

- 1 -

**NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
HELICOBACTER PYLORI AND VACCINE COMPOSITIONS THEREOF**

Background of the Invention

5 *Helicobacter pylori* is a gram-negative, S-shaped, microaerophilic bacterium that was discovered and cultured from a human gastric biopsy specimen. (Warren, J.R. and B. Marshall, (1983) *Lancet* 1: 1273-1275; and Marshall et al., (1984) *Microbios Lett.* 25: 83-88). *H. pylori* has been strongly linked to chronic gastritis and duodenal ulcer disease. (Rathbone et. al., (1986) *Gut* 27: 635-641). Moreover, evidence is
10 accumulating for an etiologic role of *H. pylori* in nonulcer dyspepsia, gastric ulcer disease, and gastric adenocarcinoma. (Blaser M. J., (1993) *Trends Microbiol.* 1: 255-260). Transmission of the bacteria occurs via the oral route, and the risk of infection increases with age. (Taylor, D.N. and M. J. Blaser, (1991) *Epidemiol. Rev* 13: 42-50). *H. pylori* colonizes the human gastric mucosa, establishing an infection that usually
15 persists for decades. Infection by *H. pylori* is prevalent worldwide. Developed countries have infection rates over 50% of the adult population, while developing countries have infection rates reaching 90% of the adults over the age of 20. (Hopkins R. J. and J. G. Morris (1994) *Am. J. Med.* 97: 265-277).

 The bacterial factors necessary for colonization of the gastric environment, and
20 for virulence of this pathogen, are poorly understood. Examples of the putative virulence factors include the following: urease, an enzyme that may play a role in neutralizing gastric acid pH (Eaton et al., (1991) *Infect. Immunol.* 59: 2470-2475; Ferrero, R.L. and A. Lee (1991) *Microb. Ecol. Hlth. Dis.* 4: 121-134; Labigne et al., (1991) *J. Bacteriol.* 173: 1920-1931); the bacterial flagellar proteins responsible for
25 motility across the mucous layer. (Hazell et al., (1986) *J. Inf. Dis.* 153: 658-663; Leying et al., (1992) *Mol. Microbiol.* 6: 2863-2874; and Haas et al., (1993) *Mol. Microbiol.* 8: 753-760); Vac A, a bacterial toxin that induces the formation of intracellular vacuoles in epithelial cells (Schmitt, W. and R. Haas, (1994) *Molecular Microbiol.* 12(2): 307-319); and several gastric tissue-specific adhesins. (Boren et al., (1993) *Science* 262: 1892-
30 1895; Evans et al., (1993) *J. Bacteriol.* 175: 674-683; and Falk et al., (1993) *Proc. Natl. Acad. Sci. USA* 90: 2035-203).

 Numerous therapeutic agents are currently available that eradicate *H. pylori* infections *in vitro*. (Huesca et. al., (1993) *Zbl. Bakt.* 280: 244-252; Hopkins, R. J. and J. G. Morris, supra). However, many of these treatments are suboptimally effective *in vivo*
35 because of bacterial resistance, altered drug distribution, patient non-compliance or poor drug availability. (Hopkins, R. J. and J. G. Morris, supra). Treatment with antibiotics combined with bismuth are part of the standard regime used to treat *H. pylori* infection.

(Malfertheiner, P. and J. E. Dominguez-Munoz (1993) *Clinical Therapeutics* 15 Supp. B: 37-48). Recently, combinations of a proton pump inhibitors and a single antibiotic have been shown to ameliorate duodenal ulcer disease. (Malfertheiner, P. and J. E. Dominguez-Munoz supra). However, methods employing antibiotic agents can have the problem of the emergence of bacterial strains which are resistant to these agents. (Hopkins, R. J. and J. G. Morris, supra). These limitations demonstrate that new more effective methods are needed to combat *H. pylori* infections *in vivo*. In particular, the design of new vaccines that may prevent infection by this bacterium is highly desirable.

Summary of the Invention

This invention relates to novel genes, e.g., genes encoding polypeptides such as bacterial surface proteins, from the organism *Helicobacter pylori* (*H. pylori*), and other related genes, their products, and uses thereof. The nucleic acids and peptides of the present invention have utility for diagnostic and therapeutics for *H. pylori* and other *Helicobacter* species. They can also be used to detect the presence of *H. pylori* and other *Helicobacter* species in a sample; and for use in screening compounds for the ability to interfere with the *H. pylori* life cycle or to inhibit *H. pylori* infection. More specifically, this invention features compositions of nucleic acids corresponding to entire coding sequences of *H. pylori* proteins, including surface or secreted proteins or parts thereof, nucleic acids capable of binding mRNA from *H. pylori* proteins to block protein translation, and methods for producing *H. pylori* proteins or parts thereof using peptide synthesis and recombinant DNA techniques. This invention also features antibodies and nucleic acids useful as probes to detect *H. pylori* infection. In addition, vaccine compositions and methods for the protection or treatment of infection by *H. pylori* are within the scope of this invention.

Detailed Description of the Drawings

Figure 1 depicts an amino acid sequence alignment of five *H. pylori* proteins (depicted in the single letter amino acid code and designated by their amino acid Sequence ID Numbers; shown N-terminal to C-terminal, left to right).

Figure 2 depicts the N-terminal portion of three *H. pylori* proteins (depicted in the single letter amino acid code and designated by their amino acid Sequence ID Numbers; shown N-terminal to C-terminal, left to right).

Detailed Description of the Invention

In one aspect, the invention features a recombinant or substantially pure preparation of *H. pylori* polypeptide of SEQ ID NO: 98. The invention also includes substantially pure nucleic acid encoding an *H. pylori* polypeptide of SEQ ID NO: 98, such nucleic acid is contained in SEQ ID NO: 1. The *H. pylori* polypeptide sequences of the invention described herein are contained in the Sequence Listing, and the nucleic acids encoding *H. pylori* polypeptides of the invention are contained in the Sequence Listing.

10 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 99, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 2.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 100, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 3.

15 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 101, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 4.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 102, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 5.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 103, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 6.

25 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 104, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 7.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 105, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 8.

30 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 106, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 9.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 107, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 10.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 108, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 11.

5 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 109, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 12.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 110, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 13.

10 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 111, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 14.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 112, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 15.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 113, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 16.

20 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 114, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 17.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 115, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 18.

25 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 116, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 19.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 117, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 20.

30 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 118, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 21.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 119, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 22.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 120, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 23.

5 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 121, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 24.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 122, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 25.

10 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 123, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 26.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 124, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 27.

15 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 125, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 28.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 126, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 29.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 127, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 30.

25 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 128, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 31.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 129, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 32.

30 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 130, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 33.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 131, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 34.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 132, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 35.

5 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 133, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 36.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 134, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 37.

10 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 135, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 38.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 136, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 39.

15 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 137, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 40.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 138, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 41.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 139, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 42.

25 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 140, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 43.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 141, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 44.

30 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 142, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 45.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 143, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 46.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 144, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 47.

5 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 145, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 48.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 146, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 49.

10 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 147, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 50.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 148, 15 such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 51.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 149, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 52.

20 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 150, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 53.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 151, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 54.

25 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 152, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 55.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 153, 30 such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 56.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 154, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 57.

35 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 155, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 58.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 156, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 59.

5 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 157, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 60.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 158, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 61.

10 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 159, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 62.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 160, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 63.

15 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 161, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 64.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 162, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 65.

20 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 163, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 66.

25 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 164, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 67.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 165, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 68.

30 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 166, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 69.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 167, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 70.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 168, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 71.

5 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 169, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 72.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 170, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 73.

10 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 171, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 74.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 172, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 75.

15 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 173, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 76.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 174, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 77.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 175, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 78.

25 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 176, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 79.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 177, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 80.

30 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 178, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 81.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 179, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 82.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 180, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 83.

5 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 181, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 84.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 182, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 85.

10 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 183, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 86.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 184, 15 such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 87.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 185, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 88.

20 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 186, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 89.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 187, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 90.

25 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 188, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 91.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 189, 30 such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 92.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 190, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 93.

35 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 191, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 94.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 192, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 95.

5 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 193, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 96.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO: 194, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 97.

10 In another aspect, the invention features an isolated nucleic acid having a nucleotide sequence encoding an *H. pylori* polypeptide at least about 60% homologous to an amino acid sequence selected from the group consisting of SEQ ID NO: 98-SEQ ID NO: 194. In a preferred embodiment, the isolated nucleic acid includes a nucleotide sequence selected from the group consisting of SEQ ID NO: 1-SEQ ID NO: 97, or a
15 complement thereof.

In another aspect, the invention features an isolated nucleic acid having a nucleotide sequence encoding an *H. pylori* polypeptide selected from the group consisting of SEQ ID NO: 98-SEQ ID NO: 194.

20 In another aspect, the invention features an isolated nucleic acid which encodes an *H. pylori* polypeptide, having a nucleotide sequence at least about 60% homologous to a nucleotide sequence selected from the group consisting of SEQ ID NO: 1-SEQ ID NO: 97, or a complement thereof.

In another aspect, the invention features an isolated nucleic acid molecule encoding an *H. pylori* polypeptide, having a nucleotide sequence which hybridizes under
25 stringent hybridization conditions to a nucleic acid molecule having the nucleotide sequence selected from the group consisting of SEQ ID NO: 1-SEQ ID NO: 97, or a complement thereof.

In another aspect, the invention features an isolated nucleic acid having a nucleotide sequence of at least 8 nucleotides in length, wherein the sequence hybridizes
30 under stringent hybridization conditions to a nucleic acid having a nucleotide sequence selected from the group consisting of SEQ ID NO: 1-SEQ ID NO: 97, or a complement thereof.

Particularly preferred is an isolated nucleic acid having a nucleotide sequence encoding an *H. pylori* cell envelope polypeptide or a fragment thereof, the nucleic acid
35 selected from the group consisting of SEQ ID NO: 63, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 50, SEQ ID NO: 51, SEQ ID NO: 61,

SEQ ID NO: 79, SEQ ID NO: 80, SEQ ID NO: 84, SEQ ID NO: 85, SEQ ID NO: 91, SEQ ID NO: 94, SEQ ID NO: 5, SEQ ID NO: 11, SEQ ID NO: 26, SEQ ID NO: 36, SEQ ID NO: 42, SEQ ID NO: 52, SEQ ID NO: 22, SEQ ID NO: 29, SEQ ID NO: 30, SEQ ID NO: 65, SEQ ID NO: 66, SEQ ID NO: 48, SEQ ID NO: 49, SEQ ID NO: 17,
5 SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 43, SEQ ID NO: 44, SEQ ID NO: 38, SEQ ID NO: 39, SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 6, SEQ ID NO: 34, SEQ ID NO: 35, SEQ ID NO: 60, SEQ ID NO: 69, and SEQ ID NO: 83, or a complement thereof.

10 In one embodiment, the *H. pylori* cell envelope polypeptide or a fragment thereof is an *H. pylori* flagella-associated polypeptide or a fragment thereof encoded by a nucleic acid having a nucleotide sequence of SEQ ID NO: 63, or a complement thereof.

In another embodiment, the *H. pylori* cell envelope polypeptide or a fragment thereof is an *H. pylori* inner membrane polypeptide or a fragment thereof encoded by a nucleic acid selected from the group consisting of SEQ ID NO: 48, SEQ ID NO: 49,
15 SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 43, SEQ ID NO: 44, SEQ ID NO: 38, and SEQ ID NO: 39, or a complement thereof.

In another embodiment, the *H. pylori* inner membrane polypeptide or a fragment thereof is an *H. pylori* polypeptide or a fragment thereof involved in transport encoded by a nucleic acid selected from the group consisting of SEQ ID NO: 48, SEQ ID NO:
20 49, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 43, and SEQ ID NO: 44, or a complement thereof.

In another embodiment, the *H. pylori* cell envelope polypeptide or a fragment thereof is an *H. pylori* outer membrane polypeptide or a fragment thereof encoded by a nucleic acid selected from the group consisting of SEQ ID NO: 7, SEQ ID NO: 8, SEQ
25 ID NO: 9, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 50, SEQ ID NO: 51, SEQ ID NO: 61, SEQ ID NO: 79, SEQ ID NO: 80, SEQ ID NO: 84, SEQ ID NO: 85, SEQ ID NO: 91, SEQ ID NO: 94, SEQ ID NO: 5, SEQ ID NO: 11, SEQ ID NO: 26, SEQ ID NO: 36, SEQ ID NO: 42, SEQ ID NO: 52, SEQ ID NO: 22, SEQ ID NO: 29, SEQ ID NO: 30, SEQ ID
30 NO: 65, and SEQ ID NO: 66, or a complement thereof.

In another embodiment, the *H. pylori* outer membrane polypeptide or a fragment thereof is an *H. pylori* polypeptide having a terminal phenylalanine residue or a fragment thereof encoded by a nucleic acid selected from the group consisting of SEQ
35 ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 36, SEQ ID NO: 42, SEQ ID NO: 50, SEQ ID NO: 51, SEQ ID

NO: 52, SEQ ID NO: 61, SEQ ID NO: 79, SEQ ID NO: 80, SEQ ID NO: 84, SEQ ID NO: 85, SEQ ID NO: 91, and SEQ ID NO: 94, or a complement thereof.

In another embodiment, *H. pylori* outer membrane polypeptide or a fragment thereof is an *H. pylori* polypeptide having a terminal phenylalanine residue and a C-terminal tyrosine cluster or a fragment thereof encoded by a nucleic acid selected from the group consisting of SEQ ID NO: 11, SEQ ID NO: 26, SEQ ID NO: 36, SEQ ID NO: 42, and SEQ ID NO: 52, or a complement thereof.

Particularly preferred is an isolated nucleic acid having a nucleotide sequence encoding an *H. pylori* cell envelope polypeptide or a fragment thereof selected from the group consisting of SEQ ID NO: 160, SEQ ID NO: 104, SEQ ID NO: 105, SEQ ID NO: 106, SEQ ID NO: 110, SEQ ID NO: 111, SEQ ID NO: 120, SEQ ID NO: 121, SEQ ID NO: 124, SEQ ID NO: 125, SEQ ID NO: 147, SEQ ID NO: 148, SEQ ID NO: 158, SEQ ID NO: 176, SEQ ID NO: 177, SEQ ID NO: 181, SEQ ID NO: 182, SEQ ID NO: 188, SEQ ID NO: 191, SEQ ID NO: 102, SEQ ID NO: 108, SEQ ID NO: 123, SEQ ID NO: 133, SEQ ID NO: 139, SEQ ID NO: 149, SEQ ID NO: 119, SEQ ID NO: 126, SEQ ID NO: 127, SEQ ID NO: 162, SEQ ID NO: 163, SEQ ID NO: 145, SEQ ID NO: 146, SEQ ID NO: 114, SEQ ID NO: 115, SEQ ID NO: 116, SEQ ID NO: 140, SEQ ID NO: 141, SEQ ID NO: 135, SEQ ID NO: 136, SEQ ID NO: 98, SEQ ID NO: 99, SEQ ID NO: 103, SEQ ID NO: 131, SEQ ID NO: 132, SEQ ID NO: 157, SEQ ID NO: 166, and SEQ ID NO: 180.

In another embodiment, the *H. pylori* cell envelope polypeptide or a fragment thereof is an *H. pylori* flagella-associated polypeptide or a fragment thereof having an amino acid sequence of SEQ ID NO: 160.

In another embodiment, the *H. pylori* cell envelope polypeptide or a fragment thereof is an *H. pylori* inner membrane polypeptide or a fragment thereof selected from the group consisting of SEQ ID NO: 145, SEQ ID NO: 146, SEQ ID NO: 114, SEQ ID NO: 115, SEQ ID NO: 116, SEQ ID NO: 140, SEQ ID NO: 141, SEQ ID NO: 135, and SEQ ID NO: 136.

In another embodiment, the *H. pylori* inner membrane polypeptide or a fragment thereof is an *H. pylori* polypeptide or a fragment thereof involved in transport selected from the group consisting of SEQ ID NO: 145, SEQ ID NO: 146, SEQ ID NO: 114, SEQ ID NO: 115, SEQ ID NO: 116, SEQ ID NO: 140, and SEQ ID NO: 141.

In another embodiment, the *H. pylori* cell envelope polypeptide or a fragment thereof is an *H. pylori* outer membrane polypeptide or a fragment thereof selected from the group consisting of SEQ ID NO: 104, SEQ ID NO: 105, SEQ ID NO: 106, SEQ ID NO: 110, SEQ ID NO: 111, SEQ ID NO: 120, SEQ ID NO: 121, SEQ ID NO: 124, SEQ ID NO: 125, SEQ ID NO: 147, SEQ ID NO: 148, SEQ ID NO: 158, SEQ ID NO: 176,

SEQ ID NO: 177, SEQ ID NO: 181, SEQ ID NO: 182, SEQ ID NO: 188, SEQ ID NO: 191, SEQ ID NO: 102, SEQ ID NO: 108, SEQ ID NO: 123, SEQ ID NO: 133, SEQ ID NO: 139, SEQ ID NO: 149, SEQ ID NO: 119, SEQ ID NO: 126, SEQ ID NO: 127, SEQ ID NO: 162, and SEQ ID NO: 163.

5 In another embodiment, the *H. pylori* outer membrane polypeptide or a fragment thereof is an *H. pylori* polypeptide having a terminal phenylalanine residue or a fragment thereof selected from the group consisting of SEQ ID NO: 104, SEQ ID NO: 105, SEQ ID NO: 106, SEQ ID NO: 108, SEQ ID NO: 110, SEQ ID NO: 111, SEQ ID NO: 120, SEQ ID NO: 121, SEQ ID NO: 123, SEQ ID NO: 124, SEQ ID NO: 125, SEQ ID NO: 133, SEQ ID NO: 139, SEQ ID NO: 147, SEQ ID NO: 148, SEQ ID NO: 149, SEQ ID NO: 158, SEQ ID NO: 176, SEQ ID NO: 177, SEQ ID NO: 181, SEQ ID NO: 182, SEQ ID NO: 188, and SEQ ID NO: 191.

15 In another embodiment, the *H. pylori* outer membrane polypeptide or a fragment thereof is an *H. pylori* polypeptide having a terminal phenylalanine residue and a C-terminal tyrosine cluster or a fragment thereof selected from the group consisting of SEQ ID NO: 108, SEQ ID NO: 123, SEQ ID NO: 133, SEQ ID NO: 139, and SEQ ID NO: 149.

20 Particularly preferred is an isolated nucleic acid having a nucleotide sequence encoding an *H. pylori* cytoplasmic polypeptide or a fragment thereof, wherein the nucleic acid is selected from the group consisting of SEQ ID NO: 57, SEQ ID NO: 58, SEQ ID NO: 86, SEQ ID NO: 87, SEQ ID NO: 88, SEQ ID NO: 89, SEQ ID NO: 92, and SEQ ID NO: 93, or a complement thereof.

25 In one embodiment, the *H. pylori* cytoplasmic polypeptide or a fragment thereof is an *H. pylori* polypeptide or a fragment thereof involved in mRNA translation, wherein the nucleic acid is selected from the group consisting of SEQ ID NO: 57 and SEQ ID NO: 58, or a complement thereof.

30 In another embodiment, the *H. pylori* cytoplasmic polypeptide or a fragment thereof is an *H. pylori* polypeptide or a fragment thereof involved in genome replication, transcription, recombination and repair, wherein the nucleic acid is selected from the group consisting of SEQ ID NO: 86, SEQ ID NO: 87, or a complement thereof.

