Nucleic Acids Research, 1999, Vol. 27, No. 16 **3325–3333**

Conservation, localization and expression of HopZ, a protein involved in adhesion of *Helicobacter pylori*

Birgit Peck, Martina Ortkamp, Klaus D. Diehl, Erika Hundt and Bernhard Knapp*

Chiron Behring GmbH & Co., Preclinical Research, PO Box 1630, D-35006 Marburg, Germany

Received April 30, 1999; Revised and Accepted July 6, 1999

DDBJ/EMBL/GenBank accession nos Y18980–Y18993

ABSTRACT

From a sarkosyl-insoluble outer membrane fraction prepared from the *Helicobacter pylori* strain ATCC 43504, 19 proteins could be sequenced N-terminally by Edman degradation. Oligonucleotides were deduced and used for screening of a genomic library. From the isolated genes, five code for different members of a *H.pylori* outer membrane protein (Hop) family. Among these, the *hopZ* gene was characterized in more detail. It encodes a protein which was shown to be located at the bacterial surface by immunofluorescence studies. Sequence analysis of the *hopZ* gene from 15 different *H.pylori* strains revealed the existence of two alleles and the possible regulation of *hopZ* expression by slipped-strand mispairing within a CT dinucleotide repeat motif located in the signal-peptide coding region. Among the different strains, the influence of this region on the expression of HopZ was analyzed on a translational level by western blot analysis of bacterial extracts and immunofluorescence studies on intact cells. The protein is expressed only in those strains in which the number of the CT dinucleotide repeats allow for an open reading frame encoding the complete protein. Addionally the function of HopZ was investigated in an adhesion assay. The wild-type strain ATCC 43504 adhered to human gastric epithel cells whereas a knockout mutant strain showed significantly reduced binding to the cells.

INTRODUCTION

Helicobacter pylori is a Gram-negative, microaerophilic, spiral bacterium which colonizes the mucosa of the human stomach. The bacterium is the main cause of chronic active gastritis and peptic ulcer, in particular duodenal ulcer, and plays a role in the development of gastric carcinoma.

Its helical shape and motility, due to four to six flagellae, enables the bacterium to migrate through the gastric mucus in order to reach the boundary layer between the mucus and the mucosa. The bacterium adheres to the endothelial cells of the stomach which protects it from the extreme acidity of the gastric lumen and displacement from the stomach by forces such as those generated by peristalsis and gastric emptying.

A consequence of chronic colonization of the mucosa can be an inflammatory granulocytic and subsequently monocytic infiltration of the epithelium which in turn, by mediators of inflammation, contributes to tissue destruction. Infection stimulates both a local and a systemic humoral immune response which are, however, unable to eliminate the pathogen effectively. Immunization is the conventional way of preventing infectious diseases. Therefore it is important to examine this option with regard to controlling a *H.pylori* infection.

Bacterial virulence factors or structures accessible to the immune system are candidates for the development of subunit vaccines. It is generally assumed that antigens of this nature are present in the outer membrane of the bacterium. Proteins of 48, 49, 50 and 67 kDa (1) and 31 kDa (2), called HopA, HopB, HopC, HopD and HopE, respectively, have been shown to function as porins in *H.pylori*. Recently, Tomb *et al.* (3) identified these proteins among an outer membrane protein family consisting of 32 members.

Three further members of this family were identified as adhesin BabA (4), AlpA and AlpB (5). BabA was demonstrated to bind to the blood group antigen LewisB (4). There are two alleles: $babA_1$ and $babA_2$, differing in the presence of a 10 bp repeat motif which results in the elimination of the start codon and the lack of Le^b antigen-binding activity. Furthermore, the existence of repetitive sequences in the signal-sequence encoding regions of five further OMP genes with the potential for slipped-strand mispairing and recombination which in turn may cause genotype variation has been described (3). Here we report on the isolation of one of these proteins, HopZ, and describe its localization, its genetic heterogeneity within different *H.pylori* strains as well as its putative function as an adhesin.

MATERIALS AND METHODS

Preparation of an outer membrane protein fraction

The *H.pylori* strain ATCC 43504 was grown for 48 h at 37°C under microaerophilic conditions on Columbia agar, and was subsequently cultivated in Columbia broth containing 7% fetal calf serum for 48 h at 37°C. The bacteria were harvested by centrifugation at 10 000 r.p.m. and washed twice with 10 mM Tris, pH 7.4. Cells were disrupted in a glass bead homogenizer (IMA, Heidelberg, Germany) at 4°C and 4000 r.p.m. for 15 min.

*To whom correspondence should be addressed. Tel: +49 6421 39 2850; Fax: +49 6421 39 4709; Email: knapp2@mbg.chiron-behring.com

Unbroken cells were removed by centrifugation. Membrane fractions were then collected by centrifugation (100 000 *g*, 60 min, 4°C), suspended in 7 mM EDTA, 10 mM Tris, pH 7.4, containing 1% (w/v) Sarkosyl® (Ciba Geigy AG, Basel, Switzerland), as described by Blaser *et al.* (6), and incubated at 37°C for 20 min. The Sarkosyl®-insoluble fraction was pelleted by centrifugation for 60 min at 4°C and the pellet was resuspended in water, aliquoted, and stored at –20°C until use.

N-terminal amino acid sequence analysis of proteins

The outer membrane protein fraction was separated by SDS–PAGE using 8, 10 or 16% acrylamide depending on the molecular weight of the proteins to be separated. The bands were electrophoretically transferred to Immobilon membranes (7) with a trans blot cell (Bio-Rad, Hercules, CA). To complete the transfer of the proteins, 0.005% SDS was added to the cathode buffer. Membranes were stained with amido black to visualize the protein bands which were then excised. Amino acid sequencing was performed with a sequenator model 477 A (Applied Biosystems, Foster City, CA). Phenylthiohydantoin derivatives were separated with an on-line analyzer model 120A (Applied Biosystems). Data were handled using the control data analysis module 900 (Applied Biosystems).

Construction of an *H.pylori* genomic library

Helicobacter pylori genomic DNA of strain ATCC 43504 was prepared using the DNAzol™ reagent (Gibco BRL, Rockville, MD). Genomic DNA was partially digested with the restriction enzyme *Sau*3AI and separated by agarose gel electrophoresis. DNA fragments between 0.5 and 5.0 kb were electroeluted (7). The *Sau*3AI treated DNA was cloned into the λZAP vector (Predigested ZAP Express™ *Bam*HI/CIAP Vector Cloning Kit, Stratagene, La Jolla, CA).

