, 33). Changes of this sequence have been shown to diminish or completely disturb OMP assembly (31, 33, 34). For example, deletion of the C-terminal phenylalanine of the *Escherichia coli* porin PhoE disabled BamA activity, while a synthetic peptide consisting of the last 12 amino acids was sufficient to activate BamA (31). So far, the BAM complex has not been analyzed in *H. pylori*. However, BamA is conserved in *H. pylori* (35), and it has been found in a recent analysis of the cell surface proteome of *H. pylori* (28). Thus, we assume that recognition of the truncated BabA protein by the BAM complex is hampered due to the missing C terminus, which might subsequently lead to rapid protein degradation to avoid accumulation in the periplasmic space. Interestingly, BabA belongs to the Y-Hop subgroup of *H. pylori* OMPs that encode a tyrosine residue at the final position (6). Future studies should concentrate on the impact of this amino acid on OMP assembly.

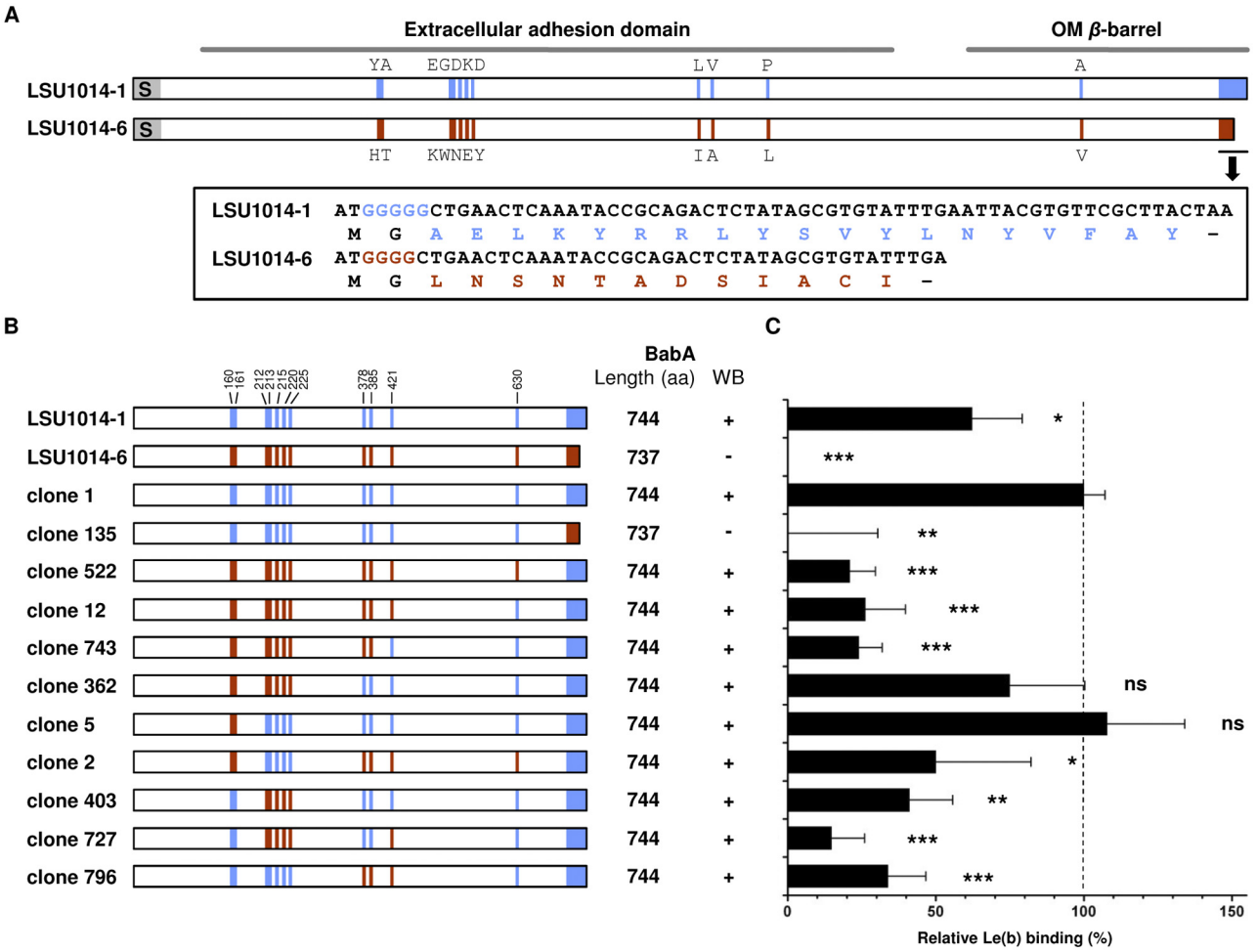

FIG 4 Analysis of *babA* mosaic alleles. (A) Schematic representation of differences in the deduced BabA protein sequence of strain pair LSU1014-1/-6. The signal sequence (S) is indicated as well as the domain structure as predicted by Phyre2 analysis (44). BabA of strain LSU1014-6 differs in 11 amino acids and in addition contains a distinct and shortened C terminus due to a frameshift mutation. (B) Schematic representation of BabA sequences from the generated mosaic alleles based on strain pair LSU1014-1/-6. The numbers above the schematic representation of BabA in strain LSU1014-1 indicate the positions of amino acid changes. The length of the BabA protein (in amino acids [aa]) and result of BabA expression determined by Western blotting (WB) are indicated to the right of the schematic representations of the Bab sequences. (C) Le(b) binding was determined by ELISA and is represented as mean + SD from at least four independent experiments. All strains except LSU1014-6 and clone 135 had Le(b)/BSA ratios of >1.5. Le(b) binding of recombinant clone 1 was set at 100%, and statistical significance versus the value for clone 1 was determined by Welch's *t* test. Statistical significance is indicated as follows: *, $P \leq 0.05$; **, $P \leq 0.01$; ***, $P \leq 0.001$; ns, not statistically significant.

