JOURNAL OF BACTERIOLOGY, Jan. 1991, p. 505–513
0021-9193/91/020505-09$02.00/0
Copyright © 1991, American Society for Microbiology
Vol. 173, No. 2

Isolation and Biochemical and Molecular Analyses of a Species-Specific Protein Antigen from the Gastric Pathogen *Helicobacter pylori*

PAUL W. O'TOOLE,[1]* SUSAN M. LOGAN,[2] MAGDALENA KOSTRZYNSKA,[2] TORKEL WADSTRÖM,[1] AND TREVOR J. TRUST[2]

Department of Medical Microbiology, University of Lund, Sölvegatan 23, S-223 62 Lund, Sweden,[1] and Department of Biochemistry and Microbiology, University of Victoria, Victoria, British Columbia V8W 3P6, Canada[2]

Received 25 May 1990/Accepted 22 October 1990

A protein of M_r 26,000 which was present in large quantities in extracts of cells of *Helicobacter pylori* was purified to homogeneity by ammonium sulfate precipitation followed by gel filtration and reversed-phase chromatography or anion-exchange chromatography. The protein appeared to be associated with the soluble fraction of the cell, and antibodies raised against the protein were reactive with whole-cell lysates of a variety of *H. pylori* strains in a simple immunodot blot assay. This reaction was species specific. Protein sequence determination of the amino terminus and internal cyanogen bromide fragments and amino acid composition analysis were performed. An oligonucleotide derived from these data was used to clone a fragment encoding most of the coding sequence. Expression in *Escherichia coli* was dependent on vector promoters. The DNA sequence of the fragment was determined. DNA probes derived from the cloned fragment hybridized to genomic DNA of all *H. pylori* strains tested, but not to DNAs of *Helicobacter mustelae*, *Wolinella succinogenes*, various *Campylobacter* species, and a panel of gram-negative enteric bacteria. The apparent uniqueness of this protein may be exploited for the development of species-specific diagnostics for this gastric pathogen.

Helicobacter (*Campylobacter*) *pylori* (11) is a curved or spiral gram-negative microaerophilic bacterium which was first isolated from a human gastric biopsy specimen in 1983 (46). Since this first isolation, it has become apparent that this organism may be one of the most common bacterial pathogens of humans. Epidemiological evidence has shown that *H. pylori* colonizes the upper gastrointestinal tract of more than one in two individuals during their lifespan; in many of these persons, the organism is associated with disease of the gastrointestinal tract (12, 13, 22). Indeed, *H. pylori* appears to be causally associated with active and chronic gastritis as well as with peptic and duodenal ulcers (4, 5, 7, 12, 22) and may also be associated with carcinoma of the stomach (17, 21). Additional evidence for the pathogenic activity of *H. pylori* has been provided by studies with gnotobiotic and barrier-born pigs (18, 20) and studies with two human volunteers (29, 35).

Because of the clinical importance of this pathogen and the large number of laboratory identifications being routinely undertaken around the world on a daily basis, there is a need for the development of rapid diagnostic test protocols. Liberation of ^{14}C from labeled urea by the potent bacterial urease is reliable for detection of an active *H. pylori* infection (30, 33) but has disadvantages in cost, time, and exposure to undesirable radiation. Cell extracts or purified cell-surface components have been evaluated as antigens in the development of serological tests for infection (9, 36). DNA-based approaches have included restriction endonuclease analysis for typing (28, 42) and the development of specific probes (6, 34). We have identified a protein that appears to be unique to *H. pylori*. We have isolated this protein and characterized it biochemically. In this paper we report on the biochemical characteristics of the protein and show that antibodies to the protein can be used to serologically identify *H. pylori*. We have also cloned and sequenced part of the gene coding for this protein. This is among the first gene sequences reported for *H. pylori*, and we also present evidence that this sequence may provide the basis for a species-specific DNA-based identification test.

MATERIALS AND METHODS

Bacterial strains. Bacterial strains used in this study are listed in Table 1. *H. pylori* strains were grown on chocolate blood agar (CBA) plates containing 0.001% vancomycin, 0.005% trimethoprim, and 0.00075% polymyxin in an atmosphere containing 10% carbon dioxide. Cultures of *Campylobacter fetus*, *Campylobacter jejuni*, *Campylobacter coli*, and *Campylobacter laridis* were also grown in this atmosphere on CBA. *Campylobacter cinaedi*, *Campylobacter fennelliae*, and *Helicobacter mustelae* strains were cultured on Diagnostic Sensitivity Test agar supplemented with 7% horse blood in a microaerophilic environment (83% nitrogen, 2% hydrogen, 5% oxygen, 10% carbon dioxide). *Wolinella succinogenes* was cultured anaerobically on gonococcal agar base (BBL Microbiology Systems, Cockeysville, Md.) incorporating 0.05% (wt/vol) cyteine hydrochloride, 8.5% hematinized horse blood, 10% heat-inactivated horse serum, and 0.35% IsoVitaleX (BBL). All other strains were grown on Trypticase soy agar (BBL). *Escherichia coli* TB1 (48) was used in cloning experiments as the host for recombinant plasmids. All cultures were grown at 37°C. Stock cultures were maintained at −70°C in 15% (vol/vol) glycerol–Trypticase soy broth.

Isolation and purification of antigen. Cells were harvested from 12 culture plates (85 mm, vented) that had been heavily inoculated with *H. pylori* 915 and incubated for 60 h. The

* Corresponding author.

TABLE 1. Bacterial strains

Strain	Source or reference
Helicobacter pylori	
17874	CCUG[a], Sweden (NCTC 11637)
915	Clinical isolate, Lund University Hospital
52	Clinical isolate, Lund University Hospital
66	Clinical isolate, Lund University Hospital
95	Clinical isolate, Lund University Hospital
253	Clinical isolate, Lund University Hospital
1139	Clinical isolate, Lund University Hospital
5155	A. Lee, University South Wales, Australia
5294	A. Lee, University South Wales, Australia
5442	A. Lee, University South Wales, Australia
5646	H. Lior, NERC,[b] Ottawa, Canada
5650	H. Lior, NERC, Ottawa, Canada
5790	H. Lior, NERC, Ottawa, Canada
Helicobacter mustelae	
23950	CCUG, Sweden
23951	CCUG, Sweden
25715	CCUG, Sweden
Campylobacter fetus	
VC1	University of Victoria, Canada
VC78	University of Victoria, Canada
VC202	University of Victoria, Canada
Campylobacter coli	
VC143	University of Victoria, Canada
VC146	University of Victoria, Canada
VC212	University of Victoria, Canada
Campylobacter jejuni	
VC26	University of Victoria, Canada
VC27	University of Victoria, Canada
VC28	University of Victoria, Canada
VC36	University of Victoria, Canada
VC74	University of Victoria, Canada
VC209	University of Victoria, Canada
Campylobacter laridis	
VC81	University of Victoria, Canada
VC213	University of Victoria, Canada
VC214	University of Victoria, Canada
Campylobacter cinaedi	
18818	CCUG, Sweden
18819	CCUG, Sweden
19218	CCUG, Sweden
Campylobacter fennelliae	
18820	CCUG, Sweden
Wolinella succinogenes	
12550	CCUG, Sweden
13145	CCUG, Sweden
Escherichia coli	
1593	University of Victoria, Canada
E1049-a-13	University of Victoria, Canada
E438	University of Victoria, Canada
TB1	48
Salmonella enteritidis 2b.1	University of Victoria, Canada
Salmonella illinois S.1093	University of Victoria, Canada
Shigella flexneri Sh1a1	University of Victoria, Canada