Particularly preferred is an isolated nucleic acid having a nucleotide sequence encoding an *H. pylori* cytoplasmic polypeptide or a fragment thereof selected from the group consisting of SEQ ID NO: 154, SEQ ID NO: 155, SEQ ID NO: 183, SEQ ID NO: 184, SEQ ID NO: 185, SEQ ID NO: 186, SEQ ID NO: 189, and SEQ ID NO: 190.

35 In one embodiment, the *H. pylori* cytoplasmic polypeptide or a fragment thereof is an *H. pylori* polypeptide or a fragment thereof involved in mRNA translation selected from the group consisting of SEQ ID NO: 154 and SEQ ID NO: 155.

In another embodiment, the *H. pylori* cytoplasmic polypeptide or a fragment thereof is an *H. pylori* polypeptide or a fragment thereof involved in genome replication, transcription, recombination and repair selected from the group consisting of SEQ ID NO: 183 and SEQ ID NO: 184.

5 Particularly preferred is an isolated nucleic acid having a nucleotide sequence encoding an *H. pylori* secreted polypeptide or a fragment thereof, the nucleic acid selected from the group consisting of SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 20, SEQ ID NO: 25, SEQ ID NO: 31, SEQ ID NO: 32, SEQ ID NO: 45, SEQ ID NO: 46, SEQ ID NO: 53, SEQ ID NO: 64, SEQ ID NO: 67,
10 SEQ ID NO: 70, SEQ ID NO: 77, SEQ ID NO: 78, SEQ ID NO: 81, SEQ ID NO: 82, SEQ ID NO: 90, SEQ ID NO: 95, and SEQ ID NO: 97, or a complement thereof.

Particularly preferred is an isolated nucleic acid having a nucleotide sequence encoding an *H. pylori* secreted polypeptide or a fragment thereof selected from the group consisting of SEQ ID NO: 100, SEQ ID NO: 101, SEQ ID NO: 107, SEQ ID NO: 109,
15 SEQ ID NO: 117, SEQ ID NO: 122, SEQ ID NO: 128, SEQ ID NO: 129, SEQ ID NO: 142, SEQ ID NO: 143, SEQ ID NO: 150, SEQ ID NO: 161, SEQ ID NO: 164, SEQ ID NO: 167, SEQ ID NO: 174, SEQ ID NO: 175, SEQ ID NO: 178, SEQ ID NO: 179, SEQ ID NO: 187, SEQ ID NO: 192, and SEQ ID NO: 194.

Particularly preferred is an isolated nucleic acid having a nucleotide sequence
20 encoding an *H. pylori* cellular polypeptide or a fragment thereof, the nucleic acid selected from the group consisting of SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 21, SEQ ID NO: 33, SEQ ID NO: 37, SEQ ID NO: 40, SEQ ID NO: 41, SEQ ID NO: 47, SEQ ID NO: 54, SEQ ID NO: 55, SEQ ID NO: 56, SEQ ID NO: 59, SEQ ID NO: 62, SEQ ID NO: 68, SEQ ID NO: 71, SEQ ID NO: 72, SEQ ID NO: 73, SEQ ID NO: 74,
25 SEQ ID NO: 75, SEQ ID NO: 76, and SEQ ID NO: 96, or a complement thereof.

Particularly preferred is an isolated nucleic acid having a nucleotide sequence encoding an *H. pylori* cellular polypeptide or a fragment thereof selected from the group consisting of SEQ ID NO: 112, SEQ ID NO: 113, SEQ ID NO: 118, SEQ ID NO: 130, SEQ ID NO: 134, SEQ ID NO: 137, SEQ ID NO: 138, SEQ ID NO: 144, SEQ ID NO:
30 151, SEQ ID NO: 152, SEQ ID NO: 153, SEQ ID NO: 156, SEQ ID NO: 159, SEQ ID NO: 165, SEQ ID NO: 168, SEQ ID NO: 169, SEQ ID NO: 170, SEQ ID NO: 171, SEQ ID NO: 172, SEQ ID NO: 173, and SEQ ID NO: 193.

In another aspect, the invention features a probe having a nucleotide sequence consisting of at least 8 nucleotides of a nucleotide sequence selected from the group
35 consisting of SEQ ID NO: 1-SEQ ID NO: 97, or a complement thereof.

In another aspect, the invention features an isolated *H. pylori* polypeptide having an amino acid sequence at least about 60% homologous to an *H. pylori* polypeptide selected from the group consisting of SEQ ID NO: 98-SEQ ID NO: 194.

5 In another aspect, the invention features an isolated *H. pylori* polypeptide which is encoded by a nucleic acid having a nucleotide sequence at least about 60% homologous to a nucleotide sequence selected from the group consisting of SEQ ID NO: 1-SEQ ID NO: 97. In one embodiment, the isolated *H. pylori* polypeptide is encoded by a nucleotide sequence selected from the group consisting of SEQ ID NO: 1-SEQ ID NO: 97.

10 In another aspect, the invention features an isolated *H. pylori* polypeptide which is encoded by a nucleic acid which hybridizes under stringent hybridization conditions to a nucleic acid selected from the group consisting of SEQ ID NO: 1-SEQ ID NO: 97, or a complement thereof.

In another aspect, the invention features an isolated *H. pylori* polypeptide having
15 an amino acid sequence selected from the group consisting of SEQ ID NO: 97-SEQ ID NO: 194.

Particularly preferred is an isolated *H. pylori* cell envelope polypeptide or a fragment thereof, wherein the polypeptide is selected from the group consisting of SEQ ID NO: 160, SEQ ID NO: 104, SEQ ID NO: 105, SEQ ID NO: 106, SEQ ID NO: 110,
20 SEQ ID NO: 111, SEQ ID NO: 120, SEQ ID NO: 121, SEQ ID NO: 124, SEQ ID NO: 125, SEQ ID NO: 147, SEQ ID NO: 148, SEQ ID NO: 158, SEQ ID NO: 176, SEQ ID NO: 177, SEQ ID NO: 181, SEQ ID NO: 182, SEQ ID NO: 188, SEQ ID NO: 191, SEQ ID NO: 102, SEQ ID NO: 108, SEQ ID NO: 123, SEQ ID NO: 133, SEQ ID NO: 139, SEQ ID NO: 149, SEQ ID NO: 119, SEQ ID NO: 126, SEQ ID NO: 127, SEQ ID NO:
25 162, SEQ ID NO: 163, SEQ ID NO: 145, SEQ ID NO: 146, SEQ ID NO: 114, SEQ ID NO: 115, SEQ ID NO: 116, SEQ ID NO: 140, SEQ ID NO: 141, SEQ ID NO: 135, SEQ ID NO: 136, SEQ ID NO: 98, SEQ ID NO: 99, SEQ ID NO: 103, SEQ ID NO: 131, SEQ ID NO: 132, SEQ ID NO: 157, SEQ ID NO: 166, and SEQ ID NO: 180.

In one embodiment, the *H. pylori* cell envelope polypeptide or a fragment thereof
30 is an *H. pylori* flagella-associated polypeptide or a fragment thereof having an amino acid sequence of SEQ ID NO: 160.

In another embodiment, the *H. pylori* cell envelope polypeptide or a fragment thereof is an *H. pylori* inner membrane polypeptide or a fragment thereof selected from the group consisting of SEQ ID NO: 145, SEQ ID NO: 146, SEQ ID NO: 114, SEQ ID
35 NO: 115, SEQ ID NO: 116, SEQ ID NO: 140, SEQ ID NO: 141, SEQ ID NO: 135, and SEQ ID NO: 136.

In another embodiment, the *H. pylori* inner membrane polypeptide or a fragment thereof is an *H. pylori* polypeptide or a fragment thereof involved in transport selected from the group consisting of SEQ ID NO: 145, SEQ ID NO: 146, SEQ ID NO: 114, SEQ ID NO: 115, SEQ ID NO: 116, SEQ ID NO: 140, SEQ ID NO: 141, SEQ ID NO: 135, and SEQ ID NO: 136.

In another embodiment, the *H. pylori* cell envelope polypeptide or a fragment thereof is an *H. pylori* outer membrane polypeptide or a fragment thereof selected from the group consisting of SEQ ID NO: 104, SEQ ID NO: 105, SEQ ID NO: 106, SEQ ID NO: 110, SEQ ID NO: 111, SEQ ID NO: 120, SEQ ID NO: 121, SEQ ID NO: 124, SEQ ID NO: 125, SEQ ID NO: 147, SEQ ID NO: 148, SEQ ID NO: 158, SEQ ID NO: 176, SEQ ID NO: 177, SEQ ID NO: 181, SEQ ID NO: 182, SEQ ID NO: 188, SEQ ID NO: 191, SEQ ID NO: 102, SEQ ID NO: 108, SEQ ID NO: 123, SEQ ID NO: 133, SEQ ID NO: 139, SEQ ID NO: 149, SEQ ID NO: 119, SEQ ID NO: 126, SEQ ID NO: 127, SEQ ID NO: 162, and SEQ ID NO: 163.

In another embodiment, the *H. pylori* outer membrane polypeptide or a fragment thereof is an *H. pylori* polypeptide having a terminal phenylalanine residue or a fragment thereof selected from the group consisting of SEQ ID NO: 104, SEQ ID NO: 105, SEQ ID NO: 106, SEQ ID NO: 108, SEQ ID NO: 110, SEQ ID NO: 111, SEQ ID NO: 120, SEQ ID NO: 121, SEQ ID NO: 123, SEQ ID NO: 124, SEQ ID NO: 125, SEQ ID NO: 133, SEQ ID NO: 139, SEQ ID NO: 147, SEQ ID NO: 148, SEQ ID NO: 149, SEQ ID NO: 158, SEQ ID NO: 176, SEQ ID NO: 177, SEQ ID NO: 181, SEQ ID NO: 182, SEQ ID NO: 188, and SEQ ID NO: 191.

In another embodiment, the *H. pylori* outer membrane polypeptide or a fragment thereof is an *H. pylori* polypeptide having a terminal phenylalanine residue and a C-terminal tyrosine cluster or a fragment thereof selected from the group consisting of SEQ ID NO: 108, SEQ ID NO: 123, SEQ ID NO: 133, SEQ ID NO: 139, and SEQ ID NO: 149.

Particularly preferred is an isolated *H. pylori* cell envelope polypeptide or a fragment thereof, wherein the polypeptide is encoded by a nucleic acid selected from the group consisting of SEQ ID NO: 63, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 50, SEQ ID NO: 51, SEQ ID NO: 61, SEQ ID NO: 79, SEQ ID NO: 80, SEQ ID NO: 84, SEQ ID NO: 85, SEQ ID NO: 91, SEQ ID NO: 94, SEQ ID NO: 5, SEQ ID NO: 11, SEQ ID NO: 26, SEQ ID NO: 36, SEQ ID NO: 42, SEQ ID NO: 52, SEQ ID NO: 22, SEQ ID NO: 29, SEQ ID NO: 30, SEQ ID NO: 65, SEQ ID NO: 66, SEQ ID NO: 48, SEQ ID NO: 49, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 43, SEQ ID NO: 44, SEQ ID NO: 38, SEQ ID NO: 39,

SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 6, SEQ ID NO: 34, SEQ ID NO: 35, SEQ ID NO: 60, and SEQ ID NO: 69, SEQ ID NO: 83.

In one embodiment, the *H. pylori* cell envelope polypeptide or a fragment thereof is an *H. pylori* flagella-associated polypeptide or a fragment thereof encoded by a nucleic acid having a nucleotide sequence of SEQ ID NO: 63.

In another embodiment, the *H. pylori* cell envelope polypeptide or a fragment thereof is an *H. pylori* inner membrane polypeptide or a fragment thereof encoded by a nucleic acid selected from the group consisting of SEQ ID NO: 48, SEQ ID NO: 49, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 43, SEQ ID NO: 44, SEQ ID NO: 38, and SEQ ID NO: 39.

In another embodiment, the *H. pylori* inner membrane polypeptide or a fragment thereof is an *H. pylori* polypeptide or a fragment thereof involved in transport encoded by a nucleic acid selected from the group consisting of SEQ ID NO: 48, SEQ ID NO: 49, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 43, and SEQ ID NO: 44.

In another embodiment, the *H. pylori* cell envelope polypeptide or a fragment thereof is an *H. pylori* outer membrane polypeptide or a fragment thereof encoded by a nucleic acid selected from the group consisting of SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 50, SEQ ID NO: 51, SEQ ID NO: 61, SEQ ID NO: 79, SEQ ID NO: 80, SEQ ID NO: 84, SEQ ID NO: 85, SEQ ID NO: 91, SEQ ID NO: 94, SEQ ID NO: 5, SEQ ID NO: 11, SEQ ID NO: 26, SEQ ID NO: 36, SEQ ID NO: 42, SEQ ID NO: 52, SEQ ID NO: 22, SEQ ID NO: 29, SEQ ID NO: 30, SEQ ID NO: 65, and SEQ ID NO: 66.

In another embodiment, the *H. pylori* outer membrane polypeptide or a fragment thereof is an *H. pylori* polypeptide having a terminal phenylalanine residue or a fragment thereof encoded by a nucleic acid selected from the group consisting of SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 36, SEQ ID NO: 42, SEQ ID NO: 50, SEQ ID NO: 51, SEQ ID NO: 52, SEQ ID NO: 61, SEQ ID NO: 79, SEQ ID NO: 80, SEQ ID NO: 84, SEQ ID NO: 85, SEQ ID NO: 91, and SEQ ID NO: 94.

In another embodiment, the *H. pylori* outer membrane polypeptide or a fragment thereof is an *H. pylori* polypeptide having a terminal phenylalanine residue and a C-terminal tyrosine cluster or a fragment thereof encoded by a nucleic acid selected from the group consisting of SEQ ID NO: 11, SEQ ID NO: 26, SEQ ID NO: 36, SEQ ID NO: 42, and SEQ ID NO: 52.

Particularly preferred is an isolated *H. pylori* cytoplasmic polypeptide or a fragment thereof, wherein the polypeptide is selected from the group consisting of SEQ ID NO: 154, SEQ ID NO: 155, SEQ ID NO: 183, SEQ ID NO: 184, SEQ ID NO: 185, SEQ ID NO: 186, SEQ ID NO: 189, and SEQ ID NO: 190.

5 In another embodiment, the *H. pylori* cytoplasmic polypeptide or a fragment thereof is an *H. pylori* polypeptide or a fragment thereof involved in mRNA translation selected from the group consisting of SEQ ID NO: 154 and SEQ ID NO: 155.

In another embodiment, the *H. pylori* cytoplasmic polypeptide or a fragment thereof is an *H. pylori* polypeptide or a fragment thereof involved in genome
10 replication, transcription, recombination and repair selected from the group consisting of SEQ ID NO: 183 and SEQ ID NO: 184.

Particularly preferred is an isolated *H. pylori* cytoplasmic polypeptide or a fragment thereof, wherein the polypeptide is encoded by a nucleic acid selected from the group consisting of SEQ ID NO: 57, SEQ ID NO: 58, SEQ ID NO: 86, SEQ ID NO: 87,
15 SEQ ID NO: 88, SEQ ID NO: 89, SEQ ID NO: 92, and SEQ ID NO: 93.

In one embodiment, the *H. pylori* cytoplasmic polypeptide or a fragment thereof is an *H. pylori* polypeptide or a fragment thereof involved in mRNA translation, wherein the polypeptide is encoded by a nucleic acid selected from the group consisting of SEQ ID NO: 57, and SEQ ID NO: 58.

20 In another embodiment, the *H. pylori* cytoplasmic polypeptide or a fragment thereof is an *H. pylori* polypeptide or a fragment thereof involved in genome replication, transcription, recombination and repair, wherein the polypeptide is encoded by a nucleic acid selected from the group consisting of SEQ ID NO: 86 and SEQ ID NO: 87.

Particularly preferred is an isolated *H. pylori* cellular polypeptide or a fragment
25 thereof, wherein the polypeptide is selected from the group consisting of SEQ ID NO: 112, SEQ ID NO: 113, SEQ ID NO: 118, SEQ ID NO: 130, SEQ ID NO: 134, SEQ ID NO: 137, SEQ ID NO: 138, SEQ ID NO: 144, SEQ ID NO: 151, SEQ ID NO: 152, SEQ ID NO: 153, SEQ ID NO: 156, SEQ ID NO: 159, SEQ ID NO: 165, SEQ ID NO: 168, SEQ ID NO: 169, SEQ ID NO: 170, SEQ ID NO: 171, SEQ ID NO: 172, SEQ ID NO:
30 173, and SEQ ID NO: 193.

Particularly preferred is an isolated *H. pylori* cellular polypeptide or a fragment thereof, wherein the polypeptide is encoded by a nucleic acid selected from the group consisting of SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 21, SEQ ID NO: 33, SEQ ID NO: 37, SEQ ID NO: 40, SEQ ID NO: 41, SEQ ID NO: 47, SEQ ID NO: 54, SEQ
35 ID NO: 55, SEQ ID NO: 56, SEQ ID NO: 59, SEQ ID NO: 62, SEQ ID NO: 68, SEQ ID NO: 71, SEQ ID NO: 72, SEQ ID NO: 73, SEQ ID NO: 74, SEQ ID NO: 75, SEQ ID NO: 76, and SEQ ID NO: 96.

Particularly preferred is an isolated *H. pylori* secreted polypeptide or a fragment thereof, wherein the polypeptide is selected from the group consisting of SEQ ID NO: 100, SEQ ID NO: 101, SEQ ID NO: 107, SEQ ID NO: 109, SEQ ID NO: 117, SEQ ID NO: 122, SEQ ID NO: 128, SEQ ID NO: 129, SEQ ID NO: 142, SEQ ID NO: 143, SEQ ID NO: 150, SEQ ID NO: 161, SEQ ID NO: 164, SEQ ID NO: 167, SEQ ID NO: 174, SEQ ID NO: 175, SEQ ID NO: 178, SEQ ID NO: 179, SEQ ID NO: 187, SEQ ID NO: 192, and SEQ ID NO: 194.

Particularly preferred is an isolated *H. pylori* secreted polypeptide or a fragment thereof, wherein the polypeptide is encoded by a nucleic acid selected from the group consisting of SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 20, SEQ ID NO: 25, SEQ ID NO: 31, SEQ ID NO: 32, SEQ ID NO: 45, SEQ ID NO: 46, SEQ ID NO: 53, SEQ ID NO: 64, SEQ ID NO: 67, SEQ ID NO: 70, SEQ ID NO: 77, SEQ ID NO: 78, SEQ ID NO: 81, SEQ ID NO: 82, SEQ ID NO: 90, SEQ ID NO: 95, and SEQ ID NO: 97.

In another aspect, the invention features a chimeric *H. pylori* polypeptide comprising at least two *H. pylori* polypeptides or fragments thereof, wherein the polypeptides are encoded by nucleic acid sequences selected from the group consisting of SEQ ID NO:1-SEQ ID NO:97.

In another aspect, the invention features a chimeric *H. pylori* polypeptide comprising at least two *H. pylori* polypeptides or fragments thereof, wherein the polypeptides are selected from the group consisting of SEQ ID NO:98-SEQ ID NO:194.

In another aspect, the invention features a fusion protein comprising an *H. pylori* polypeptide which comprises an amino acid sequence selected from the group consisting of SEQ ID NO: 98-SEQ ID NO: 194 operatively linked to a non-*H. pylori* polypeptide.