The BabA sequence comparison from Le(b) binding and nonbinding strains displayed a high level of variation in the predicted adhesion domain of BabA, while the C terminus was less variable. However, we were not able to identify specific amino acids essential for Le(b) binding. These finding are in agreement with a previous sequence comparison of Le(b) binding and nonbinding strains (15). On the basis of mosaic *babA* alleles, we identified some amino acids that affected the Le(b) binding strength but did not completely disable adherence. Further studies are needed to correlate protein sequence diversity and capacity of Le(b) binding.

The significantly increased binding of clone 1 compared to strain LSU1014-1 despite identical *babA* sequences is likely to be the result of integration of this allele in the different LSU1014-6 strain background of this transformant. Since only the coding *babA* sequence was used for transformation, the *babA* promoter region might differ and differentially influence gene transcription. Recently, it was demonstrated that variation in length of the thymine (T) nucleotide repeat located in the *sabA* promoter adjacent to the −35 element resulted in different transcriptional activity (36, 37). These differences led to distinct protein expression and receptor binding activity (36). The *babA* promoter of *H. pylori* strain 26695 contains a homopolymeric adenine (A) repeat between the −35 and −10 promoter elements (36), which might similarly affect gene transcription or transcript stability. Since the incorporation of BabA into the outer membrane is essential for its functional activity, differences in the functionality of the BAM complex, which is responsible for the correct assembly and membrane insertion of OMPs, might present an alternative explanation for the observed differences in Le(b) binding.

Our data demonstrate that the ability to bind to Le(b) is a relatively stable phenotype during human infection but that microevolution within *babA* occurs commonly, driven by mutation and recombination. These sequence changes can result in a complete loss of protein expression, such as the one observed in the