[a] CCUG, Culture Collection, University of Gothenburg, Sweden.
[b] NERC, National Enteric Reference Centre, Ottawa, Canada.

cells (0.8 g [wet weight]) were washed in phosphate-buffered saline (pH 7.2) and gently resuspended in 0.6% (wt/vol) octyl-β-D-glucopyranoside (octylglucoside; Serva Fine Chemicals, Heidelberg, Federal Republic of Germany). Extraction of surface proteins was allowed to take place for 60 min at the ambient temperature with gentle agitation. Cells were removed by centrifugation at 6,000 × *g* for 15 min and then reextracted. The pooled supernatants were cleared by centrifugation at 35,000 × *g* for 20 min, and ammonium sulfate added to 50% saturation. After overnight incubation at 4°C, the sample was spun at 35,000 × *g* for 30 min, and the pellet was suspended in 3 ml of 0.05 M Tris hydrochloride containing 0.145 M NaCl (Tris-saline [pH 7.5]). After dialysis against the same buffer, the sample was crudely fractionated by molecular sieving on a 70- by 1.5-cm Sephadex G100 column (Pharmacia, Uppsala, Sweden), equilibrated, and run with Tris-saline buffer. Fractions were examined by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) under reducing conditions, and those containing a subunit M_r-26,000 protein were pooled, lyophilized, and suspended in Tris-saline. Purification was achieved by using a ProRPC HR5/10 column (Pharmacia). Chromatography conditions were as follows. Solvent A was 0.1% trifluoroacetic acid in high-pressure liquid chromatography quality water and solvent B was 100% acetonitrile. The flow rate was 0.7 ml/min, and chromotography was performed at room temperature. Elution of proteins was achieved by using a gradient of 0 to 60% acetonitrile in 60 min.

As an alternative to reversed-phase liquid chromatography (RPLC), pooled Sepharose G100 fractions containing the M_r-26,000 protein were subjected to ion-exchange chromatography on a MonoQ column (Pharmacia) that had been preequilibrated with 0.05 M Tris hydrochloride (pH 7.5). After application of the sample, the column was washed with the equilibration buffer and developed with a gradient from 0 to 1.0 M NaCl in the same buffer. The peak eluting around 0.14 M NaCl contained the protein of interest.

Purification of peptides. Cyanogen bromide (CNBr) cleavage of purified proteins was performed as previously described (26). Peptides were purified by RPLC as described for the intact polypeptide above. Peptides purified in this manner were dried under vacuum on a Speed-Vac (Savant Instruments Inc., Hicksville, N.Y.) and stored at −20°C.

Antibodies. Polyclonal antiserum was raised in adult New Zealand White rabbits by intramuscular injections. The rabbits were immunized with approximately 50 μg of fast-protein liquid chromatography-purified protein in Freund complete adjuvant. Booster doses were given in Freund incomplete adjuvant on days 14 and 28. On day 42 the rabbits were exsanguinated, and the serum was collected and stored at −20°C. Control nonimmune serum was obtained before the first injection.

Electrophoresis. SDS-PAGE was performed by the

method of Laemmli (19) in a mini-slab apparatus (Bio-Rad Laboratories, Richmond, Calif.). Protein solubilized in sample buffer was stacked in a 4.5% acrylamide gel (100 V, constant voltage) and separated by using 12.5% acrylamide (200 V). Protein was stained with Coomassie brilliant blue R-250. Standard conditions for agarose gel electrophoresis of DNA (41) were employed.

When required, separated proteins or peptides were transferred from the slab to nitrocellulose paper (NCP) by the methanol-Tris-glycine system of Towbin et al. (45). Electroblotting was carried out in a Bio-Rad transblot apparatus for 18 h at 60 V. Nondenaturing isoelectric focusing gels were run in a mini-slab gel apparatus by the method of Robertson et al. (39) and were stained with Coomassie blue.

Amino acid composition analysis. The purified protein was dialyzed extensively against distilled water, lyophilized, and then hydrolyzed in 6 N HCl at 100°C for 18 h. Amino acid composition was determined on a Beckman 119CL amino acid analyzer. The method used was as described by the manufacturer for a 90-min single-column procedure.

N-terminal sequence analysis. Amino acid analysis was performed on an Applied Biosystems model 470A gas-phase sequenator (Applied Biosystems, Foster City, Calif.) running a standard operation program. Phenylthiohydantoin derivatives were separated on an IBM cyanocolumn (4.6 by 250 mm) as described by Hunkapiller and Hood (16) fitted to a Beckman model 32 high-pressure liquid chromatograph equipped with a Hewlett-Packard integrator (Hewlett-Packard Co., Palo Alto, Calif.).

Western immunoblotting. After electroblotting, unreacted sites on the NCP were blocked with a 1% solution of gelatin in 10 mM Tris hydrochloride–0.9% NaCl (pH 7.4) (GTS) for 1 h at room temperature. The NCP was then incubated with an appropriate dilution of antiserum in the same buffer for 2 h. The NCP was washed five times with GTS. Goat anti-rabbit immunoglobulin conjugated to alkaline phosphatase (Caltag Labs) was then added in GTS buffer and incubated for 1 h at room temperature. After incubation, the NCP was washed five times in GTS. The reactive bands were visualized as described by Blake et al. (3) with 5-bromo-4-chloro-3-indolyl phosphate (Boehringer Mannheim, Federal Republic of Germany) as the alkaline phosphate substrate and Nitro Blue Tetrazolium (Sigma Chemical Co., St. Louis, Mo.) as the color development reagent.

Immunodot blotting. Dot blotting was performed by transferring 10 μg of whole-cell lysate of each strain to be tested to NCP. After the NCP was dried at 37°C for 1 h, it was blocked with GTS, treated with antiserum, washed, and developed by the procedure described for Western blotting.

Cell fractionation. Glycine extraction of 48-h *Helicobacter* cells was performed by the method of McCoy et al. (31). The periplasmic fraction was prepared by the osmotic shock procedure of Willis et al. (47); cell envelopes and membranes were prepared by differential centrifugation as previously described (25). To prepare the cytoplasmic fraction, cell suspensions were passed through a French press at 16,000 lb/in^2, and intact cells were removed by two centrifugations at 3,000 × *g* for 30 min. The cell envelope was removed by centrifugation at 43,000 × *g* for 30 min, and the supernatant used as the cytoplasmic fraction. The outer membrane fraction was isolated from cell envelopes by the sodium lauryl sarcosinate procedure of Filip et al. (10). Cell cultures were also treated with trypsin and trypsin inhibitor as previously described (25).