In another aspect, the invention features a vaccine formulation for prophylactic or therapeutic treatment of an *H. pylori* infection comprising an effective amount of at least one isolated nucleic acid of the invention.

In another aspect, the invention features a vaccine formulation for prophylactic or therapeutic treatment of an *H. pylori* infection comprising an effective amount of at least one *H. pylori* polypeptide of the invention.

Preferably, the vaccine formulation of the invention further includes a pharmaceutically acceptable carrier. In one embodiment, the pharmaceutically acceptable carrier includes an adjuvant. In another embodiment, the pharmaceutically acceptable carrier includes a delivery system, e.g., a live vector, e.g., a bacteria or a virus. In another embodiment, the pharmaceutically acceptable carrier includes both an adjuvant and a delivery system.

In another aspect, the invention features a method of treating or reducing a risk of *H. pylori* infection in a subject. The method includes administering to a subject a vaccine formulation of the invention, such that treatment or reduction of risk of *H. pylori* infection occurs.

5 In another aspect, the invention features a method of producing a vaccine formulation of the invention. The method includes combining at least one isolated *H. pylori* polypeptide or a fragment thereof selected from the group consisting of SEQ ID NO: 98-SEQ ID NO: 194 with a pharmaceutically acceptable carrier to thereby form a vaccine formulation.

10 In another aspect, the invention features a method of producing a vaccine formulation of the invention. The method includes culturing a cell under condition that permit expression of an *H. pylori* polypeptide or a fragment thereof selected from the group consisting of SEQ ID NO: 98-SEQ ID NO: 194; isolating the *H. pylori* polypeptide from the cell; and combining at least one isolated *H. pylori* polypeptide or a
15 fragment thereof with a pharmaceutically acceptable carrier to thereby form a vaccine formulation.

In another aspect, the invention pertains to any individual *H. pylori* polypeptide member or nucleic acid encoding such a member from the above-identified groups of *H. pylori* polypeptides.

20 In another aspect, the invention features nucleic acids capable of binding mRNA of *H. pylori*. Such nucleic acid is capable of acting as antisense nucleic acid to control the translation of mRNA of *H. pylori*. A further aspect features a nucleic acid which is capable of binding specifically to an *H. pylori* nucleic acid. These nucleic acids are also referred to herein as complements and have utility as probes and as capture reagents.

25 In another aspect, the invention features an expression system comprising an open reading frame corresponding to *H. pylori* nucleic acid. The nucleic acid further comprises a control sequence compatible with an intended host. The expression system is useful for making polypeptides corresponding to *H. pylori* nucleic acid.

30 In another aspect, the invention features a cell transformed with the expression system to produce *H. pylori* polypeptides.

In another aspect, the invention features a method of generating antibodies against *H. pylori* polypeptides which are capable of binding specifically to *H. pylori* polypeptides. Such antibodies have utility as reagents for immunoassays to evaluate the abundance and distribution of *H. pylori*-specific antigens.

35 In another aspect, the invention features a method of generating vaccines for immunizing an individual against *H. pylori*. The vaccination method includes: immunizing a subject with at least one *H. pylori* polypeptide according to the present

- 22 -

invention, e.g., a surface or secreted polypeptide, or active portion thereof, and a pharmaceutically acceptable carrier. Such vaccines have therapeutic and/or prophylactic utilities.

5 In another aspect, the invention provides a method for generating a vaccine comprising a modified immunogenic *H. pylori* polypeptide, e.g., a surface or secreted polypeptide, or active portion thereof, and a pharmacologically acceptable carrier.

In another aspect, the invention features a method of evaluating a compound, e.g. a polypeptide, e.g., a fragment of a host cell polypeptide, for the ability to bind an *H. pylori* polypeptide. The method includes: contacting the candidate compound with an
10 *H. pylori* polypeptide and determining if the compound binds or otherwise interacts with an *H. pylori* polypeptide. Compounds which bind *H. pylori* are candidates as activators or inhibitors of the bacterial life cycle. These assays can be performed *in vitro* or *in vivo*.

In another aspect, the invention features a method of evaluating a compound, e.g.
15 a polypeptide, e.g., a fragment of a host cell polypeptide, for the ability to bind an *H. pylori* nucleic acid, e.g., DNA or RNA. The method includes: contacting the candidate compound with an *H. pylori* nucleic acid and determining if the compound binds or otherwise interacts with an *H. pylori* polypeptide. Compounds which bind *H. pylori* are candidates as activators or inhibitors of the bacterial life cycle. These assays can be
20 performed *in vitro* or *in vivo*.

The invention features *H. pylori* polypeptides, preferably a substantially pure preparation of an *H. pylori* polypeptide, or a recombinant *H. pylori* polypeptide. In preferred embodiments: the polypeptide has biological activity; the polypeptide has an amino acid sequence at least 60%, 70%, 80%, 90%, 95%, 98%, or 99% identical or
25 homologous to an amino acid sequence of the invention contained in the Sequence Listing, preferably it has about 65% sequence identity with an amino acid sequence of the invention contained in the Sequence Listing, and most preferably it has about 92% to about 99% sequence identity with an amino acid sequence of the invention contained in the Sequence Listing; the polypeptide has an amino acid sequence essentially the same
30 as an amino acid sequence of the invention contained in the Sequence Listing; the polypeptide is at least 5, 10, 20, 50, 100, or 150 amino acid residues in length; the polypeptide includes at least 5, preferably at least 10, more preferably at least 20, more preferably at least 50, 100, or 150 contiguous amino acid residues of the invention contained in the Sequence Listing. In yet another preferred embodiment, the amino acid
35 sequence which differs in sequence identity by about 7% to about 8% from the *H. pylori* amino acid sequences of the invention contained in the Sequence Listing is also encompassed by the invention.

In preferred embodiments: the *H. pylori* polypeptide is encoded by a nucleic acid of the invention contained in the Sequence Listing, or by a nucleic acid having at least 60%, 70%, 80%, 90%, 95%, 98%, or 99% homology with a nucleic acid of the invention contained in the Sequence Listing.

5 In a preferred embodiment, the subject *H. pylori* polypeptide differs in amino acid sequence at 1, 2, 3, 5, 10 or more residues from a sequence of the invention contained in the Sequence Listing. The differences, however, are such that the *H. pylori* polypeptide exhibits an *H. pylori* biological activity, e.g., the *H. pylori* polypeptide retains a biological activity of a naturally occurring *H. pylori* polypeptide.

10 In preferred embodiments, the polypeptide includes all or a fragment of an amino acid sequence of the invention contained in the Sequence Listing; fused, in reading frame, to additional amino acid residues, preferably to residues encoded by genomic DNA 5' or 3' to the genomic DNA which encodes a sequence of the invention contained in the Sequence Listing.

15 In yet other preferred embodiments, the *H. pylori* polypeptide is a recombinant fusion protein having a first *H. pylori* polypeptide portion and a second polypeptide portion, e.g., a second polypeptide portion having an amino acid sequence unrelated to *H. pylori*. The second polypeptide portion can be, e.g., any of glutathione-S-transferase, a DNA binding domain, or a polymerase activating domain. In preferred embodiment
20 the fusion protein can be used in a two-hybrid assay.

Polypeptides of the invention include those which arise as a result of alternative transcription events, alternative RNA splicing events, and alternative translational and postranslational events.

The invention also encompasses an immunogenic component which includes at
25 least one *H. pylori* polypeptide in an immunogenic preparation; the immunogenic component being capable of eliciting an immune response specific for the *H. pylori* polypeptide, e.g., a humoral response, an antibody response, or a cellular response. In preferred embodiments, the immunogenic component comprises at least one antigenic determinant from a polypeptide of the invention contained in the Sequence Listing.

30 In another aspect, the invention provides a substantially pure nucleic acid having a nucleotide sequence which encodes an *H. pylori* polypeptide. In preferred embodiments: the encoded polypeptide has biological activity; the encoded polypeptide has an amino acid sequence at least 60%, 70%, 80%, 90%, 95%, 98%, or 99% homologous to an amino acid sequence of the invention contained in the Sequence
35 Listing; the encoded polypeptide has an amino acid sequence essentially the same as an amino acid sequence of the invention contained in the Sequence Listing; the encoded polypeptide is at least 5, 10, 20, 50, 100, or 150 amino acids in length; the encoded

polypeptide comprises at least 5, preferably at least 10, more preferably at least 20, more preferably at least 50, 100, or 150 contiguous amino acids of the invention contained in the Sequence Listing.

5 In preferred embodiments: the nucleic acid of the invention is that contained in the Sequence Listing; the nucleic acid is at least 60%, 70%, 80%, 90%, 95%, 98%, or 99% homologous with a nucleic acid sequence of the invention contained in the Sequence Listing.

10 In a preferred embodiment, the encoded *H. pylori* polypeptide differs (e.g., by amino acid substitution, addition or deletion of at least one amino acid residue) in amino acid sequence at 1, 2, 3, 5, 10 or more residues, from a sequence of the invention contained in the Sequence Listing. The differences, however, are such that: the *H. pylori* encoded polypeptide exhibits a *H. pylori* biological activity, e.g., the encoded *H. pylori* enzyme retains a biological activity of a naturally occurring *H. pylori*.

15 In preferred embodiments, the encoded polypeptide includes all or a fragment of an amino acid sequence of the invention contained in the Sequence Listing; fused, in reading frame, to additional amino acid residues, preferably to residues encoded by genomic DNA 5' or 3' to the genomic DNA which encodes a sequence of the invention contained in the Sequence Listing.

20 In preferred embodiments, the subject *H. pylori* nucleic acid will include a transcriptional regulatory sequence, e.g. at least one of a transcriptional promoter or transcriptional enhancer sequence, operably linked to the *H. pylori* gene sequence, e.g., to render the *H. pylori* gene sequence suitable for expression in a recombinant host cell.

25 In yet a further preferred embodiment, the nucleic acid which encodes an *H. pylori* polypeptide of the invention, hybridizes under stringent conditions to a nucleic acid probe corresponding to at least 8 consecutive nucleotides of the invention contained in the Sequence Listing; more preferably to at least 12 consecutive nucleotides of the invention contained in the Sequence Listing; more preferably to at least 20 consecutive nucleotides of the invention contained in the Sequence Listing; more preferably to at least 40 consecutive nucleotides of the invention contained in the Sequence Listing.

30 In a preferred embodiment, the nucleic acid encodes a peptide which differs by at least one amino acid residue from the sequences of the invention contained in the Sequence Listing.

35 In a preferred embodiment, the nucleic acid differs by at least one nucleotide from a nucleotide sequence of the invention contained in the Sequence Listing which encodes amino acids of the invention contained in the Sequence Listing.

In another aspect, the invention encompasses: a vector including a nucleic acid which encodes an *H. pylori* polypeptide or an *H. pylori* polypeptide variant as described

herein; a host cell transfected with the vector; and a method of producing a recombinant *H. pylori* polypeptide or *H. pylori* polypeptide variant; including culturing the cell, e.g., in a cell culture medium, and isolating the *H. pylori* or *H. pylori* polypeptide variant, e.g., from the cell or from the cell culture medium.

- 5 In another aspect, the invention features, a purified recombinant nucleic acid having at least 50%, 60%, 70%, 80%, 90%, 95%, 98%, or 99% homology with a sequence of the invention contained in the Sequence Listing.

- The invention also provides a probe or primer which includes a substantially purified oligonucleotide. The oligonucleotide includes a region of nucleotide sequence which hybridizes under stringent conditions to at least 8 consecutive nucleotides of sense or antisense sequence of the invention contained in the Sequence Listing, or naturally occurring mutants thereof. In preferred embodiments, the probe or primer further includes a label group attached thereto. The label group can be, e.g., a radioisotope, a fluorescent compound, an enzyme, and/or an enzyme co-factor.
- 10 15 Preferably the oligonucleotide is at least 8 and less than 10, 20, 30, 50, 100, or 150 nucleotides in length.

The invention also provides an isolated *H. pylori* polypeptide which is encoded by a nucleic acid which hybridizes under stringent hybridization conditions to a nucleic acid contained in the Sequence Listing.

- 20 The invention further provides nucleic acids, e.g., RNA or DNA, encoding a polypeptide of the invention. This includes double stranded nucleic acids as well as coding and antisense single strands.

- The *H. pylori* strain, from which genomic sequences have been sequenced, has been deposited in the American Type Culture Collection (ATCC # 55679; deposited by Genome Therapeutics Corporation, 100 Beaver Street, Waltham, MA 02154) as strain HP-J99.
- 25

- Included in the invention are: allelic variations; natural mutants; induced mutants; proteins encoded by DNA that hybridizes under high or low stringency conditions to a nucleic acid which encodes a polypeptide of the invention contained in the Sequence Listing (for definitions of high and low stringency see Current Protocols in Molecular Biology, John Wiley & Sons, New York, 1989, 6.3.1 - 6.3.6 and 6.4.1-6.4.10, hereby incorporated by reference); and, polypeptides specifically bound by antisera to *H. pylori* polypeptides, especially by antisera to an active site or binding domain of *H. pylori* polypeptide. The invention also includes fragments, preferably biologically active fragments. These and other polypeptides are also referred to herein as *H. pylori* polypeptide analogs or variants.
- 30 35

Putative functions have been determined for several of the *H. pylori* polypeptides of the invention, as shown in Table 1.

Accordingly, uses of the claimed *H. pylori* polypeptides based on these identified functions, as well as other functions as described herein, are also within the scope of the invention.

In addition, the present invention encompasses *H. pylori* polypeptides characterized as shown in Table 1 below, including: *H. pylori* cell envelope proteins, *H. pylori* secreted proteins, *H. pylori* cytoplasmic proteins and *H. pylori* cellular proteins. Members of these groups were identified by BLAST homology searches and by searches for secretion signal or transmembrane protein motifs. Polypeptides related by significant homology to the polypeptides of Table 1 are also considered to be classified in the manner of the homologs shown in Table 1.

TABLE 1

	ntSeqID[PCT]	aaSeqID[PCT]
A. CELL ENVELOPE		
A.1 Flagella-associated		
hp1p13939_24322162_f3_17	63	160
A.2 Outer membrane		
A.2.1 Terminal phe residue		
02ge10116_23462_f2_43	7	104
02ge10116_804550_f2_44	8	105
02ge41622_14875000_c2_65	9	106
01cp20708_214843_c2_49	13	110
01cp20708_4960952_c1_43	14	111
06ae11016_4729625_c3_68	23	120
06ep10615_49068_c2_87	24	121
06gp71906_35158328_f3_85	27	124
06gp71906_3941642_f2_70	28	125
13ae10610_156411_c3_33	50	147
13ae10610_6522827_c3_37	51	148
hp4e53394_11798952_c2_101	61	158
06ge20501_4298568_c3_53	79	176
11ae12004_3367666_c2_41	80	177
hp7e10433_5345837_c3_13	84	181
14ce61516_2460911_f2_9	85	182
11ap20714_2077_c3_103	91	188
02cp10615_21908138_f1_4	94	191
A.2.2 No terminal phe residue		
07gp11909_26460892_f2_6	5	102
A.2.3 Phe and Tyr cluster at C-terminus		
02ge41622_34176513_c1_50	11	108

06gp71906_20486556_f2_65	26	123
hp7e10520_14728137_f1_1	36	133
02ae31010_417818_f3_29	42	139
13ae10610_26855313_f3_15	52	149
A.2.4 Via homology		
hp5p15212_13729635_c3_35	22	119
07ee11402_1046877_c3_100	29	126
14ee41924_1046877_c3_104	30	127
hp1p13939_21641016_f1_1	65	162
hp4p62853_4766691_f3_23	66	163
A.3 Inner membrane		
A.3.1 Proteins involved in transport		
06cp30603_664083_c1_94	48	145
09cp10713_36359687_c1_119	49	146
04ep41903_16667055_c1_37	17	114
04ep41903_19689182_c1_43	18	115
14ce31519_24650009_c1_17	19	116
09ce10413_26734687_f3_23	43	140
hp6p10904_6726062_f3_13	44	141
A.3.2 Other inner membrane proteins		
02ae31010_16679640_f2_21	38	135
07ee50709_16679640_f3_60	39	136
A.4 Other cell envelope proteins		
01ce61016_1056562_c3_123	1	98
09cp61003_16619192_c2_83	2	99
02ge10116_15632000_c2_114	6	103
04ae61517_12345837_f2_4	34	131
04ae61517_21744091_f3_5	35	132
hp4e13394_26750068_c3_113	60	157
hp5p15575_1053590_c1_35	69	166
hp7e10433_5345837_c2_8	83	180
B. CYTOPLASMIC PROTEINS		
B.1. Proteins involved in mRNA translation		
hp3e10946_32609412_f3_4	57	154
hp3e10946_34175837_f3_3	58	155
B.2 Proteins involved in genome replication, transcription, recombination and repair		
14ce61516_12600937_f2_11	86	183
14cp11908_25402267_c3_104	87	184
B.3 Other cytoplasmic proteins		
05ce10910_23712780_c1_4	88	185
hp7e10192_23712780_f2_5	89	186
11ap20714_34663910_f3_29	92	189
hp8e10065_4962812_f2_18	93	190
C. SECRETED PROTEINS		
01ce61016_23593955_c3_140	3	100

- 28 -

09cp61003_23593955_c1_79	4	101
02ge41622_20730462_f1_19	10	107
01cp20708_10628177_c2_50	12	109
05ae30220_24415693_c3_175	20	117
06gp10409_4015687_f2_11	25	122
hp2e10911_10213593_c1_73	31	128
hp2e10911_35567005_c2_88	32	129
09ze10333_1457137_f3_11	45	142
06cp30603_10744075_c3_136	46	143
12ae10622_30273255_f1_13	53	150
05ce10208_4707035_c2_17	64	161
06ep30223_176437_c2_134	67	164
hp5p15575_26016387_f2_16	70	167
hp6p12244_4881375_c3_97	77	174
06ce20610_34647187_c2_33	78	175
hp7e10433_36339535_f3_3	81	178
hp7e10433_36339535_f3_3	82	179
hp7e10420_24391078_f1_3	90	187
02ce71018_35720091_c3_4	95	192
hp6e10363_30517031_f3_3	97	194
D. OTHER CELLULAR PROTEINS		
01ae11010_26437877_c2_52	15	112
hp4p33322_5891077_c2_45	16	113
hp3p21118_54628_c3_3	21	118
02ae31010_1064125_f1_11	33	130
hp2e10911_15680337_c3_105	37	134
hp2e10911_24804577_c3_104	40	137
hp2e10911_32234750_c1_68	41	138
06cp30603_26070252_c3_140	47	144
03ae10804_235286_f3_19	54	151
09ge11604_4804692_c1_8	55	152
hp2p10610_21987687_c2_5	56	153
hp4e13394_26182793_f2_45	59	156
hp4e53394_2082126_c2_102	62	159
06ep30223_25402187_c1_112	68	165
hp6e10491_12712706_f3_12	71	168
hp6p12129_12542880_c3_29	72	169
hp6p12129_17067265_c3_29	73	170
hp6p12129_214055_f1_2	74	171
hp6p12129_214055_f3_17	75	172
hp6p12244_33492712_c3_88	76	173
hp1e13054_22360653_f2_4	96	193

[In Table 1, "nt" represents nucleotide Seq. ID number and "aa" represents amino acid Seq. ID number]

Definitions

The terms "purified polypeptide" and "isolated polypeptide" and "a substantially pure preparation of a polypeptide" are used interchangeably herein and, as used herein, mean a polypeptide that has been substantially, and preferably completely, separated
5 from other proteins, lipids, and nucleic acids with which it naturally occurs. Preferably, the polypeptide is also separated from substances, e.g., antibodies or gel matrix, e.g., polyacrylamide, which are used to purify it. Preferably, the polypeptide constitutes at least 10, 20, 50 70, 80 or 95% dry weight of the purified preparation. Preferably, the preparation contains: sufficient polypeptide to allow protein sequencing; at least 1, 10,
10 or 100 µg of the polypeptide; at least 1, 10, or 100 mg of the polypeptide. Furthermore, the terms "purified polypeptide" and "isolated polypeptide" and "a substantially pure preparation of a polypeptide," as used herein, refer to both a polypeptide obtained from nature or produced by recombinant DNA techniques as described herein.