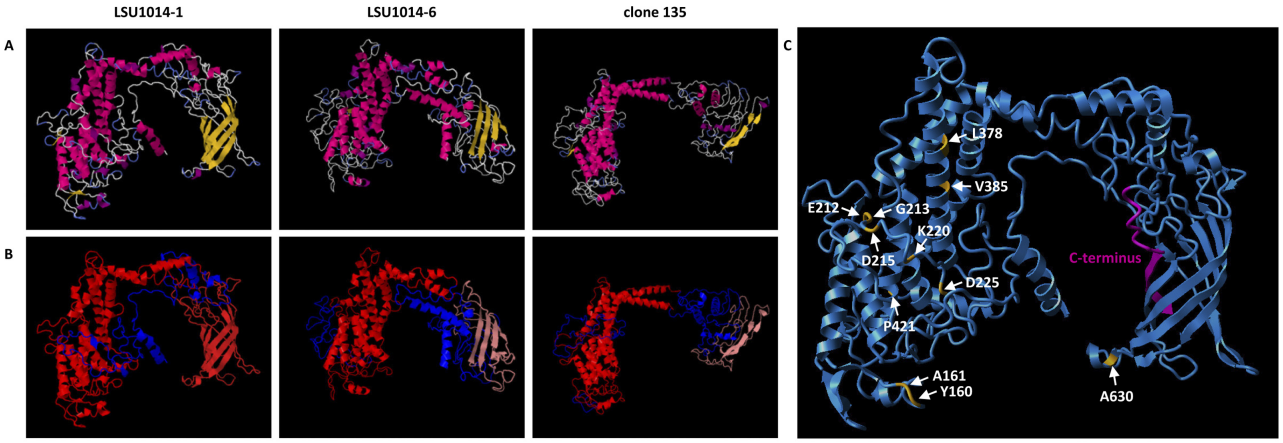

FIG 5 Phyre2 structure prediction of BabA. The BabA sequences of strains LSU1014-1 and LSU1014-6 and clone 135 were subjected to Phyre2 analysis for structure prediction (see Materials and Methods). (A and B) The models are colored showing the secondary structure (A) or confidence of the predicted structure (B), ranging from high (red) to low (blue). (C) Structural model generated with the BabA sequence of strain LSU1014-1 [Le(b) binding] with the amino acids that differed in strains LSU1014-1 and LSU1014-6 (not binding) indicated in orange. Additionally, the location of the distinct C terminus is colored in pink.

sequence encoding the C terminus, or in gradual changes in binding properties, such as those caused by the amino acid replacements in the coding sequence other than the C terminus (Fig. 5C). It seems likely that these changes are the result of selection *in vivo*, reflecting changing conditions in the stomach. Our study is a step toward a better understanding of the correlations between BabA sequence and Le(b) binding phenotype, which will be greatly enhanced once BabA crystal structures together with its ligand Le(b) become available, permitting a mapping of the functional information provided here to specific parts of the interaction surface in the BabA structure. A detailed understanding of the critical amino acid residues involved in Le(b) binding will also be essential in guiding efforts toward pharmacological inhibition of BabA-Le(b) interactions in a therapeutic or prophylactic setting.

MATERIALS AND METHODS

***H. pylori* strains and culture conditions.** The sequential *H. pylori* isolates analyzed in this study were obtained during two clinical trials performed in Louisiana in the United States (21) and Colombia (22). These trials compared the development of gastric lesions in groups treated with antibiotics, sucralfate, or a dietary supplement. Only five of the 16 individuals from whom the LSU strains were isolated (LSU1013, LSU1014, LSU1021, LSU1040, and LSU1067), and none of the Colombian trial participants (NQ strains) had received antibiotics that might have caused a selective bottleneck in the *H. pylori* population despite the failure to achieve eradication. The genetic relationships of the sequential isolates were previously analyzed by extended multilocus sequence analysis (MLSA) (38). *H. pylori* strain J99 (39) and an isogenic J99*babA* mutant were used as controls in the Lewis b binding assay. For a complete list of strains, see Table S1 in the supplemental material. Bacteria were cultured from frozen stocks on blood agar plates (Columbia agar base II; Oxoid, Germany) supplemented with 10% horse blood (Oxoid), and the antibiotics vancomycin (10 mg/liter), polymyxin B (2500 U/liter), amphotericin B (4 mg/liter), and trimethoprim (5 mg/liter). Plates for the J99*babA* mutant carrying a chloramphenicol resistance cassette were additionally supplemented with chloramphenicol (20 mg/liter). Strains were incubated under microaerobic conditions (5% O_2, 10% CO_2, and 85% N_2) at 37°C.

Generation of an isogenic J99*babA* mutant. The J99*babA* mutant was constructed by natural transformation-mediated allelic exchange as described previously (40). Oligonucleotide sequences are provided in Table S2 in the supplemental material. Briefly, *babA* was amplified by PCR and ligated into pUC18 (pSUS2703). This plasmid was used as a template for an inverse PCR, designed to almost completely delete the *babA* gene. Finally, a chloramphenicol (*cat*) resistance cassette (41) was introduced. The resulting plasmid (pSUS2705) was used as a suicide plasmid for natural transformation of *H. pylori* strain J99. The successful chromosomal replacement of *babA* with the *cat* cassette via allelic exchange (double crossover) was verified by PCR using suitable primer combinations. PCR with a pUC18-specific primer was performed to exclude possible integration of the complete plasmid into the genome (single crossover).

