# Cloning and Genetic Characterization of the Helicobacter pylori and Helicobacter mustelae flaB Flagellin Genes and Construction of H. pylori flaA- and flaB-Negative Mutants by Electroporation-Mediated Allelic Exchange

SEBASTIAN SUERBAUM,<sup>1,2\*</sup> CHRISTINE JOSENHANS,<sup>2</sup> AND AGNÈS LABIGNE<sup>1</sup>

Unité des Entérobactéries, Institut Pasteur, INSERM U199, F-75724 Paris Cedex 15, France,<sup>1\*</sup> and Medizinische Mikrobiologie und Immunologie, Ruhr-Universität Bochum, D-4630 Bochum 1, Germany<sup>2</sup>

## Received 9 December 1992/Accepted 16 March 1993

Helicobacter pylori is one of the most common human pathogens. It causes chronic gastritis and is involved in the pathogenesis of gastroduodenal ulcer disease and possibly gastric carcinoma. Helicobacter mustelae is a bacterium closely related to H. pylori that causes gastritis and ulcer disease in ferrets and is therefore considered an important animal model of gastric Helicobacter infections. Motility, even in a viscous environment, is conferred to the bacteria by several sheathed flagella and is regarded as one of their principal virulence factors. The flagellar filament of H. pylori consists of two different flagellin species expressed in different amounts. The gene (flaA) encoding the major flagellin has recently been cloned and sequenced. Here we report the cloning and sequencing of two highly homologous new flagellin genes from H. pylori 85P and H. mustelae NCTC 12032. The nucleotide sequence of the H. pylori gene proved that it encoded the second flagellin molecule found in H. pylori flagellar filaments. The genes were named flaB. The H. mustelae and H. pylori flaB genes both coded for proteins with 514 amino acids and molecular masses of 54.0 and 53.9 kDa, respectively. The proteins shared \$1.7% identical amino acids. The degree of conservation between H. pylori FlaB and the H. pylori FlaA major flagellin was much lower (58%). Both flaB genes were preceded by  $\sigma^{54}$ -like promoter sequences. Mapping of the transcription start site for the H. pylori flaB gene by a primer extension experiment confirmed the functional activity of the  $\sigma^{54}$  promoter. To evaluate the importance of both genes for motility, flaA- and flaB-disrupted mutants of H. pylori N6 were constructed by electroporation-mediated allelic exchange and characterized by Western blot (immunoblot) analysis and motility testing. Both mutations selectively abolished the expression of the targeted gene without affecting the synthesis of the other flagellin molecule. Whereas flaA mutants were completely nonmotile, flaB mutants retained motility.

Helicobacter pylori (20) is a gram-negative, spiral-shaped bacterium that is now generally accepted to be the causative agent of chronic, active type B gastritis in humans. *H. pylori* infection is also regarded as a prerequisite for the occurrence of duodenal ulcers and a cofactor in the pathogenesis of gastric ulcer and gastric carcinoma (38, 41; for reviews see references 1 and 42).

Several potential virulence factors have been suggested to play a role in Helicobacter pathogenesis, those studied in most detail being the production of abundant amounts of urease, motility, the expression of cytotoxins and other extracellular enzymes, and adhesion to tissue-specific receptors (for reviews see references 1, 2, and 6). In regards to urease expression and motility, experiments using the gnotobiotic piglet model of H. pylori infection have supported the importance of these properties for the onset of colonization (7-9). A high degree of motility, even in a viscous environment (24), is conferred to *H. pylori* by a bundle of three to six flagella that extends from one pole of the bacterium. The flagella consist of a basal part that contains the flagellar motor and the hook structure, the central filament, and a membranous sheath that envelops each filament (16, 21). No details are known about the basal parts of Helicobacter flagella. As for the sheath, there is only

Recently, Kostrzynska et al. (26) have demonstrated the presence of a second flagellin molecule of slightly higher molecular mass in *H. pylori* flagellar filaments. It is expressed in much smaller amounts than the major flagellin and has been shown to be located primarily in the proximal part of the filament. The function of this protein and its importance for motility are as yet unknown. Research in the field of *Helicobacter* pathogenesis has been hampered by the difficulties encountered in the establishment of animal models. The three most promising animal models are the infection of gnotobiotic piglets with *H. pylori* (7–9), the colonization of the murine or rat stomach with *Helicobacter felis* (31), and the *Helicobacter mustelae* infection of the ferret. *H. mustelae* is a very common pathogen of ferrets, in which, like *H. pylori* in humans, it colonizes the gastric mucosa and

relatively basic biochemical information available. It is a bilayered membrane that contains proteins, lipids, and lipopolysaccharides and thus closely resembles the outer membrane, although differences, especially in fatty acid composition and protein pattern, have been noted (18). The sheath has been suggested to play a role in protecting the filaments from the acidic environment of the gastric lumen. The filaments have been the subject of several detailed investigations. The major constituent of the *H. pylori* flagellar filament is a 53-kDa flagellin molecule. The gene encoding this major flagellin has recently been cloned and sequenced by Leying et al. (32) and is called *flaA*.

<sup>\*</sup> Corresponding author.

TABLE 1. Vectors and hybrid plasmids used in this study

Plasmid	Vector	Size (kb)	Characteristics <sup>a</sup>	Reference
	pIC20R2	2.7	Rep <sub>Eo</sub> , Ap <sup>r</sup>	36
	pILL570	5.3	$\operatorname{Rep}_{F_{c1}}$ Mob, $\operatorname{Sp}^{r}$	29
	pILL575	10	Rep <sub>Ec</sub> , Rep <sub>Ci</sub> , Mob, Cos, Km <sup>r</sup>	29
pILL600	pBR322	5.7	Ap <sup>r</sup> Km <sup>r</sup> ; source of kanamycin cassette	30
pILL683	pILL575	45	Km <sup>r</sup> ; cosmid containing H. pylori flaB	This study
pHL319-2-4	pIC20R2	6.3	Ap <sup>r</sup> ; plasmid containing H. pylori flaA	32
pSUS10	pIC20R2	7.7	$Ap^{r}$ Km <sup>r</sup> ; H. pylori flaA $\Omega$ Km	This study
pSUS19	pILL570	8.8	Sp <sup>r</sup> ; subclone of pILL683, contains H. pylori flaB	This study
pSUS22	pILL570	10.2	$Sp^{r}$ Km <sup>r</sup> ; H. pylori flaB $\Omega$ Km	This study
pSUS100	pILL570	12.3	Sp <sup>r</sup> ; plasmid containing H. mustelae flaB	This study
pSUS101	pILL570	8.3	Sp <sup>r</sup> ; subclone of pSUS100, contains H. mustelae flaB	This study

<sup>a</sup> Rep<sub>Ec</sub> and Rep<sub>Cj</sub>, plasmids capable of replicating in *E. coli* and *C. jejuni*, respectively; Mob, conjugative plasmid due to the presence of OriT; Ap<sup>r</sup>, Km<sup>r</sup>, and Sp<sup>r</sup>, resistance to ampicillin, kanamycin, and spectinomycin, respectively; Cos, presence of lambda cos site.

causes chronic gastritis and ulcer disease (13-15). In contrast to the other animal models, which primarily allow studies concerning colonization, the *H. mustelae* infection of the ferret represents a disease model and seems thus of special interest for the study of pathogenesis and the role of host immunity in gastric *Helicobacter* infections.

*H. mustelae* flagella are ultrastructurally very similar to those of *H. pylori*; however, they extend from both bacterial poles and from the lateral parts of the bacteria (20, 40). Only one flagellin molecule has been purified to date. The aminoterminal sequence of this protein was strongly homologous with that of the major *H. pylori* flagellin (49). So far, none of the virulence-associated genes of this organism has been cloned. However, genetic studies in this organism are of interest, not only because sequence information will show which domains of a protein are characteristic for the gastric *Helicobacter* species and, thus, could be of importance for pathogenicity, but also because cloned antigens and genetically defined mutants will provide important tools for pathogenesis and vaccination studies in this model.

We report here on the cloning and sequencing of two homologous new flagellin genes of H. pylori and H. mustelae that code for the minor flagellin proteins and are called flaB. H. pylori mutants in the flaA and flaB genes were constructed by means of reverse genetics to test the effect of those mutations on motility.