For example, an "isolated" or "purified" protein or biologically active portion
15 thereof is substantially free of cellular material or other contaminating proteins from the cell or tissue source from which the *H. pylori* protein is derived, or substantially free from chemical precursors or other chemicals when chemically synthesized. The language "substantially free of cellular material" includes preparations of *H. pylori* protein in which the protein is separated from cellular components of the cells from
20 which it is isolated or recombinantly produced. In one embodiment, the language "substantially free of cellular material" includes preparations of *H. pylori* protein having less than about 30% (by dry weight) of non-*H. pylori* protein (also referred to herein as a "contaminating protein"), more preferably less than about 20% of non-*H. pylori* protein, still more preferably less than about 10% of non-*H. pylori* protein, and most preferably
25 less than about 5% non-*H. pylori* protein. When the *H. pylori* protein or biologically active portion thereof is recombinantly produced, it is also preferably substantially free of culture medium, i.e., culture medium represents less than about 20%, more preferably less than about 10%, and most preferably less than about 5% of the volume of the protein preparation.

30 The language "substantially free of chemical precursors or other chemicals" includes preparations of *H. pylori* protein in which the protein is separated from chemical precursors or other chemicals which are involved in the synthesis of the protein. In one embodiment, the language "substantially free of chemical precursors or other chemicals" includes preparations of *H. pylori* protein having less than about 30% (by dry
35 weight) of chemical precursors or non-*H. pylori* chemicals, more preferably less than about 20% chemical precursors or non-*H. pylori* chemicals, still more preferably less

- 30 -

than about 10% chemical precursors or non-*H. pylori* chemicals, and most preferably less than about 5% chemical precursors or non-*H. pylori* chemicals.

A purified preparation of cells refers to, in the case of plant or animal cells, an *in vitro* preparation of cells and not an entire intact plant or animal. In the case of cultured
5 cells or microbial cells, it consists of a preparation of at least 10% and more preferably 50% of the subject cells.

A purified or isolated or a substantially pure nucleic acid, e.g., a substantially pure DNA, (are terms used interchangeably herein) is a nucleic acid which is one or both of the following: not immediately contiguous with both of the coding sequences with
10 which it is immediately contiguous (i.e., one at the 5' end and one at the 3' end) in the naturally-occurring genome of the organism from which the nucleic acid is derived; or which is substantially free of a nucleic acid with which it occurs in the organism from which the nucleic acid is derived. The term includes, for example, a recombinant DNA which is incorporated into a vector, e.g., into an autonomously replicating plasmid or
15 virus, or into the genomic DNA of a prokaryote or eukaryote, or which exists as a separate molecule (e.g., a cDNA or a genomic DNA fragment produced by PCR or restriction endonuclease treatment) independent of other DNA sequences. Substantially pure DNA also includes a recombinant DNA which is part of a hybrid gene encoding additional *H. pylori* DNA sequence.

20 A "contig" as used herein is a nucleic acid representing a continuous stretch of genomic sequence of an organism.

An "open reading frame", also referred to herein as ORF, is a region of nucleic acid which encodes a polypeptide. This region may represent a portion of a coding sequence or a total sequence and can be determined from a stop to stop codon or from a
25 start to stop codon.

As used herein, a "coding sequence" is a nucleic acid which is transcribed into messenger RNA and/or translated into a polypeptide when placed under the control of appropriate regulatory sequences. The boundaries of the coding sequence are determined by a translation start codon at the five prime terminus and a translation stop
30 code at the three prime terminus. A coding sequence can include but is not limited to messenger RNA, synthetic DNA, and recombinant nucleic acid sequences.

A "complement" of a nucleic acid as used herein refers to an anti-parallel or antisense sequence that participates in Watson-Crick base-pairing with the original sequence.

35 A "gene product" is a protein or structural RNA which is specifically encoded by a gene.

- 31 -

As used herein, the term "probe" refers to a nucleic acid, peptide or other chemical entity which specifically binds to a molecule of interest. Probes are often associated with or capable of associating with a label. A label is a chemical moiety capable of detection. Typical labels comprise dyes, radioisotopes, luminescent and chemiluminescent moieties, fluorophores, enzymes, precipitating agents, amplification sequences, and the like. Similarly, a nucleic acid, peptide or other chemical entity which specifically binds to a molecule of interest and immobilizes such molecule is referred herein as a "capture ligand". Capture ligands are typically associated with or capable of associating with a support such as nitro-cellulose, glass, nylon membranes, beads, particles and the like. The specificity of hybridization is dependent on conditions such as the base pair composition of the nucleotides, and the temperature and salt concentration of the reaction. These conditions are readily discernable to one of ordinary skill in the art using routine experimentation.

Homologous refers to the sequence similarity or sequence identity between two polypeptides or between two nucleic acid molecules. When a position in both of the two compared sequences is occupied by the same base or amino acid monomer subunit, e.g., if a position in each of two DNA molecules is occupied by adenine, then the molecules are homologous at that position. The percent of homology between two sequences is a function of the number of matching or homologous positions shared by the two sequences divided by the number of positions compared x 100. For example, if 6 of 10 of the positions in two sequences are matched or homologous then the two sequences are 60% homologous. By way of example, the DNA sequences ATTGCC and TATGGC share 50% homology. Generally, a comparison is made when two sequences are aligned to give maximum homology.

Nucleic acids are hybridizable to each other when at least one strand of a nucleic acid can anneal to the other nucleic acid under defined stringency conditions. Stringency of hybridization is determined by: (a) the temperature at which hybridization and/or washing is performed; and (b) the ionic strength and polarity of the hybridization and washing solutions. Hybridization requires that the two nucleic acids contain complementary sequences; depending on the stringency of hybridization, however, mismatches may be tolerated. Typically, hybridization of two sequences at high stringency (such as, for example, in a solution of 0.5X SSC, at 65° C) requires that the sequences be essentially completely homologous. Conditions of intermediate stringency (such as, for example, 2X SSC at 65 ° C) and low stringency (such as, for example 2X SSC at 55° C), require correspondingly less overall complementarity between the hybridizing sequences. (1X SSC is 0.15 M NaCl, 0.015 M Na citrate). A preferred, non-limiting example of stringent hybridization conditions are hybridization in 6X

sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2 X SSC, 0.1% SDS at 50-65°C.

The terms peptides, proteins, and polypeptides are used interchangeably herein.

As used herein, the term "surface protein" refers to all surface accessible
5 proteins, e.g. inner and outer membrane proteins, proteins adhering to the cell wall, and secreted proteins.

A polypeptide has *H. pylori* biological activity if it has one, two and preferably more of the following properties: (1) if when expressed in the course of an *H. pylori* infection, it can promote, or mediate the attachment of *H. pylori* to a cell; (2) it has an
10 enzymatic activity, structural or regulatory function characteristic of an *H. pylori* protein; (3) the gene which encodes it can rescue a lethal mutation in an *H. pylori* gene; (4) or it is immunogenic in a subject. A polypeptide has biological activity if it is an antagonist, agonist, or super-agonist of a polypeptide having one of the above-listed properties.

15 A biologically active fragment or analog is one having an *in vivo* or *in vitro* activity which is characteristic of the *H. pylori* polypeptides of the invention contained in the Sequence Listing, or of other naturally occurring *H. pylori* polypeptides, e.g., one or more of the biological activities described herein. Especially preferred are fragments which exist *in vivo*, e.g., fragments which arise from post transcriptional processing or
20 which arise from translation of alternatively spliced RNA's. Fragments include those expressed in native or endogenous cells as well as those made in expression systems, e.g., in CHO cells. Because peptides such as *H. pylori* polypeptides often exhibit a range of physiological properties and because such properties may be attributable to different portions of the molecule, a useful *H. pylori* fragment or *H. pylori* analog is one
25 which exhibits a biological activity in any biological assay for *H. pylori* activity. Most preferably the fragment or analog possesses 10%, preferably 40%, more preferably 60%, 70%, 80% or 90% or greater of the activity of *H. pylori*, in any *in vivo* or *in vitro* assay.

Analogues can differ from naturally occurring *H. pylori* polypeptides in amino acid sequence or in ways that do not involve sequence, or both. Non-sequence modifications
30 include changes in acetylation, methylation, phosphorylation, carboxylation, or glycosylation. Preferred analogs include *H. pylori* polypeptides (or biologically active fragments thereof) whose sequences differ from the wild-type sequence by one or more conservative amino acid substitutions or by one or more non-conservative amino acid substitutions, deletions, or insertions which do not substantially diminish the biological
35 activity of the *H. pylori* polypeptide. Conservative substitutions typically include the substitution of one amino acid for another with similar characteristics, e.g., substitutions within the following groups: valine, glycine; glycine, alanine; valine, isoleucine,

leucine; aspartic acid, glutamic acid; asparagine, glutamine; serine, threonine; lysine, arginine; and phenylalanine, tyrosine. Other conservative substitutions can be made in view of the table below.

5

TABLE 2
CONSERVATIVE AMINO ACID REPLACEMENTS

For Amino Acid	Code	Replace with any of
Alanine	A	D-Ala, Gly, beta-Ala, L-Cys, D-Cys
Arginine	R	D-Arg, Lys, D-Lys, homo-Arg, D-homo-Arg, Met, Ile, D-Met, D-Ile, Orn, D-Orn
Asparagine	N	D-Asn, Asp, D-Asp, Glu, D-Glu, Gln, D-Gln
Aspartic Acid	D	D-Asp, D-Asn, Asn, Glu, D-Glu, Gln, D-Gln
Cysteine	C	D-Cys, S-Me-Cys, Met, D-Met, Thr, D-Thr
Glutamine	Q	D-Gln, Asn, D-Asn, Glu, D-Glu, Asp, D-Asp
Glutamic Acid	E	D-Glu, D-Asp, Asp, Asn, D-Asn, Gln, D-Gln
Glycine	G	Ala, D-Ala, Pro, D-Pro, β -Ala, Acp
Isoleucine	I	D-Ile, Val, D-Val, Leu, D-Leu, Met, D-Met
Leucine	L	D-Leu, Val, D-Val, Leu, D-Leu, Met, D-Met
Lysine	K	D-Lys, Arg, D-Arg, homo-Arg, D-homo-Arg, Met, D-Met, Ile, D-Ile, Orn, D-Orn
Methionine	M	D-Met, S-Me-Cys, Ile, D-Ile, Leu, D-Leu, Val, D-Val
Phenylalanine	F	D-Phe, Tyr, D-Thr, L-Dopa, His, D-His, Trp, D-Trp, Trans-3,4, or 5-phenylproline, cis-3,4, or 5-phenylproline
Proline	P	D-Pro, L-I-thioazolidine-4-carboxylic acid, D-or L-1-oxazolidine-4-carboxylic acid
Serine	S	D-Ser, Thr, D-Thr, allo-Thr, Met, D-Met, Met(O), D-Met(O), L-Cys, D-Cys
Threonine	T	D-Thr, Ser, D-Ser, allo-Thr, Met, D-Met, Met(O), D-Met(O), Val, D-Val
Tyrosine	Y	D-Tyr, Phe, D-Phe, L-Dopa, His, D-His
Valine	V	D-Val, Leu, D-Leu, Ile, D-Ile, Met, D-Met

Other analogs within the invention are those with modifications which increase peptide stability; such analogs may contain, for example, one or more non-peptide bonds (which replace the peptide bonds) in the peptide sequence. Also included are: analogs that include residues other than naturally occurring L-amino acids, e.g., D-amino acids or non-naturally occurring or synthetic amino acids, e.g., β or γ amino acids; and cyclic analogs.

As used herein, the term "fragment", as applied to an *H. pylori* analog, will ordinarily be at least about 20 residues, more typically at least about 40 residues,

- 34 -

preferably at least about 60 residues in length. Fragments of *H. pylori* polypeptides can be generated by methods known to those skilled in the art. The ability of a candidate fragment to exhibit a biological activity of *H. pylori* polypeptide can be assessed by methods known to those skilled in the art as described herein. Also included are *H. pylori* polypeptides containing residues that are not required for biological activity of the peptide or that result from alternative mRNA splicing or alternative protein processing events.

An "immunogenic component" as used herein is a moiety, such as an *H. pylori* polypeptide, analog or fragment thereof, that is capable of eliciting a humoral and/or cellular immune response in a host animal alone or in combination with an adjuvant.

An "antigenic component" as used herein is a moiety, such as an *H. pylori* polypeptide, analog or fragment thereof, that is capable of binding to a specific antibody with sufficiently high affinity to form a detectable antigen-antibody complex.

As used herein, the term "transgene" means a nucleic acid (encoding, e.g., one or more polypeptides), which is partly or entirely heterologous, i.e., foreign, to the transgenic animal or cell into which it is introduced, or, is homologous to an endogenous gene of the transgenic animal or cell into which it is introduced, but which is designed to be inserted, or is inserted, into the cell's genome in such a way as to alter the genome of the cell into which it is inserted (e.g., it is inserted at a location which differs from that of the natural gene or its insertion results in a knockout). A transgene can include one or more transcriptional regulatory sequences and any other nucleic acid, such as introns, that may be necessary for optimal expression of the selected nucleic acid, all operably linked to the selected nucleic acid, and may include an enhancer sequence.

As used herein, the term "transgenic cell" refers to a cell containing a transgene. As used herein, a "transgenic animal" is any animal in which one or more, and preferably essentially all, of the cells of the animal includes a transgene. The transgene can be introduced into the cell, directly or indirectly by introduction into a precursor of the cell, by way of deliberate genetic manipulation, such as by a process of transformation of competent cells or by microinjection or by infection with a recombinant virus. This molecule may be integrated within a chromosome, or it may be extrachromosomally replicating DNA.

The term "antibody" as used herein is intended to include fragments thereof which are specifically reactive with *H. pylori* polypeptides.

As used herein, the term "cell-specific promoter" means a DNA sequence that serves as a promoter, i.e., regulates expression of a selected DNA sequence operably linked to the promoter, and which effects expression of the selected DNA sequence in specific cells of a tissue. The term also covers so-called "leaky" promoters, which

regulate expression of a selected DNA primarily in one tissue, but cause expression in other tissues as well.

Misexpression, as used herein, refers to a non-wild type pattern of gene expression. It includes: expression at non-wild type levels, i.e., over or under
5 expression; a pattern of expression that differs from wild type in terms of the time or stage at which the gene is expressed, e.g., increased or decreased expression (as compared with wild type) at a predetermined developmental period or stage; a pattern of expression that differs from wild type in terms of decreased expression (as compared with wild type) in a predetermined cell type or tissue type; a pattern of expression that
10 differs from wild type in terms of the splicing size, amino acid sequence, post-translational modification, or biological activity of the expressed polypeptide; a pattern of expression that differs from wild type in terms of the effect of an environmental stimulus or extracellular stimulus on expression of the gene, e.g., a pattern of increased or decreased expression (as compared with wild type) in the presence of an increase or
15 decrease in the strength of the stimulus.

As used herein, "host cells" and other such terms denoting microorganisms or higher eukaryotic cell lines cultured as unicellular entities refers to cells which can become or have been used as recipients for a recombinant vector or other transfer DNA, and include the progeny of the original cell which has been transfected. It is understood
20 by individuals skilled in the art that the progeny of a single parental cell may not necessarily be completely identical in genomic or total DNA compliment to the original parent, due to accident or deliberate mutation.

As used herein, the term "control sequence" refers to a nucleic acid having a base sequence which is recognized by the host organism to effect the expression of encoded
25 sequences to which they are ligated. The nature of such control sequences differs depending upon the host organism; in prokaryotes, such control sequences generally include a promoter, ribosomal binding site, terminators, and in some cases operators; in eukaryotes, generally such control sequences include promoters, terminators and in some instances, enhancers. The term control sequence is intended to include at a
30 minimum, all components whose presence is necessary for expression, and may also include additional components whose presence is advantageous, for example, leader sequences.

As used herein, the term "operably linked" refers to sequences joined or ligated to function in their intended manner. For example, a control sequence is operably linked
35 to coding sequence by ligation in such a way that expression of the coding sequence is achieved under conditions compatible with the control sequence and host cell.

- 36 -

The metabolism of a substance, as used herein, means any aspect of the, expression, function, action, or regulation of the substance. The metabolism of a substance includes modifications, e.g., covalent or non-covalent modifications of the substance. The metabolism of a substance includes modifications, e.g., covalent or non-covalent modification, the substance induces in other substances. The metabolism of a substance also includes changes in the distribution of the substance. The metabolism of a substance includes changes the substance induces in the distribution of other substances.

A "sample" as used herein refers to a biological sample, such as, for example, tissue or fluid isolated from an individual (including without limitation plasma, serum, cerebrospinal fluid, lymph, tears, saliva and tissue sections) or from *in vitro* cell culture constituents, as well as samples from the environment.

The practice of the invention will employ, unless otherwise indicated, conventional techniques of chemistry, molecular biology, microbiology, recombinant DNA, and immunology, which are within the skill of the art. Such techniques are explained fully in the literature. See e.g., Sambrook, Fritsch, and Maniatis, *Molecular Cloning; Laboratory Manual* 2nd ed. (1989); *DNA Cloning*, Volumes I and II (D.N Glover ed. 1985); *Oligonucleotide Synthesis* (M.J. Gait ed, 1984); *Nucleic Acid Hybridization* (B.D. Hames & S.J. Higgins eds. 1984); the series, *Methods in Enzymology* (Academic Press, Inc.), particularly Vol. 154 and Vol. 155 (Wu and Grossman, eds.) and *PCR-A Practical Approach* (McPherson, Quirke, and Taylor, eds., 1991).

I. Isolation of Nucleic Acids of *H. pylori* and Uses Therefor

H. pylori Genomic Sequence

This invention provides nucleotide sequences of the genome of *H. pylori* which thus comprises a DNA sequence library of *H. pylori* genomic DNA. The detailed description that follows provides nucleotide sequences of *H. pylori*, and also describes how the sequences were obtained and how ORFs and protein-coding sequences were identified. Also described are methods of using the disclosed *H. pylori* sequences in methods including diagnostic and therapeutic applications. Furthermore, the library can be used as a database for identification and comparison of medically important sequences in this and other strains of *H. pylori*.