***bab* genotyping and full-length sequencing of *babA*.** Genomic DNA from low-passage-number single colonies of bacterial strains was isolated using the QIAamp DNA minikit (Qiagen, Germany). PCRs were performed according to standard protocols, and amplification products were purified using the QIAquick PCR purification kit (Qiagen). PCR amplicons were sequenced bidirectionally by the Sanger method using the BigDye terminator v1.1 cycle sequencing kit and a 3130xl genetic analyzer (Applied Biosystems). Sequence data were analyzed using BioNumerics v6.01 (Applied Maths NV, Sint-Martens-Latem, Belgium). *bab* genotypes in loci A and B were determined using combinations of locus- and *bab* gene-specific primers as previously described (14). Analysis of locus C was performed by PCR amplification of the complete locus (16). Using BlastX search, the partial gene sequences were identified as either *babA* or *babB*, while no *babC* genes were found (42). Identified *babA* genes were completely sequenced using overlapping PCR fragments of locus-specific primers and *bab* gene-specific primers located in the middle variable region. See Table S2 in the supplemental material for a complete list of primer sequences. The identity of sequenced *babA* genes to their orthologue in strain 26695 based on BlastX search ranged between 85 and 92%. Pairwise sequence alignments were analyzed for the numbers of isolated single nucleotide polymorphisms (SNPs) and of clusters of nucleotide polymorphisms (CNPs), which comprise polymorphisms that are $<$200 bp apart from each other and are an indication of recombination.

Analysis of BabA and BabB expression in *H. pylori* strains. *H. pylori* strains were grown on blood agar plates for up to 48 h. Bacterial cells were harvested in phosphate-buffered saline (PBS) by centrifugation (5,000 × *g*, 4°C, 10 min). Equal amounts of bacteria were resuspended in sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) sample buffer, incubated for 10 min at 95°C, separated by SDS-PAGE (8%), and analyzed by Western blotting. Primary antibodies against BabA and BabB raised in rabbits were kindly provided by Thomas Borén. Peroxidase-labeled AffiniPure goat anti-rabbit IgG (H+L) (Jackson ImmunoResearch Laboratories) was used as a secondary antibody. Western blots were developed with SuperSignal West Pico chemiluminescent substrate

(Thermo Scientific, Rockford) and exposed to Amersham Hyperfilm ECL film (GE Healthcare, Little Chalfont, United Kingdom).

Analysis of Lewis b binding ability of *H. pylori* strains. *H. pylori* strains were tested for their Le(b) binding ability using an enzyme-linked immunosorbent assay (ELISA) as previously described (43). Briefly, bacterial strains were grown on plates for 24 h and harvested in PBS by centrifugation (2,795 × *g*, 5 min, 4°C). For surface labeling, 2×10^8 live bacteria were incubated with 125 μg/25 μl NHS-LC-biotin [succinimidyl-6-(biotinamido)hexanoate] for 50 min at room temperature (RT) in the dark. Labeled bacteria were washed twice in PBS, collected by centrifugation (2,795 × *g*, 10 min, 4°C), and finally resuspended in 1 ml of PBS with 5% bovine serum albumin (BSA). The wells of a 96-well Universal Covalent microtiter plate (Corning Costar, Cambridge, MA) were coated in two replicate plates for each experimental condition with either 50 μl PBS, 250 ng BSA, or 250 ng Le(b)-BSA for 1 h at RT in the dark. After removal of supernatants, the plate was exposed to UV light for 30 s in a Stratalinker (Stratagene, Germany). The remaining binding sites on the plates were blocked by incubation with 5% BSA in PBS for 1 h at RT. Biotinylated bacteria (50 μl/well) were coincubated in the coated plates for 1 h at RT in the dark. The plates were washed three times with PBS, and adherent bacteria were fixed with 100 μl paraformaldehyde (2% in 100 mM potassium phosphate [pH 7]) for 1 h at RT. After three washing steps with 0.05% Tween 20 in PBS, the plate was blocked with 10% fetal calf serum (FCS) in PBS for 1 h at RT. The plate was washed five times with 0.05% Tween 20 in PBS and incubated with neutravidin-horseradish peroxidase (HRP)-conjugate in PBS plus 10% FCS for 90 min at RT to detect bound biotinylated bacteria. After seven washes with PBS containing 0.05% Tween 20, the plate was incubated with 100 μl/well 3,3′,5,5′-tetramethylbenzidine (TMB) (BD Biosciences) for 30 min at RT. The reaction was stopped by the addition of 1 M H_3PO_4 (50 μl/well), and the extinction at 450 nm was measured using a microplate reader. A Le(b)/BSA binding ratio of >1.5 was indicative of Le(b) binding. Le(b) binding of all strains in duplicate samples was determined in at least three independent experiments.