# MATERIALS AND METHODS

Bacteria, plasmids, and culture conditions. The cloning experiments were performed with genomic DNA prepared from H. pylori 85P (29) and H. mustelae NCTC 12032 (National Collection of Type Cultures, Colindale, London, United Kingdom). H. pylori N6 (11) was used as the recipient strain for the electroporation experiments because of its favorable transformability. A urease-negative mutant of H. pylori N6 disrupted in the ureB gene (11) was used as a control in the motility tests. Escherichia coli MC1061 (3) was used as the host for cloning and subcloning experiments. Vectors and recombinant plasmids used in this study are listed in Table 1. H. pylori and H. mustelae strains were grown on horse blood or sheep blood agar plates, respectively, supplemented with vancomycin (10 mg/liter), polymyxin B (2,500 U/liter), trimethoprim (5 mg/liter), and amphotericin B (4 mg/liter). Plates were incubated at 37°C under microaerobic conditions. E. coli strains were grown in L broth without glucose (10 g of tryptone, 5 g of yeast extract, and 5 g of NaCl per liter; pH 7.0) or on L agar plates (1.5% agar) at 37°C. Antibiotic concentrations for the selection of recombinant clones were as follows (in milligrams per liter): kanamycin, 20; spectinomycin, 100; carbenicillin, 100.

**Preparation of DNA.** Genomic DNA from *H. mustelae* was prepared by the method of Majewski and Goodwin (35). Cosmid and plasmid DNAs were prepared by an alkaline lysis procedure and then purified in cesium chloride-ethidium bromide gradients as previously described (11).

**DNA analysis and cloning methodology.** Restriction endonucleases, T4 DNA ligase, and DNA polymerase I large (Klenow) fragment were purchased from Amersham or Boehringer Mannheim, T4 DNA polymerase was from Biolabs, and calf intestinal phosphatase was from Pharmacia. All enzymes were used as described in the instructions of the manufacturers. DNA fragments were separated on agarose gels run in Tris-acetate buffer. The 1-kb ladder from Bethesda Research Laboratories was used as a fragment size standard. When necessary, DNA fragments were isolated by electroelution from agarose gels as previously described (28) and recovered from the migration buffer by means of an Elutip-d minicolumn (Schleicher & Schuell, Dassel, Germany). Basic DNA manipulations were performed by the protocols described by Sambrook et al. (44).

Hybridization. For Southern blot hybridizations, DNA fragments were transferred from agarose gels to nylon membranes (Hybond N; Amersham Corp.) by capillary blotting (47) and baked for 2 h at 80°C. Southern blot hybridizations were performed under low-stringency conditions (2× SSPE-0.05% sodium dodecyl sulfate [SDS], 45 to 55°C)  $(1 \times SSPE)$ is 0.18 M NaCl, 10 mM NaPO<sub>4</sub>, and 1 mM EDTA [pH 7.5]). Colony blots for screening of the H. pylori cosmid bank and identification of subclones were prepared on nitrocellulose membranes (Schleicher & Schuell) by the protocol of Sambrook et al. (44). Hybridizations were performed under low-stringency conditions (2× SSC-0.1% SDS-30% formamide, 37°C) (1× SSC is 0.15 M NaCl plus 15 mM sodium citrate [pH 7.0]). Radioactive labelling of purified restriction fragments and polymerase chain reaction (PCR) products was performed by random priming (10). Hybridization was revealed by autoradiography with Amersham Hyperfilm-MP.

**Cosmid cloning.** The construction of the cosmid gene bank of *H. pylori* 85P in *E. coli* HB101, which was used for the cloning of the *H. pylori flaB* gene, has been described previously (29).

**DNA sequencing.** Appropriate fragments of plasmid DNA were subcloned into M13mp18/19 vectors (37). Single-stranded DNA was prepared by phage infection of *E. coli* 

TABLE 2. Oligonucleotides used for PCR and primer extension experiments

Name	Position <sup>a</sup>	Strand	Length	Sequence (5'-3')	Reference
OLpBR322-1	2,085-2,104	+	20	GTGATGACGGTGAAAACCTC	50
OLpBR322-2	2,834-2,853	_	20	GGTTGGACTCAAGACGATAG	50
OLKm-1	699-724	+	25	CTGCTAAGGTATATAAGCTGGTGGG	53
OLKm-2	1,305-1,329	_	25	CATACTGTTCTTCCCCGATATCCTC	53
OLHMflaB-2	51-75	+	25	AATCAAAGATAAGGAATTAAGCGGC	This study
OLHMflaB-3	b	-	25	CTGAAATGAAAATAACCCCCAAAGTC	This study
OLHPflaA-1	225-249	+	25	GGAGAGATTGAGTTCAGGTTTAAGG	32
OLHPflaA-2	1,602-1,626	-	25	ACTCATCGCATAGCTGCCTGATTGC	32
OLHPflaB-2	775–798	-	24	CAGCGCTCGTAGAAATGCGCACGG	This study
OLHPflaB-7	_	+	21	GTTTATACCTATTAATGAATG	This study
OLHPflaAPE-2	169–198	-	30	GAGTTCGGATTGCACATGCGCATTCATCGC	32
OLHPflaBPE-1	184-210	-	27	GGTCTCTGTTGTTTTGAACCCCTACCG	This study

<sup>a</sup> Numbers refer to nucleotides in the sequences of pBR322 (for OLpBR322-1 and -2), the kanamycin resistance gene (for OLKm-1 and -2), the *H. pylori flaA* gene (for OLHPflaA-1, -2, and OLHPflaAPE-2), the *H. mustelae flaB* gene (for OLHMflaB-2), a sequence close to the rightmost *Bam*HI site in pSUS101 (for OLHMflaB-3), the *H. pylori flaB* gene (for OLHPflaB-2 and OLHPflaBPE-1), and a sequence upstream the reported sequence for *H. pylori flaB* (for OLHPflaB-7), respectively.

<sup>b</sup> —, not published.

JM101 (54). Sequencing was performed by the dideoxynucleotide chain termination method (45) by using the U.S. Biochemicals Sequenase kit. Both the M13 universal primer and additional specific primers were used to sequence both the coding and noncoding DNA strands. Sequencing of double-stranded DNA was performed as described previously (5).

Extraction of RNA and primer extension. Total RNA was extracted from H. pylori 85P by the method described by Leying et al. (32), which itself is a modification of a protocol described by Chirgwin et al. (4). For primer extension, oligonucleotides OLHPflaBPE-1 and OLHPflaAPE-2 (Table 2) were end labelled by using the 5'-DNA terminus labelling system of Bethesda Research Laboratories. Twenty micrograms of RNA was precipitated and resuspended in 25 µl of avian myoblastosis virus buffer (1× avian myeloblastosis virus buffer is 50 mM Tris-HCl [pH 8.3], 40 mM KCl, 5 mM MgCl<sub>2</sub>, and 5 mM dithiothreitol). A 1.5-pmol amount of labelled primer was added, and the mixture was heated to 95°C for 3 min and then allowed to cool to 60°C over a period of 1 h. After the annealing period, deoxynucleoside triphosphates (dNTPs) (final concentration, 0.1 mM of each of the four dNTPs) and 28 U of avian myeloblastosis virus reverse transcriptase (Boehringer Mannheim) were added, and the samples were incubated at 42°C for 1 h. To demonstrate that RNA and not DNA was the template for extension, control samples were incubated with RNase (1 mg/ml) for 10 min at 37°C before the reverse transcriptase was added. After extension, samples were extracted once with phenol-chloroform, precipitated, and resuspended in 10 µl of sequencing gel-loading buffer. Sequencing reactions were performed on double-stranded templates by using the same oligonucleotides as those used for the primer extension experiment. Samples were analyzed on a 9% sequencing gel.

**Electroporation of H. pylori.** To construct H. pylori mutants, appropriate plasmid constructions carrying the targeted gene disrupted by a cassette containing a *Campylobacter* kanamycin resistance gene were transformed into H. pylori N6 by means of electroporation as described previously (11). After electroporation, bacteria were grown on nonselective plates for a period of 48 h to allow for the expression of the antibiotic resistance and then transferred onto kanamycin-containing plates. The selective plates were incubated for up to 6 days. Described mutant strains are descendants of single H. pylori colonies.

**PCR.** PCRs were carried out by using a Perkin-Elmer Cetus thermal cycler with the GeneAmp kit (Perkin-Elmer Cetus). One hundred picomoles of each primer and at least 5 pmol of target DNA were used. The DNA was denatured at 94°C for 1 min, primers were annealed at temperatures between 42 and 55°C (depending on the calculated melting temperatures of the primers) for 2 min, and products were extended at 72°C for 2 min. Twenty-five cycles were performed.

Western blotting (immunoblotting). Flagella were enriched as previously described (49). SDS-polyacrylamide gel electrophoresis (SDS-PAGE) was carried out as described by Lugtenberg et al. (34). The concentration of acrylamide in the gel was 11%. Sixty micrograms of protein was applied to each slot of the gel. For Western blotting, proteins were transferred to nitrocellulose sheets (Schleicher & Schuell) by the method of Towbin et al. (52). The blots were incubated with rabbit antiserum AK 179 (32) raised against purified flagellar filaments, and bound antibody was visualized by a peroxidase-coupled goat anti-rabbit antibody (Biogenzia, Bochum, Germany).