Screening of the *H.pylori* genomic library and isolation of the complete genes

From the N-terminal amino acid sequences oligonucleotides were deduced using the species-specific codon usage of *H.pylori*, which had been determined from 19 known *H.pylori* genes (8). The oligonucleotides were labeled with digoxigenin (DIG) using the DIG Oligonucleotide 3'-End Labeling Kit (Boehringer Mannheim, Germany) and employed for screening of the λZAP *H.pylori* genomic library under standard conditions, but using a hybridization temperature of 32°C (9). Inserts showing hybridization were excised and sequenced by the dideoxynucleotide chain termination method (10) using the DNA sequencer 373A (Applied Biosystems). From the boundaries of the *Sau*3AI fragments encoding specific amino acid sequences identical to those determined from the proteins by Edman degradation oligonucleotides were deduced, DIG labeled and used for screening of the ZAP Express™ library in order to determine the complete sequences of the isolated genes.

Amplification of *hopZ* from different *H.pylori* strains

Genomic DNA from the *H.pylori* strains 25392, 25144, pa1, pa2 (obtained from P. Malfertheiner, University of Magdeburg, Germany), 1811a, 5060d(A), G50, Tx30A, 342, 326/01 (obtained from A. Covacci, Chiron Vaccines S. p. A., Siena, Italy), ATCC 43504, ATCC 51110, ATCC 49503 and CCUG 17874 were prepared as described above. For the amplification of the complete *hopZ* gene of all strains, the oligonucleotides GCCTGATATGGGTGGCATGGG (forward) and CTGAGC-TAACGGGTCTCACAAAAA (reverse) were used. PCR products were purified using the PCR purification kit (Qiagen, Hilden, Germany) and sequenced as described above.

Sequence analysis

Sequence analysis including the alignments of the sequences of the different strains was performed using the WISCONSIN PACKAGE, UNIX, v.8, Genetics Computer Group. The genomic DNA sequence from the *H.pylori* strain 26695 is accessible via http://www.tigr.org/tdb/mdb/hpdb/hpdb.html . The DNA sequences of the *hopZ* genes of all other *H.pylori* strains analyzed are accessible under the accession numbers Y18980–Y18993.

Expression and purification of HopZ and HopZ/aa126–281

The *hopZ* gene fragments encoding the putative mature proteins of the *H.pylori* strains ATCC 43504 and 5060d(A) as well as the subfragment HopZ/aa126–281 of strain ATCC 43504 were amplified from genomic DNA by PCR following standard procedures and inserted into the vector pQE30 (Qiagen) to obtain plasmids pQE30/hopZ, pQE30/hopZ5060d(A) and pQE30/hopZ/aa126–281, respectively. The plasmids were transformed in *Escherichia coli* XL1-Blue cells (Stratagene). After induction of the protein expression by IPTG, cells were disrupted and the proteins were purified from the insoluble fraction using nickel-chelate affinity chromatography according to standard protocols (Qiagen).

Western blot analysis on whole bacterial extracts

The recombinant proteins representing the mature HopZ of strain ATCC 43504 as well as the subfragment HopZ/aa126–281 were used to raise antibodies in rabbits following standard procedures. Antibodies directed against *E.coli* components were removed according to the method of Gruber and Zingales (11). *Helicobacter pylori* cells of the strains indicated above were grown for 48 h in Müller–Hinton broth containing 0.1% β-cyclodextrin, harvested by centrifugation and dissolved in Tris-glycin SDS sample buffer (Novex, San Diego, CA). Dilutions of the antisera were used for western blot analysis (12) on these *H.pylori* extracts employing the vectastain ABC kit (Serva, Heidelberg, Germany). *Escherichia coli* extracts containing the HopZ protein from strain 5060d(A) were used as positive control.

Immunofluorescence studies

Immunofluorescence studies on intact bacterial *H.pylori* cells, grown as described above, were performed essentially as described earlier (13). For the assay, 2×10^8 *H.pylori* cells were incubated with antisera raised against the three recombinant proteins, respectively, at a dilution of 1:160 using preimmune sera as a control.

Construction of a knockout mutant strain

In order to disrupt the *hopZ* gene the kanamycin resistance cassette of pILL600 (14) was cloned into the single *Pst*I site of the gene using plasmid pQE30/hopZ and plasmid pQE30/hopZkancas was obtained. This plasmid was used to naturally transform *H.pylori* strain ATCC 43504 cells according to the procedure described earlier (15). Cells were incubated for 24 h

under microaerophilic conditions and then transferred to Columbia agar plates containing 5% horse blood and 20 μg/ml kanamycin. After 48 h, single colonies were visible and after three more passages genomic DNA was prepared as described above. The insertion of the kanamycin resistance cassette was confirmed by PCR using different combinations of primers.

Adhesion assay

1×10^5 AGS CRL-1739 cells (human adeno gastric carcinoma) were incubated for 24 h at 37°C under 5% CO_2, prepared as described by Nilius *et al.* (16) and then co-incubated with 2×10^8 *H.pylori* cells strains ATCC 43504, ATCC 43504Δ*hopZ*, 342 and 326/01, respectively, for 2 h at 37°C under 5% CO_2. *Escherichia coli* XL1 Blue was used as negative control. The cells were fixed with 2% glutaraldehyde, 0.1 M Na-cacodylate, 3% (w/v) sucrose pH 7.2 at 4°C overnight. The cells were then stained using 0.005% (w/v) acridine orange, 0.1 M Na-cacodylate and 3% (w/v) sucrose.

RESULTS

Isolation and characterization of the *hopZ* gene from *H.pylori*

The preparation of a Sarkosyl®-insoluble membrane fraction of *H.pylori* strain ATCC 43504, its separation using SDS–PAGE and subsequent N-terminal sequencing of the bands allowed the identification of 19 proteins (data not shown). Oligo-nucleotides were deduced and used to screen a λZAP library carrying genomic DNA fragments of *H.pylori*.

One of the sequences was identical to the N-terminus of HopC, a member of the *H.pylori* outer membrane family (1). The gene encoding HopC could not be isolated from the screening approach because its nucleotide sequence differs from that of the oligonucleotide deduced by the N-terminal amino acid sequence obtained. However, this oligonucleotide hybridized with two different DNA fragments coding for sequences homologous to HopC. One was shown to contain two homologous genes arranged in tandem; these were termed *hopX* and *hopY*; the gene on the second fragment, which was termed *hopZ*, was selected for further analysis.