Generation of *babA* mosaic alleles. Strain pair LSU1014-1/-6 was selected for generation of mosaic *babA* genes. The nonadherent strain LSU1014-6 was transformed with PCR products of *babA* from the adherent strain LSU1014-1 or from strain LSU1014-6 [restoration of intact C terminus with primer babA_1014-1_2186(5G)] by natural transformation or electroporation (see Table S3 in the supplemental material). Transformants were enriched for either Le(b) binding or BabA expression. Enrichment of strains with restored Le(b) binding was performed using a biopanning procedure as previously described (17). Briefly, transformed LSU1014-6 was mixed with biotinylated Le(b) conjugate and streptavidin-coated magnetic beads. Le(b) binding transformants were captured with a magnet and subsequently cultured on plates. BabA-expressing clones were enriched using a pulldown assay. Briefly, protein A magnetic beads were incubated with 10 μl anti-BabA antiserum for 1 h at 4°C. The beads were then incubated with 1.5×10^8 bacteria for 1 h at 4°C. BabA-expressing transformants were captured with a magnet and subsequently cultured. Recombinant clones were single colony purified and then analyzed for *babA* gene sequence, BabA expression, and Le(b) binding as described above.

Structure prediction by Phyre2. Deduced BabA sequences of strain pair LSU1014-1/-6 and the generated *babA* mosaic alleles were subjected to the Phyre2 web server (44) for structure prediction using the intensive modeling mode. The structural model of the adhesion domain of BabA was based on similarities to the extracellular adhesion domain of SabA from *H. pylori* (25). The modeling of the β-barrel domain was based on similarities to different templates (c2k0lA, d1g90a, and c2x27X).

Nucleotide sequence accession numbers. The *babA* sequences have been deposited in the European Nucleotide Archive (ENA) and are available from the ENA browser at http://www.ebi.ac.uk/ena/data/view/LN650055-LN650087.

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at http://mbio.asm.org/lookup/suppl/doi:10.1128/mBio.02281-14/-/DCSupplemental.

Figure S1, TIF file, 0.9 MB.
Figure S2, PDF file, 0.2 MB.
Figure S3, TIF file, 1.1 MB.
Table S1, DOC file, 0.1 MB.
Table S2, DOC file, 0.3 MB.
Table S3, DOC file, 0.1 MB.

ACKNOWLEDGMENTS

This work was supported by grant DFG SFB 900/A1 from the German Research Foundation to S. Suerbaum and grant DFG SFB 900/B6 to C. Josenhans. L. Kennemann was supported by a Ph.D. training grant in the framework of International Research Training Group IRTG 1273.

We thank Thomas Borén for kindly providing BabA and BabB antisera and Birgit Brenneke, Jessika Schulze, and Kerstin Ellrott for excellent technical assistance.