Slide agglutination. The ability of antisera to agglutinate live *Helicobacter* cells was tested by the method of Lior et al. (23). The presence or absence of agglutination was determined after 1 min.

Electron microscopy. Samples were routinely negatively stained with a solution of 1% (wt/vol) ammonium molybdate and 1% (vol/vol) glycerol on Formvar-coated grids. For immunogold electron microscopy, a grid covered with Formvar film was coated on a 50-μl drop of bacterial cells suspended in 100 mM Tris hydrochloride (pH 7.5) containing 150 mM NaCl, 0.05% Tween 20, and 0.5% bovine serum albumin (Tris-NaCl-TW-BSA) for 5 min. The grid was then removed from the drop and floated on a drop of Tris-NaCl-TW-BSA containing 10% (vol/vol) antiserum. After 1 h of incubation, the grid was removed, and nonspecifically bound immunoglobulin was removed by floating the grid on two drops of Tris-NaCl-TW-BSA. The grid was then floated on a drop of Tris-NaCl-TW-BSA containing a 1:10 dilution of 15-nm colloidal gold particles coated with protein A (Jannsen Biotech, Olen, Belgium). After incubation for 1 h, the nonspecifically bound colloidal gold particles were removed by floating the grid on two drops of Tris-NaCl-TW-BSA and one drop of distilled water. The grids were stained by floating on a drop of 1% (wt/vol) ammonium molybdate containing 0.1% glycerol (pH 7.5) and were examined in a Philips EM-300 electron microscope (Philips Electronic Instruments, Inc., Mahwah, N.J.) operated at an accelerating voltage of 60 kV. Images were recorded on 70-mm Fine Grain Release Film (Kodak, Rochester, N.Y.).

Indirect fluorescent antibody testing. Indirect fluorescent antibody testing was performed as previously described (26).

Recombinant DNA procedures. Reagents and restriction enzymes used for molecular cloning were from Boehringer or Sigma. Standard techniques were used for molecular cloning, transformation in *E. coli*, and hybridization (41). The plasmid vectors pK18/19 (38), pUC18/19 (48), and pKK233-2 (1) were employed.

Extraction of genomic DNA from *H. pylori*. The complete range of *H. pylori* strains examined showed varied sensitivities to lysozyme, ranging from good lytic activity to apparent resistance. We therefore adapted the following procedure, which, being independent of enzymatic murolytic attack, was applicable with equal success to all isolates tested. Cells were harvested by scraping from five culture plates (0.3 g [wet weight]), washed twice in phosphate-buffered saline, and suspended gently in 1 ml of 50 mM Tris hydrochloride (pH 8.0). After the addition of 100 μl of 10% (wt/vol) SDS, 100 μl of Tween lysing mix (2% [vol/vol] Tween 20–62 mM EDTA in 50 mM Tris hydrochloride [pH 8.0]), and 50 μl of proteinase K (50 mg/ml in water), the sample was incubated for 16 h at 50°C. Effective proteolytic digestion was marked by a reduction in coloration and viscosity. The volume of the sample was increased to 4.5 ml with TE buffer (10 mM Tris hydrochloride, 1 mM EDTA), and 4.5 g of cesium chloride was added and gently mixed. The sample was spun at 55,000 rpm for 14 h in a Beckman VTi65.1 rotor, after which the chromosomal DNA could be visualized as a white mass. The DNA was removed, dialyzed extensively against TE buffer, and frozen in aliquots at −20°C until required. It was found that several gentle phenol extractions could be substituted for ultracentrifugation, but the average mass of the resulting product, although suitable for hybridization studies or plasmid cloning, was too small for cosmid cloning.

Oligonucleotide synthesis. Synthetic oligonucleotides were generated on a Biosearch Bio Sam One DNA synthesizer (Biosearch, San Rafael, Calif.) and purified by PAGE.

DNA sequence determination. Double-stranded chain-ter-

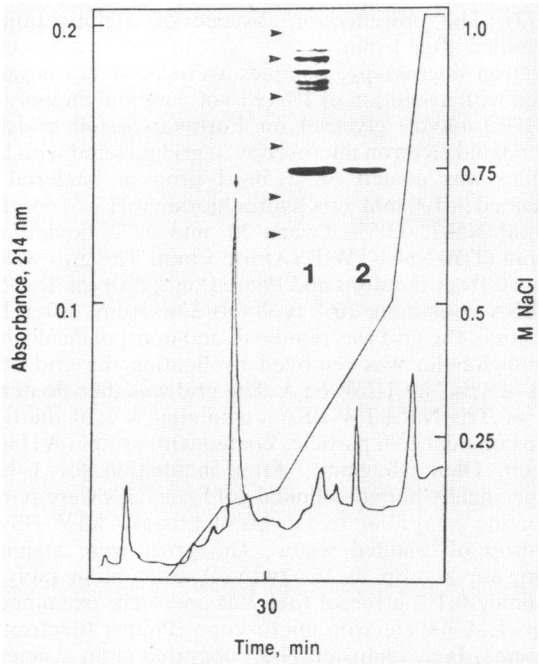

FIG. 1. Elution profile of the M_r-26,000 protein (arrow) from a MonoQ ion-exchange column. The 0 to 1.0 M NaCl gradient is indicated. The inset shows Coomassie blue-stained 12.5% SDS–PAGE gels of the fraction eluted from Sephadex G-100 column containing partially purified M_r 26,000 protein (lane 1) and the M_r-26,000 protein purified by ion-exchange chromatography on MonoQ Sepharose (lane 2) (peak in elution profile). Arrowheads on left indicate standards at M_r 97,400, 66,200, 45,000, 31,000, 21,500, and 14,400.

mination DNA sequencing was performed after alkali denaturation by using [^{35}S]dATP (Amersham Corp., United Kingdom) and Sequenase or Taq polymerase (U.S. Biochemical Corp., Cleveland, Ohio) according to the manufacturer's instructions. Template plasmid DNA was generated by the method of Misra (32) with exonuclease BAL 31 to generate ordered sets of deletions in a 900-bp fragment carrying the gene for the M_r-26,000 protein. Template DNA was purified by ultracentrifugation in ethidium bromide-cesium chloride.

Hybridization analysis. Filters for hybridization were prepared by the method of Southern (43) after standard electrophoresis of digested genomic DNA. For dot blot analysis, standardized suspensions of the bacterial strains were applied to NCP in a dot blot apparatus. After gravity filtration of the buffer, the wells were dried and the bacteria were lysed by the method of Grunstein and Hogness (14). Filters were then probed with either the cloned *Hin*dIII fragment or an intragenic fragment extending from an *Xba*I site at residue 589 (see Fig. 5) downstream to the *Hin*dIII site. Both probes were purified by excision from agarose gels and labeled by the digoxigenin antigen system from Boehringer. After hybridization, filters were subjected to two 5-min washes in 2× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate)–0.1% SDS at room temperature and two 15-min washes in 0.1% SSC–0.1% SDS at 68°C. Filters were blocked and developed as recommended by the manufacturers.