Motility testing. For motility testing, brain heart infusion agar plates supplemented with 10% fetal calf serum and antibiotics as described above were used. Two-layered plates were poured (the lower layer containing 1.5% agar, the soft overlay containing 0.5% agar). Plates were inoculated by placing small slices of blood agar plates, densely grown with the strain to be tested, on the surface of the motility plate, the grown side of the slice facing the soft layer. Plates were incubated for 3 days under microaerobic conditions.

Nucleotide sequence accession numbers. The nucleotide sequences reported in this article appear in the GenBank data base under the accession numbers L08907 (*H. pylori*) and L08908 (*H. mustelae*).

#### RESULTS

Cloning of a flagellin gene from *H. mustelae* NCTC 12032. To clone a flagellin gene from *H. mustelae*, genomic DNA was prepared from *H. mustelae* NCTC 12032 and restricted with various endonucleases. Restriction fragments were hybridized with a 2.1-kb *BglII-Bam*HI restriction fragment of plasmid pHL319-2-4 (which contains the 5' end of the *H. pylori flaA* gene encoding the *H. pylori* major flagellin) under A)



FIG. 1. Linear restriction maps of the recombinant plasmids pSUS100 and pSUS101 (A) and pSUS19 (B), harboring the *H. mustelae* and *H. pylori flaB* genes, respectively. Only the inserts are depicted. Plasmid pSUS101 is a subclone of the 3.0-kb *ClaI* restriction fragment of pSUS100. The location of the ORFs is indicated by the open boxes, and the sense of transcription is indicated by arrows below the boxes. Asterisks mark restriction sites of the vector no longer present in the recombinant plasmids.

low-stringency conditions. Hybridizations of this probe with a complete BglII digest of genomic H. mustelae DNA revealed two hybridizing bands of 7 and 7.5 kb. Hybridizations with probes corresponding to other fragments of the H. pylori flaA gene suggested that the homologous gene was not restricted by BglII but that each of the two hybridizing fragments contained one complete copy of a homologous gene. Chromosomal DNA fragments of 6.5 to 8 kb were electroeluted from a preparative agarose gel, ligated into the Bg/II-restricted and dephosphorylated pILL570 plasmid vector, and transformed into *E. coli* MC1061 cells. Plasmid DNA was prepared from 50 recombinant clones, cut with BglII, and subjected to Southern blot hybridization with the 2.1-kb H. pylori flaA probe. Three of the 50 clones contained identical inserts that hybridized with the probe; one of them (pSUS100) was selected for further analysis. Figure 1A shows the restriction map of pSUS100. Further Southern blot analysis indicated that the hybridizing DNA segment was internal to a 3.0-kb ClaI restriction fragment that was subcloned into the plasmid vector pILL570 and mapped in more detail. The physical map of the resulting plasmid pSUS101, which contained the presumed flagellin gene, is also shown in Fig. 1A.

Sequencing of the H. mustelae flaB flagellin gene. Appropri-

ate restriction fragments of pSUS101 were cloned into M13mp18/19 vectors and sequenced by the dideoxynucleotide chain termination method. The nucleotide sequence obtained is shown in Fig. 2A. The sequenced DNA fragment contained one large open reading frame (ORF) of 1,542 nucleotides coding for a protein with a predicted molecular mass of 54.0 kDa. The deduced protein, in its amino-terminal and carboxy-terminal parts, shared marked homologies with several bacterial flagellins. The highest degrees of homology were found with Campylobacter flagellins and the H. pylori FlaA (Fig. 3). The N-terminal amino acid sequence of the protein deduced from the cloned gene was not identical to the sequence previously determined for the major H. mustelae flagellin (49) (Table 3). It shared, however, very strong homology with the N-terminal amino acid sequence of the minor flagellin of *H. pylori* purified by Kostrzynska et al. (26). The gene was named flaB so that the name flaA, in analogy to the nomenclature in H. pylori, can be given to the gene coding for the major flagellin. It is preceded by a ribosome-binding site similar to the consensus sequence for *E. coli* (46). A stem-loop structure (free energy,  $\Delta \hat{G} = -10.8$ kcal [ca. -45.2 kJ]/mol) consistent with a rho-independent transcriptional stop signal (43) is located downstream of the ORF. Upstream of the ORF there is a putative  $\sigma^{54}$ -depenļ

	+++xxx	+++xxx	<b>o</b> 54
1	CTGGAATCGGGTGGGTTGCT	AATTCTACAACATTTTTTTT	CAATTTTTTTAAATCAAAGAT AAGGAATTAAGCGGCAAAAA G <u>TGGAACGCTTGTTGCT</u> TTA
•		rbs M S	FRINTNI AALNAH SIGVQTN
101	CTTGGCTTGTTTAGTTCAAA	TGCAAAGGATGCAACATGAG	TTTTAGGATAAATACAAATA TCGCGGCGCTAAATGCACAT TCTATAGGTGTTCAGACAAA
	RNIAGSL	EKLSSG	L R I N K A A D D A S G M A I A D S L R
201	TAGAAATATTGCAGGCTCTC	TAGAAAAGCTTAGCTCTGGT	CTAAGAATCAACAAAGCAGC AGATGATGCTTCAGGGATGG CAATTGCAGATAGTCTGAGA
	SQSESLG	QAVRNAN	DAIGMI QIADKAM DEQLKIL
301	AGCCAGAGTGAGAGTTTGGG	ACAGGCAGTGAGAAATGCCA	ATGATGCCATTGGTATGATA CAAATTGCTGATAAGGCGAT GGATGAGCAACTAAAAATCT
	DTIKAK	AIQAAQD	G Q S Q E S R R S L Q S D I R R L M E E
401	TAGATACCATCAAGGCAAAG	GCCATCCAAGCAGCCCAAGA	TGGTCAAAGCCAAGAATCCA GAAGATCGCTTCAAAGTGAT ATTAGGAGATTGATGGAGGA
	LDNIANT	TSFNGQ	QMLSGAF TNKEFQI GAYSNT
501	GCTTGATAATATCGCAAATA	CGACGAGCTTCAATGGTCAG	CAAATGCTCTCAGGTGCATT TACAAATAAAGAATTTCAAA TTGGTGCGTATTCTAACACT
	T V K A S I G	PTSSDKI	GHIRME TASFSGV GMLASAG
601	ACTGTAAAGGCATCCATTGG	CCCCACAAGCTCAGATAAAA	TCGGACATATTAGGATGGAA ACTGCATCATTTAGCGGTGT GGGCATGCTTGCTAGTGCTG
	GNNLTE	VALNFKA	T D G V N S F E L E N V R I S T S A G T
701	GGGGAAATAACCTCACAGAA	GTGGCGCTAAATTTTAAGGC	GACAGATGGGGTGAATAGTT TTGAACTAGAAAACGTAAGA ATTTCTACTTCAGCGGGCAC
	GIGALSE	VINRFS	DKLGIRA TYN VMAT GTSPVM
801	GGGGATTGGCGCTTTGAGTG	AGGTCATCAACCGCTTCTCT	GACAAATTGGGCATACGAGC TACTTATAATGTTATGGCAA CAGGGACAAGTCCTGTGATG
	SGTVRGL	VINGVRI	GTVNEV RKNDSDG RLINAIN
901	TCTGGGACGGTACGCGGACT	TGTGATTAATGGTGTGAGAA	TTGGGACGGTTAATGAAGTG CGCAAAAATGACTCAGATGG AAGGCTTATCAATGCGATTA
	SVKNQT	GVEASLD	ITGRINL VSLDGRAISVHAD
1001	ACTCAGTCAAAAATCAAACA	GGCGTGGAAGCATCACTGGA	TATCACAGGAAGAATCAATC TTGTTTCTTTGGATGGAAGG GCGATCTCTGTGCATGCAGA
	GEASHVF	GEGNFT	G I S G N N H A I V G R L T L I R T D A
1101	TGGCGAGGCCTCTCATGTTT	TTGGAGAGGGGGAATTTCACT	GGAATCTCTGGCAATAACCA TGCCATCGTAGGACGTCTCA CACTCATCCGCACAGATGCT
	RDIIVSG	VNFSHIG	LHSAQGVAETTAN LRQLRGM
1201	AGAGACATCATCGTAAGTGG	CGTGAATTTCTCTCACATTG	GTCTTCACTCCGCTCAAGGG GTGGCTGAGACCACAGCCAA CCTCAGACAACTCAGGGGAA
	FGADIA	SAAGANA	N K A Q A D I N R Q G I G A G V T S L K
1301	TGTTTGGTGCAGATATTGCA	TCTGCTGCTGGCGCCAATGC	CAACAAAGCCCAGGCAGATA TCAATAGACAGGGGATTGGC GCTGGGGTGACCAGTCTCAA
1 4 0 1	G A M I V M D		T Q L D K V R S D M G S V Q I Q L V S T
1401	GGGTGCAATGATTGTCATGG	ATATGGTGGACTCTGCGCGC	ACTUAGUTAGATAAAGTUUG UTUAGAUATGGGTTUTGTGU AGATUCAACTGGTGTUTACU
1501			
1201		CAGGIAAATGTCAAGGCTG	CAGAMICACAAAIIAGAGAT GTGGATTTTGCTGCAGAATC TGCAAACTTCTCCAAAAACA
	традъс	S F A L A Q A	ичаббиа ткттб.