The homology of the deduced N-terminal amino acid sequence of HopZ to the corresponding sequences of the other members of the OMP family as well as its hydrophobicity suggested the presence of a signal sequence of 18 amino acid residues cleaved between an alanine and a glutamic acid residue. The *hopZ* gene from strain ATCC 43504 encodes a mature protein of 681 amino acid residues with a calculated molecular weight of 74.2 kDa and an isoelectric point of 8.4.

Allelic variation of *hopZ*

The *hopZ* genes of another 13 *H.pylori* strains were PCR-amplified and sequenced in order to analyze the variability of the HopZ protein. The deduced amino acid sequences are shown in Figure 1. Comparison of the amino acid sequences revealed two allelic variants of the gene. Strains ATCC 43504, 25392, 25144, 1811a and 5060d(A) carry allele I whereas pa1, pa2, ATCC 51110, ATCC 49503, G50, Tx30A, 342, CCUG 17874, 326/01 and 26695 contain allele II. The most prominent difference between both alleles is a 20 amino acid region present only in allele I. At the boundaries of this region, a significant variability is found between both alleles. The amino acid homology within these variable regions encoded by the different alleles is only in the range of 50%. In contrast, the N-terminal 100 amino acids and the C-terminal 400 amino acid residues are conserved between HopZ of all the different strains with 98–100% identity. These regions correspond to the one N-terminal and the seven C-terminal domains which are common to all members of the OMP family (3). Therefore we conclude that these sequences are conserved not only between the different proteins of the OMP family but also between different strains.

Localization of HopZ at the surface of *H.pylori*

The complete *hopZ* gene as well as the subfragment encoding amino acid residues 126–281 of strain ATCC 43504, representing the variable region, were expressed in *E.coli* and the histidine-tagged proteins were purified by nickel-chelate chromatography and used for the preparation of antisera in rabbits.

Western blot analysis of the total cellular proteins of *H.pylori* cultures using antisera raised against the two different proteins revealed a protein band, which corresponds to the expected molecular weight (Fig. 2, lane 1). Only a slight cross-reaction with other proteins of the Hop family was observed—also using antisera raised against the entire protein (data not shown)—suggesting that the major epitopes are located in the region specific for HopZ. A *hopZ* knockout mutant strain, ATCC 43504Δ*hopZ*, constructed by insertion of the kanamycin resistance cassette of *Campylobacter coli* into the *hopZ* gene of *H.pylori* strain ATCC 43504, was used as a control. Indeed the antisera did not recognize HopZ in the mutant strain (lane 2).

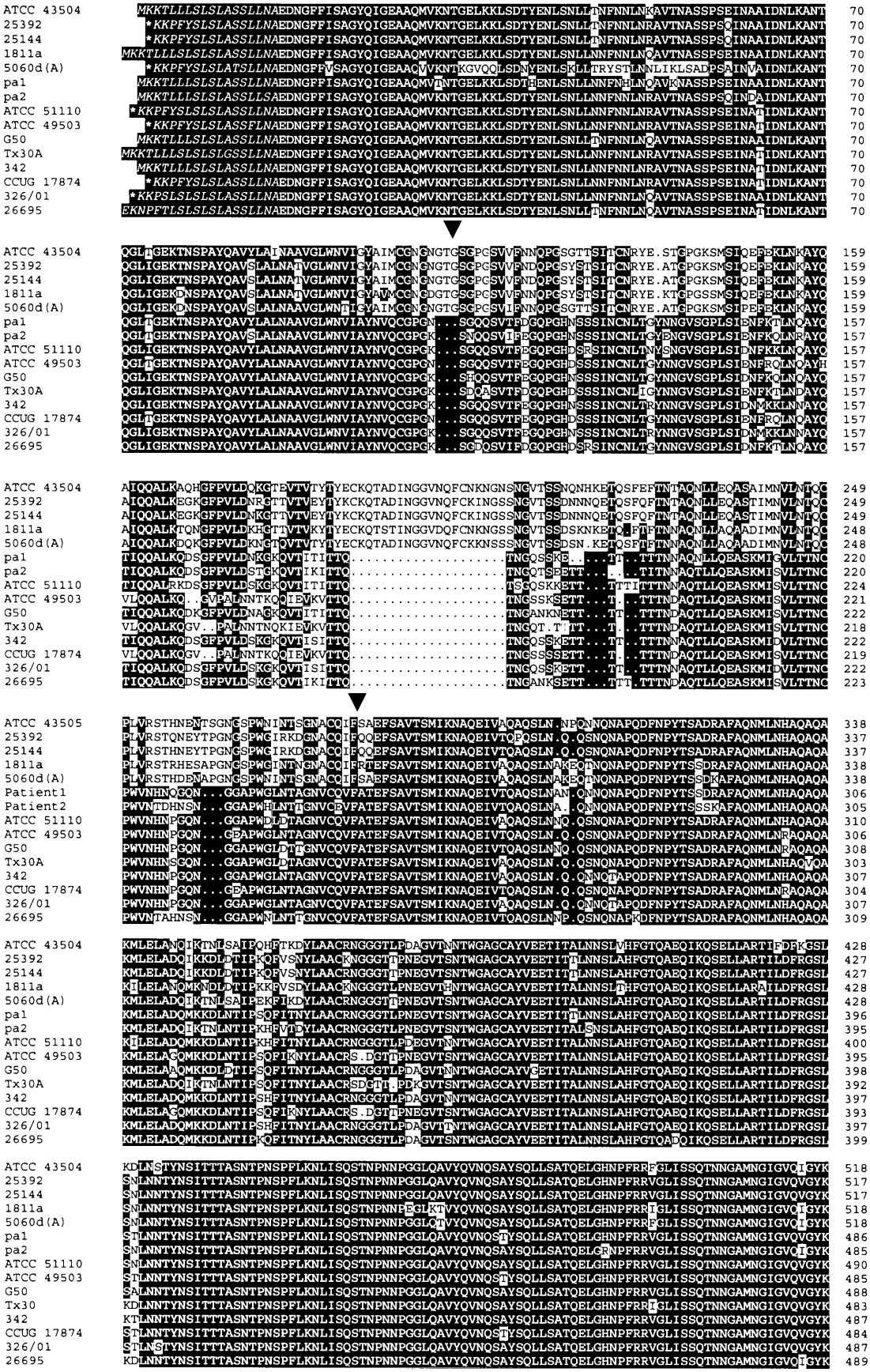
Immunofluorescence studies were performed to investigate whether HopZ is exposed on the surface of *H.pylori*. Figure 3 shows the immunofluorescence of *H.pylori* strain ATCC 43504 probed with antisera raised against the total membrane fraction (A), as well as against the recombinant HopZ protein (B) and its fragment HopZ/aa126–281 (C). The antisera raised against the complete HopZ protein and the region of amino acids 126–281 showed similar intensities and percentages of immunofluorescence positive cells (>90%) as the antiserum raised against the total membrane fraction. The mutant strain ATCC 43504Δ*hopZ* which was used as a negative control, did not show any fluorescence.