Nucleotide sequence accession number. The coding sequence shown in Fig. 5 has been submitted to the GenBank data base under accession number M55507.

RESULTS

Purification of *H. pylori* 915 protein. Initial studies showed that relatively large amounts of a protein with an apparent M_r of 26,000 were readily removed from cells of *H. pylori* 915 by extraction with 0.2 M glycine hydrochloride (pH 2.2) or mild detergent extraction with 0.6% octylglucoside (8). This ability to remove the protein from cells with octylglucoside was exploited for preparative scale isolation. After octylglucoside extraction and concentration by 0 to 50% ammonium sulfate precipitation, the M_r-26,000 subunit protein was initially separated from contaminating proteins by molecular sieving. Under the nondenaturing conditions used for the Sephadex G100 fractionation, the M_r-26,000 subunit eluted in the void volume. SDS-PAGE analysis of this material showed a major polypeptide of apparent subunit M_r 26,000, with small amounts of a M_r-55,000 protein and several other minor bands. The M_r-26,000 protein was purified to homogeneity by RPLC with a ProRPC Hr5/10 column on a Pharmacia fast-protein liquid chromatography apparatus or by anion-exchange chromatography with a MonoQ HR 5/5 column (Fig. 1).

Biochemical characterization. Amino acid composition analysis of the M_r-26,000 protein purified from *H. pylori* showed that the protein contained approximately 232 residues per molecule (Table 2). The predicted M_r of 25,304 is close to the value of 26,000 estimated by SDS-PAGE. Asparagine, glutamine, lysine, glycine, alanine, valine, and leucine were prominent. The calculated relative hydrophobicity was approximately 39.2%. When isoelectric focusing was performed in a minigel system under nondenaturing conditions, the protein focused into two bands with pIs of 5.9 and 6.0.

The electrophoretic mobility of the protein was not affected by the temperature of solubilization in SDS. The M_r-26,000 protein appeared to be fully solubilized at all temperatures between 20 and 100°C. However, in the absence of β-mercaptoethanol in the solubilization buffer, the protein migrated with an apparent M_r of 46,000, compared

TABLE 2. Amino acid composition of the M_r-26,000 antigen of *H. pylori* 915[a]

Amino acid	No. of residues
Asx	29
Thr	7
Ser	13
Glx	26
Pro	8
Gly	20
Ala	19
Val	19
Met	6
Ile	11
Leu	18
Tyr	4
Phe	13
His	9
Lys	21
Arg	11
Cys	NT[b]
Try	NT

[a] There was a total of 232 residues per mole. The M_r was 25,304. The percentage of hydrophobic residues (Val, Met, Ile, Leu, Ala, Phe, Trp, and Pro) was 39.2%.
[b] NT, Not tested.

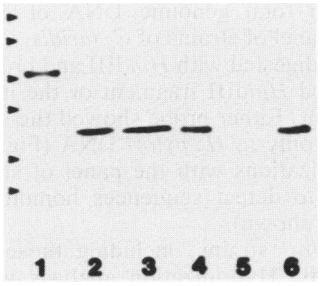

FIG. 2. SDS-PAGE analysis the M_r-26,000 protein. Lanes 1 through 3 contained protein purified by fast-protein liquid chromatography and stained with coomassie blue: 1, protein solubilized in sample buffer without β-mercaptoethanol for 10 min at 100°C; 2, protein solubilized in sample buffer with β-mercaptoethanol for 10 min at 20°C; 3, protein solubilized in sample buffer with β-mercaptoethanol for 10 min at 100°C. Lanes 4 through 6 were immunoblotted with a 1:5,000 dilution of antiserum raised against the M_r-26,000 protein of strain 915: 4, whole-cell lysate of strain 915; 5, cell envelopes; 6, supernatant of cells disrupted by passage through a French pressure cell after sedimentation of cell envelopes (cytoplasm). Arrows on left indicate standards at M_r 97,400, 66,200, 45,000, 31,000, 21,500, and 14,400.

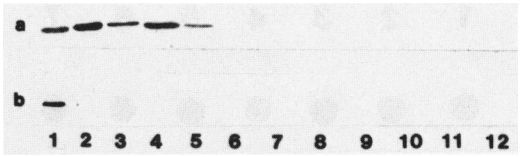

FIG. 3. Western blot analysis of whole-cell lysates reacted with a 1:5,000 dilution of antiserum raised against RPLC-purified M_r-26,000 protein of *H. pylori* 915. (A) Lanes: 1, *H. pylori* 915; 2, *H. pylori* 17874; 3, *H. pylori* 5294; 4, *H. pylori* 5442; 5, *E. coli* TB1 harboring pK26K; 6, *E. coli* TB1 harboring pKK233-2 (vector plasmid); 7, *C. fetus* VC78; 8, *C. fetus* VC202; 9, *C. jejuni* VC36; 10, *C. jejuni* VC74; 11, *C. coli* VC212; 12, *C. laridis* VC81. (B) Lanes: 1, *H. pylori* 5650; 2, *H. mustelae* 23950; 3, *H. mustelae* 23951; 4, *H. mustelae* 25715; 5, *C. cinaedi* 18818; 6, *C. cinaedi* 18819; 7, *C. cinaedi* 19218; 8, *C. fennelliae* 18820; 9, *W. succinogenes* 12550; 10, *W. succinogenes* 13145; 11, *E. coli* 1593.

with the single band at M_r 26,000 under reducing conditions, suggesting the presence of interchain disulfide bonds (Fig. 2). Aggregation properties were also apparent upon storage and dialysis.

The N-terminal amino acid sequence of the protein from strain 915 was determined for the first 46 residues (Table 3). A search of the National Biomedical Research Foundation and the Swiss protein sequence libraries failed to reveal any sequences with significant homology to the N terminus of the M_r-26,000 protein (24), suggesting that it may be a protein unique to *H. pylori*. The protein was susceptible to chemical cleavage by CNBr because of methionine residues. Of the digest products, the three most easily purified were isolated by RPLC in sufficient quantity to allow N-terminal sequence analysis (Table 3). The N-terminal sequence analysis of the intact protein and CNBr peptides provided sequence information on 113 amino acids (47% of the total sequence of the protein).