To determine the genomic sequence of *H. pylori*, DNA was isolated from a strain of *H. pylori* (ATCC # 55679; deposited by Genome Therapeutics Corporation, 100 Beaver Street, Waltham, MA 02154) and mechanically sheared by nebulization to a

- 37 -

median size of 2 kb. Following size fractionation by gel electrophoresis, the fragments were blunt-ended, ligated to adapter oligonucleotides, and cloned into each of 20 different pMPX vectors (Rice et al., abstracts of Meeting of Genome Mapping and Sequencing, Cold Spring Harbor, NY, 5/11-5/15, 1994, p. 225) to construct a series of
5 "shotgun" subclone libraries.

DNA sequencing was achieved using multiplex sequencing procedures essentially as disclosed in Church et al., 1988, *Science* 240:185; U.S. Patents No. 4,942,124 and 5,149,625). DNA was extracted from pooled cultures and subjected to chemical or enzymatic sequencing. Sequencing reactions were resolved by
10 electrophoresis, and the products were transferred and covalently bound to nylon membranes. Finally, the membranes were sequentially hybridized with a series of labelled oligonucleotides complimentary to "tag" sequences present in the different shotgun cloning vectors. In this manner, a large number of sequences could be obtained from a single set of sequencing reactions. The cloning and sequencing procedures are
15 described in more detail in the Exemplification.

Individual sequence reads obtained in this manner were assembled using the FALCON™ program (Church et al., 1994, *Automated DNA Sequencing and Analysis*, J.C. Venter, ed., Academic Press) and PHRAP (P. Green, Abstracts of DOE Human Genome Program Contractor-Grantee Workshop V, Jan. 1996, p.157). The average
20 contig length was about 3-4 kb.

A variety of approaches are used to order the contigs so as to obtain a continuous sequence representing the entire *H. pylori* genome. Synthetic oligonucleotides are designed that are complementary to sequences at the end of each contig. These oligonucleotides may be hybridized to libraries of *H. pylori* genomic DNA in, for
25 example, lambda phage vectors or plasmid vectors to identify clones that contain sequences corresponding to the junctional regions between individual contigs. Such clones are then used to isolate template DNA and the same oligonucleotides are used as primers in polymerase chain reaction (PCR) to amplify junctional fragments, the nucleotide sequence of which is then determined.

The *H. pylori* sequences were analyzed for the presence of open reading frames (ORFs) comprising at least 180 nucleotides. As a result of the analysis of ORFs based on stop-to-stop codon reads, it should be understood that these ORFs may not correspond to the ORF of a naturally-occurring *H. pylori* polypeptide. These ORFs may contain start codons which indicate the initiation of protein synthesis of a naturally-
30 occurring *H. pylori* polypeptide. Such start codons within the ORFs provided herein can be identified by those of ordinary skill in the relevant art, and the resulting ORF and the encoded *H. pylori* polypeptide is within the scope of this invention. For example, within

the ORFs a codon such as AUG or GUG (encoding methionine or valine) which is part of the initiation signal for protein synthesis can be identified and the ORF modified to correspond to a naturally-occurring *H. pylori* polypeptide. The predicted coding regions were defined by evaluating the coding potential of such sequences with the program

5 GENEMARK™ (Borodovsky and McIninch, 1993, *Comp. Chem.* 17:123).

Other *H. pylori* Nucleic Acids

The nucleic acids of this invention may be obtained directly from the DNA of the above referenced *H. pylori* strain by using the polymerase chain reaction (PCR). See

10 "PCR, A Practical Approach" (McPherson, Quirke, and Taylor, eds., IRL Press, Oxford, UK, 1991) for details about the PCR. High fidelity PCR can be used to ensure a faithful DNA copy prior to expression. In addition, the authenticity of amplified products can be checked by conventional sequencing methods. Clones carrying the desired sequences described in this invention may also be obtained by screening the libraries by means of

15 the PCR or by hybridization of synthetic oligonucleotide probes to filter lifts of the library colonies or plaques as known in the art (see, e.g., Sambrook et al., *Molecular Cloning, A Laboratory Manual* 2nd edition, 1989, Cold Spring Harbor Press, NY).

It is also possible to obtain nucleic acids encoding *H. pylori* polypeptides from a cDNA library in accordance with protocols herein described. A cDNA encoding an *H.*

20 *pylori* polypeptide can be obtained by isolating total mRNA from an appropriate strain. Double stranded cDNAs can then be prepared from the total mRNA. Subsequently, the cDNAs can be inserted into a suitable plasmid or viral (e.g., bacteriophage) vector using any one of a number of known techniques. Genes encoding *H. pylori* polypeptides can also be cloned using established polymerase chain reaction techniques in accordance

25 with the nucleotide sequence information provided by the invention. The nucleic acids of the invention can be DNA or RNA. Preferred nucleic acids of the invention are contained in the Sequence Listing.

The nucleic acids of the invention can also be chemically synthesized using standard techniques. Various methods of chemically synthesizing polydeoxynucleotides

30 are known, including solid-phase synthesis which, like peptide synthesis, has been fully automated in commercially available DNA synthesizers (See e.g., Itakura et al. U.S. Patent No. 4,598,049; Caruthers et al. U.S. Patent No. 4,458,066; and Itakura U.S. Patent Nos. 4,401,796 and 4,373,071, incorporated by reference herein).

Nucleic acids isolated or synthesized in accordance with features of the present

35 invention are useful, by way of example, without limitation, as probes, primers, capture ligands, antisense genes and for developing expression systems for the synthesis of proteins and peptides corresponding to such sequences. As probes, primers, capture

ligands and antisense agents, the nucleic acid normally consists of all or part (approximately twenty or more nucleotides for specificity as well as the ability to form stable hybridization products) of the nucleic acids of the invention contained in the Sequence Listing. These uses are described in further detail below.

5 Probes

A nucleic acid isolated or synthesized in accordance with the sequence of the invention contained in the Sequence Listing can be used as a probe to specifically detect *H. pylori*. With the sequence information set forth in the present application, sequences of twenty or more nucleotides are identified which provide the desired inclusivity and
10 exclusivity with respect to *H. pylori*, and extraneous nucleic acids likely to be encountered during hybridization conditions. More preferably, the sequence will comprise at least twenty to thirty nucleotides to convey stability to the hybridization product formed between the probe and the intended target molecules.

Sequences larger than 1000 nucleotides in length are difficult to synthesize but
15 can be generated by recombinant DNA techniques. Individuals skilled in the art will readily recognize that the nucleic acids, for use as probes, can be provided with a label to facilitate detection of a hybridization product.

Nucleic acid isolated and synthesized in accordance with the sequence of the invention contained in the Sequence Listing can also be useful as probes to detect
20 homologous regions (especially homologous genes) of other *Helicobacter* species using appropriate stringency hybridization conditions as described herein.

Capture Ligand

For use as a capture ligand, the nucleic acid selected in the manner described above with respect to probes, can be readily associated with a support. The manner in
25 which nucleic acid is associated with supports is well known. Nucleic acid having twenty or more nucleotides in a sequence of the invention contained in the Sequence Listing have utility to separate *H. pylori* nucleic acid from the nucleic acid of each other and other organisms. Nucleic acid having twenty or more nucleotides in a sequence of the invention contained in the Sequence Listing can also have utility to separate other
30 *Helicobacter* species from each other and from other organisms. Preferably, the sequence will comprise at least twenty nucleotides to convey stability to the hybridization product formed between the probe and the intended target molecules. Sequences larger than 1000 nucleotides in length are difficult to synthesize but can be generated by recombinant DNA techniques.

- 40 -

Primers

Nucleic acid isolated or synthesized in accordance with the sequences described herein have utility as primers for the amplification of *H. pylori* nucleic acid. These nucleic acids may also have utility as primers for the amplification of nucleic acids in other *Helicobacter* species. With respect to polymerase chain reaction (PCR) techniques, nucleic acid sequences of ≥ 10 -15 nucleotides of the invention contained in the Sequence Listing have utility in conjunction with suitable enzymes and reagents to create copies of *H. pylori* nucleic acid. More preferably, the sequence will comprise twenty or more nucleotides to convey stability to the hybridization product formed between the primer and the intended target molecules. Binding conditions of primers greater than 100 nucleotides are more difficult to control to obtain specificity. High fidelity PCR can be used to ensure a faithful DNA copy prior to expression. In addition, amplified products can be checked by conventional sequencing methods.

The copies can be used in diagnostic assays to detect specific sequences, including genes from *H. pylori* and/or other *Helicobacter* species. The copies can also be incorporated into cloning and expression vectors to generate polypeptides corresponding to the nucleic acid synthesized by PCR, as is described in greater detail herein.

Antisense

Nucleic acid or nucleic acid-hybridizing derivatives isolated or synthesized in accordance with the sequences described herein have utility as antisense agents to prevent the expression of *H. pylori* genes. These sequences also have utility as antisense agents to prevent expression of genes of other *Helicobacter* species.

In one embodiment, nucleic acid or derivatives corresponding to *H. pylori* nucleic acids is loaded into a suitable carrier such as a liposome or bacteriophage for introduction into bacterial cells. For example, a nucleic acid having twenty or more nucleotides is capable of binding to bacteria nucleic acid or bacteria messenger RNA. Preferably, the antisense nucleic acid is comprised of 20 or more nucleotides to provide necessary stability of a hybridization product of non-naturally occurring nucleic acid and bacterial nucleic acid and/or bacterial messenger RNA. Nucleic acid having a sequence greater than 1000 nucleotides in length is difficult to synthesize but can be generated by recombinant DNA techniques. Methods for loading antisense nucleic acid in liposomes is known in the art as exemplified by U.S. Patent 4,241,046 issued December 23, 1980 to Papahadjopoulos et al.

II. Expression of *H. pylori* Nucleic Acids

Nucleic acid isolated or synthesized in accordance with the sequences described herein have utility to generate polypeptides. The nucleic acid of the invention exemplified in the Sequence Listing or fragments of the nucleic acid encoding active portions of *H. pylori* polypeptides can be cloned into suitable vectors or used to isolate nucleic acid. The isolated nucleic acid is combined with suitable DNA linkers and cloned into a suitable vector.

The function of a specific gene or operon can be ascertained by expression in a bacterial strain under conditions where the activity of the gene product(s) specified by the gene or operon in question can be specifically measured. Alternatively, a gene product may be produced in large quantities in an expressing strain for use as an antigen, an industrial reagent, for structural studies, etc. This expression can be accomplished in a mutant strain which lacks the activity of the gene to be tested, or in a strain that does not produce the same gene product(s). This includes, but is not limited to other *Helicobacter* strains, or other bacterial strains such as *E. coli*, *Norcardia*, *Corynebacterium*, *Campylobacter*, and *Streptomyces* species. In some cases the expression host will utilize the natural *Helicobacter* promoter whereas in others, it will be necessary to drive the gene with a promoter sequence derived from the expressing organism (e.g., an *E. coli* beta-galactosidase promoter for expression in *E. coli*).

To express a gene product using the natural *H. pylori* promoter, a procedure such as the following can be used. A restriction fragment containing the gene of interest, together with its associated natural promoter element and regulatory sequences (identified using the DNA sequence data) is cloned into an appropriate recombinant plasmid containing an origin of replication that functions in the host organism and an appropriate selectable marker. This can be accomplished by a number of procedures known to those skilled in the art. It is most preferably done by cutting the plasmid and the fragment to be cloned with the same restriction enzyme to produce compatible ends that can be ligated to join the two pieces together. The recombinant plasmid is introduced into the host organism by, for example, electroporation and cells containing the recombinant plasmid are identified by selection for the marker on the plasmid. Expression of the desired gene product is detected using an assay specific for that gene product.

In the case of a gene that requires a different promoter, the body of the gene (coding sequence) is specifically excised and cloned into an appropriate expression plasmid. This subcloning can be done by several methods, but is most easily accomplished by PCR amplification of a specific fragment and ligation into an

expression plasmid after treating the PCR product with a restriction enzyme or exonuclease to create suitable ends for cloning.

A suitable host cell for expression of a gene can be any procaryotic or eucaryotic cell. For example, an *H. pylori* polypeptide can be expressed in bacterial cells such as *E. coli*, insect cells (baculovirus), yeast, or mammalian cells such as Chinese hamster ovary cell (CHO). Other suitable host cells are known to those skilled in the art.

Expression in eucaryotic cells such as mammalian, yeast, or insect cells can lead to partial or complete glycosylation and/or formation of relevant inter- or intra-chain disulfide bonds of a recombinant peptide product. Examples of vectors for expression in yeast *S. cerevisiae* include pYepSec1 (Baldari et al., (1987) *Embo J.* 6:229-234), pMFa (Kurjan and Herskowitz, (1982) *Cell* 30:933-943), pJRY88 (Schultz et al., (1987) *Gene* 54:113-123), and pYES2 (Invitrogen Corporation, San Diego, CA). Baculovirus vectors available for expression of proteins in cultured insect cells (SF 9 cells) include the pAc series (Smith et al., (1983) *Mol. Cell Biol.* 3:2156-2165) and the pVL series (Lucklow, V.A., and Summers, M.D., (1989) *Virology* 170:31-39). Generally, COS cells (Gluzman, Y., (1981) *Cell* 23:175-182) are used in conjunction with such vectors as pCDM 8 (Aruffo, A. and Seed, B., (1987) *Proc. Natl. Acad. Sci. USA* 84:8573-8577) for transient amplification/expression in mammalian cells, while CHO (dhfr⁻ Chinese Hamster Ovary) cells are used with vectors such as pMT2PC (Kaufman et al. (1987), *EMBO J.* 6:187-195) for stable amplification/expression in mammalian cells. Vector DNA can be introduced into mammalian cells via conventional techniques such as calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, or electroporation. Suitable methods for transforming host cells can be found in Sambrook et al. (Molecular Cloning: A Laboratory Manual, 2nd Edition, Cold Spring Harbor Laboratory press (1989)), and other laboratory textbooks.

Expression in procaryotes is most often carried out in *E. coli* with either fusion or non-fusion inducible expression vectors. Fusion vectors usually add a number of NH₂ terminal amino acids to the expressed target gene. These NH₂ terminal amino acids often are referred to as a reporter group. Such reporter groups usually serve two purposes: 1) to increase the solubility of the target recombinant protein; and 2) to aid in the purification of the target recombinant protein by acting as a ligand in affinity purification. Often, in fusion expression vectors, a proteolytic cleavage site is introduced at the junction of the reporter group and the target recombinant protein to enable separation of the target recombinant protein from the reporter group subsequent to purification of the fusion protein. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase. Typical fusion expression vectors include pGEX (Amrad Corp., Melbourne, Australia), pMAL (New England

- 43 -

Biolabs, Beverly, MA) and pRIT5 (Pharmacia, Piscataway, NJ) which fuse glutathione S-transferase, maltose E binding protein, or protein A, respectively, to the target recombinant protein. A preferred reporter group is poly(His), which may be fused to the amino or carboxy terminus of the protein and which renders the recombinant fusion protein easily purifiable by metal chelate chromatography.

Inducible non-fusion expression vectors include pTrc (Amann et al., (1988) *Gene* 69:301-315) and pET11d (Studier et al., Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, California (1990) 60-89). While target gene expression relies on host RNA polymerase transcription from the hybrid trp-lac fusion promoter in pTrc, expression of target genes inserted into pET11d relies on transcription from the T7 gn10-lac 0 fusion promoter mediated by coexpressed viral RNA polymerase (T7 gn1). This viral polymerase is supplied by host strains BL21(DE3) or HMS174(DE3) from a resident λ prophage harboring a T7 gn1 under the transcriptional control of the lacUV 5 promoter.

For example, a host cell transfected with a nucleic acid vector directing expression of a nucleotide sequence encoding an *H. pylori* polypeptide can be cultured under appropriate conditions to allow expression of the polypeptide to occur. The polypeptide may be secreted and isolated from a mixture of cells and medium containing the peptide. Alternatively, the polypeptide may be retained cytoplasmically and the cells harvested, lysed and the protein isolated. A cell culture includes host cells, media and other byproducts. Suitable media for cell culture are well known in the art.

Polypeptides of the invention can be isolated from cell culture medium, host cells, or both using techniques known in the art for purifying proteins including ion-exchange chromatography, gel filtration chromatography, ultrafiltration, electrophoresis, and immunoaffinity purification with antibodies specific for such polypeptides.

Additionally, in many situations, polypeptides can be produced by chemical cleavage of a native protein (e.g., tryptic digestion) and the cleavage products can then be purified by standard techniques.

In the case of membrane bound proteins, these can be isolated from a host cell by contacting a membrane-associated protein fraction with a detergent forming a solubilized complex, where the membrane-associated protein is no longer entirely embedded in the membrane fraction and is solubilized at least to an extent which allows it to be chromatographically isolated from the membrane fraction. Several different criteria are used for choosing a detergent suitable for solubilizing these complexes. For example, one property considered is the ability of the detergent to solubilize the *H. pylori* protein within the membrane fraction at minimal denaturation of the membrane-associated protein allowing for the activity or functionality of the membrane-associated

- 44 -

protein to return upon reconstitution of the protein. Another property considered when selecting the detergent is the critical micelle concentration (CMC) of the detergent in that the detergent of choice preferably has a high CMC value allowing for ease of removal after reconstitution. A third property considered when selecting a detergent is the hydrophobicity of the detergent. Typically, membrane-associated proteins are very hydrophobic and therefore detergents which are also hydrophobic, e.g., the triton series, would be useful for solubilizing the hydrophobic proteins. Another property important to a detergent can be the capability of the detergent to remove the *H. pylori* protein with minimal protein-protein interaction facilitating further purification. A fifth property of the detergent which should be considered is the charge of the detergent. For example, if it is desired to use ion exchange resins in the purification process then preferably detergent should be an uncharged detergent. Chromatographic techniques which can be used in the final purification step are known in the art and include hydrophobic interaction, lectin affinity, ion exchange, dye affinity and immunoaffinity.

One strategy to maximize recombinant *H. pylori* peptide expression in *E. coli* is to express the protein in a host bacteria with an impaired capacity to proteolytically cleave the recombinant protein (Gottesman, S., Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, California (1990) 119-128). Another strategy would be to alter the nucleic acid encoding an *H. pylori* peptide to be inserted into an expression vector so that the individual codons for each amino acid would be those preferentially utilized in highly expressed *E. coli* proteins (Wada et al., (1992) *Nuc. Acids Res.* 20:2111-2118). Such alteration of nucleic acids of the invention can be carried out by standard DNA synthesis techniques.

The nucleic acids of the invention can also be chemically synthesized using standard techniques. Various methods of chemically synthesizing polydeoxynucleotides are known, including solid-phase synthesis which, like peptide synthesis, has been fully automated in commercially available DNA synthesizers (See, e.g., Itakura et al. U.S. Patent No. 4,598,049; Caruthers et al. U.S. Patent No. 4,458,066; and Itakura U.S. Patent Nos. 4,401,796 and 4,373,071, incorporated by reference herein).

III. *H. pylori* Polypeptides

This invention encompasses isolated *H. pylori* polypeptides encoded by the disclosed *H. pylori* genomic sequences, including the polypeptides of the invention contained in the Sequence Listing. Polypeptides of the invention are preferably at least 5 amino acid residues in length. Using the DNA sequence information provided herein, the amino acid sequences of the polypeptides encompassed by the invention can be deduced using methods well-known in the art. It will be understood that the sequence of

- 45 -

an entire nucleic acid encoding an *H. pylori* polypeptide can be isolated and identified based on an ORF that encodes only a fragment of the cognate protein-coding region. This can be achieved, for example, by using the isolated nucleic acid encoding the ORF, or fragments thereof, to prime a polymerase chain reaction with genomic *H. pylori* DNA as template; this is followed by sequencing the amplified product.