The immunofluorescence studies clearly demonstrate the surface exposition of HopZ. Particularly segments of the highly variable region spanning amino acids 126–281 of this outer membrane protein are located at the cell surface.

Genotypic variation of *hopZ* and consequences for its expression

Tomb *et al.* (3) described dinucleotide repeats located in the region encoding the signal sequences of five members of the OMP family. These repetitive sequences were assumed to generate phenotypic variation by a slipped-strand mispairing mechanism (17). Also for the *hopZ* gene (HP0009) of strain 26695, a repeat of 11 CT dinucleotides, causing a frameshift in the signal-sequence coding region, was identified. In contrast, the *hopZ* gene of strain ATCC 43504 contains only 7 CT dinucleotide repeats in frame with the residual gene. The sequencing of the *hopZ* gene of 13 more *H.pylori* strains allowed a detailed analysis of the gene status within different strains.

The DNA sequence encoding the signal peptide of the HopZ protein contains stretches of CT dinucleotide repeats which


ATCC 43504
25392
25144
1811a
5060d(A)
pa1
pa2
ATCC 51110
ATCC 49503
G50
Tx30A
342
CCUG 17874
326/01
26695
Patient1
Patient2

```
ATCC 43504   QFFGEKRKWGARYYGFFDYNHAYIKSSFFNSASDVFTYGVGTDVLYNFINDKATKNNKISFGVFGGIALAGTSWLNSQYVNLATFNNFYS  608
25392        QFFGEKRRWGLRYYGFFDYNHAYIKSSFFNSASDVFTYGVGTDVLYNFINDKATKNSKISFGVFGGIALAGTSWLNSQYVNLATFNNFYS  607
25144        QFFGEKRRWGLRYYGFFDYNHAYIKSSFFNSASDVFTYGVGTDVLYNFINDKATKNSKISFGVFGGIALAGTSWLNSQYVNLATFNNFYS  607
1811a        QFFGEKKRWGARYYGFFDYNHAYIKSSFFNSASDVFTYGVGTDVLYNFINDKATKNSKISFGVFGGIALAGTSWLNSQYVNLATFNNFYS  608
5060d(A)     QFFGEKKRWGARYYGFFDYNHAYIKSSFFNSASDVFTYGVGTDVLYNFINDKATKNSKISFGVFGGIAHAGTSWLNSQYVNLATFNNFYS  608
pa1          QFFGEKRRWGLRYYGFFDYNHAYIKSSFFNSASDVFTYGVGTDVLYNFINDKATKNNKISFGVFGGIALAGTSWLNSQYVNLATFNNFYS  576
pa2          QFFGEKRRWGLRYYGFFDYNHAYIKSSFFNSASDVFTYGVGTDVLYNFINDKATKNNKISFGVFGGIALAGTSWLNSQYVNLATFNNFYS  575
ATCC 51110   QFFGEKRRWGLRYYGFFDYNHAYIKSSFFNSASDVFTYGVGTDVLYNFINDKATKNSKISFGVFGGIALAGTSWLNSQYVNLATFNNFYS  580
ATCC 49503   QFFGEKRRWGLRYYGFFDYNHAYIKSSFFNSASDVFTYGVGTDVLYNFINDKATKNNKISFGVFGGIALAGTSWLNSQYVNLATFNNFYS  575
G50          QFFGEKRRWGLRYYGFFDYNHAYIKSSFFNSASDVFTYGVGTDVLYNFINDKATKNNKISFGVFGGIALAGTSWLNSQYVNLATFNNFYS  578
Tx30A        QFFGEKRRWGLRYYGFFDYNHAYIKSSFFNSASDVFTYGVGTDVLYNFINDKATKNNKISFGVFGGIALAGTSWLNSQYVNLATFNNFYS  572
342          QFFGEKRKWGLRYYGFFDYNHAYIKSSFFNSASDVFTYGVGTDVLYNFINDKATKNNKISFGVFGGIALAGTSWLNSQYVNLATFNNFYS  577
CCUG 17874   QFFGEKRRWGLRYYGFFDYNHAYIKSSFFNSASDVFTYGVGTDVLYNFINDKATKNNKISFGVFGGIALAGTSWLNSQYVNLATFNNFYS  573
326/01       QFFGEKRRWGLRYYGFFDYNHAYIKSSFFNSASDVFTYGVGTDVLYNFINDKATKNNKISFGVFGGIALAGTSWLNSQYVNLATFNNFYS  577
26695        QFFGEKRRWGLRYYGFFDYNHAYIKSSFFNSASDVFTYGVGTDVLYNFINDKATKNNKISFGVFGGIALAGTSWLNSQYVNLATFNNFYS  579

ATCC 43504   AKMNVANFQFLFNLGLRMNLAKNKKKASDHAAQHGVELGVKIPTINTNYYSLLGTQLEYRRLYSVYLNYVFAY*  681
25392        AKMNVANFQFLFNLGLRMNLAKNKKKASDHAAQHGVELGVKIPTINTNYYSLLGTQLEYRRLYSVYLNYVFAY*  680
25144        AKMNVANFQFLFNLGLRMNLAKNKKKASDHAAQHGVELGVKIPTINTNYYSLLGTQLEYRRLYSVYLNYVFAY*  680
1811a        AKMNVANFQFLFNLGLRMNLARNKKKASDHAAQHGMELGVKIPTINTNYYSLLGTQLQYRRLYSVYLNYVFAY*  681
5060d(A)     AKMNVANFQFLFNLGLRMNLARNKKKASDHAAQHGMELGVKIPTINTNYYSLLGTQLQYRRLYSVYLNYVFAY*  681
pa1          AKMNVANFQFLFNLGLRMNLAKNKKKASDHVAQHGVELGVKIPTINTNYYSLLGTQLEYRRLYSVYLNYVFAY*  649
pa2          AKMNVANFQFLFNLGLRMNLAKNKKKASDHAAQHGVELGVKIPTINTNYYSLLGAQLQYRRLYSVYLNYVFAY*  648
ATCC 51110   AKMNVANFQFLFNLGLRMNLAKNKKKASDHAAQHGVELGVKIPTINTNYYSLLGTQLQYRRLYSVYLNYVFAY*  653
ATCC 49503   AKMNVANFQFLFNLGLRMNLAKNKKKASDHAAQHGVELGVKIPTINTNYYSLLGTQLQYRRLYSVYLNYVFAY*  647
G50          AKMNVANFQFLFNLGLRMNLAKNKKKASDHVAQHGVELGVKIPTINTNYYSLLGTQLQYRRLYSVYLNYVFAY*  651
Tx30         AKMNVANFQFLFNLGLRMNLAKNKKKASDHAAQHGVELGVKIPTINTNYYSLLGTQLQYRRLYSVYLNYVFAY*  645
342          AKMNVANFQFLFNLGLRMNLAKNKKKASDHAAQHGVELGVKIPTINTNYYSLLGTQLQYRRLYSVYLNYVFAY*  650
CCUG 17874   AKMNVANFQFLFNLGLRMNLAKNKKKASDHAAQHGVELGVKIPTINTNYYSLLGTQLQYRRLYSVYLNYVFAY*  646
326/01       AKMNVANFQFLFNLGLRMNLAKNKKKASDHAAQHGVELGVKIPTINTNYYSLLGTQLQYRRLYSVYLNYVFAY*  650
26695        AKMNVANFQFLFNLGLRMNLAKNKKKASDHVAQHGVELGVKIPTINTNYYSLLGTQLQYRRLYSVYLNYVFAY*  652
```