REFERENCES

1. **Kusters JG, van Vliet AH, Kuipers EJ.** 2006. Pathogenesis of *Helicobacter pylori* infection. Clin. Microbiol. Rev. **19:**449–490. http://dx.doi.org/10.1128/CMR.00054-05.
2. **Schreiber S, Konradt M, Groll C, Scheid P, Hanauer G, Werling HO, Josenhans C, Suerbaum S.** 2004. The spatial orientation of *Helicobacter pylori* in the gastric mucus. Proc. Natl. Acad. Sci. U. S. A. **101:**5024–5029. http://dx.doi.org/10.1073/pnas.0308386101.
3. **Howitt MR, Lee JY, Lertsethtakarn P, Vogelmann R, Joubert LM, Ottemann KM, Amieva MR.** 2011. ChePep controls *Helicobacter pylori* infection of the gastric glands and chemotaxis in the *Epsilonproteobacteria*. mBio **2**(4):e00098. http://dx.doi.org/10.1128/mBio.00098-11.
4. **Bücker R, Azevedo-Vethacke M, Groll C, Garten D, Josenhans C, Suerbaum S, Schreiber S.** 2012. *Helicobacter pylori* colonization critically depends on postprandial gastric conditions. Sci. Rep. **2:**994. http://dx.doi.org/10.1038/srep00994.
5. **Moore ME, Borén T, Solnick JV.** 2011. Life at the margins: modulation of attachment proteins in *Helicobacter pylori*. Gut Microbes **2:**42–46. http://dx.doi.org/10.4161/gmic.2.1.14626.
6. **Alm RA, Bina J, Andrews BM, Doig P, Hancock RE, Trust TJ.** 2000. Comparative genomics of *Helicobacter pylori*: analysis of the outer membrane protein families. Infect. Immun. **68:**4155–4168. http://dx.doi.org/10.1128/IAI.68.7.4155-4168.2000.
7. **Mahdavi J, Sondén B, Hurtig M, Olfat FO, Forsberg L, Roche N, Angstrom J, Larsson T, Teneberg S, Karlsson KA, Altraja S, Wadström T, Kersulyte D, Berg DE, Dubois A, Petersson C, Magnusson KE, Norberg T, Lindh F, Lundskog BB, Arnqvist A, Hammarström L, Borén T.** 2002. *Helicobacter pylori* SabA adhesin in persistent infection and chronic inflammation. Science **297:**573–578. http://dx.doi.org/10.1126/science.1069076.
8. **Rossez Y, Gosset P, Boneca IG, Magalhaes A, Ecobichon C, Reis CA, Cieniewski-Bernard C, Joncquel Chevalier Curt M, Leonard R, Maes E, Sperandio B, Slomianny C, Sansonetti PJ, Michalski JC, Robbe-Masselot C.** 2014. The LacdiNAc specific adhesin LabA mediates adhesion of *Helicobacter pylori* to human gastric mucosa. J. Infect. Dis. **15:**1286–1295. http://dx.doi.org/10.1093/infdis/jiu239.
9. **Senkovich OA, Yin J, Ekshyyan V, Conant C, Traylor J, Adegboyega P, McGee DJ, Rhoads RE, Slepenkov S, Testerman TL.** 2011. *Helicobacter pylori* AlpA and AlpB bind host laminin and influence gastric inflammation in gerbils. Infect. Immun. **79:**3106–3116. http://dx.doi.org/10.1128/IAI.01275-10.
10. **Borén T, Falk P, Roth KA, Larson G, Normark S.** 1993. Attachment of *Helicobacter pylori* to human gastric epithelium mediated by blood group antigens. Science **262:**1892–1895. http://dx.doi.org/10.1126/science.8018146.
11. **Ilver D, Arnqvist A, Ogren J, Frick IM, Kersulyte D, Incecik ET, Berg DE, Covacci A, Engstrand L, Borén T.** 1998. *Helicobacter pylori* adhesin binding fucosylated histo-blood group antigens revealed by retagging. Science **279:**373–377. http://dx.doi.org/10.1126/science.279.5349.373.
12. **Tomb JF, White O, Kerlavage AR, Clayton RA, Sutton GG, Fleischmann RD, Ketchum KA, Klenk HP, Gill S, Dougherty BA, Nelson K,**

Quackenbush J, Zhou L, Kirkness EF, Peterson S, Loftus B, Richardson D, Dodson R, Khalak HG, Glodek A, McKenney K, Fitzegerald LM, Lee N, Adams MD, Hickey EK, Berg DE, Gocayne JD, Utterback TR, Peterson JD, Kelley JM, Cotton MD, Weidman JM, Fujii C, Bowman C, Watthey L, Wallin E, Hayes WS, Borodovsky M, Karp PD, Smith HO, Fraser CM, Venter JC. 1997. The complete genome sequence of the gastric pathogen *Helicobacter pylori*. Nature **388:**539–547. http://dx.doi.org/10.1038/41483.