The polypeptides of the invention can be isolated from wild-type or mutant *H. pylori* cells or from heterologous organisms or cells (including, but not limited to, bacteria, yeast, insect, plant and mammalian cells) into which an *H. pylori* nucleic acid has been introduced and expressed. In addition, the polypeptides can be part of recombinant fusion proteins.

H. pylori polypeptides of the invention can be chemically synthesized using commercially automated procedures such as those referenced herein.

H. pylori polypeptides of the invention are also intended to include chimeric proteins and truncated proteins as described herein.

Chimeric *H. pylori* proteins

H. pylori chimeric polypeptides comprise one or more *H. pylori* polypeptides fused together. These combined sequences can be made by combining two or more genes, or two or more polypeptide encoding sequences, or at least one gene and at least one polypeptide encoding sequence in tandem, and the subsequent expression of the encoded proteins by conventional molecular biological techniques. The combined nucleotide sequences may be composed of a combination of either full length *H. pylori* nucleotide sequences or fragments of such sequences, e.g., fragments which contain immunologically relevant portions of the encoded *H. pylori* protein. These chimeric *H. pylori* proteins then contain the combined or synergistic vaccine potential of each individual *H. pylori* protein sequence and can be used in vaccine formulations of the invention.

Truncated gene expression and protein production

H. pylori proteins encoded by a given nucleotide sequence can also be used in a biologically active truncated form. Such truncation can be produced, for example, by the elimination of either 5' and/or 3' regions of the encoding nucleotide sequence. These truncations can affect recombinant expression of the encoded protein and/or subsequent purification of the protein. For example, truncation of a nucleotide sequence encoding a predicted export sequence of a specific protein may alter expression of the protein. Alternatively, C-terminal truncation of an *H. pylori* polypeptide by elimination of the 3' end of the nucleic acid coding region may also improve protein expression and

subsequent purification and use, as is outlined in Example VIII below. Deletion of nucleic acid regions encoding internal *H. pylori* protein regions can also result in improved protein expression, purification and/or efficacy as a vaccine candidate.

5 IV. Identification of Nucleic Acids Encoding Vaccine Components and Targets for Agents Effective Against *H. pylori*

The disclosed *H. pylori* genome sequence includes segments that direct the synthesis of ribonucleic acids and polypeptides, as well as origins of replication, promoters, other types of regulatory sequences, and intergenic nucleic acids. The invention encompasses nucleic acids encoding immunogenic components of vaccines and targets for agents effective against *H. pylori*. Identification of said immunogenic components involved in the determination of the function of the disclosed sequences can be achieved using a variety of approaches. Non-limiting examples of these approaches are described briefly below.

15 Homology to known sequences: Computer-assisted comparison of the disclosed *H. pylori* sequences with previously reported sequences present in publicly available databases is useful for identifying functional *H. pylori* nucleic acid and polypeptide sequences. It will be understood that protein-coding sequences, for example, may be compared as a whole, and that a high degree of sequence homology between two proteins (such as, for example, >80-90%) at the amino acid level indicates that the two proteins also possess some degree of functional homology, such as, for example, among enzymes involved in metabolism, DNA synthesis, or cell wall synthesis, and proteins involved in transport, cell division, etc. In addition, many structural features of particular protein classes have been identified and correlate with specific consensus sequences, such as, for example, binding domains for nucleotides, DNA, metal ions, and other small molecules; sites for covalent modifications such as phosphorylation, acylation, and the like; sites of protein:protein interactions, etc. These consensus sequences may be quite short and thus may represent only a fraction of the entire protein-coding sequence. Identification of such a feature in an *H. pylori* sequence is therefore useful in determining the function of the encoded protein and identifying useful targets of antibacterial drugs.

Of particular relevance to the present invention are structural features that are common to secretory, transmembrane, and surface proteins, including secretion signal peptides and hydrophobic transmembrane domains. *H. pylori* proteins identified as containing putative signal sequences and/or transmembrane domains are useful as immunogenic components of vaccines.

Identification of essential genes: Nucleic acids that encode proteins essential for growth or viability of *H. pylori* are preferred drug targets. *H. pylori* genes can be tested for their biological relevance to the organism by examining the effect of deleting and/or disrupting the genes, i.e., by so-called gene "knockout", using techniques known to those skilled in the relevant art. In this manner, essential genes may be identified.

Strain-specific sequences: Because of the evolutionary relationship between different *H. pylori* strains, it is believed that the presently disclosed *H. pylori* sequences are useful for identifying, and/or discriminating between, previously known and new *H. pylori* strains. It is believed that other *H. pylori* strains will exhibit at least 70% sequence homology with the presently disclosed sequence. Systematic and routine analyses of DNA sequences derived from samples containing *H. pylori* strains, and comparison with the present sequence allows for the identification of sequences that can be used to discriminate between strains, as well as those that are common to all *H. pylori* strains. In one embodiment, the invention provides nucleic acids, including probes, and peptide and polypeptide sequences that discriminate between different strains of *H. pylori*. Strain-specific components can also be identified functionally by their ability to elicit or react with antibodies that selectively recognize one or more *H. pylori* strains.

In another embodiment, the invention provides nucleic acids, including probes, and peptide and polypeptide sequences that are common to all *H. pylori* strains but are not found in other bacterial species.

Specific Example: Determination Of Candidate Protein Antigens For Antibody And Vaccine Development

The selection of candidate protein antigens for vaccine development can be derived from the nucleic acids encoding *H. pylori* polypeptides. First, the ORF's can be analyzed for homology to other known exported or membrane proteins and analyzed using the discriminant analysis described by Klein, et al. (Klein, P., Kanehsia, M., and DeLisi, C. (1985) *Biochimica et Biophysica Acta* 815, 468-476) for predicting exported and membrane proteins.

Homology searches can be performed using the BLAST algorithm contained in the Wisconsin Sequence Analysis Package (Genetics Computer Group, University Research Park, 575 Science Drive, Madison, WI 53711) to compare each predicted ORF amino acid sequence with all sequences found in the current GenBank, SWISS-PROT and PIR databases. BLAST searches for local alignments between the ORF and the databank sequences and reports a probability score which indicates the probability of finding this sequence by chance in the database. ORF's with significant homology (e.g. probabilities lower than 1×10^{-6} that the homology is only due to random chance) to

membrane or exported proteins represent protein antigens for vaccine development. Possible functions can be provided to *H. pylori* genes based on sequence homology to genes cloned in other organisms.

Discriminant analysis (Klein, et al. supra) can be used to examine the ORF amino acid sequences. This algorithm uses the intrinsic information contained in the ORF amino acid sequence and compares it to information derived from the properties of known membrane and exported proteins. This comparison predicts which proteins will be exported, membrane associated or cytoplasmic. ORF amino acid sequences identified as exported or membrane associated by this algorithm are likely protein antigens for vaccine development.

Surface exposed outer membrane proteins are likely to represent the best antigens to provide a protective immune response against *H. pylori*. Among the algorithms that can be used to aid in prediction of these outer membrane proteins include the presence of an amphipathic beta-sheet region at their C-terminus. This region which has been detected in a large number of outer membrane proteins in Gram negative bacteria is often characterized by hydrophobic residues (Phe or Tyr) approximately at positions 1, 3, 5, 7 and 9 from the C-terminus (e.g., see Figure 1, block F). Importantly, these sequences have not been detected at the C-termini of periplasmic proteins, thus allowing preliminary distinction between these classes of proteins based on primary sequence data. This phenomenon has been reported previously by Struyve et al. (*J. Mol. Biol.* 218:141-148, 1991).

Also illustrated in Figure 1 are additional amino acid sequence motifs found in many outer membrane proteins of *H. pylori*. The amino acid sequence alignment in Figure 1 depicts portions of the sequence of five *H. pylori* proteins (depicted in the single letter amino acid code) labeled with their amino acid Sequence ID Numbers and shown N-terminal to C-terminal, left to right. Six distinct blocks (labeled A through F) of similar amino acid residues are found including the distinctive hydrophobic residues (Phe or Tyr; F or Y according to the single letter code for amino acid residues) frequently found at positions near the C-terminus of outer membrane proteins. The presence of several shared motifs clearly establishes the similarity between members of this group of proteins.

In addition, outer membrane proteins isolated from *H. pylori* frequently share a motif near the mature N-terminus (i.e., after processing to remove the secretion signal) as illustrated in the blocked amino acid residues in Figure 2. Figure 2 depicts the N-terminal portion of three *H. pylori* proteins (designated by their amino acid Sequence ID Numbers and shown N-terminal to C-terminal, left to right).

- 49 -

One skilled in the art would know that these shared sequence motifs are highly significant and establish a similarity among this group of proteins.

Infrequently it is not possible to distinguish between multiple possible nucleotides at a given position in the nucleic acid sequence. In those cases the ambiguities are denoted by an extended alphabet as follows:

These are the official IUPAC-IUB single-letter base codes

Code	Base Description	
G	Guanine	
A	Adenine	
T	Thymine	
C	Cytosine	
R	Purine	(A or G)
Y	Pyrimidine	(C or T or U)
M	Amino	(A or C)
K	Ketone	(G or T)
S	Strong interaction	(C or G)
W	Weak interaction	(A or T)
H	Not-G	(A or C or T)
B	Not-A	(C or G or T)
V	Not-T (not-U)	(A or C or G)
D	Not-C	(A or G or T)
N	Any	(A or C or G or T)

The amino acid translations of this invention account for the ambiguity in the nucleic acid sequence by translating the ambiguous codon as the letter "X". In all cases, the permissible amino acid residues at a position are clear from an examination of the nucleic acid sequence based on the standard genetic code.

V. Production of Fragments and Analogs of *H. pylori* Nucleic Acids and Polypeptides

Based on the discovery of the *H. pylori* gene products of the invention provided in the Sequence Listing, one skilled in the art can alter the disclosed structure (of *H. pylori* genes), e.g., by producing fragments or analogs, and test the newly produced structures for activity. Examples of techniques known to those skilled in the relevant art which allow the production and testing of fragments and analogs are discussed below. These, or analogous methods can be used to make and screen libraries of polypeptides,

e.g., libraries of random peptides or libraries of fragments or analogs of cellular proteins for the ability to bind *H. pylori* polypeptides. Such screens are useful for the identification of inhibitors of *H. pylori*.

5 Generation of Fragments

Fragments of a protein can be produced in several ways, e.g., recombinantly, by proteolytic digestion, or by chemical synthesis. Internal or terminal fragments of a polypeptide can be generated by removing one or more nucleotides from one end (for a terminal fragment) or both ends (for an internal fragment) of a nucleic acid which
10 encodes the polypeptide. Expression of the mutagenized DNA produces polypeptide fragments. Digestion with "end-nibbling" endonucleases can thus generate DNA's which encode an array of fragments. DNA's which encode fragments of a protein can also be generated by random shearing, restriction digestion or a combination of the above-discussed methods.

15 Fragments can also be chemically synthesized using techniques known in the art such as conventional Merrifield solid phase f-Moc or t-Boc chemistry. For example, peptides of the present invention may be arbitrarily divided into fragments of desired length with no overlap of the fragments, or divided into overlapping fragments of a desired length.

20

Alteration of Nucleic Acids and Polypeptides: Random Methods

Amino acid sequence variants of a protein can be prepared by random mutagenesis of DNA which encodes a protein or a particular domain or region of a protein. Useful methods include PCR mutagenesis and saturation mutagenesis. A
25 library of random amino acid sequence variants can also be generated by the synthesis of a set of degenerate oligonucleotide sequences. (Methods for screening proteins in a library of variants are elsewhere herein).

(A) PCR Mutagenesis

30 In PCR mutagenesis, reduced Taq polymerase fidelity is used to introduce random mutations into a cloned fragment of DNA (Leung et al., 1989, *Technique* 1:11-15). The DNA region to be mutagenized is amplified using the polymerase chain reaction (PCR) under conditions that reduce the fidelity of DNA synthesis by Taq DNA polymerase, e.g., by using a dGTP/dATP ratio of five and adding Mn^{2+} to the PCR
35 reaction. The pool of amplified DNA fragments are inserted into appropriate cloning vectors to provide random mutant libraries.

(B) Saturation Mutagenesis

Saturation mutagenesis allows for the rapid introduction of a large number of single base substitutions into cloned DNA fragments (Mayers et al., 1985, *Science* 229:242). This technique includes generation of mutations, e.g., by chemical treatment or irradiation of single-stranded DNA *in vitro*, and synthesis of a complimentary DNA strand. The mutation frequency can be modulated by modulating the severity of the treatment, and essentially all possible base substitutions can be obtained. Because this procedure does not involve a genetic selection for mutant fragments both neutral substitutions, as well as those that alter function, are obtained. The distribution of point mutations is not biased toward conserved sequence elements.

(C) Degenerate Oligonucleotides

A library of homologs can also be generated from a set of degenerate oligonucleotide sequences. Chemical synthesis of a degenerate sequences can be carried out in an automatic DNA synthesizer, and the synthetic genes then ligated into an appropriate expression vector. The synthesis of degenerate oligonucleotides is known in the art (see for example, Narang, SA (1983) *Tetrahedron* 39:3; Itakura et al. (1981) *Recombinant DNA, Proc 3rd Cleveland Sympos. Macromolecules*, ed. AG Walton, Amsterdam: Elsevier pp273-289; Itakura et al. (1984) *Annu. Rev. Biochem.* 53:323; Itakura et al. (1984) *Science* 198:1056; Ike et al. (1983) *Nucleic Acid Res.* 11:477. Such techniques have been employed in the directed evolution of other proteins (see, for example, Scott et al. (1990) *Science* 249:386-390; Roberts et al. (1992) *PNAS* 89:2429-2433; Devlin et al. (1990) *Science* 249: 404-406; Cwirla et al. (1990) *PNAS* 87: 6378-6382; as well as U.S. Patents Nos. 5,223,409, 5,198,346, and 5,096,815).

Alteration of Nucleic Acids and Polypeptides: Methods for Directed Mutagenesis

Non-random or directed, mutagenesis techniques can be used to provide specific sequences or mutations in specific regions. These techniques can be used to create variants which include, e.g., deletions, insertions, or substitutions, of residues of the known amino acid sequence of a protein. The sites for mutation can be modified individually or in series, e.g., by (1) substituting first with conserved amino acids and then with more radical choices depending upon results achieved, (2) deleting the target residue, or (3) inserting residues of the same or a different class adjacent to the located site, or combinations of options 1-3.

(A) Alanine Scanning Mutagenesis

Alanine scanning mutagenesis is a useful method for identification of certain residues or regions of the desired protein that are preferred locations or domains for mutagenesis, Cunningham and Wells (*Science* 244:1081-1085, 1989). In alanine scanning, a residue or group of target residues are identified (e.g., charged residues such as Arg, Asp, His, Lys, and Glu) and replaced by a neutral or negatively charged amino acid (most preferably alanine or polyalanine). Replacement of an amino acid can affect the interaction of the amino acids with the surrounding aqueous environment in or outside the cell. Those domains demonstrating functional sensitivity to the substitutions are then refined by introducing further or other variants at or for the sites of substitution. Thus, while the site for introducing an amino acid sequence variation is predetermined, the nature of the mutation per se need not be predetermined. For example, to optimize the performance of a mutation at a given site, alanine scanning or random mutagenesis may be conducted at the target codon or region and the expressed desired protein subunit variants are screened for the optimal combination of desired activity.

(B) Oligonucleotide-Mediated Mutagenesis

Oligonucleotide-mediated mutagenesis is a useful method for preparing substitution, deletion, and insertion variants of DNA, see, e.g., Adelman et al., (*DNA* 2:183, 1983). Briefly, the desired DNA is altered by hybridizing an oligonucleotide encoding a mutation to a DNA template, where the template is the single-stranded form of a plasmid or bacteriophage containing the unaltered or native DNA sequence of the desired protein. After hybridization, a DNA polymerase is used to synthesize an entire second complementary strand of the template that will thus incorporate the oligonucleotide primer, and will code for the selected alteration in the desired protein DNA. Generally, oligonucleotides of at least 25 nucleotides in length are used. An optimal oligonucleotide will have 12 to 15 nucleotides that are completely complementary to the template on either side of the nucleotide(s) coding for the mutation. This ensures that the oligonucleotide will hybridize properly to the single-stranded DNA template molecule. The oligonucleotides are readily synthesized using techniques known in the art such as that described by Crea et al. (*Proc. Natl. Acad. Sci. USA*, 75: 5765[1978]).

(C) Cassette Mutagenesis

Another method for preparing variants, cassette mutagenesis, is based on the technique described by Wells et al. (*Gene*, 34:315[1985]). The starting material is a plasmid (or other vector) which includes the protein subunit DNA to be mutated. The

- 53 -

codon(s) in the protein subunit DNA to be mutated are identified. There must be a unique restriction endonuclease site on each side of the identified mutation site(s). If no such restriction sites exist, they may be generated using the above-described oligonucleotide-mediated mutagenesis method to introduce them at appropriate locations in the desired protein subunit DNA. After the restriction sites have been introduced into the plasmid, the plasmid is cut at these sites to linearize it. A double-stranded oligonucleotide encoding the sequence of the DNA between the restriction sites but containing the desired mutation(s) is synthesized using standard procedures. The two strands are synthesized separately and then hybridized together using standard techniques. This double-stranded oligonucleotide is referred to as the cassette. This cassette is designed to have 3' and 5' ends that are comparable with the ends of the linearized plasmid, such that it can be directly ligated to the plasmid. This plasmid now contains the mutated desired protein subunit DNA sequence.

15 (D) Combinatorial Mutagenesis

Combinatorial mutagenesis can also be used to generate mutants (Ladner et al., WO 88/06630). In this method, the amino acid sequences for a group of homologs or other related proteins are aligned, preferably to promote the highest homology possible. All of the amino acids which appear at a given position of the aligned sequences can be selected to create a degenerate set of combinatorial sequences. The variegated library of variants is generated by combinatorial mutagenesis at the nucleic acid level, and is encoded by a variegated gene library. For example, a mixture of synthetic oligonucleotides can be enzymatically ligated into gene sequences such that the degenerate set of potential sequences are expressible as individual peptides, or alternatively, as a set of larger fusion proteins containing the set of degenerate sequences.

Other Modifications of *H. pylori* Nucleic Acids and Polypeptides

It is possible to modify the structure of an *H. pylori* polypeptide for such purposes as increasing solubility, enhancing stability (e.g., shelf life *ex vivo* and resistance to proteolytic degradation *in vivo*). A modified *H. pylori* protein or peptide can be produced in which the amino acid sequence has been altered, such as by amino acid substitution, deletion, or addition as described herein.

An *H. pylori* peptide can also be modified by substitution of cysteine residues preferably with alanine, serine, threonine, leucine or glutamic acid residues to minimize dimerization via disulfide linkages. In addition, amino acid side chains of fragments of

- 54 -

the protein of the invention can be chemically modified. Another modification is cyclization of the peptide.