Immunochemical analysis. The M_r-26,000 protein was a good immunogen in rabbits. Western blotting of whole-cell lysates and glycine extracts of eight *H. pylori* isolates from diverse geographic sources and with different phenotypic characteristics with antisera to the purified M_r-26,000 protein of strain 915 showed that an antigenically cross-reactive protein of approximate subunit M_r 26,000 was present in all *H. pylori* strains tested (Fig. 3). This was confirmed by immunodot blot analysis (Fig. 4). Antigenically cross-reactive proteins were not detectable by dot immunoblotting of the strains of *H. mustelae*, *C. cinaedi*, *C. fennelliae*, *W. succinogenes*, *C. fetus*, *C. jejuni*, *C. coli*, *C. laridis*, *E. coli*, *Salmonella enteritidis*, *Salmonella illinois*, and *Shigella flexneri* tested (Table 1).

Localization. The M_r-26,000 protein of *H. pylori* 915 was readily removed from the cells after relatively mild treatment by either eluting in buffer of low pH or by adding detergent, suggesting that the protein might be located on the cell surface. Although the purified M_r-26,000 protein was susceptible to treatment with trypsin, the protein was not removed from the cells by trypsin. Moreover, monospecific antiserum failed to agglutinate cells of the homologous strain, and indirect fluorescent antibody testing and immunogold electron microscopy with this antiserum failed to reveal antibody binding to the cell surface, indicating that the protein was not surface exposed. Cell fractionation experiments showed that the protein could be best recovered from the cytosol (Fig. 2).

Cloning the gene encoding the M_r-26,000 protein. The following mixture of synthetic oligonucleotides capable of encoding amino acid residues 22 to 26 of the protein was synthesized: GACGA(A/G)CA(C/T)TT(T/C)GA(A/G)TT. Complete digests of genomic DNA by *Sal*I, *Eco*RI, and *Hin*dIII were electrophoresed, blotted to NCP, and probed with the labeled oligonucleotide. A single band of approximately 900 bp in the *Hin*dIII-digested sample reacted with the probe. Fragments in this size range were excised from an agarose gel, ligated to pK18 linearized with *Hin*dIII, and transformed into strain TB1. Clones carrying the desired fragment were identified by colony hybridization. The resulting plasmid, called p26K, was shown to carry a 900-bp fragment that reacted with the oligonucleotide in Southern hybridization analysis.

Expression of the M_r-26,000 protein in *E. coli*. Cells harboring p26K were lysed by boiling in final sample buffer and

TABLE 3. N-terminal amino acid sequences of the M_r-26,000 *H. pylori*-specific protein and internal peptides

Sample	Residues[a]
Intact protein	MLVTKLAPDF KAPAVLGNNE VDEHFELSKN LGKNGVILFF XPKDFT
CNBr peptides	
CNBr 1	VADITKSISR DYDVLFEEAI ALRGAFLIDL N
CNBr 2	KVRHAVINDL PLGRNADE
CNBr 3	VDALLHFEEH GEVXPAGWR

[a] Amino acid residues are designated by the single-letter nomenclature. X, Not identified.

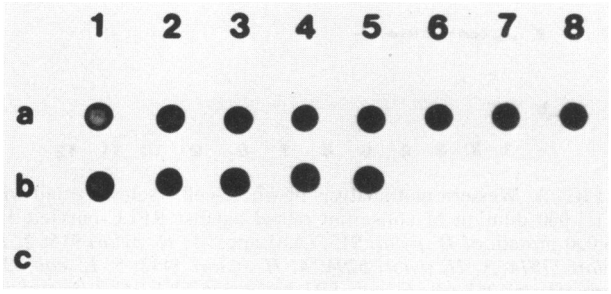

FIG. 4. Immunodot blot of whole-cell lysates reacted with 1:5,000 dilution of antiserum raised against RPLC-purified M_r-26,000 protein of *H. pylori* 915. Dots: a1, *H. pylori* 915; a2, *H. pylori* 17874; a3, *H. pylori* 5155; a4, *H. pylori* 5294; a5, *H. pylori* 5442; a6, *H. pylori* 5646; a7, *H. pylori* 5650; a8, *H. pylori* 5790; b1, *H. pylori* 52; b2, *H. pylori* 66; b3, *H. pylori* 95; b4, *H. pylori* 253; b5, *H. pylori* 1139; b6, *C. jejuni* VC36; b7, *C. jejuni* VC74; b8, *C. coli* VC212; c1, *C. laridis* VC81; c2, *E. coli* 1593; c3, *E. coli* E1049-a-13; c4, *E. coli* E438; c5, *S. enteritidis* 2b.1; c6, *S. illinois* S.1093; c7, *S. flexneri* Sh.1a1.

tested for production of the antigen by Western immunoblotting. No expression could be detected. The insert of p26K was excised with *Hin*dIII and recloned in the expression vector pKK233-2 cut with the same enzyme. The resulting plasmid, pK26K, directed expression of levels of the M_r-26,000 protein readily detectable by Western immunoblotting (Fig. 3), and the product comigrated with the polypeptide produced by *H. pylori* under the gel conditions shown.

DNA sequence determination. Ordered sets of deletions entering the cloned *H. pylori* DNA from both extremities were generated by exonuclease BAL 31 and recloned into a pUC vector cut with appropriate restriction enzymes, allowing both strands to be sequenced. Sequenase was routinely used for sequencing reactions, and Taq polymerase was employed when local secondary structure in particular clones caused problems.

A single long open reading frame composed of 225 codons was found in the cloned fragment (Fig. 5). The amino-terminal residue identified by sequencing of the protein occurs at position 28 in the translated sequence. This coding sequence was still open at the 3′ end of the anticoding strand, and it was apparent that the coding sequence had been truncated at the *Hin*dIII site used for cloning. The possibility of an in-frame fusion with sequences flanking the pK18 multiple cloning site would mean the addition of 45 amino acid residues to the *H. pylori* protein expressed in *E. coli*. Such a chimeric protein would have an M_r of 27,100.

The TGAAA-18-TAAAAT motif beginning at residue 88 resembles closely the consensus sequences for known promoters of *E. coli* (15), and the spacing of 18 residues is within the known limits for function. The ATG codon at position 262 is presumably the initiation codon. It is preceded by a putative ribosome binding site, AGGA, optimally spaced at seven nucleotides from the start codon. The 27-residue polypeptide potentially coded by the upstream sequences lacks a possible translation initiation site.

Base composition analysis indicates an overall G+C content of 36.5% in the cloned fragment and 41% within the truncated open reading frame coding the M_r-26,000 protein. These values compare with 37.7% G+C for the *C. coli* flagellin gene and 30 to 35% for the *C. coli* genome (2, 27, 37).

Hybridization analysis. Southern blot hybridization was performed with total genomic DNA of eight *H. pylori* isolates and a panel of strains of *C. laridis*, *C. jejuni*, and *C. coli*. DNA was digested with *Hin*dIII and probed with either the entire cloned *Hin*dIII fragment or the intragenic *Xba*I-*Hin*dIII fragment. Either probe showed the same specificity and hybridized only to *H. pylori* DNA (Fig. 6). Additional dot blot hybridizations with the panel of strains shown in Table 1 failed to detect sequences homologous to either probe (data not shown).