1601 ATATCTTGGCTCAAAGCGGA AGCTTCGCGCTAGCCCAGGC CAATGCTGTCCAACAAATG TCTTAAGACTTTTGCAATAA CCTTTCTCAAAGCCTTGGTT

1701 TTAGGCTTTTTTCATTGTTT TAAAATCATCCACCTCCTCC ATTTTGCCAATGCCAATGC GAGGGTTTTTTAAATTACCCA CCCCGCTTTTTTAGACTTTG

**o**54

																							σ54			
к	1	TATT	AATG	AAT	ATT	GTAGC	AT/	AGAA	TTTT	GACT	AAACGC	CCT	ATTA	AACC	ATAA	AACT	ATA	ACGO	CGCT	АААА	АТСАА	AGAG	TTGG	AACAC	CCTT	TGC
					+1				rhe		м	c	F D	т	N	r N	т	۵	ΔΤ	т	s u	۵	v	C V	0	N
	1 . 1		~~ ~ ~	~~~~			3 00		100															6 V	~~	
	101	TIGAC	JTAA	CAG		FATCT	ATC	JCAA	AGGA	IGCA	AACATG	AGT	TTTA	GAT	AAAT	ACCAA	TAT	CGCC	GCTT	TAAC	TICIC	ATGC	GGTA	GGGGT	TCAA	AAC
		N R	D	L	S	s s	L	E	ĸ	L S	SG	L	R	I	N K	Α	Α	DD	) S	S	GΜ	А	ΙA	D	S L	R
	201	AACAG	GAGA	CCTI	TCA	AGCTC	GC1	TGA	AAAG	гтаас	GCTCAG	GGC	TTAG	GATC	ААТА	AGCC	GCI	GACG	ATTC	TAGT	GGGAT	GGCG	ATCG	CTGAT	AGCT	таа
		c	0	c 1	N	T.	c	0	<b>Δ</b> Τ	D	ΝΔ	N		х т	C .	MV		ν π	Δ		λ	MD		~ T		т.
	201	00201	*	3000			0	~~~~													,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,			~~ 1	л 	
	301	GGAG	ICAA	AGCU	CGA	ATTTG	GG	CAA	GCGA	CCGG	AACGC	TAA	TGACO	SCTA	TTGG	ATGG	TTC	AAAC	CGCT	GATA.	AAGCG	ATGG	ATGA	GCAAA	TCAA	AAT
		LI	о т	1	ĸ	тк	F	A V	Q	A /	A Q	D	GQ	т	LI	E S	R	R	A L	Q	S D	I	Q	R L	L	Е
	401	CTTAC	GACA	CCAT	TAA	AACCA	AAC	SCCG	TTCA	AGCCO	GCTCAA	GAT	GGGCI	AAAC	TTTA	GAAAG	CCG	GAAGA	GCGC	TCCA	GAGCG	ATAT	TCAA	AGGTT	GTTA	GAA
		ΕL	D	N	IŻ	A N	т	т	S I	F N	GΩ	Q	м	L	S G	s	F	S N	I K	E :	FQ	I	G A	Y	S N	Т
	501	GAACT	FAGA	CAAJ	ATC	GCTAA	CAC	CAC	AAGC	TTA	ACGGCC	AAC	AAAT	GCTT	TCAG	GAAGT	TTT	TCTA	ACAA	AGAA	TTTCA	AATT	GGCG	CGTAT	тста	ACA
		т	v	K A	s	I	G	s	т s	s	DK	I	GI	ιv	R	ME	Т	s	s	FS	G	E G	м	LA	s	A
	601	CCACO	GTT	AAA	CGT	TATT	GGC	TCA	ACGA	CTC	GATAA	AAT	0000	ATG	TGCG	ATCC	- 	CTTC	- 	 ጥጥጥል	-	GAAG	ссат.	CTCC	CTAC	ccc
		A (	2 3	N	Т.	<b>π</b> F		, c	т.	N C	к	0	V N	~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	v ,			.0110 v	T E		V D	т	c		,	~
	701			~								~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~					-			1				1 5	м 	6
	101	GGCCC		CAAP		ACIG	AAG		GAIIC	AAI		CAA	GICA	1166	CGTG	ATGA	TTA	TAAG	ATTG	AAAC	CGTGC	GCAT	FTCT.	ACGAG	CGCT	GGC
		TG	1	G	A	LS	Е	T	1 1	R	FS	N	т	L	G V	R	A	S Y	N	VI	MA	т	GG	Т	P V	Q
	801	ACAGO	GGAT	TGGA	GCG	TAAG	CGF	AAT	CATC	ATCO	STTTTT	CTA	ACACI	TTA	GGCG	TAGG	GCG	TCTT	ATAA	TGTC	ATGGC	TACC	GGCG	GCACT	CCCG	TAC
		S	G	τv	'R	Е	L	т	I N	G	VE	I	G	r v	N	D V	н	K	N	DA	D	G R	L	T N	Α	I
	901	AATCA	AGGA	ACTO	TTA	GGAG	CTT	ACC	ATTA	TGGG	GTAGA	AAT	TGGG	ACCG	TGAAT	GATG	TGC	ATAA	AAAC	GACG	CTGAC	GGGA	GGTT	GACTA	ATGC	GAT
		N S	s v	K	D	R T	G	; v	Е	A S	S L	D	ΙΟ	G	R I	N	L	н	<b>S</b> I	D	GR	А	Ι	s v	н	A
1	001	CAACI	rccg	TCAA	AGAG	CAGGA	cco	GCG	TGGA	GCGA	AGCTTG	GAT	ATTC	AGG	GCGCZ	AATTA	TTT	GCAC	TCCA	TTGA	22222	6060		 TCTGT	CCAT	
		Δ ς	Δ	s	G (	v v	F	6	c (	: N	т	C	T	e 1	с т	~	U	N V	T T	C 1		- GCGC	3A11 7 m	10101		JCA N
,	101			~~~~	0							~ ~ ~							1			1	L 1	к	1 0	A
1	101	GCGAG	- JOC	GAGC	GGT	AGGT	111	TGG	GGGA	GGAA	ATT TTA	CAG	GGATI	TCT	GGA	ACAG	CAI	GCAG	TTAT	TGGGG	CGCTT	AACC	I'TGA(	CCAGA	ACCG.	ACG
		R	D	1 1	· •	s	G	V	N F	S	нν	G	FF	i s	А	Q G	v	A	E	Ү Т	v	N L	R	A V	R	G
1	201	CTAGA	GAC.	ATTA	TTG	GAGC	GGI	GTG	AATTI	TAGO	CATGT	GGG	CTTTC	CATTO	CCGCI	CAAG	GGG	TGGC	AGAA	TACAG	CCGTG	AATT	<b>FGAG</b>	AGCGG	TTAG	GGG
		IF	D	A	N	VA	S	A	А	G A	N N	A I	N G	A	QF	E	Т	N	s Q	G	ΙG	А	G'	νт	S	L
1	301	CATTI	TTG.	ATGC	GAA	GTGG	CTI	CAG	CAGCO	GGGG	GCGAAC	GCT	AATGO	CGC	ACAAC	CTGA	GAC	CAAT	TCTC	AAGG	CATAG	GGGC	rggg	STAAC	AAGC	CTT
		KG	A	м	I V	и м	D	м	A D	) S	AR	т	Q	LI	о к	I	R	S D	м	G	s v	0 1	M E	L	и т	т
1	401	AAAGG	GGC	GATG	ATTO	TGAT	GGA	CAT	GCGG	аттс	AGCGC	GCA	GCAR	TTA	SACAP	GATC	CGA	TCGG	ATAT	GGGT	TCGGT	GCAN	 • TCC	- • • • • • • •		- A A
		т	N	N T	s	v	T	0 1	J N	v	K D	Δ	F C		т				5		F	C	11 GG/		J I I M	v
1	501	CCATT	ידעמי	 		ידכידא	100	*		• •		cccr		. ¥	- 	20200	v م س م	meer				S A	14	г Э аттт		1
-	501	N T				C C	ACC	CAA	31GAP	M N	AAAGC	3000	GAAI		1AAA	AGAG	ATG	I GGA	TTTT(	JUTG/	AGAG	AGTG	-GAA(	TTTT	TAA	ATA
	c 0 1	C) ) []	. L	А 1000~	~~	3 G	3		A	m A	Q	A	N A	V	Q C	N	V	L	КГ	Ĺ	Q *					
1	601	CAATA	TTTT:	TGGC	GCAF	AGCG	GGA	GTT	TTGCI	'AT'GG	CGCAA	GCG	AATGC	GGT	SCAAC	AGAA	TGT	CTTA	AGGC	rttt <i>i</i>	ACAAT	AACA	SCCC	TTTTA.	ATTC	AAA