Figure 1. (Opposite and above) Comparison of amino acid sequences of HopZ of 15 different *H.pylori* strains. The putative signal peptides are indicated in italics. Asterisks denote the stop codons; note that the *hopZ* genes of the strains 25392, 25144, 5060d(A), ATCC 51110, ATCC 49503, CCUG 17874, 326/01 and 26695 do not encode an in frame start methionine. Amino acid positions of the mature proteins are indicated on the right side of the figure. Amino acid residues shown to be conserved between the 15 different strains of *H.pylori* are shown in white letters on black background. The amino acid residues which differ from the consensus sequence are given in black letters. The positions of the amino residues 126–281, which were expressed in *E.coli*, are indicated by arrows.

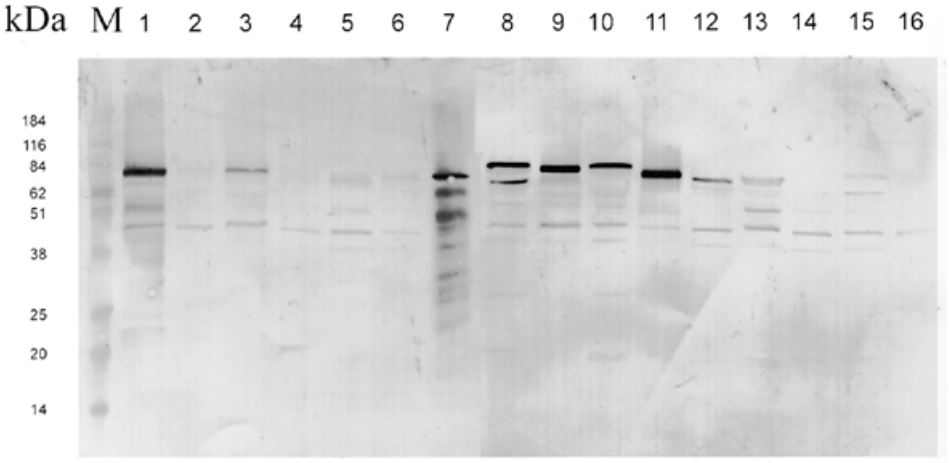


Figure 2. Western blot analysis of the bacterial extracts of the *H.pylori* strains ATCC 43504 (lane 1), 1811a (lane 3), 25392 (lane 4), 25144 (lane 5), 5060d(A) (lane 6), Tx30A (lane 8), pa1 (lane 9), pa2 (lane 10), G50 (lane 11), 342 (lane 12), 326/01 (lane 13), ATCC 49503 (lane 14), ATCC 51110 (lane 15) and CCUG 17874 (lane 16), using an antiserum raised against a recombinant fragment HopZ/aa126–281 of strain ATCC 43504. Lane 2, extract of mutant strain ATCC 43504Δ*hopZ* (negative control); lane 7, extract of *E.coli* expressing HopZ of strain 5060d(A) (positive control).

differ in their number between the 15 different *H.pylori* strains analyzed (Table 1). The presence of 7 or 10 CT dinucleotides in the signal-peptide coding region allows the expression of an intact *hopZ* translation product. A deletion or insertion of one or two dinucleotides in this repeat region leads to a frameshift mutation within the *hopZ* gene. Therefore strains carrying 6 (5060d(A), ATCC 49503, 25144, 25392, CCUG 17874), 9 (ATCC 51110), 11 (26695) or 12 (326/01) CT dinucleotide repeats are not expected to express an intact HopZ protein (Table 1).

Table 1. Signal-sequence coding region of the *hopZ* gene of different *H.pylori* strains