Many *H. pylori* strains, including those tested above, contain plasmids. Hybridization analysis with undigested genomic DNA and purified plasmid DNA showed that the gene encoding the M_r-26,000 protein was chromosomally located (data not shown).

DISCUSSION

Relatively little is known about *H. pylori* at the molecular level. We initiated the present study both to identify potential virulence determinants and to isolate species-specific DNA probes. We first became interested in this cryptic protein because of its relative abundance in cell extracts generated by relatively mild treatments. The protein appeared not to be surface expressed, however, and cell fractionation experiments showed that the protein was not associated with the particulate fraction of the cell. The protein sequence data derived biochemically concurred with that subsequently available from the DNA sequence. Computer analysis of the derived amino acid sequence (although only corresponding to approximately 70% of the mature protein) with a number of models failed to detect membrane-spanning segments or potential transmembrane helices, in agreement with subcellular localization. DNA sequence analysis also showed that the signal sequence typically found in exported proteins was absent. Under nondenaturing conditions, the protein appeared to exist as a dimer with interchain disulfide bridges, presumably resulting from one or more of the cysteine residues present in the amino or carboxy terminal of the polypeptide. It is apparent from its behavior in gel filtration experiments that the protein also forms large aggregates in the buffer conditions used. The relevance, if any, of this to biological function is as yet unclear.

The protein appears to be antigenically conserved in *H. pylori* and antigenically unique to the species. Importantly, the protein can be detected immunologically in whole-cell lysates in a simple dot blot assay with monospecific polyclonal antibody. This assay is easy to perform and may be a useful addition to those currently employed in the rapid identification of these important pathogens.

The cloned open reading frame for the M_r-26,000 protein was expressed in *E. coli* only when a strong promoter was provided. This infers that the translation initiation signals of this *H. pylori* gene are functional in *E. coli* but that the promoter sequences are either missing or not recognized by *E. coli* RNA polymerase. The former is unlikely, since more than 260 residues of the upstream flanking sequences were present in the cloned fragment. We have other unpublished data suggesting lack of expression of *H. pylori* genes in genomic libraries in phage lambda replacement vectors. Clayton et al. (6) reported poor expression of *H. pylori* genes in *E. coli* and that production of urease components was dependent on promoter sequences in the plasmid vector. Curiously, a sequence with strong homology to consensus promoter sequences of *E. coli* was present upstream of the gene encoding the M_r-26,000 protein. Analysis of transcripts

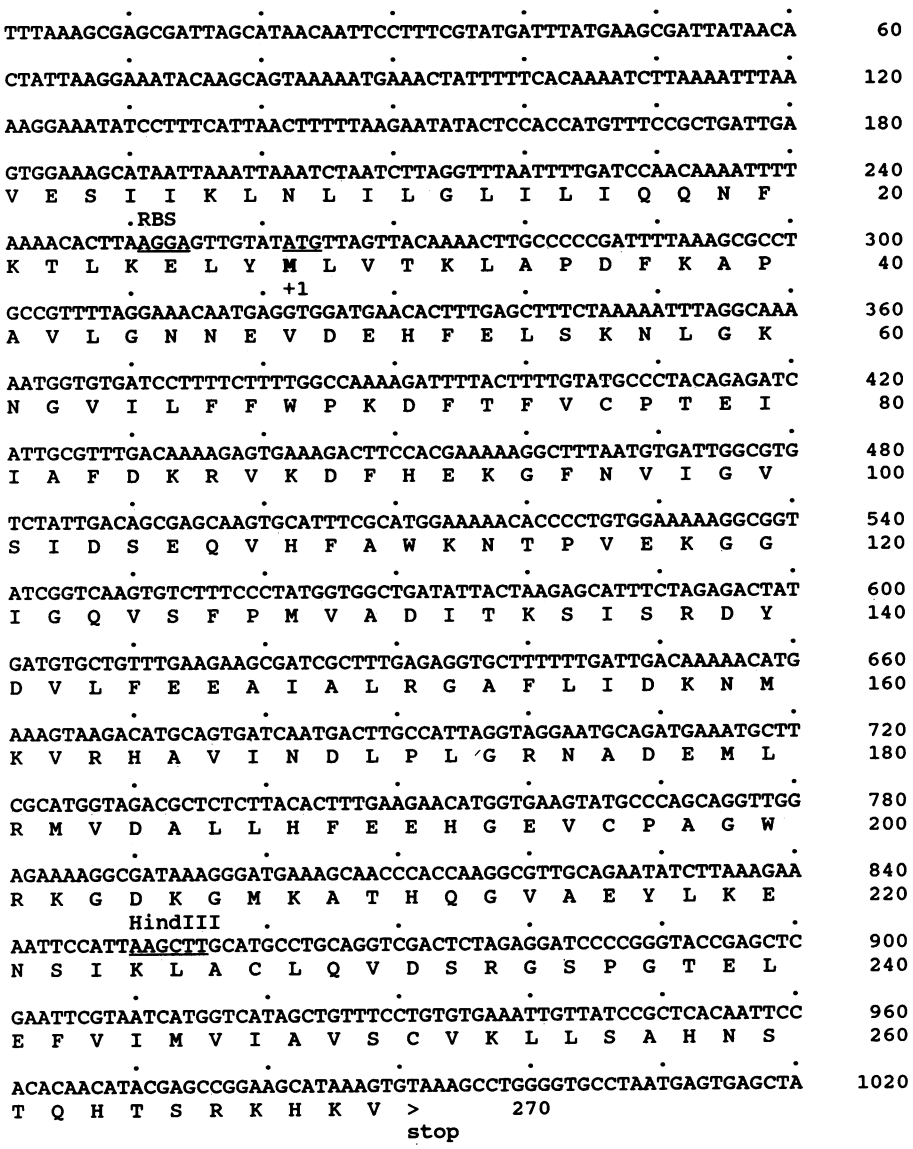

FIG. 5. DNA sequence and translation product of a fragment encoding part of the M_r-26,000 protein. Residues after position 854 are from the pK18 vector sequence. The methionine at position 28 is the N-terminal residue identified by chemical protein sequence determination.

and definitive promoter mapping will investigate the unlikely possibility that mRNA is transcribed from this point in *E. coli* but not translated. It is also noteworthy that the protein produced in *E. coli* comigrated with that purified from *H. pylori*, when the size of the protein predicted from the DNA sequence is only M_r 22,000. We are presently purifying the protein from *E. coli* lysates to determine whether the 27-residue hydrophobic leader is retained in this host, which would yield a polypeptide of M_r 25,400. This is a possible explanation for the observed behavior in electrophoresis.

Probes derived from the structural gene for the M_r-26,000 polypeptide hybridized to all *H. pylori* strains tested so far. Genomic DNA samples from many of these have striking restriction fragment length polymorphisms readily visible directly from the gels (data not shown), in accordance with a previous report of considerable subspecies variation (28). Nevertheless, the gene for the M_r-26,000 protein is present in all strains, suggesting that it is involved in some essential function. There were at least three groups of strains based upon the size of the *Hin*dIII digest fragment carrying the 5′ end of the gene, but we have not been able to correlate this subdivision with known phenotypic markers such as lectin specificity.