1701 AGGGCGTTAGCCCTTTTTAT CAGTTATTTTTATAAGTTAG AATGATGGATATTTATCAAA AAAACTTACAAGCTCTTTTC AAAAAAGACCCTCTTTTGTT

FIG. 2. Nucleotide sequences of the *H. mustelae* (A) and *H. pylori* (B) *flaB* flagellin genes. Numbers on the left represent the nucleotide positions of the first base in each line. The deduced amino acid sequences are given by letters positioned above the first nucleotide of the codons. Asterisks mark the stop codons. Bases marked with an + or an × in panel A belong to two putative upstream enhancing elements located directly adjacent to each other. Sixteen underlined bases marked  $\sigma^{54}$  represent the  $\sigma^{54}$ -dependent promoters. The ribosome-binding sites (rbs) are indicated by double underlining, and putative transcription terminators are indicated by arrows. In panel B, the transcription start site is indicated by double underlining and marked  $\pm 1$ start site is indicated by double underlining and marked +1.

HmB HpB HpA	**************************************	91 91 91
CcA	G <u>***</u> **V* <u>*</u> *NAK*NSDNLS*A*D <u>0**</u> SR <u>******</u> S <u>***</u> VA <u>***</u> ** <u>*****</u> ANT <u>****</u> *S <u>*G**</u> *L <u>*IL**</u> ******L <u>*****</u> L	91
HmB	**A**I******\$Q****S*****R**M*************************	181
HpB	IKTKAVQAAQDGQTLESRRALQSDIQRLLEELDNIANTTSFNGQQMLSGSFSNKEFQIGAYSNTTVKASIGSTSSDKIGHVRMETSSFSG	181
HpA	V*V**T*****T****T***K*I***V**IQG****G***TY***AL***0*T****V****OSI*****T****O*TA*GALIT	181
CcA	* <u>*</u> * <u>*</u> T <u>*****</u> S*KT <u>*</u> TM* <u>*A</u> **N**M** <u>****</u> ** <u>***</u> ** <u>**</u> K*L <u>***G</u> *T <u>*O***</u> *SS <u>**O</u> *I <u>*****A*O*S***</u> VT <u>*</u> LNRFTKF—	180
HmB	V*****G*N****A***ATD***SFEL*N***************DK**I**T****TS**M****-G*V****R**	270
НрВ	EGMLASAAGANLTEVGLNFKQVNGVNDYKIETVRISTSAGTGIGALSEIINRFSNTLGVRASYNVMATGGTPVQSGTVR-ELTINGVEIG	270
HpA	ASGDIS*T****D****VTL*S*KV*S*****V*A*V*KN**RT**K*YAS*IT*SDVA****SLS_N**L**IHL*	260
CcA	SGT** <u>*</u> TI <u>*</u> NS* <u>*</u> IE*F*FDS <u>*</u> V* <u>*</u> * <u>V***L</u> <u>**A</u> * <u>E</u> <u>**</u> *NADKT <u>*</u> I* <u>*</u> TFDLKSV*AYAIKA <u>*</u> NTSQDFA* <u>**</u> *V* <u>*</u>	260
HmB	***E*R***S****I*****NQ*******T****V*L*****DGEASH***E*****NN**IV*****I*********	358
HpB	TVNDVHKNDADGRLTNAINSVKDRTGVEASLDIQGRINLHSIDGRAISVHAASASGQVFGGGNFTGISGTQHAVI-GRLTLTRTDARDI	358
HpA	NIA*IK***S****VA***A*TSE****YT*QK**L**R****G*EIKTD*V*NGPSALTMVN*GODLTKGSTNY***S***L**KS*	350
CcA	K**YSDG <u>*EN*S*IS***A</u> ***T <u>***Q*</u> *K <u>*</u> EN <u>*</u> KLV <u>*</u> T <u>*A***G*</u> KI-TG*IGVGAGILHTENY <u>***</u> S <u>*</u> VKN <u>*</u> G** <u>*</u>	337
HmB	********_*I*L********T*A***QL**M*G*DI******************************	402
НрВ	IVSGVNFS-HVGFHSAQGVAEYTVNLRAVRGIFDANVASAAGANA	402
НрА	N*VSASD*Q*L**TAIGFGESQ-***T****D*T*N*N***K**	399
CcA	NI**T <u>**</u> SAIGMGATDMISQSS*S <u>**</u> ESK <u>*QIS*</u> AN*D <u>*</u> MGFNAYNGGGAKQIIFASSIAGFMSQAGSGFSAGSGFSVGS <u>*K*</u> Y	427
HmB	*K***DI*R****	453
НрВ	NGAQAETNSQGIG-AGVTSLKGAMIVMDMADSARTQLDKIRSDMGSVQMELV	453
НрА	*AVI*SG*QSLGS***T*R***V*I*I*E***NQMI	449
CcA	SAILSASIQIVSSAASISSTYVVSTGSGFSAGSGNSQFAALRISTVSAHDET* <u>***</u> T <u>****</u> A <u>***</u> I*ET <u>*</u> I*N <u>**Q</u> * <u>A</u> * <u>I</u> ****NQIT	511
HmB	S*****T*******************************	
HpB	TTINNISVTQVNVKAAESQIRDVDFAEESANFSKYNILAQSGSFAMAQANAVQQNVLRLLQ 514	
НрА	S*V****I**************************N*N*N******	
CcA	S <u>*****</u> T* <u>*****</u> S <u>*****</u> S <u>*****</u> S <u>*****</u> S <u>****</u> Y* <u>*A</u> <u>******</u> Y <u>*****</u> SS <u>****</u> *572 54% identity	

FIG. 3. Comparison of the deduced amino acid sequence of the *H. pylori flaB* gene with those deduced from the *H. mustelae flaB*, *H. pylori flaA*, and *C. coli flaA* genes. The amino-terminal methionine residues have been omitted in correspondence with the situation in the mature proteins. The letters on the left represent the sources of the sequence (HmB, *H. mustelae flaB*; HpB, *H. pylori flaB*; HpA, *H. pylori flaA*; CcA, *C. coli flaA*); the numbers on the right represent the positions of the last amino acid in each line. Asterisks mark amino acids identical with those in the *H. pylori flaB* sequence. Underlining indicates residues conserved in all four compared sequences.

dent promoter sequence (14 of 16 nucleotides identical with the *E. coli* consensus sequence) (51). About 100 nucleotides upstream of the ribosome-binding site, two potential upstream enhancing sequences (GGT-N10-ACA and TGC-N10-ACA) were found.

Cloning and sequencing of the flaB flagellin gene from H. pylori 85P. To clone the gene coding for the minor H. pylori flagellin, we used the cloned H. mustelae flaB gene as a probe to screen the previously described cosmid gene bank of H. pylori 85P (29). To obtain a probe free of vector contamination, two oligonucleotides (OLHMflaB-2 and -3) that permitted the amplification of the entire gene from the genomic DNA of H. mustelae by PCR were synthesized. The purified PCR product was used to screen 400 cosmid clones by low-stringency colony blot hybridization. Seven cosmids were found to hybridize, and one of them (pILL683) was used to subclone the *flaB*-related gene. Fragments with sizes of 3 to 4 kb were generated by partial restriction of the cosmid DNA with endonuclease *Sau3A*, purified, and ligated into the *Bgl*II site of plasmid vector pILL570. One hundred clones were screened with the *H. mustelae flaB* probe, and 1 of the 12 hybridizing clones (pSUS19) was mapped in detail (Fig. 1) and sequenced. The nucleotide sequence is shown in Fig. 2B. The clone contained one large ORF of 1,542 nucleotides coding for a protein of 53.9 kDa. The N-terminal amino acid sequence of the deduced protein was completely

 TABLE 3. Alignment of the N-terminal amino acid sequence deduced from the cloned H. mustelae flagellin gene with available sequences obtained by protein sequencing

Source N-terminal amino acid sequence <sup>a</sup>				
H. mustelae FlaB	MSFRINTNIAALNAHSIGV			
H. pylori FlaB <sup>b</sup>	SFRINTNIAALTSHAVGV			
H. mustelae FlaA <sup>c</sup>	A F Q V N T N I N A L T T X A * * * * * * * * * * * * *			
H. pylori FlaA <sup>d</sup>	AFQVNTNINAMNAHVQSA			

<sup>a</sup> Asterisks indicate residues identical with those of the cloned H. mustelae flagellin gene.