H. pylori strain	sequence of the signal-peptide coding region	number of CT repeats	gene status	allele
5060d(A)	ATGAAAAAAACCCTTTTACTCTCTCTCTCT.....CGCTACATCGCTTTTAAACGCTGA	6	off	I
	M K K T L L L S L S R Y I A F K R *			
ATCC 49503	ATGAAAAAAACCCTTTTACTCTCTCTCTCT.....CGCTTCATCGTTTTTAAACGCTGA	6	off	II
	M K K T L L L S L S R F I V F K R *			
25144	ATGAAAAAAACCCTTTTACTCTCTCTCTCT.....CGCTTCATCGCTTTTAAACGCTGA	6	off	I
	M K K T L L L S L S R F I A F K R *			
25392	ATGAAAAAAACCCTTTTACTCTCTCTCTCT.....CGCTTCATCGCTTTTAAACGCTGA	6	off	I
	M K K T L L L S L S R F I A F K R *			
CCUG 17874	ATGAAAAAAACCCTTTTACTCTCTCTCTCT.....CGCTTCATCGCTTTTAAACGCTGA	6	off	II
	M K K T L L L S L S R F I A F K R *			
ATCC 43504	ATGAAAAAAACCCTTTTACTCTCTCTCTCTCT.....CGCTTCATCGCTTTTAAACGCT	7	on	I
	M K K T L L L S L S L A S S L L N A			
G50	ATGAAAAAAACCCTTTTACTCTCTCTCTCTCT.....CGCTTCATCGCTTTTAAACGCT	7	on	II
	M K K T L L L S L S L A S S L L N A			
pa1	ATGAAAAAAACCCTTTTACTCTCTCTCTCTCT.....CGCTTCATCGCTTTTAAACGCT	7	on	II
	M K K T L L L S L S L A S S L L N A			
pa2	ATGAAAAAAACCCTTTTACTCTCTCTCTCTCT.....CGCTTCATCGCTTTTAAGCGCT	7	on	II
	M K K T L L L S L S L A S S L L S A			
342	ATGAAAAAAACCCTTTTACTCTCTCTCTCTCT.....CGCTTCATCGCTTTTAAACGCT	7	on	II
	M K K T L L L S L S L A S S L L N A			
ATCC 51110	ATGAAAAAAACCCTTTTACTCTCTCTCTCTCTCTCT.....CGCTTCATCGCTTTTAAACGCTGA	9	off	II
	M K K T L L L S L S L S R F I A F K R *			
1811a	ATGAAAAAAACCCTTTTACTCTCTCTCTCTCTCTCT....CGCTTCATCGCTTTTAAACGCT	10	on	I
	M K K T L L L S L S L S L A S S L L N A			
Tx30A	ATGAAAAAAACCCTTTTACTCTCTCTCTCTCTCTCT....CGGTTCATCGCTTTTAAACGCT	10	on	II
	M K K T L L L S L S L S L G S S L L N A			
26695	ATGAAAAAAACCCTTTTACTCTCTCTCTCTCTCTCT..CGCTTCATCGCTTTTAA	11	off	II
	M K K T L L L S L S L S L S L H R F *			
326/01	ATGAAAAAAACC.....CTCTCTCTCTCTCTCTCTCTCTCTCGCTTCATCGCTTTTAAACGCTGA	12	off	II
	M K K T L S L S L S L S R F I A F K R *			

Nucleotide and deduced amino acid sequences of the signal peptide-coding region of the *hopZ* gene of 15 different *H.pylori* strains. Asterisks denote the stop codons of the *hopZ* gene of the strains whose translational start at the designated methionine residue leads to a truncated product. The numbers of the CT dinucleotide repeats, the gene status and the allele type are given on the right.

To verify the role of the slipped-strand mispairing mechanism in the regulation of expression of the *hopZ* gene, western blot analysis of bacterial extracts as well as immunofluorescence studies of different *H.pylori* strains were performed using the antiserum raised against the specific subfragment HopZ/ aa126–281 of strain ATCC 43504. Western blot analysis showed that the antiserum raised against a subfragment of the allele I strain ATCC 43504 reacts with HopZ protein from the allele I and II type (Fig. 2). The HopZ-specific protein band was detected only in the allele I strains ATCC 43504 and 1811a and the allele II strains Tx30A, pa1, pa2, G50 and 342. In contrast, no HopZ-specific protein band was found for the mutant strain ATCC 43504Δ*hopZ*, for the allele I strains 25392, 25144 and 5060d(A) and for the allele II strains ATCC 49503 and CCUG 17874. The allele II strains 326/01 and ATCC 51110 show a weak reaction with a protein of ~70 kDa which is not assumed to be HopZ.

To confirm the results of the western blot analysis, immunofluorescence studies were performed (Fig. 3). No fluorescence could be detected for the strains 326/01 and ATCC 51110 as well as for the other strains which show a frameshift mutation in the signal-sequence coding region. Immunofluorescence indicating HopZ expression was demonstrated only for those strains found to be clearly positive in western blot analysis and exhibiting no frameshift mutation. Therefore the slipped-strand mispairing mechanism in the signal-sequence coding region is suggested to be responsible to switch on or off the expression of HopZ. RT–PCR analysis showed the presence of *hopZ*-specific mRNA in all 14 *H.pylori* strains analyzed (data not shown).

Function of the protein HopZ

The adherence of the wild-type strain ATCC 43504 and the mutant strain ATCC 43504Δ*hopZ* to human gastric epithelial cells was compared. Furthermore the *H.pylori* strains 342 and 326/01 were tested. The wildtype strain strongly adheres to the epithelial cells whereas only a few cells of the mutant strain show adhesion. Additionally strain 342 expressing *hopZ* is