13. **Gerhard M, Lehn N, Neumayer N, Borén T, Rad R, Schepp W, Miehlke S, Classen M, Prinz C.** 1999. Clinical relevance of the *Helicobacter pylori* gene for blood-group antigen-binding adhesin. Proc. Natl. Acad. Sci. U. S. A. **96:**12778–12783. http://dx.doi.org/10.1073/pnas.96.22.12778.
14. **Colbeck JC, Hansen LM, Fong JM, Solnick JV.** 2006. Genotypic profile of the outer membrane proteins BabA and BabB in clinical isolates of *Helicobacter pylori*. Infect. Immun. **74:**4375–4378. http://dx.doi.org/10.1128/IAI.00485-06.
15. **Hennig EE, Mernaugh R, Edl J, Cao P, Cover TL.** 2004. Heterogeneity among *Helicobacter pylori* strains in expression of the outer membrane protein BabA. Infect. Immun. **72:**3429–3435. http://dx.doi.org/10.1128/IAI.72.6.3429-3435.2004.
16. **Hennig EE, Allen JM, Cover TL.** 2006. Multiple chromosomal loci for the *babA* gene in *Helicobacter pylori*. Infect. Immun. **74:**3046–3051. http://dx.doi.org/10.1128/IAI.74.5.3046-3051.2006.
17. **Bäckström A, Lundberg C, Kersulyte D, Berg DE, Borén T, Arnqvist A.** 2004. Metastability of *Helicobacter pylori bab* adhesin genes and dynamics in Lewis b antigen binding. Proc. Natl. Acad. Sci. U. S. A. **101:**16923–16928. http://dx.doi.org/10.1073/pnas.0404817101.
18. **Solnick JV, Hansen LM, Salama NR, Boonjakuakul JK, Syvanen M.** 2004. Modification of *Helicobacter pylori* outer membrane protein expression during experimental infection of rhesus macaques. Proc. Natl. Acad. Sci. U. S. A. **101:**2106–2111. http://dx.doi.org/10.1073/pnas.0308573100.
19. **Styer CM, Hansen LM, Cooke CL, Gundersen AM, Choi SS, Berg DE, Benghezal M, Marshall BJ, Peek RM, Jr, Borén T, Solnick JV.** 2010. Expression of the BabA adhesin during experimental infection with *Helicobacter pylori*. Infect. Immun. **78:**1593–1600. http://dx.doi.org/10.1128/IAI.01297-09.
20. **Ohno T, Vallström A, Rugge M, Ota H, Graham DY, Arnqvist A, Yamaoka Y.** 2011. Effects of blood group antigen-binding adhesin expression during *Helicobacter pylori* infection of Mongolian gerbils. J. Infect. Dis. **203:**726–735. http://dx.doi.org/10.1093/infdis/jiq090.
21. **Taylor NS, Fox JG, Akopyants NS, Berg DE, Thompson N, Shames B, Yan L, Fontham E, Janney F, Hunter FM.** 1995. Long-term colonization with single and multiple strains of *Helicobacter pylori* assessed by DNA fingerprinting. J. Clin. Microbiol. **33:**918–923.
22. **Correa P, Fontham ET, Bravo JC, Bravo LE, Ruiz B, Zarama G, Realpe JL, Malcom GT, Li D, Johnson WD, Mera R.** 2000. Chemoprevention of gastric dysplasia: randomized trial of antioxidant supplements and anti-*Helicobacter pylori* therapy. J. Natl. Cancer Inst. **92:**1881–1888. http://dx.doi.org/10.1093/jnci/92.23.1881.
23. **Sheu SM, Sheu BS, Chiang WC, Kao CY, Wu HM, Yang HB, Wu JJ.** 2012. *H. pylori* clinical isolates have diverse *babAB* genotype distributions over different topographic sites of stomach with correlation to clinical disease outcomes. BMC Microbiol. **12:**89. http://dx.doi.org/10.1186/1471-2180-12-89.
24. **Garcia-Boronat M, Diez-Rivero CM, Reinherz EL, Reche PA.** 2008. PVS: a web server for protein sequence variability analysis tuned to facilitate conserved epitope discovery. Nucleic Acids Res. **36:**W35–W41. http://dx.doi.org/10.1093/nar/gkn211.
25. **Pang SS, Nguyen ST, Perry AJ, Day CJ, Panjikar S, Tiralongo J, Whisstock JC, Kwok T.** 2014. The three-dimensional structure of the extracellular adhesion domain of the sialic acid-binding adhesin SabA from *Helicobacter pylori*. J. Biol. Chem. **289:**6332–6340.
26. **Kennemann L, Brenneke B, Andres S, Engstrand L, Meyer TF, Aebischer T, Josenhans C, Suerbaum S.** 2012. *In vivo* sequence variation in HopZ, a phase-variable outer membrane protein of *Helicobacter pylori*. Infect. Immun. **80:**4364–4373. http://dx.doi.org/10.1128/IAI.00977-12.
27. **Kennemann L, Didelot X, Aebischer T, Kuhn S, Drescher B, Droege M, Reinhardt R, Correa P, Meyer TF, Josenhans C, Falush D, Suerbaum S.** 2011. *Helicobacter pylori* genome evolution during human infection. Proc. Natl. Acad. Sci. U. S. A. **108:**5033–5038. http://dx.doi.org/10.1073/pnas.1018444108.
28. **Voss BJ, Gaddy JA, McDonald WH, Cover TL.** 2014. Analysis of surface-exposed outer membrane proteins in *Helicobacter pylori*. J. Bacteriol. **196:**2455–2471. http://dx.doi.org/10.1128/JB.01768-14.