<sup>b</sup> From Kostrzynska et al. (26).

<sup>c</sup> From Suerbaum et al. (49).

<sup>d</sup> From Leying et al. (32). Residues 1 to 13 were obtained by protein sequencing, and residues 14 to 18 were deduced from the nucleotide sequence.



FIG. 4. Mapping of the transcription start point of the *H. pylori* flaA (A) and flaB (B) genes by primer extension. Lanes: T, C, G, and A, DNA sequencing reactions performed with the same primer as that used for primer extension; 1, primer extension reactions; 2, control reactions with addition of RNase prior to extension. Arrows mark the major primer extension products.

identical to the N-terminal sequence of the purified minor *H.* pylori flagellin as reported by Kostrzynska et al. (26). The gene was named *flaB*. Upstream of the ORF, a Shine-Dalgarno site and a  $\sigma^{54}$  promoter-like sequence (14 of 16 nucleotides matching the *E. coli* consensus) were found. Downstream, a putative transcription terminator (free energy,  $\Delta G = -15.8$  kcal [ca. -66.1 kJ]/mol) was identified.

Mapping of the transcription start point of the *H. pylori flaB* gene. Since the functional activity of  $\sigma^{54}$ -like promoters had not yet been demonstrated in *Helicobacter* species, a primer extension experiment was performed to determine the transcription start point for the *H. pylori flaB* gene. Total RNA was extracted from *H. pylori* 85P, and the mRNA start point of the *flaB* gene was determined by extension of oligonucleotide OLHPflaBPE-1 (Fig. 4B). The transcription of the *flaB* gene starts at the adenine residue (nucleotide 113 of the sequence, as shown in Fig. 2B) 25 nucleotides upstream of the ATG start codon. This residue is positioned 12 nucleotides downstream of the end of the proposed promoter. This result therefore confirms the functional activity of the proposed promoter. The amount of extended product was very low when compared with the amount of product obtained in a parallel experiment using the *flaA*-specific primer OLHPflaAPE-2 (Fig. 4A). No extension product was obtained in the control reactions in which RNase was added prior to extension.

Construction of isogenic flaA and flaB mutants of H. pylori N6. Isogenic mutants of H. pylori N6 deficient in expression of major or minor flagellin were constructed by the following approach: for inactivation of flaA, a 1.4-kb BamHI restriction fragment of plasmid pILL600 containing a gene encoding resistance to kanamycin (aph3'-III) was cloned into the unique BamHI site of plasmid pHL319-2-4, which is situated in the middle of the H. pylori flaA gene, giving rise to plasmid pSUS10. For inactivation of *flaB*, the same kanamycin cassette (cut out with SmaI) was inserted into the unique SacI site of pSUS19, situated in the flaB gene, after treatment of the linearized pSUS19 with T4 polymerase. The resulting plasmid was called pSUS22. H. pylori N6 was transformed with either pSUS10 or pSUS22 by electroporation. More than 50 independent transformants were obtained in transformation experiments with both plasmids. Several mutant strains (defined as descendants of a single colony of kanamycin-resistant H. pylori) were analyzed by a PCR approach (Fig. 5). The strains were tested (i) for the absence of vector sequences by using oligonucleotides OLpBR322-1 and -2 derived from the nucleotide sequence of pBR322 shared by pILL570 (not shown), (ii) for the presence of the kanamycin cassette by using primers OLKm-1 and -2 derived from the nucleotide sequence of the kanamycin resistance gene (Fig. 5A), and (iii) for integration of the cassette into the targeted gene by using primers specific for sequences of the targeted gene upstream and downstream of the site of disruption (primers OLHPflaA-1 and -2 for the detection of flaA [Fig. 5B]) and OLHPflaB-2 and -7 for amplification of a part of *flaB* [Fig. 5C]). In the mutant strains, under the experimental conditions used, disruption of the targeted gene led to a negative amplification reaction for the particular gene because of the augmentation in size of



FIG. 5. Genotypic characterization of *H. pylori* wild-type strain N6 and corresponding isogenic *flaA* and *flaB* mutants by PCR. (A) Detection of the kanamycin resistance gene (predicted PCR product, 630 bp); (B) detection of the *flaA* gene (predicted PCR product, 1,400 bp); (C) detection of the *flaB* gene (predicted PCR product, 750 bp). Lanes: 1, wild-type strain *H. pylori* N6; 2, N6 *flaA* mutant; 3, N6 *flaB* mutant; S, DNA size markers.



FIG. 6. Western blot of flagella partially purified from *H. pylori* N6 and the isogenic *flaA* and *flaB* mutants developed with an antiserum raised against *H. pylori* flagellar filaments.

the fragment to be amplified (fragments over 2 kb could not be amplified in our system that uses unpurified bacterial lysates as a source of template DNA). Correct integration was also verified by using combinations of one primer specific for the targeted gene and one specific for the kanamycin cassette (data not shown). For all tested strains, the analysis confirmed that a double crossover event had taken place, leading to the complete elimination of the vector and to rescue of the antibiotic resistance by allelic replacement of the intact copy of the targeted gene by the disrupted copy introduced by electroporation.

**Phenotypic characterization of H.** pylori flaA and flaB mutants. The flaA and flaB mutant strains were characterized by Western blots of partially purified flagella developed with antiserum AK179 (raised against purified flagellar filaments) and by motility testing. The results of the Western blots (Fig. 6) showed disappearance of the 53-kDa FlaA band in the N6 flaA mutant and disappearance of a weak band of slightly higher molecular weight in the flaB mutant, showing that the introduced mutations selectively abolished the expression of only the targeted flagellin gene, but did not seem to affect the expression of the other gene to a detectable extent.

The motility of the mutant strains was tested by demonstration of swarming on soft agar plates (Fig. 7). Since the plates contained kanamycin to assure the stability of the mutants, a previously described kanamycin-resistant *ureB*disrupted mutant of *H. pylori* N6 (11) was used as a control. *flaA* mutants were completely nonmotile. *flaB* mutants, however, maintained motility, and the degree of their motility compared with that of the *ureB* mutant appeared unaffected by the mutation.

#### DISCUSSION

Motility is considered a major virulence factor of H. pylori and (together with urease expression) is one of the two



FIG. 7. Motility of isogenic *flaA* and *flaB* mutant strains of *H. pylori* N6 compared with urease-negative *ureB* mutant. The figure shows swarming of the *ureB* and *flaB* strains in a kanamycin-containing soft agar plate. The *flaA* strain exhibits no swarming.

bacterial properties for which this hypothesis is supported by experiments in an animal model (7–9).

The H. pylori flagellum was therefore among the first structures of this bacterium to be subjected to a detailed molecular analysis. In 1989, Geis et al. (16) reported the purification of H. pylori flagellin as a single 51-kDa protein. The N-terminal amino acid sequence of this protein was subsequently determined and proved that the purified molecule was indeed a flagellin (19). Later, Kostrzynska et al. (26) published a more detailed ultrastructural and biochemical study of the H. pylori flagellar filament. These investigators found that the filament was composed of two different flagellin species, a predominant one with a molecular mass of 56 kDa and a minor one with a molecular mass of 57 kDa that was expressed in much lower amounts. Comparison of the N-terminal sequences showed that the 51-kDa protein described by Geis et al. and the 56-kDa major flagellin of Kostrzynska et al. were the same proteins. The gene coding for this major flagellin was recently cloned and sequenced by Leying et al. (32) by using an oligonucleotide derived from the N-terminal amino acid sequence of the purified protein as a probe. The predicted molecular mass of the gene product is 53.2 kDa (it is, however, not excluded that the molecular mass of the native protein is higher if it is posttranslationally modified as has been described for Campylobacter coli flagellins [33]).