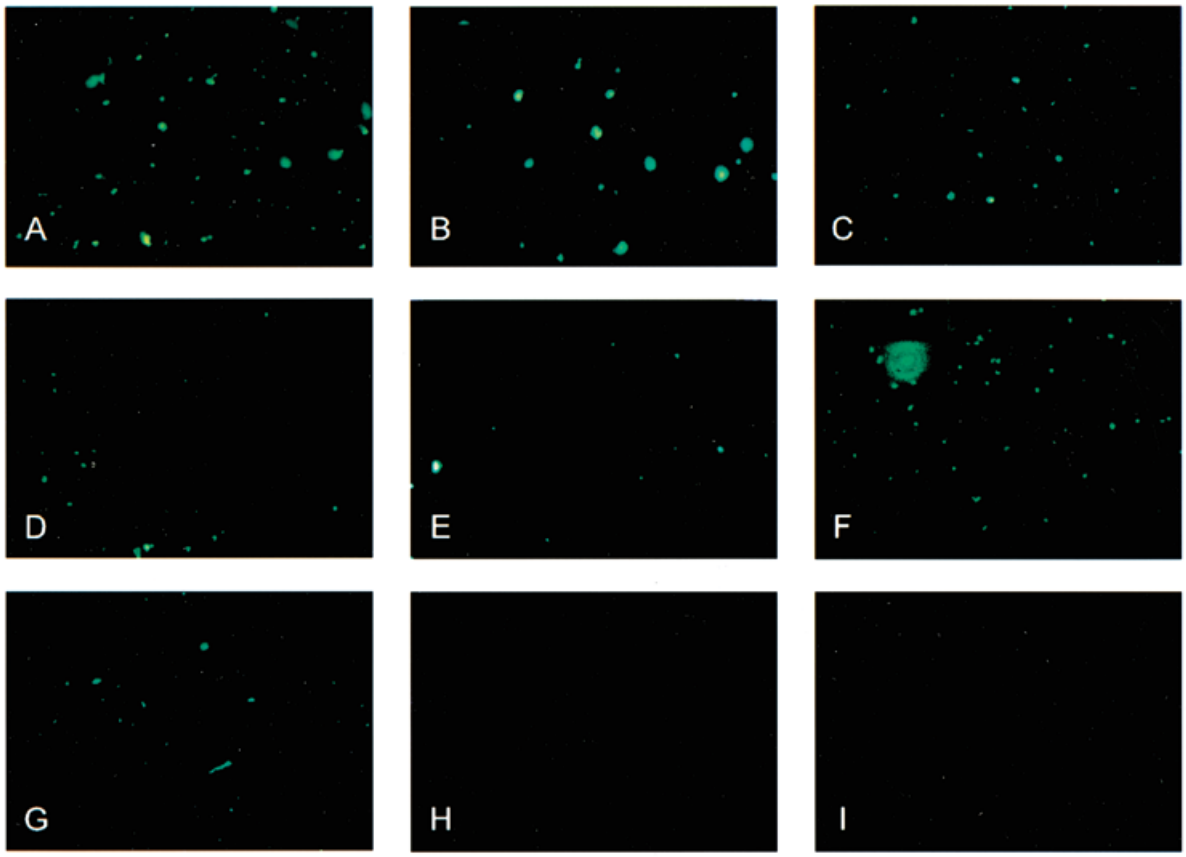


Figure 3. Immunofluorescence studies on intact *H.pylori* cells. Strain ATCC 43504 was probed with the antiserum raised against the whole membrane fraction (**A**), the mature HopZ protein (**B**), the subfragment amino acids 126–281 (**C**). Strains, Tx30 (**D**), 1811a (**E**), pa2 (**F**), pa1(**G**), ATCC 43504Δ*hopZ* (**H**) and 326/01 (**I**) were analyzed using the antiserum directed against the subfragment amino acids 126–281.

able to adhere to the epithelial cells whereas strain 326/01 lacking HopZ adheres to a lower extent (Fig. 4). These data demonstrate that the lack of HopZ results in a significant reduction of adherence. Therefore we conclude that HopZ is involved in adhesion.

DISCUSSION

A further member of the OMP family of *H.pylori*, the protein HopZ, was identified and characterized with respect to its conservation, localization, expression and function.

Hancock *et al*. (18) described an amino acid identity value of >90% between the members of the OMP family of two different *H.pylori* strains. Indeed, the proteins HopV, HopW, HopX and HopY of the *H.pylori* strain ATCC 43504 are highly similar to proteins predicted from the sequenced genome of strain 26695 (3) encoded by the open reading frames HP1469, HP1501, HP0252 and HP0254, respectively, with identity values of 96.3, 94.5, 97.2 and 98.4%. We could confirm the high conservation of these proteins in 13 further strains (manuscript in preparation). In contrast, the HopZ protein analyzed here shows much more variability: e.g. the identity between the protein of strain ATCC 43504 and that of strain 26695 is only 79.1%. The sequence analysis of another 13 different strains shows that HopZ is encoded by two alleles with prominent differences in the region located between the 100 N-terminal and the 400 C-terminal highly conserved amino acids. The corresponding region was also described to be variable between the different members of the OMP family of strain 26695 (3). The HopZ protein is the first example within the OMP family for which a significant variability was found between different *H.pylori* strains.

Immunofluorescence studies using antisera raised against the total HopZ protein as well as against the HopZ-specific region of amino acids 126–281 demonstrated the presence of HopZ on the surface of intact *H.pylori* cells. Thus HopZ is indeed an outer membrane protein and especially the variable region, amino acids 126–281, is exposed on the cell surface.

The expression of the *hopZ* gene is regulated on the level of translation by a CT dinucleotide repeat motif of different length in different strains and localized at the 5' end of the coding region. An analogous control mechanism was described for a gene of *Haemophilus influenzae* encoding a lipopolysaccharide (LPS) synthetase involved in the phenotypic variation of surface LPS (17). This gene can carry a different number of tandem repeat units at its 5' end, which is the result of a slipped-strand mispairing

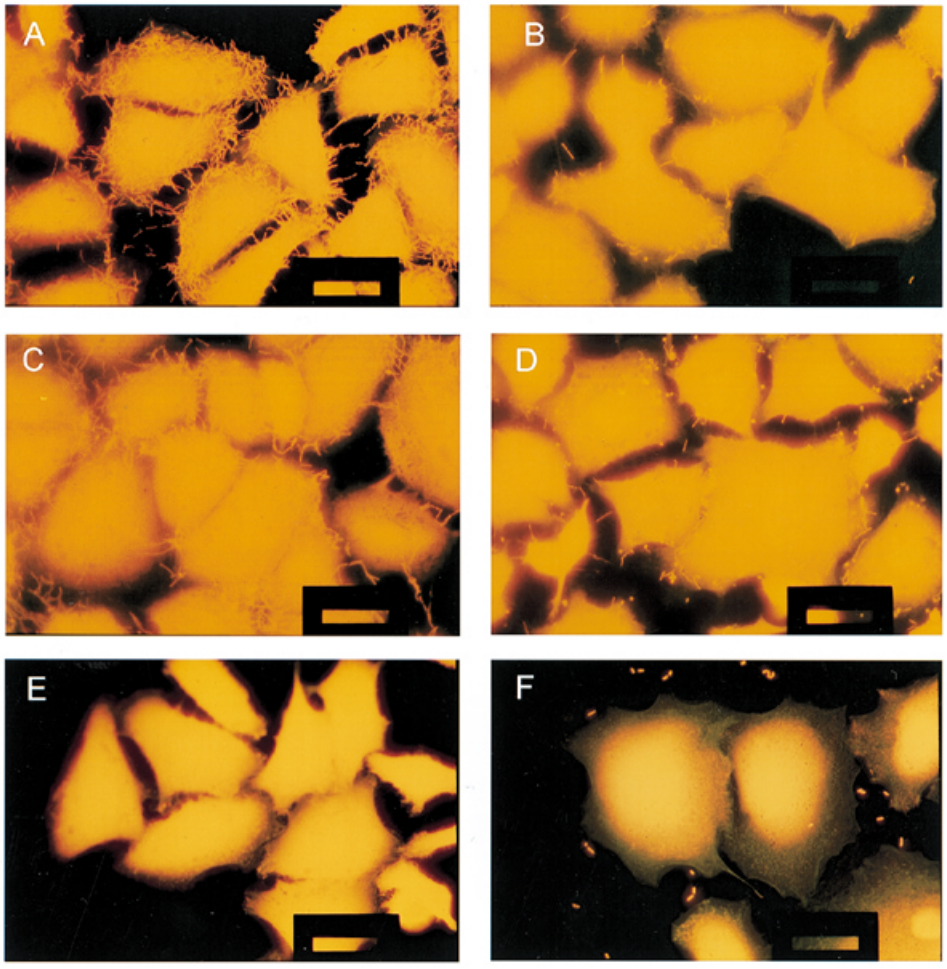


Figure 4. Adherence of *H.pylori* strains ATCC 43504 (**A**), ATCC 43504Δ*hopZ* (**B**), 342 (**C**) and 326/01 (**D**) to AGS CRL-1739 cells (human gastric epithelial cells). Cells without bacteria (**E**) and cells incubated with *E.coli* XL1 Blue (**F**) were used as controls.

mechanism. Depending on the number of repeats, the gene can either be translated in its full length or the translation product is truncated. Tomb *et al.* (3) described eight genes from *H.pylori* which contain dinucleotide repeats in their signal-sequence coding regions. Genotypic variation caused by a slipped-strand mispairing mechanism within these repeats was postulated to switch on or off the expression of these genes; however, it has not been demonstrated. Appelmelk *et al*. (19) showed that two *H.pylori* genes encoding fucosyltransferases can be switched on or off by slipped-strand mispairing within C-tracts in their 5' coding regions.