None of the other bacterial species tested was detected with these probes. It has been shown previously that the most closely related species are *W. succinogenes*, *C. fennelliae*, and *C. cinaedi* and that the true campylobacters are genetically more distantly related (40, 44). *H. mustelae* was included in the same genus as *H. pylori* on the basis of similarity of five major taxonomic features (11), but DNA-DNA hybridization was sufficiently low to warrant separate species. Thompson and co-workers (44) have argued that the weight of evidence of 16S rRNA sequence data and phenotypic characterization is sufficient to justify inclusion of *C. cinaedi* and *C. fennelliae* in the same genus as *W. succinogenes* and *H. pylori*, but this has been disputed by later,

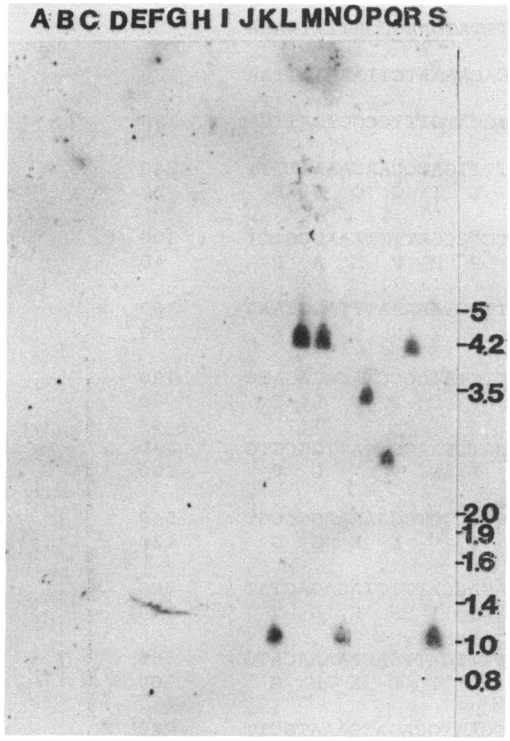

FIG. 6. Southern blot hybridization of DNA from *Campylobacter* spp. and *H. pylori* with the gene encoding the M_r-26,000 protein. DNA was digested with *Hin*dIII, separated on a 1% agarose gel, and transferred to NCP. Hybridization was with the cloned *Hin*dIII fragment carrying most of the coding sequence for the M_r-26,000 protein. Lanes: A through F, *C. jejuni* VC26, VC27, VC28, VC36, VC74, and VC209, respectively; G through I, *C. coli* VC143, VC146, and VC212, respectively; J and K, *C. laridis* VC213 and VC214, respectively; L through R, *H. pylori* 915, 17874, 52, 66, 95, 253, and 1139, respectively; S, *Hin*dIII digest of p26K. Size markers (in kilobase pairs) are indicated on the right.

more comprehensive investigations (11). It is apparent from our study that the gene we have cloned is derived from sequences truly unique to *H. pylori*; although this fails to elucidate taxonomical questions, it supports the applicability of the sequence as a specific probe. Since we have not found homologous sequences, either protein or DNA, in the various data bases, it is likely that the function of the M_r-26,000 protein is also peculiar to this species, with the reservation that there may be homologous domains in the carboxy terminus missing from our sequence.

DNA-based diagnostic procedures are finding widespread routine application in bacteriological laboratories. We are currently evaluating oligonucleotides based on the DNA sequence of the M_r-26,000 protein for the specific detection of *H. pylori* in clinical samples by using the polymerase chain reaction technique.

ACKNOWLEDGMENTS

This work was supported by grants from the Swedish Medical Research Council (project B90-16X-09085-01), the Medical Faculty of Lund University, and the Kock Foundation to P.W.O. and by grants from the Medical Research Council of Canada and the British Columbia Health Care Research Foundation to T.J.T. T.J.T. was the recipient of a Visiting Scientist Fellowship from the Swedish Medical Research Council.

J. Austin is thanked for electron microscopy.

REFERENCES

1. **Amann, E., and J. Brosius.** 1985. 'ATG' vectors for the regulated high-level expression of cloned genes in *Escherichia coli*. Gene **40**:183–190.
2. **Belland, R. J., and T. J. Trust.** 1982. Deoxyribonucleic acid relatedness between thermophilic members of the genus *Campylobacter*. J. Gen. Microbiol. **128**:2515–2522.
3. **Blake, M. S., K. H. Johnston, G. R. Russell-Jones, and E. C. Gotschlich.** 1984. A rapid sensitive method for detection of alkaline phosphatase conjugated anti-antibodies on Western blots. Anal. Biochem. **136**:175–181.
4. **Blaser, M. J.** 1987. Gastric *Campylobacter*-like organisms, gastritis and peptic ulcer disease. Gastroenterology **93**:371–383.
5. **Buck, G. E., W. K. Gourley, W. K. Lee, K. Subramanyam, J. M. Latimer, and R. A. DiNuzzo.** 1986. Relation of *Campylobacter pyloridis* to gastritis and peptic ulcer. J. Infect. Dis. **153**:664–669.
6. **Clayton, C. L., B. W. Wren, P. Mullany, A. Topping, and S. Tabaqchali.** 1989. Molecular cloning and expression of *Campylobacter pylori* species-specific antigens in *Escherichia coli* K-12. Infect. Immun. **57**:623–629.
7. **Dooley, C. P., and H. Cohen.** 1988. The clinical significance of *Campylobacter pylori*. Ann. Intern. Med. **108**:70–79.
8. **Evans, D. G., D. J. Evans, Jr., J. J. Moulds, and D. Y. Graham.** 1988. *N*-Acetylneuraminyllactose-binding fibrillar hemagglutinin of *Campylobacter pylori*: a putative colonization factor antigen. Infect. Immun. **56**:2896–2906.
9. **Evans, D. J., Jr., D. G. Evans, D. Y. Graham, and P. D. Klein.** 1989. A sensitive and specific serologic test for detection of *Campylobacter pylori* infection. Gastroenterology **96**:1004–1008.
10. **Filip, C., G. Fletcher, J. L. Wulff, and C. F. Earhart.** 1973. Solubilization of the cytoplasmic membrane of *Escherichia coli* by the ionic detergent sodium lauryl sarcosinate. J. Bacteriol. **115**:717–722.
11. **Goodwin, C. S., J. A. Armstrong, T. Chilvers, M. Peters, M. D. Collind, L. Sly, W. McConnell, and W. E. S. Harper.** 1989. Transfer of *Campylobacter pylori* and *Campylobacter mustelae* to *Helicobacter* gen. nov. as *Helicobacter* comb. nov. and *Helicobacter mustelae* comb. nov., respectively. Int. J. Syst. Bacteriol. **39**:397–405.
12. **Graham, D. Y., and P. D. Klein.** 1987. *Campylobacter pyloridis* gastritis: the past, the present and speculations about the future. Am. J. Gastrol. **82**:283–286.
13. **Graham, D. Y., P. D. Klein, A. R. Opekun, and T. W. Boutton.** 1988. Effect of age on the frequency of active *Campylobacter pylori* infections diagnosed by the [^{13}C]-urea breath test in normal subjects and patients with peptic ulcer disease. J. Infect. Dis. **157**:777–780.
14. **Grunstein, M., and D. S. Hogness.** 1975. Colony hybridization: a method for the isolation of cloned DNAs that contain a specific gene. Proc. Natl. Acad. Sci. USA **72**:3961–3965.
15. **Hawley, D. K., and W. R. McClure.** 1983. Compilation and analysis of *Escherichia coli* promoter DNA sequences. Nucleic Acids Res. **11**:2237–2255.
16. **Hunkapiller, M. W., and L. E. Hood.** 1983. Analysis of phenylthiohydantoins by ultrasensitive gradient high-performance liquid chromatography. Methods Enzymol. **91**:486–493.
17. **Jiang, S. J., W. Z. Liu, D. Z. Zhang, Y. Shi, S. D. Xiao, Z. N. Zhang, and D. Y. Liu.** 1987. Campylobacter-like organisms in chronic gastritis, peptic ulcer and gastric carcinoma. Scand. J. Gastroenterol. **22**:553–558.
18. **Krakowska, S., D. R. Morgan, W. G. Kraft, and R. D. Leunk.** 1987. Establishment of the gastric pathogen *Campylobacter pylori* infection in neonatal piglets. Infect. Immun. **55**:2789–2796.
19. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
20. **Lambert, J. R., M. Borromeo, K. J. Pinkard, H. Turner, C. B. Chapman, and M. L. Smith.** 1987. Colonisation of gnotobiotic