In this study, the cloned flaA gene from H. pylori was used as a probe to clone a flagellin gene from the closely related animal pathogen H. mustelae. The gene that was cloned and sequenced, although clearly a flagellin gene, did not encode the major H. mustelae flagellin, which we had purified previously (49). The N-terminal amino acid sequence of the deduced gene product exhibited, however, strong homologies with that of the minor *H. pylori* flagellin as described by Kostrzynska et al. Since the gene encoding the minor H. pylori flagellin had not been described to date, we used the cloned H. mustelae gene to screen a cosmid gene bank of H. pylori 85P and to clone and sequence the homologous H. pylori flaB gene. The N-terminal amino acid sequence of the protein deduced from the nucleotide sequence of this gene was exactly identical with that of the purified minor flagellin. Also, its amino acid composition was in good agreement overall with the results of chemical determinations reported by Kostrzynska et al. (26), except for the proline content, which was significantly lower.

The H. mustelae flaB gene is the first entire structural gene to be cloned from this organism. It is very similar to the H. pylori flaB gene. The deduced proteins share 420 (81.7%) identical amino acids. This level of conservation is similar to that found for the ureA and ureB genes of H. pylori and H. felis (12) and might be representative of the degree of conservation of functionally homologous proteins within the gastric Helicobacter sp. The degree of homology with other bacterial flagellins is much lower. In particular, the flaA and flaB genes of H. pylori share only 58% identical amino acids, indicating that both proteins may by functionally different. The relatively low relatedness of H. pylori FlaA and FlaB is remarkable because it represents a striking difference from the situation in the closely related genus Campylobacter. C. coli and Campylobacter jejuni both have two flagellin genes. The two genes in these species are, however, almost identical (more than 90% identity on both the amino acid and nucleotide levels) and located almost directly adjacent to each other (22, 39). In H. pylori, not only are the genes much less homologous but also there was no evidence for a genetic linkage of the genes, since they were found on completely unrelated cosmids (data not shown). Hybridization studies have not provided evidence for the existence of further flagellin genes homologous with flaA or flaB (data not shown).

Both genes are preceded by typical AGGA Shine-Dalgarno sequences (46) and begin with an ATG start codon. Upstream of both the H. mustelae and H. pylori flaB genes, there are sequences which closely (14 of 16 nucleotides) resemble the consensus sequence for the so-called  $\sigma^{54}$ dependent promoters (5'-TGGYRYR-N4-TTGCA-3', with Y = T or C and R = G or A [51]). Since the activity of  $\sigma^{54}$ -like promoter sequences had not yet been demonstrated in H. pylori, a primer extension experiment that showed that the synthesis of *flaB* mRNA starts 12 nucleotides after the end of the proposed promoter was performed. Since this distance is typical for  $\sigma^{54}$ -dependent promoters (also called -12/-24 promoters), the result confirms that transcription of the *H. pylori flaB* gene is in fact under the control of the proposed promoter. The small amount of extended product in comparison with the amount of product obtained in a parallel experiment using a *flaA*-specific primer (that otherwise confirmed the position of the *flaA* mRNA start point as reported by Leying et al. [32]) suggests that the low expression level of FlaB compared with that of FlaA is due to regulation at the transcriptional level.

 $\sigma^{54}$ -dependent promoters were initially found to regulate the expression of genes involved in the assimilation of nitrogen. More recently,  $\sigma^{54}$ -dependent promoters have also been found associated with genes serving other physiological functions, including pilin genes in *Pseudomonas aeruginosa* and *Neisseria gonorrhoeae*, the *H. pylori* urease, and several flagellar genes in different species (5, 29; for reviews, see references 27 and 51).

Although the majority of bacterial flagellin genes have  $\sigma^{28}$ -like promoter elements (including the *H. pylori* and *Campylobacter flaA* genes),  $\sigma^{54}$ -like promoters have been found upstream of several *Caulobacter crescentus* flagellar genes and the *C. coli* and *C. jejuni flaB* genes (25). The precise role of these promoters in the regulation of flagellar assembly has yet to be studied.

Also notable is the relatively high G+C content of both the *H. mustelae* and the *H. pylori flaB* genes (45.3 and 46.8%, respectively) when compared with the reported overall G+C

content of the genomes. This is a further difference from the *Campylobacter* flagellin genes, which have G+C contents matching the low (about 35% for *C. jejuni*) G+C contents of these species.

The construction of unflagellated and nonmotile mutants of *H. pylori* by disruption of the *flaA* gene with a chloramphenicol acetyltransferase gene cassette was first reported by Haas et al. (23). The *flaA* mutants reported in this study have been constructed by a very similar approach, and the same flagellin clone (pHL319-2-4) was used to construct the disruption. As expected, the results for the *flaA* mutants of strain N6 were identical to those reported earlier: *flaA* mutants were incapable of synthesizing FlaA as visualized by SDS-PAGE (not shown) and Western blotting (Fig. 6), and mutants were nonmotile and did not have flagella when analyzed by electron microscopy (48).

As for the minor flagellin, Kostrzynska et al. (26) described that it was primarily located in the proximal part of the filament. Our study demonstrates that flaB mutants retained motility, and the minor flagellin thus does not seem to be required for motility. Preliminary electron microscopy studies (48) have confirmed that the bacteria are indeed flagellated, but an in-depth study of the morphology and other properties of the *flaB* mutants has yet to be performed and will be published separately. So far, the role of the minor flagellin is unclear and can be only an object of hypotheses. Because of their different promoters, the expression of both flagellins can probably be regulated independently, and it seems therefore conceivable that the relative contribution of both proteins to the filament may vary in vivo and permit an adaptation of the physicochemical properties of the filament to different environmental conditions such as the viscosity of the medium. Antibodies against the major flagellin are produced in most patients with H. pylori infection (17), and although it is as yet unknown if those antibodies prevent motility, it seems conceivable that an increase of the relative amount of FlaB in the filament can facilitate bacterial evasion from host immunity. It also remains to be investigated whether one or both *H. pylori* flagellins, by certain domains, interact with the flagellar sheath and contribute to its assembly. Studies are now under way in our laboratories to test these hypotheses and to gain further insight into the relative roles of the two flagellin molecules and the regulation of their expression. Likewise, cloning of the H. mustelae flaA gene and of the H. felis flagellin genes as well as the construction of mutants in these organisms are under way. Studies of our H. pylori flaA and flaB mutants in the gnotobiotic piglet model as well as of nonmotile H. mustelae and H. felis mutants in the respective animal models will permit clarification of the role of flagella and motility in the pathogenesis of gastric Helicobacter infections and also evaluation of the usefulness of flagellin mutants for vaccination purposes.

#### ACKNOWLEDGMENTS

S.S. thanks Wolfgang Opferkuch for his continuous help and encouragement, Hermann Leying for the gift of pHL319-2-4 and antiserum AK179 as well as for his support of this work, and Susanne Wendt for technical assistance with the Western blot.

S.S. was supported by grant Su 133/1-1 from the Deutsche Forschungsgemeinschaft. The work in the laboratory of A.L. was supported by research grants from the Fondation del'Avenir and the Fondation Recherche et Partage.

### REFERENCES

1. Blaser, M. J. 1990. *Helicobacter pylori* and the pathogenesis of gastroduodenal inflammation. J. Infect. Dis. 161:626-633.