29. **Tommassen J.** 2010. Assembly of outer-membrane proteins in bacteria and mitochondria. Microbiology **156:**2587–2596. http://dx.doi.org/10.1099/mic.0.042689-0.
30. **Kim KH, Aulakh S, Paetzel M.** 2012. The bacterial outer membrane beta-barrel assembly machinery. Protein Sci. **21:**751–768. http://dx.doi.org/10.1002/pro.2069.
31. **Robert V, Volokhina EB, Senf F, Bos MP, Van Gelder P, Tommassen J.** 2006. Assembly factor Omp85 recognizes its outer membrane protein substrates by a species-specific C-terminal motif. PLoS Biol. **4:**e377. http://dx.doi.org/10.1371/journal.pbio.0040377.
32. **Paramasivam N, Habeck M, Linke D.** 2012. Is the C-terminal insertional signal in Gram-negative bacterial outer membrane proteins species-specific or not? BMC Genomics **13:**510–513. http://dx.doi.org/10.1186/1471-2164-13-510.
33. **Struyvé M, Moons M, Tommassen J.** 1991. Carboxy-terminal phenylalanine is essential for the correct assembly of a bacterial outer membrane protein. J. Mol. Biol. **218:**141–148. http://dx.doi.org/10.1016/0022-2836(91)90880-F.
34. **de Cock H, Struyvé M, Kleerebezem M, van der Krift T, Tommassen J.** 1997. Role of the carboxy-terminal phenylalanine in the biogenesis of outer membrane protein PhoE of *Escherichia coli* K-12. J. Mol. Biol. **269:**473–478. http://dx.doi.org/10.1006/jmbi.1997.1069.
35. **Liechti G, Goldberg JB.** 2012. Outer membrane biogenesis in *Escherichia coli*, *Neisseria meningitidis*, and *Helicobacter pylori*: paradigm deviations in *H. pylori*. Front. Cell. Infect. Microbiol. **2:**29. http://dx.doi.org/10.3389/fcimb.2012.00029.
36. **Aberg A, Gideonsson P, Vallstrom A, Olofsson A, Ohman C, Rakhimova L, Boren T, Engstrand L, Brannstrom K, Arnqvist A.** 2014. A repetitive DNA element regulates expression of the *Helicobacter pylori* sialic acid binding adhesin by a rheostat-like mechanism. PLoS Pathog. **10:**e1004234. http://dx.doi.org/10.1371/journal.ppat.1004234.
37. **Kao CY, Sheu SM, Sheu BS, Wu JJ.** 2012. Length of thymidine homopolymeric repeats modulates promoter activity of *sabA* in *Helicobacter pylori*. Helicobacter **17:**203–209. http://dx.doi.org/10.1111/j.1523-5378.2012.00936.x.
38. **Falush D, Kraft C, Taylor NS, Correa P, Fox JG, Achtman M, Suerbaum S.** 2001. Recombination and mutation during long-term gastric colonization by *Helicobacter pylori*: estimates of clock rates, recombination size, and minimal age. Proc. Natl. Acad. Sci. U. S. A. **98:**15056–15061. http://dx.doi.org/10.1073/pnas.251396098.
39. **Alm RA, Ling LS, Moir DT, King BL, Brown ED, Doig PC, Smith DR, Noonan B, Guild BC, deJonge BL, Carmel G, Tummino PJ, Caruso A, Uria-Nickelsen M, Mills DM, Ives C, Gibson R, Merberg D, Mills SD, Jiang Q, Taylor DE, Vovis GF, Trust TJ.** 1999. Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori*. Nature **397:**176–180. http://dx.doi.org/10.1038/16495.
40. **Kulick S, Moccia C, Kraft C, Suerbaum S.** 2008. The *Helicobacter pylori mutY* homologue HP0142 is an antimutator gene that prevents specific C to A transversions. Arch. Microbiol. **189:**263–270. http://dx.doi.org/10.1007/s00203-007-0315-9.
41. **Ge Z, Hiratsuka K, Taylor DE.** 1995. Nucleotide sequence and mutational analysis indicate that two *Helicobacter pylori* genes encode a P-type ATPase and a cation-binding protein associated with copper transport. Mol. Microbiol. **15:**97–106. http://dx.doi.org/10.1111/j.1365-2958.1995.tb02224.x.
42. **Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25:**3389–3402. http://dx.doi.org/10.1093/nar/25.17.3389.
43. **Aspholm M, Kalia A, Ruhl S, Schedin S, Arnqvist A, Lindén S, Sjöström R, Gerhard M, Semino-Mora C, Dubois A, Unemo M, Danielsson D, Teneberg S, Lee WK, Berg DE, Borén T.** 2006. *Helicobacter pylori* adhesion to carbohydrates. Methods Enzymol. **417:**293–339. http://dx.doi.org/10.1016/S0076-6879(06)17020-2.
44. **Kelley LA, Sternberg MJ.** 2009. Protein structure prediction on the Web: a case study using the Phyre server. Nat. Protoc. **4:**363–371. http://dx.doi.org/10.1038/nprot.2009.2.
45. **Edgar RC.** 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. **32:**1792–1797. http://dx.doi.org/10.1093/nar/gkh340.