In order to enhance stability and/or reactivity, an *H. pylori* polypeptide can be modified to incorporate one or more polymorphisms in the amino acid sequence of the protein resulting from any natural allelic variation. Additionally, D-amino acids, non-natural amino acids, or non-amino acid analogs can be substituted or added to produce a modified protein within the scope of this invention. Furthermore, an *H. pylori* polypeptide can be modified using polyethylene glycol (PEG) according to the method of A. Schon and co-workers (Wie et al., supra) to produce a protein conjugated with PEG. In addition, PEG can be added during chemical synthesis of the protein. Other modifications of *H. pylori* proteins include reduction/alkylation (Tarr, *Methods of Protein Microcharacterization*, J. E. Silver ed., Humana Press, Clifton NJ 155-194 (1986)); acylation (Tarr, supra); chemical coupling to an appropriate carrier (Mishell and Shiigi, eds, *Selected Methods in Cellular Immunology*, WH Freeman, San Francisco, CA (1980), U.S. Patent 4,939,239; or mild formalin treatment (Marsh, (1971) *Int. Arch. of Allergy and Appl. Immunol.*, 41: 199 - 215).

To facilitate purification and potentially increase solubility of an *H. pylori* protein or peptide, it is possible to add an amino acid fusion moiety to the peptide backbone. For example, hexa-histidine can be added to the protein for purification by immobilized metal ion affinity chromatography (Hochuli, E. et al., (1988) *Bio/Technology*, 6: 1321 - 1325). In addition, to facilitate isolation of peptides free of irrelevant sequences, specific endoprotease cleavage sites can be introduced between the sequences of the fusion moiety and the peptide.

To potentially aid proper antigen processing of epitopes within an *H. pylori* polypeptide, canonical protease sensitive sites can be engineered between regions, each comprising at least one epitope via recombinant or synthetic methods. For example, charged amino acid pairs, such as KK or RR, can be introduced between regions within a protein or fragment during recombinant construction thereof. The resulting peptide can be rendered sensitive to cleavage by cathepsin and/or other trypsin-like enzymes which would generate portions of the protein containing one or more epitopes. In addition, such charged amino acid residues can result in an increase in the solubility of the peptide.

Primary Methods for Screening Polypeptides and Analogs

Various techniques are known in the art for screening generated mutant gene products. Techniques for screening large gene libraries often include cloning the gene library into replicable expression vectors, transforming appropriate cells with the

- 55 -

resulting library of vectors, and expressing the genes under conditions in which detection of a desired activity, e.g., in this case, binding to *H. pylori* polypeptide or an interacting protein, facilitates relatively easy isolation of the vector encoding the gene whose product was detected. Each of the techniques described below is amenable to

5 high through-put analysis for screening large numbers of sequences created, e.g., by random mutagenesis techniques.

(A) Two Hybrid Systems

Two hybrid assays such as the system described above (as with the other

10 screening methods described herein), can be used to identify polypeptides, e.g., fragments or analogs of a naturally-occurring *H. pylori* polypeptide, e.g., of cellular proteins, or of randomly generated polypeptides which bind to an *H. pylori* protein. (The *H. pylori* domain is used as the bait protein and the library of variants are expressed as fish fusion proteins.) In an analogous fashion, a two hybrid assay (as with the other

15 screening methods described herein), can be used to find polypeptides which bind a *H. pylori* polypeptide.

(B) Display Libraries

In one approach to screening assays, the candidate peptides are displayed on the

20 surface of a cell or viral particle, and the ability of particular cells or viral particles to bind an appropriate receptor protein via the displayed product is detected in a "panning assay". For example, the gene library can be cloned into the gene for a surface membrane protein of a bacterial cell, and the resulting fusion protein detected by panning (Ladner et al., WO 88/06630; Fuchs et al. (1991) *Bio/Technology* 9:1370-1371; and Goward et al. (1992) *TIBS* 18:136-140). In a similar fashion, a detectably labeled

25 ligand can be used to score for potentially functional peptide homologs. Fluorescently labeled ligands, e.g., receptors, can be used to detect homologs which retain ligand-binding activity. The use of fluorescently labeled ligands, allows cells to be visually inspected and separated under a fluorescence microscope, or, where the morphology of

30 the cell permits, to be separated by a fluorescence-activated cell sorter.

A gene library can be expressed as a fusion protein on the surface of a viral particle. For instance, in the filamentous phage system, foreign peptide sequences can be expressed on the surface of infectious phage, thereby conferring two significant benefits. First, since these phage can be applied to affinity matrices at concentrations

35 well over 10^{13} phage per milliliter, a large number of phage can be screened at one time. Second, since each infectious phage displays a gene product on its surface, if a particular phage is recovered from an affinity matrix in low yield, the phage can be amplified by

another round of infection. The group of almost identical *E. coli* filamentous phages M13, fd., and f1 are most often used in phage display libraries. Either of the phage gIII or gVIII coat proteins can be used to generate fusion proteins without disrupting the ultimate packaging of the viral particle. Foreign epitopes can be expressed at the NH₂-terminal end of pIII and phage bearing such epitopes recovered from a large excess of phage lacking this epitope (Ladner et al. PCT publication WO 90/02909; Garrard et al., PCT publication WO 92/09690; Marks et al. (1992) *J. Biol. Chem.* 267:16007-16010; Griffiths et al. (1993) *EMBO J* 12:725-734; Clackson et al. (1991) *Nature* 352:624-628; and Barbas et al. (1992) *PNAS* 89:4457-4461).

A common approach uses the maltose receptor of *E. coli* (the outer membrane protein, LamB) as a peptide fusion partner (Charbit et al. (1986) *EMBO* 5, 3029-3037). Oligonucleotides have been inserted into plasmids encoding the LamB gene to produce peptides fused into one of the extracellular loops of the protein. These peptides are available for binding to ligands, e.g., to antibodies, and can elicit an immune response when the cells are administered to animals. Other cell surface proteins, e.g., OmpA (Schorr et al. (1991) *Vaccines* 91, pp. 387-392), PhoE (Agterberg, et al. (1990) *Gene* 88, 37-45), and PAL (Fuchs et al. (1991) *Bio/Tech* 9, 1369-1372), as well as large bacterial surface structures have served as vehicles for peptide display. Peptides can be fused to pilin, a protein which polymerizes to form the pilus-a conduit for interbacterial exchange of genetic information (Thiry et al. (1989) *Appl. Environ. Microbiol.* 55, 984-993). Because of its role in interacting with other cells, the pilus provides a useful support for the presentation of peptides to the extracellular environment. Another large surface structure used for peptide display is the bacterial motive organ, the flagellum. Fusion of peptides to the subunit protein flagellin offers a dense array of many peptide copies on the host cells (Ku wajima et al. (1988) *Bio/Tech.* 6, 1080-1083). Surface proteins of other bacterial species have also served as peptide fusion partners. Examples include the *Staphylococcus* protein A and the outer membrane IgA protease of *Neisseria* (Hansson et al. (1992) *J. Bacteriol.* 174, 4239-4245 and Klauser et al. (1990) *EMBO J.* 9, 1991-1999).

In the filamentous phage systems and the LamB system described above, the physical link between the peptide and its encoding DNA occurs by the containment of the DNA within a particle (cell or phage) that carries the peptide on its surface. Capturing the peptide captures the particle and the DNA within. An alternative scheme uses the DNA-binding protein LacI to form a link between peptide and DNA (Cull et al. (1992) *PNAS USA* 89:1865-1869). This system uses a plasmid containing the LacI gene with an oligonucleotide cloning site at its 3'-end. Under the controlled induction by arabinose, a LacI-peptide fusion protein is produced. This fusion retains the natural

ability of LacI to bind to a short DNA sequence known as LacO operator (LacO). By installing two copies of LacO on the expression plasmid, the LacI-peptide fusion binds tightly to the plasmid that encoded it. Because the plasmids in each cell contain only a single oligonucleotide sequence and each cell expresses only a single peptide sequence, the peptides become specifically and stably associated with the DNA sequence that directed its synthesis. The cells of the library are gently lysed and the peptide-DNA complexes are exposed to a matrix of immobilized receptor to recover the complexes containing active peptides. The associated plasmid DNA is then reintroduced into cells for amplification and DNA sequencing to determine the identity of the peptide ligands.

As a demonstration of the practical utility of the method, a large random library of dodecapeptides was made and selected on a monoclonal antibody raised against the opioid peptide dynorphin B. A cohort of peptides was recovered, all related by a consensus sequence corresponding to a six-residue portion of dynorphin B. (Cull et al. (1992) *Proc. Natl. Acad. Sci. U.S.A.* 89-1869)

This scheme, sometimes referred to as peptides-on-plasmids, differs in two important ways from the phage display methods. First, the peptides are attached to the C-terminus of the fusion protein, resulting in the display of the library members as peptides having free carboxy termini. Both of the filamentous phage coat proteins, pIII and pVIII, are anchored to the phage through their C-termini, and the guest peptides are placed into the outward-extending N-terminal domains. In some designs, the phage-displayed peptides are presented right at the amino terminus of the fusion protein. (Cwirla, et al. (1990) *Proc. Natl. Acad. Sci. U.S.A.* 87, 6378-6382) A second difference is the set of biological biases affecting the population of peptides actually present in the libraries. The LacI fusion molecules are confined to the cytoplasm of the host cells.

The phage coat fusions are exposed briefly to the cytoplasm during translation but are rapidly secreted through the inner membrane into the periplasmic compartment, remaining anchored in the membrane by their C-terminal hydrophobic domains, with the N-termini, containing the peptides, protruding into the periplasm while awaiting assembly into phage particles. The peptides in the LacI and phage libraries may differ significantly as a result of their exposure to different proteolytic activities. The phage coat proteins require transport across the inner membrane and signal peptidase processing as a prelude to incorporation into phage. Certain peptides exert a deleterious effect on these processes and are underrepresented in the libraries (Gallop et al. (1994) *J. Med. Chem.* 37(9):1233-1251). These particular biases are not a factor in the LacI display system.

The number of small peptides available in recombinant random libraries is enormous. Libraries of 10^7 - 10^9 independent clones are routinely prepared. Libraries as

- 58 -

large as 10^{11} recombinants have been created, but this size approaches the practical limit for clone libraries. This limitation in library size occurs at the step of transforming the DNA containing randomized segments into the host bacterial cells. To circumvent this limitation, an *in vitro* system based on the display of nascent peptides in polysome
5 complexes has recently been developed. This display library method has the potential of producing libraries 3-6 orders of magnitude larger than the currently available phage/phagemid or plasmid libraries. Furthermore, the construction of the libraries, expression of the peptides, and screening, is done in an entirely cell-free format.

In one application of this method (Gallop et al. (1994) *J. Med. Chem.*
10 37(9):1233-1251), a molecular DNA library encoding 10^{12} decapeptides was constructed and the library expressed in an *E. coli* S30 *in vitro* coupled transcription/translation system. Conditions were chosen to stall the ribosomes on the mRNA, causing the accumulation of a substantial proportion of the RNA in polysomes and yielding complexes containing nascent peptides still linked to their encoding RNA.
15 The polysomes are sufficiently robust to be affinity purified on immobilized receptors in much the same way as the more conventional recombinant peptide display libraries are screened. RNA from the bound complexes is recovered, converted to cDNA, and amplified by PCR to produce a template for the next round of synthesis and screening. The polysome display method can be coupled to the phage display system. Following
20 several rounds of screening, cDNA from the enriched pool of polysomes was cloned into a phagemid vector. This vector serves as both a peptide expression vector, displaying peptides fused to the coat proteins, and as a DNA sequencing vector for peptide identification. By expressing the polysome-derived peptides on phage, one can either continue the affinity selection procedure in this format or assay the peptides on
25 individual clones for binding activity in a phage ELISA, or for binding specificity in a completion phage ELISA (Barret, et al. (1992) *Anal. Biochem* 204,357-364). To identify the sequences of the active peptides one sequences the DNA produced by the phagemid host.

30 Secondary Screening of Polypeptides and Analogs

The high through-put assays described above can be followed by secondary
screens in order to identify further biological activities which will, e.g., allow one skilled in the art to differentiate agonists from antagonists. The type of a secondary screen used will depend on the desired activity that needs to be tested. For example, an assay can be
35 developed in which the ability to inhibit an interaction between a protein of interest and its respective ligand can be used to identify antagonists from a group of peptide fragments isolated through one of the primary screens described above.

Therefore, methods for generating fragments and analogs and testing them for activity are known in the art. Once the core sequence of interest is identified, it is routine for one skilled in the art to obtain analogs and fragments.

5 Peptide Mimetics of *H. pylori* Polypeptides

The invention also provides for reduction of the protein binding domains of the subject *H. pylori* polypeptides to generate mimetics, e.g. peptide or non-peptide agents. The peptide mimetics are able to disrupt binding of a polypeptide to its counter ligand, e.g., in the case of an *H. pylori* polypeptide binding to a naturally occurring ligand. The
10 critical residues of a subject *H. pylori* polypeptide which are involved in molecular recognition of a polypeptide can be determined and used to generate *H. pylori*-derived peptidomimetics which competitively or noncompetitively inhibit binding of the *H. pylori* polypeptide with an interacting polypeptide (see, for example, European patent applications EP-412,762A and EP-B31,080A).

15 For example, scanning mutagenesis can be used to map the amino acid residues of a particular *H. pylori* polypeptide involved in binding an interacting polypeptide, peptidomimetic compounds (e.g. diazepam or isoquinoline derivatives) can be generated which mimic those residues in binding to an interacting polypeptide, and which therefore can inhibit binding of an *H. pylori* polypeptide to an interacting polypeptide
20 and thereby interfere with the function of *H. pylori* polypeptide. For instance, non-hydrolyzable peptide analogs of such residues can be generated using benzodiazepine (e.g., see Freidinger et al. in *Peptides: Chemistry and Biology*, G.R. Marshall ed., ESCOM Publisher: Leiden, Netherlands, 1988), azepine (e.g., see Huffman et al. in *Peptides: Chemistry and Biology*, G.R. Marshall ed., ESCOM Publisher: Leiden,
25 Netherlands, 1988), substituted gamma lactam rings (Garvey et al. in *Peptides: Chemistry and Biology*, G.R. Marshall ed., ESCOM Publisher: Leiden, Netherlands, 1988), keto-methylene pseudopeptides (Ewenson et al. (1986) *J Med Chem* 29:295; and Ewenson et al. in *Peptides: Structure and Function* (Proceedings of the 9th American Peptide Symposium) Pierce Chemical Co. Rockland, IL, 1985), β -turn dipeptide cores (Nagai et al. (1985) *Tetrahedron Lett* 26:647; and Sato et al. (1986) *J Chem Soc Perkin Trans*
30 1:1231), and β -aminoalcohols (Gordon et al. (1985) *Biochem Biophys Res Commun* 126:419; and Dann et al. (1986) *Biochem Biophys Res Commun* 134:71).

VI. Vaccine Formulations for *H. pylori* Nucleic Acids and Polypeptides

35 This invention also features vaccine compositions or formulations (used interchangeably herein) for protection against infection by *H. pylori* or for treatment of *H. pylori* infection. As used herein, the term "treatment of *H. pylori* infection" refers to

- 60 -

therapeutic treatment of an existing or established *H. pylori* infection. The terms "protection against *H. pylori* infection" or "prophylactic treatment" refer to the use of *H. pylori* vaccine formulation for reducing the risk of or preventing an infection in a subject at risk for *H. pylori* infection. In one embodiment, the vaccine compositions contain one or more immunogenic components, such as a surface protein, from *H. pylori*, or portion thereof, and a pharmaceutically acceptable carrier. For example, in one embodiment, the vaccine formulations of the invention contain at least one or combination of *H. pylori* polypeptides or fragments thereof, from same or different *H. pylori* antigens. Nucleic acids and *H. pylori* polypeptides for use in the vaccine formulations of the invention include the nucleic acids and polypeptides set forth in the Sequence Listing, preferably those *H. pylori* nucleic acids that encode surface proteins and surface proteins or fragments thereof. For example, a preferred nucleic acid and *H. pylori* polypeptide for use in a vaccine composition of the invention is selected from the group of nucleic acids which encode cell envelope proteins and *H. pylori* cell envelope proteins as set forth in Table 1. However, any nucleic acid encoding an immunogenic *H. pylori* protein and *H. pylori* polypeptide, or portion thereof, can be used in the present invention. These vaccines have therapeutic and/or prophylactic utilities.

One aspect of the invention provides a vaccine composition for protection against infection by *H. pylori* which contains at least one immunogenic fragment of an *H. pylori* protein and a pharmaceutically acceptable carrier. Preferred fragments include peptides of at least about 10 amino acid residues in length, preferably about 10-20 amino acid residues in length, and more preferably about 12-16 amino acid residues in length.

Immunogenic components of the invention can be obtained, for example, by screening polypeptides recombinantly produced from the corresponding fragment of the nucleic acid encoding the full-length *H. pylori* protein. In addition, fragments can be chemically synthesized using techniques known in the art such as conventional Merrifield solid phase f-Moc or t-Boc chemistry.

In one embodiment, immunogenic components are identified by the ability of the peptide to stimulate T cells. Peptides which stimulate T cells, as determined by, for example, T cell proliferation or cytokine secretion are defined herein as comprising at least one T cell epitope. T cell epitopes are believed to be involved in initiation and perpetuation of the immune response to the protein allergen which is responsible for the clinical symptoms of allergy. These T cell epitopes are thought to trigger early events at the level of the T helper cell by binding to an appropriate HLA molecule on the surface of an antigen presenting cell, thereby stimulating the T cell subpopulation with the relevant T cell receptor for the epitope. These events lead to T cell proliferation, lymphokine secretion, local inflammatory reactions, recruitment of additional immune

- 61 -

cells to the site of antigen/T cell interaction, and activation of the B cell cascade, leading to the production of antibodies. A T cell epitope is the basic element, or smallest unit of recognition by a T cell receptor, where the epitope comprises amino acids essential to receptor recognition (e.g., approximately 6 or 7 amino acid residues). Amino acid sequences which mimic those of the T cell epitopes are within the scope of this invention.

In another embodiment, immunogenic components of the invention are identified through genomic vaccination. The basic protocol is based on the idea that expression libraries consisting of all or parts of a pathogen genome, e.g., an *H. pylori* genome, can confer protection when used to genetically immunize a host. This expression library immunization (ELI) is analogous to expression cloning and involves reducing a genomic expression library of a pathogen, e.g., *H. pylori*, into plasmids that can act as genetic vaccines. The plasmids can also be designed to encode genetic adjuvants which can dramatically stimulate the humoral response. These genetic adjuvants can be introduced at remote sites and act as well extracellularly as intracellularly.

This is a new approach to vaccine production that has many of the advantages of live/attenuated pathogens but no risk of infection. An expression library of pathogen DNA is used to immunize a host thereby producing the effects of antigen presentation of a live vaccine without the risk. For example, in the present invention, random fragments from the *H. pylori* genome or from cosmid or plasmid clones, as well as PCR products from genes identified by genomic sequencing, can be used to immunize a host. The feasibility of this approach has been demonstrated with *Mycoplasma pulmonis* (Barry et al., *Nature* 377:632-635, 1995), where even partial expression libraries of *Mycoplasma pulmonis*, a natural pathogen in rodents, provided protection against challenge from the pathogen.

ELI is a technique that allows for production of a non-infectious multipartite vaccine, even when little is known about pathogen's biology, because ELI uses the immune system to screen candidate genes. Once isolated, these genes can be used as genetic vaccines or for development of recombinant protein vaccines. Thus, ELI allows for production of vaccines in a systematic, largely mechanized fashion.