Slipped-strand mispairing events within the repetitive sequences of the *hopZ* gene of *H.pylori* strains would allow the OMP expression to be readily switched on or off. Western blot analysis on *H.pylori* extracts from different strains as well as immunfluorescence studies demonstrated that HopZ is expressed only in those strains in which the *hopZ* gene is not interrupted in its reading frame, although *hopZ*-specific mRNA could be detected also in those strains which do not express the HopZ protein. Therefore the expression of the *hopZ* gene seems to be regulated on the translational level. However, the fact that the bacteria are able to abandon *hopZ* expression revealed that this protein is not essential for *H.pylori* at least *in vitro*.

Recently, five porins have been characterized among the members of the OMP family of *H.pylori* (1,2). However, we have evidence that HopZ is not able to form pores in bilayer experiments, suggesting that HopZ does not function as a porin (data not shown). A further member of the *H.pylori* OMP family, BabA, was characterized as an adhesin, binding to the histo-blood group antigen Lewis B (4). Moreover the adhesins AlpA and AlpB have been identified (5). We have shown that HopZ is also involved in adhesion of *H.pylori*. This was demonstrated by comparing the adherence of the wild-type strain ATCC 43504 and the mutant strain ATCC 43504Δ*hopZ* to human epithelial cells. The data suggest that HopZ functions as a major adhesin because knocking out *hopZ* gene results in significant decrease of adhesive bacteria. Furthermore, strain 342 expressing *hopZ* adhered to the epithelial cells whereas strain 326/01 lacking HopZ showed reduction in the number of

adhesive bacteria. However, some bacteria of the mutant strain as well as strain 326/01 are able to bind the epithelial cells. This adhesion could be mediated by other adhesins of *H.pylori*. Thus HopZ is the fourth protein of the OMP family which is involved in adhesion of *H.pylori* to gastric endothelial cells.

ACKNOWLEDGEMENTS

We are very grateful to A. Labigne, Institute Pasteur, Paris, for providing the plasmid pILL600 and U. Nau, Chiron-Behring GmbH & Co. for the preparation of the DNA library. We thank K.-H. Schmidt, Hoechst Marion Roussell, for protein sequencing, T. Weimer and C. Nerlich, Centeon GmbH, for oligonucleotide synthesis and DNA sequencing and E. Spangenberg-Beyer for technical help in the adhesion assay. We are also grateful to M. Dietrich and G. Dietrich for critical reading of the manuscript.

REFERENCES

1. Exner,M.M., Doig,P., Trust,T.J. and Hancock,R.E.W. (1995) *Infect. Immun.*, **63**, 1567–1572.
2. Doig,P., Exner,M.M., Hancock,R.E.W. and Trust,T.J. (1995) *J. Bacteriol.*, **177**, 5447–5452.
3. Tomb,J.-F.,White,O., Kerlavage,A.R., Clayton,R.A., Sutton,G.G., Fleischmann,R.D., Ketchum,K.A., Klenk,H.P., Gill,S., Dougherty,B.A., Nelson,K., Quackenbush,J., Zhou,L., Kirkness,E.F., Peterson,S., Loftus,B., Richardson,D., Dodson,R., Khalak,H.G., Glodek,A., McKenney,K., Fitzegerald,L.M., Lee,N., Adams,M.D., Venter,J.C *et al.* (1997) *Nature*, **388**, 539–547.
4. Ilver,D., Arnqvist,A., Ögren,J., Frick,I.-M., Kersulyte,D., Incecik,E.T., Berg,D.E., Covacci,A., Engstrand,L. and Borén,T. (1998) *Science*, **279**, 373–377.
5. Odenbreit,S., Till,M., Hofreuter,D., Faller,G. and Haas,R. (1999) *Mol. Microbiol.*, **31**, 1537–1548.
6. Blaser,M.J., Hopkins,J.A., Berka,R.M., Vasil,M.L. and Wang,W.-L.L. (1983) *Infect. Immunol.*, **42**, 276–284.
7. Matsudaira,P.J. (1987) *J. Biol. Chem.*, **262**, 10035–10038.
8. Peck,B. (1997) Diploma Thesis, Fachbereich Chemie der Philipps Universität Marburg, Marburg, Germany.
9. Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
10. Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl Acad. Sci. USA*, **74**, 5463–5467.
11. Gruber,A. and Zingales,B. (1995) *Biotechniques*, **19**, 28–30.
12. Johnson,D.A., Gautsch,J.W., Sportsman,J.R. and Edler,J.H. (1984) *Gene Anal. Technol.*, **1**, 3–8.
13. Schorr,J., Knapp,B., Hundt,E., Küpper,H.A. and Amann,E. (1991) *Vaccine*, **9**, 675–681.
14. Labigne,A., Courcoux,P. and Tompkins,L. (1992) *Res. Microbiol.*, **143**, 15–26.
15. Doig,P., O'Toole,P.W. and Trust,J.T. (1996) In Clayton,C.L. and Mobley,H.L.T. (eds), *Helicobacter pylori Protocols. Molecular Characterization of H.pylori Surface Antigens.* Humana Press, Totowa, NJ, pp. 177–189.
16. Nilius,M., Bode,G., Büchler,M. and Malfertheiner,P. (1994) *Eur. J. Clin. Invest.*, **24**, 454–459.
17. Moxon,E.R., Rainey,P.B., Nowak,M.A. and Lenski,R.E. (1994) *Curr. Biol.*, **4**, 24–33
18. Hancock,R.E.W., Alm,R., Bina,J. and Trust,T. (1998) *Nature Biotechnol.*, **16**, 216–217.
19. Appelmelk,B.J., Martin,S.L., Kusters,J.G., van denEijnden,D.H. and Vandenbrouke-Grauls,C.M.J.E. (1998) *Gut*, **43**, A13.