- Casadaban, M., and S. N. Cohen. 1980. Analysis of gene control signals by DNA fusion and cloning in *E. coli*. J. Mol. Biol. 138:179-207.
- Chirgwin, J. M., A. E. Przybyla, R. J. MacDonald, and W. J. Rutter. 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry 18:5294– 5299.
- Cussac, V., R. L. Ferrero, and A. Labigne. 1992. Expression of Helicobacter pylori urease genes in Escherichia coli grown under nitrogen-limiting conditions. J. Bacteriol. 174:2466-2473.
- Dick, J. D. 1990. Helicobacter (Campylobacter) pylori: a new twist to an old disease. Annu. Rev. Microbiol. 44:249-269.
- Eaton, K. A., C. L. Brooks, D. R. Morgan, and S. Krakowka. 1991. Essential role of urease in pathogenesis of gastritis induced by *Helicobacter pylori* in gnotobiotic piglets. Infect. Immun. 59:2470-2475.
- Eaton, K. A., D. R. Morgan, and S. Krakowka. 1989. Campylobacter pylori virulence factors in gnotobiotic piglets. Infect. Immun. 57:1119–1125.
- Eaton, K. A., D. R. Morgan, and S. Krakowka. 1992. Motility as a factor in the colonisation of gnotobiotic piglets by *Helicobac*ter pylori. J. Med. Microbiol. 37:123–127.
- Feinberg, A., and B. Vogelstein. 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. 132:6–13.
- Ferrero, R. L., V. Cussac, P. Courcoux, and A. Labigne. 1992. Construction of isogenic urease-negative mutants of *Helicobacter pylori* by allelic exchange. J. Bacteriol. 174:4212–4217.
- 12. Ferrero, R. L., and A. Labigne. Mol. Microbiol., in press.
- Fox, J. G., P. Correa, N. S. Taylor, A. Lee, G. Otto, J. C. Murphy, and R. Rose. 1990. *Helicobacter mustelae*-associated gastritis in ferrets. An animal model of *Helicobacter pylori* gastritis in humans. Gastroenterology 99:352-361.
- Fox, J. G., G. Otto, J. C. Murphy, N. S. Taylor, and A. Lee. 1991. Gastric colonization of the ferret with *Helicobacter* species: natural and experimental infections. Rev. Infect. Dis. 13(Suppl. 8):S671–S680.
- Fox, J. G., G. Otto, N. S. Taylor, W. Rosenblad, and J. C. Murphy. 1991. *Helicobacter mustelae*-induced gastritis and elevated gastric pH in the ferret (*Mustela putorius furo*). Infect. Immun. 59:1875-1880.
- Geis, G., H. Leying, S. Suerbaum, U. Mai, and W. Opferkuch. 1989. Ultrastructure and chemical analysis of *Campylobacter pylori* flagella. J. Clin. Microbiol. 27:436–441.
- 17. Geis, G., and S. Suerbaum. Unpublished data.
- Geis, G., S. Suerbaum, B. Forsthoff, H. Leying, and W. Opferkuch. Ultrastructure and biochemical studies of the *Helicobac*ter pylori flagellar sheath. J. Med. Microbiol., in press.
- Geis, G., S. Suerbaum, H. Leying, B. Forsthoff, H. P. Kroll, and W. Opferkuch. 1990. Investigations on ultrastructure and biochemistry of *Helicobacter pylori* flagellar filament and sheath. Rev. Esp. Enferm. Apar. Dig. 78(Suppl. 1):33.
- Goodwin, C. S., J. A. Armstrong, T. Chilvers, M. Peters, M. D. Collins, L. Sly, W. McConnell, and W. E. S. Harper. 1989. Transfer of *Campylobacter pylori* and *Campylobacter mustelae* to *Helicobacter* gen. nov. as *Helicobacter pylori* comb. nov. and *Helicobacter mustelae* comb. nov., respectively. Int. J. Syst. Bacteriol. 39:397-405.
- Goodwin, C. S., R. K. McCulloch, J. A. Armstrong, and S. H. Wee. 1985. Unusual cellular fatty acids and distinctive ultrastructure in a spiral bacterium (*Campylobacter pyloridis*) from the human gastric mucosa. J. Med. Microbiol. 19:257–267.
- Guerry, P., S. M. Logan, S. Thornton, and T. J. Trust. 1990. Genomic organization and expression of *Campylobacter* flagellin genes. J. Bacteriol. 172:1853–1860.
- Haas, R., H. Leying, T. F. Meyer, G. Geis, S. Suerbaum, and W. Opferkuch. 1991. Cloning of a *H. pylori* flagellin gene and construction of a non-flagellated mutant by transformation-mediated allelic exchange. Ital. J. Gastroenterol. 23(Suppl. 2):31.

- 24. Hazell, S. L., A. Lee, L. Brady, and W. Hennessy. 1986. Campylobacter pyloridis and gastritis: association with intracellular spaces and adaptation to an environment of mucus as important factors in colonization of the gastric epithelium. J. Infect. Dis. 153:658–663.
- Helman, J. D. 1991. Alternative sigma factors and the regulation of flagellar gene expression. Mol. Microbiol. 5:2875–2882.
- Kostrzynska, M., J. D. Betts, J. W. Austin, and T. J. Trust. 1991. Identification, characterization, and spatial localization of two flagellin species in *Helicobacter pylori* flagella. J. Bacteriol. 173:937-946.
- 27. Kustu, S., E. Santero, J. Keener, D. Popham, and D. Weiss. 1989. Expression of  $\sigma^{54}$  (*ntrA*)-dependent genes is probably united by a common mechanism. Microbiol. Rev. **53**:367–376.
- Labigne, A., P. Courcoux, and L. Tompkins. 1992. Cloning of Campylobacter jejuni genes required for leucine biosynthesis, and construction of *leu*-negative mutant of *C. jejuni* by shuttle transposon mutagenesis. Res. Microbiol. 143:15–26.
- Labigne, A., V. Cussac, and P. Courcoux. 1991. Shuttle cloning and nucleotide sequences of *Helicobacter pylori* genes responsible for urease activity. J. Bacteriol. 173:1920–1931.
- Labigne-Roussel, A., P. Courcoux, and L. Tompkins. 1988. Gene disruption and replacement as a feasible approach for mutagenesis of *Campylobacter jejuni*. J. Bacteriol. 170:1704–1708.
- Lee, A., J. G. Fox, G. Otto, and J. Murphy. 1990. A small animal model of human *Helicobacter pylori* active chronic gastritis. Gastroenterology 99:1315–1323.
- 32. Leying, H., S. Suerbaum, G. Geis, and R. Haas. 1992. Cloning and genetic characterization of a *Helicobacter pylori* flagellin gene. Mol. Microbiol. 6:2863–2874.
- Logan, S. M., T. J. Trust, and P. Guerry. 1989. Evidence for posttranslational modification and gene duplication of *Campylobacter* flagellin. J. Bacteriol. 171:3031–3038.
- 34. Lugtenberg, B., J. Meijers, R. Peters, P. van der Hoek, and L. van Alphen. 1975. Electrophoretic resolution of the "major outer membrane protein" of *Escherichia coli* K12 into four bands. FEBS Lett. 58:254–258.
- 35. Majewski, S. I., and C. S. Goodwin. 1988. Restriction endonuclease analysis of the genome of *Campylobacter pylori* with a rapid extraction method: evidence for considerable genomic variation. J. Infect. Dis. 157:465-471.
- 36. Marsh, J. L., M. Erfle, and E. J. Wykes. 1984. The pIC plasmid and phage vectors with versatile cloning sites for recombinant selection by insertional inactivation. Gene 32:481-485.
- Messing, J., and J. Vieira. 1982. A new pair of M13 vectors for selecting either DNA strand of double-digest restriction fragments. Gene 19:269-276.
- Nomura, A., G. N. Stemmermann, P.-H. Chyou, I. Kato, G. I. Perez-Perez, and M. J. Blaser. 1991. *H. pylori* infection and gastric carcinoma among Japanese Americans in Hawaii. N. Engl. J. Med. 325:1132–1136.
- Nuijten, P. J., F. J. van Asten, W. Gaastra, and B. A. van der Zeijst. 1990. Structural and functional analysis of two Campylobacter jejuni flagellin genes. J. Biol. Chem. 265:17798-17804.
- O'Rourke, J., A. Lee, and J. G. Fox. 1992. An ultrastructural study of *Helicobacter mustelae* and evidence of a specific association with gastric mucosa. J. Med. Microbiol. 36:420-427.
- Parsonnet, J., G. D. Friedman, D. P. Vandersteen, Y. Chang, J. H. Vogelman, N. Orentreich, and R. K. Sibley. 1991. *Helicobacter pylori* infection and the risk of gastric carcinoma. N. Engl. J. Med. 325:1127-1131.
- 42. Peterson, W. L. 1991. *Helicobacter pylori* and peptic ulcer disease. N. Engl. J. Med. 324:1043-1048.
- 43. Rosenberg, M., and D. Court. 1979. Regulatory sequences involved in the promotion and termination of RNA transcription. Annu. Rev. Genet. 13:319–353.
- 44. Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74:5463-5467.
- 46. Shine, J., and L. Dalgarno. 1974. The 3'-terminal sequence of

*Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. Proc. Natl. Acad. Sci. USA **71:**1342–1346.

- Southern, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98:503-517.
- 48. Suerbaum, S., and G. Geis. Unpublished data.
- Suerbaum, S., G. Geis, C. Josenhans, and W. Opferkuch. 1992. Biochemical studies of *Helicobacter mustelae* fatty acid composition and flagella. Infect. Immun. 60:1695–1698.
- Sutcliffe, J. G. 1979. Complete nucleotide sequence of the Escherichia coli plasmid pBR322. Cold Spring Harbor Symp. Quant. Biol. 43:77-86.
- 51. Thöny, B., and H. Hennecke. 1989. The -24/-12 promoter comes of age. FEMS Microbiol. Rev. 63:341-358.
- Towbin, H., T. Staehelin, and J. Gordon. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc. Natl. Acad. Sci. USA 76:4350-4354.
- Trieu-Cuot, P., G. Gerbaud, T. Lambert, and P. Courvalin. 1985. In vivo transfer of genetic information between Grampositive and Gram-negative bacteria. EMBO J. 4:3583–3587.
- 54. Yanisch-Perron, C., J. Vieira, and J. Messing. 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene 33:103–199.