pigs with *Campylobacter pylori*—an animal model? J. Infect. Dis. **155**:1344.
21. **Lambert, J. R., K. A. Dunn, E. R. Eaves, M. G. Korman, and J. Hansky.** 1986. *Campylobacter pyloridis* in diseases of the human upper gastrointestinal tract. Gastroenterology **90**:1509.
22. **Lee, A., and S. L. Hazell.** 1988. *Campylobacter pylori* in health and disease: an ecological perspective. Microb. Ecol. Health Dis. **1**:1–16.
23. **Lior, H., D. L. Woodward, J. A. Edgar, L. J. Laroche, and P. Gill.** 1982. Serotyping of *Campylobacter jejuni* by slide agglutination based on heat-labile antigenic factors. J. Clin. Microbiol. **15**:761–768.
24. **Lipman, D. J., and W. R. Pearson.** 1985. Rapid and sensitive protein similarity searches. Science **227**:1435–1441.
25. **Logan, S. M., and T. J. Trust.** 1983. Molecular identification of surface protein antigens of *Campylobacter jejuni*. Infect. Immun. **42**:675–682.
26. **Logan, S. M., and T. J. Trust.** 1986. Location of epitopes on *Campylobacter jejuni* flagella. J. Bacteriol. **168**:739–745.
27. **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.
28. **Majewski, S. L. H., 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.
29. **Marshall, B. J., J. A. Armstrong, D. B. McGechie, and R. J. Glancy.** 1985. Attempt to fulfill Koch's postulates for pyloric *Campylobacter*. Med. J. Aust. **142**:436–439.
30. **Marshall, B. J., and I. Surveyor.** 1988. Carbon-14 urea breath test for the diagnosis of *Campylobacter pylori* associated with gastritis. J. Nucl. Med. **29**:11–16.
31. **McCoy, E. C., D. Doyle, K. Burda, L. B. Corbeil, and A. J. Winter.** 1975. Superficial antigens of *Campylobacter* (*Vibrio*) *fetus*: characterization of an antiphagocytic component. Infect. Immun. **11**:517–525.
32. **Misra, T. K.** 1985. A strategy to create ordered deletions for rapid nucleotide sequencing. Gene **34**:263–268.
33. **Mobley, H. L. T., and R. P. Hausinger.** 1989. Microbial ureases: significance, regulation, and molecular characterization. Microbiol. Rev. **53**:85–108.
34. **Morotomi, M., S. Hoshina, P. Green, H. Neu, P. LoGerfo, I. Watanabe, M. Mutai, and I. B. Weinstein.** 1989. Oligonucleotide probe for detection and identification of *Campylobacter pylori*. J. Clin. Microbiol. **27**:2652–2655.
35. **Morris, A., and G. Nicholson.** 1987. Ingestion of *Campylobacter pylori* causes gastritis and raised fasting gastric pH. Am. J. Gastroenterol. **82**:192–199.
36. **Newell, D. G., and B. J. Rathbone.** 1989. The serodiagnosis of *Campylobacter pylori* infection. Serodiag. Immunother. Infect. Dis. **3**:1–5.
37. **Owen, R. J.** 1983. Nucleic acids in the classification of campylobacters. Eur. J. Microbiol. **2**:367–377.
38. **Pridmore, R. D.** 1987. New and versatile cloning vectors with kanamycin-resistance marker. Gene **56**:309–312.
39. **Robertson, E. F., H. K. Dannelly, P. J. Molloy, and H. C. Reeves.** 1987. Rapid electrofocussing in a vertical polyacrylamide minigel system. Anal. Biochem. **167**:290–294.
40. **Romaniuk, P. J., B. Zoltowska, T. J. Trust, D. J. Lane, G. J. Olsen, N. R. Pace, and D. A. Stahl.** 1987. *Campylobacter pylori*, the spiral bacterium associated with human gastritis, is not a true *Campylobacter* sp. J. Bacteriol. **169**:2137–2141.
41. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
42. **Simor, A. E., B. Shames, B. Drumm, P. Sherman, D. E. Low, and J. L. Penner.** 1990. Typing of *Campylobacter pylori* by bacterial DNA restriction endonuclease analysis and determination of plasmid profile. J. Clin. Microbiol. **28**:83–86.
43. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.
44. **Thompson, L. M., R. M. Smibert, J. L. Johnson, and N. R. Krieg.** 1988. Phylogenetic study of the genus *Campylobacter*. Int. J. Syst. Bacteriol. **38**:190–200.
45. **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 **74**:4350–4354.
46. **Warren, J. R., and B. Marshall.** 1983. Unidentified curved bacilli on gastric epithelium in active chronic gastritis. Lancet **i**:1273–1275.
47. **Willis, R. C., R. G. Morris, C. Cirakogiu, G. D. Schellenberg, N. H. Gerber, and C. E. Furlong.** 1974. Preparation of the periplasmic binding proteins from *Salmonella typhimurium* and *Escherichia coli*. Arch. Biochem. Biophys. **161**:64–75.
48. **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**:103–119.