Screening immunogenic components can be accomplished using one or more of several different assays. For example, *in vitro*, peptide T cell stimulatory activity is assayed by contacting a peptide known or suspected of being immunogenic with an antigen presenting cell which presents appropriate MHC molecules in a T cell culture. Presentation of an immunogenic *H. pylori* peptide in association with appropriate MHC molecules to T cells in conjunction with the necessary costimulation has the effect of transmitting a signal to the T cell that induces the production of increased levels of

- 62 -

cytokines, particularly of interleukin-2 and interleukin-4. The culture supernatant can be obtained and assayed for interleukin-2 or other known cytokines. For example, any one of several conventional assays for interleukin-2 can be employed, such as the assay described in *Proc. Natl. Acad. Sci USA*, 86: 1333 (1989) the pertinent portions of which are incorporated herein by reference. A kit for an assay for the production of interferon is also available from Genzyme Corporation (Cambridge, MA).

Alternatively, a common assay for T cell proliferation entails measuring tritiated thymidine incorporation. The proliferation of T cells can be measured *in vitro* by determining the amount of ³H-labeled thymidine incorporated into the replicating DNA of cultured cells. Therefore, the rate of DNA synthesis and, in turn, the rate of cell division can be quantified.

Vaccine compositions or formulations of the invention containing one or more immunogenic components (e.g., *H. pylori* polypeptide or fragment thereof or nucleic acid encoding an *H. pylori* polypeptide or fragment thereof) preferably include a pharmaceutically acceptable carrier. The term "pharmaceutically acceptable carrier" is intended to include any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. Suitable pharmaceutically acceptable carriers include, for example, one or more of water, saline, phosphate buffered saline, dextrose, glycerol, ethanol and the like, as well as combinations thereof. Pharmaceutically acceptable carriers may further comprise minor amounts of auxiliary substances such as wetting or emulsifying agents, preservatives or buffers, which enhance the shelf life or effectiveness of the *H. pylori* nucleic acid or polypeptide. For vaccine formulations of the invention containing *H. pylori* polypeptides, the polypeptide is preferably coadministered with a suitable adjuvant and/or a delivery system described herein.

It will be apparent to those of skill in the art that the therapeutically effective amount of DNA or protein of this invention will depend, *inter alia*, upon the administration schedule, the unit dose of an *H. pylori* nucleic acid or polypeptide administered, whether the protein or nucleic acid is administered in combination with other therapeutic agents, the immune status and health of the patient, and the therapeutic activity of the particular protein or nucleic acid.

Vaccine formulations are conventionally administered parenterally, e.g., by injection, either subcutaneously or intramuscularly. Methods for intramuscular immunization are described by Wolff et al. (1990) *Science* 247: 1465-1468 and by Sedegah et al. (1994) *Immunology* 91: 9866-9870. Other modes of administration include oral and pulmonary formulations, suppositories, and transdermal applications. Oral immunization is preferred over parenteral methods for inducing protection against

infection by *H. pylori*. Czinn et. al. (1993) *Vaccine* 11: 637-642. Oral formulations include such normally employed excipients as, for example, pharmaceutical grades of mannitol, lactose, starch, magnesium stearate, sodium saccharine, cellulose, magnesium carbonate, and the like.

5 In one embodiment, the vaccine formulation includes, as a pharmaceutically acceptable carrier, an adjuvant. Examples of the suitable adjuvants for use in the vaccine formulations of the invention include, but are not limited, to aluminum hydroxide; N-acetyl-muramyl--L-threonyl-D-isoglutamine (thr-MDP); N-acetyl-nor-muramyl-L-alanyl-D-isoglutamine (CGP 11637, referred to as nor-MDP); N-
10 acetylmuramyl-L-alanyl-D-isoglutaminyl-L-alanine-2-(1'-2'-dipalmitoyl-sn-glycero-3-hydroxyphosphoryloxy)-ethylamine (CGP 19835A, referred to as MTP-PE); RIBI, which contains three components from bacteria; monophosphoryl lipid A; trehalose dimycolate; cell wall skeleton (MPL + TDM + CWS) in a 2% squalene/Tween 80 emulsion; and cholera toxin. Others which may be used are non-toxic derivatives of
15 cholera toxin, including its B subunit, and/or conjugates or genetically engineered fusions of the *H. pylori* polypeptide with cholera toxin or its B subunit, procholeraenoid, fungal polysaccharides, including schizophyllan, muramyl dipeptide, muramyl dipeptide derivatives, phorbol esters, labile toxin of *E. coli*, non-*H. pylori* bacterial lysates, block polymers or saponins.

20 In another embodiment, the vaccine formulation includes, as a pharmaceutically acceptable carrier, a delivery system. Suitable delivery systems for use in the vaccine formulations of the invention include biodegradable microcapsules or immunostimulating complexes (ISCOMs), cochleates, or liposomes, genetically engineered attenuated live vectors such as viruses or bacteria, and recombinant (chimeric) virus-like
25 particles, e.g., bluetongue. In another embodiment of the invention, the vaccine formulation includes both a delivery system and an adjuvant.

Delivery systems in humans may include enteric release capsules protecting the antigen from the acidic environment of the stomach, and including *H. pylori* polypeptide in an insoluble form as fusion proteins. Suitable carriers for the vaccines of the
30 invention are enteric coated capsules and polylactide-glycolide microspheres. Suitable diluents are 0.2 N NaHCO₃ and/or saline.

Vaccines of the invention can be administered as a primary prophylactic agent in adults or in children, as a secondary prevention, after successful eradication of *H. pylori* in an infected host, or as a therapeutic agent in the aim to induce an immune response in
35 a susceptible host to prevent infection by *H. pylori*. The vaccines of the invention are administered in amounts readily determined by persons of ordinary skill in the art. Thus, for adults a suitable dosage will be in the range of 10 µg to 10 g, preferably 10 µg

- 64 -

to 100 mg, for example 50 µg to 50 mg. A suitable dosage for adults will also be in the range of 5 µg to 500 mg. Similar dosage ranges will be applicable for children.

The amount of adjuvant employed will depend on the type of adjuvant used. For example, when the mucosal adjuvant is cholera toxin, it is suitably used in an amount of 5 µg to 50 µg, for example 10 µg to 35 µg. When used in the form of microcapsules, the amount used will depend on the amount employed in the matrix of the microcapsule to achieve the desired dosage. The determination of this amount is within the skill of a person of ordinary skill in the art.

Those skilled in the art will recognize that the optimal dose may be more or less depending upon the patient's body weight, disease, the route of administration, and other factors. Those skilled in the art will also recognize that appropriate dosage levels can be obtained based on results with known oral vaccines such as, for example, a vaccine based on an *E. coli* lysate (6 mg dose daily up to total of 540 mg) and with an enterotoxigenic *E. coli* purified antigen (4 doses of 1 mg) (Schulman et al., *J. Urol.* 150:917-921 (1993)); Boedecker et al., *American Gastroenterological Assoc.* 999:A-222 (1993)). The number of doses will depend upon the disease, the formulation, and efficacy data from clinical trials. Without intending any limitation as to the course of treatment, the treatment can be administered over 3 to 8 doses for a primary immunization schedule over 1 month (Boedeker, *American Gastroenterological Assoc.* 888:A-222 (1993)).

In a preferred embodiment, a vaccine composition of the invention can be based on a killed whole *E. coli* preparation with an immunogenic fragment of an *H. pylori* protein of the invention expressed on its surface or it can be based on an *E. coli* lysate, wherein the killed *E. coli* acts as a carrier or an adjuvant.

It will be apparent to those skilled in the art that some of the vaccine compositions of the invention are useful only for preventing *H. pylori* infection, some are useful only for treating *H. pylori* infection, and some are useful for both preventing and treating *H. pylori* infection. In a preferred embodiment, the vaccine composition of the invention provides protection against *H. pylori* infection by stimulating humoral and/or cell-mediated immunity against *H. pylori*. It should be understood that amelioration of any of the symptoms of *H. pylori* infection is a desirable clinical goal, including a lessening of the dosage of medication used to treat *H. pylori*-caused disease, or an increase in the production of antibodies in the serum or mucous of patients.

VII. Antibodies Reactive With *H. pylori* Polypeptides

The invention also includes antibodies specifically reactive with the subject *H. pylori* polypeptide. Anti-protein/anti-peptide antisera or monoclonal antibodies can be made by standard protocols (See, for example, *Antibodies: A Laboratory Manual* ed. by Harlow and Lane (Cold Spring Harbor Press: 1988)). A mammal such as a mouse, a hamster or rabbit can be immunized with an immunogenic form of the peptide. Techniques for conferring immunogenicity on a protein or peptide include conjugation to carriers or other techniques well known in the art. An immunogenic portion of the subject *H. pylori* polypeptide can be administered in the presence of adjuvant. The progress of immunization can be monitored by detection of antibody titers in plasma or serum. Standard ELISA or other immunoassays can be used with the immunogen as antigen to assess the levels of antibodies.

In a preferred embodiment, the subject antibodies are immunospecific for antigenic determinants of the *H. pylori* polypeptides of the invention, e.g. antigenic determinants of a polypeptide of the invention contained in the Sequence Listing, or a closely related human or non-human mammalian homolog (e.g., 90% homologous, more preferably at least 95% homologous). In yet a further preferred embodiment of the invention, the anti-*H. pylori* antibodies do not substantially cross react (i.e., react specifically) with a protein which is for example, less than 80% percent homologous to a sequence of the invention contained in the Sequence Listing. By "not substantially cross react", it is meant that the antibody has a binding affinity for a non-homologous protein which is less than 10 percent, more preferably less than 5 percent, and even more preferably less than 1 percent, of the binding affinity for a protein of the invention contained in the Sequence Listing. In a most preferred embodiment, there is no crossreactivity between bacterial and mammalian antigens.

The term antibody as used herein is intended to include fragments thereof which are also specifically reactive with *H. pylori* polypeptides. Antibodies can be fragmented using conventional techniques and the fragments screened for utility in the same manner as described above for whole antibodies. For example, F(ab')₂ fragments can be generated by treating antibody with pepsin. The resulting F(ab')₂ fragment can be treated to reduce disulfide bridges to produce Fab' fragments. The antibody of the invention is further intended to include bispecific and chimeric molecules having an anti-*H. pylori* portion.

Both monoclonal and polyclonal antibodies (Ab) directed against *H. pylori* polypeptides or *H. pylori* polypeptide variants, and antibody fragments such as Fab' and F(ab')₂, can be used to block the action of *H. pylori* polypeptide and allow the study of the role of a particular *H. pylori* polypeptide of the invention in aberrant or unwanted

- 66 -

intracellular signaling, as well as the normal cellular function of the *H. pylori* and by microinjection of anti-*H. pylori* polypeptide antibodies of the present invention.

Antibodies which specifically bind *H. pylori* epitopes can also be used in immunohistochemical staining of tissue samples in order to evaluate the abundance and pattern of expression of *H. pylori* antigens. Anti *H. pylori* polypeptide antibodies can be used diagnostically in immuno-precipitation and immuno-blotting to detect and evaluate *H. pylori* levels in tissue or bodily fluid as part of a clinical testing procedure. Likewise, the ability to monitor *H. pylori* polypeptide levels in an individual can allow determination of the efficacy of a given treatment regimen for an individual afflicted with such a disorder. The level of an *H. pylori* polypeptide can be measured in cells found in bodily fluid, such as in urine samples or can be measured in tissue, such as produced by gastric biopsy. Diagnostic assays using anti-*H. pylori* antibodies can include, for example, immunoassays designed to aid in early diagnosis of *H. pylori* infections. The present invention can also be used as a method of detecting antibodies contained in samples from individuals infected by this bacterium using specific *H. pylori* antigens.

Another application of anti-*H. pylori* polypeptide antibodies of the invention is in the immunological screening of cDNA libraries constructed in expression vectors such as λ gt11, λ gt18-23, λ ZAP, and λ ORF8. Messenger libraries of this type, having coding sequences inserted in the correct reading frame and orientation, can produce fusion proteins. For instance, λ gt11 will produce fusion proteins whose amino termini consist of β -galactosidase amino acid sequences and whose carboxy termini consist of a foreign polypeptide. Antigenic epitopes of a subject *H. pylori* polypeptide can then be detected with antibodies, as, for example, reacting nitrocellulose filters lifted from infected plates with anti-*H. pylori* polypeptide antibodies. Phage, scored by this assay, can then be isolated from the infected plate. Thus, the presence of *H. pylori* gene homologs can be detected and cloned from other species, and alternate isoforms (including splicing variants) can be detected and cloned.

VIII. Kits Containing Nucleic Acids, Polypeptides or Antibodies of the Invention

The nucleic acid, polypeptides and antibodies of the invention can be combined with other reagents and articles to form kits. Kits for diagnostic purposes typically comprise the nucleic acid, polypeptides or antibodies in vials or other suitable vessels. Kits typically comprise other reagents for performing hybridization reactions, polymerase chain reactions (PCR), or for reconstitution of lyophilized components, such as aqueous media, salts, buffers, and the like. Kits may also comprise reagents for sample processing such as detergents, chaotropic salts and the like. Kits may also

- 67 -

comprise immobilization means such as particles, supports, wells, dipsticks and the like. Kits may also comprise labeling means such as dyes, developing reagents, radioisotopes, fluorescent agents, luminescent or chemiluminescent agents, enzymes, intercalating agents and the like. With the nucleic acid and amino acid sequence information provided
5 herein, individuals skilled in art can readily assemble kits to serve their particular purpose. Kits further can include instructions for use.

IX. Drug Screening Assays Using *H. pylori* Polypeptides

By making available purified and recombinant *H. pylori* polypeptides, the
10 present invention provides assays which can be used to screen for drugs which are either agonists or antagonists of the normal cellular function, in this case, of the subject *H. pylori* polypeptides, or of their role in intracellular signaling. Such inhibitors or potentiators may be useful as new therapeutic agents to combat *H. pylori* infections in humans. A variety of assay formats will suffice and, in light of the present inventions,
15 will be comprehended by the skilled artisan.

In many drug screening programs which test libraries of compounds and natural extracts, high throughput assays are desirable in order to maximize the number of compounds surveyed in a given period of time. Assays which are performed in cell-free systems, such as may be derived with purified or semi-purified proteins, are often
20 preferred as "primary" screens in that they can be generated to permit rapid development and relatively easy detection of an alteration in a molecular target which is mediated by a test compound. Moreover, the effects of cellular toxicity and/or bioavailability of the test compound can be generally ignored in the *in vitro* system, the assay instead being focused primarily on the effect of the drug on the molecular target as may be manifest in
25 an alteration of binding affinity with other proteins or change in enzymatic properties of the molecular target. Accordingly, in an exemplary screening assay of the present invention, the compound of interest is contacted with an isolated and purified *H. pylori* polypeptide.

Screening assays can be constructed *in vitro* with a purified *H. pylori*
30 polypeptide or fragment thereof, such as an *H. pylori* polypeptide having enzymatic activity, such that the activity of the polypeptide produces a detectable reaction product. The efficacy of the compound can be assessed by generating dose response curves from data obtained using various concentrations of the test compound. Moreover, a control assay can also be performed to provide a baseline for comparison. Suitable products
35 include those with distinctive absorption, fluorescence, or chemi-luminescence properties, for example, because detection may be easily automated. A variety of synthetic or naturally occurring compounds can be tested in the assay to identify those

- 68 -

which inhibit or potentiate the activity of the *H. pylori* polypeptide. Some of these active compounds may directly, or with chemical alterations to promote membrane permeability or solubility, also inhibit or potentiate the same activity (e.g., enzymatic activity) in whole, live *H. pylori* cells.

5

This invention is further illustrated by the following examples which should not be construed as limiting. The contents of all references and published patent applications cited throughout this application are hereby incorporated by reference.

10

EXEMPLIFICATION

I. Cloning and Sequencing of *H. pylori* DNA

H. pylori chromosomal DNA was isolated according to a basic DNA protocol outlined in Schleif R.F. and Wensink P.C., *Practical Methods in Molecular Biology*, p.98, Springer-Verlag, NY., 1981, with minor modifications. Briefly, cells were pelleted, resuspended in TE (10 mM Tris, 1 mM EDTA, pH 7.6) and GES lysis buffer (5.1 M guanidium thiocyanate, 0.1 M EDTA, pH 8.0, 0.5% N-laurylsarcosine) was added. Suspension was chilled and ammonium acetate (NH₄Ac) was added to final concentration of 2.0 M. DNA was extracted, first with chloroform, then with phenol-chloroform, and reextracted with chloroform. DNA was precipitated with isopropanol, washed twice with 70% EtOH, dried and resuspended in TE.

Following isolation whole genomic *H. pylori* DNA was nebulized (Bodenteich et al., *Automated DNA Sequencing and Analysis* (J.C. Venter, ed.), Academic Press, 1994) to a median size of 2000 bp. After nebulization, the DNA was concentrated and separated on a standard 1% agarose gel. Several fractions, corresponding to approximate sizes 900-1300 bp, 1300-1700 bp, 1700-2200 bp, 2200-2700 bp, were excised from the gel and purified by the GeneClean procedure (Bio101, Inc.).

The purified DNA fragments were then blunt-ended using T4 DNA polymerase. The healed DNA was then ligated to unique BstXI-linker adapters in 100-1000 fold molar excess. These linkers are complimentary to the BstXI-cut pMPX vectors, while the overhang is not self-complimentary. Therefore, the linkers will not concatemerize nor will the cut-vector religate itself easily. The linker-adopted inserts were separated from the unincorporated linkers on a 1% agarose gel and purified using GeneClean. The linker-adopted inserts were then ligated to each of the 20 pMPX vectors to construct a series of "shotgun" subclone libraries. The vectors contain an out-of-frame lacZ gene at the cloning site which becomes in-frame in the event that an adapter-dimer is cloned, allowing these to be avoided by their blue-color.

- 69 -

All subsequent steps were based on the multiplex DNA sequencing protocols outlined in Church G.M. and Kieffer-Higgins S., *Science* 240:185-188, 1988. Only major modifications to the protocols are highlighted. Briefly, each of the 20 vectors was then transformed into DH5 α competent cells (Gibco/BRL, DH5 α transformation
5 protocol). The libraries were assessed by plating onto antibiotic plates containing ampicillin, methicillin and IPTG/Xgal. The plates were incubated overnight at 37°C. Successful transformants were then used for plating of clones and pooling into the multiplex pools. The clones were picked and pooled into 40 ml growth medium cultures. The cultures were grown overnight at 37°C. DNA was purified using the
10 Qiagen Midi-prep kits and Tip-100 columns (Qiagen, Inc.). In this manner, 100 μ g of DNA was obtained per pool. Fifteen 96-well plates of DNA were generated to obtain a 5-10 fold sequence redundancy assuming 250-300 base average read-lengths.

These purified DNA samples were then sequenced using the multiplex DNA sequencing based on chemical degradation methods (Church G.M. and Kieffer-Higgins
15 S., *Science* 240:185-188, 1988) or by Sequithrem (Epicenter Technologies) dideoxy sequencing protocols. The sequencing reactions were electrophoresed and transferred onto nylon membranes by direct transfer electrophoresis from 40 cm gels (Richterich P. and Church G.M., *Methods in Enzymology* 218:187-222, 1993) or by electroblotting (Church, *supra*). 24 samples were run per gel. 45 successful membranes were produced
20 by chemical sequencing and 8 were produced by dideoxy sequencing. The DNA was covalently bound to the membranes by exposure to ultraviolet light, and hybridized with labeled oligonucleotides complimentary to tag sequences on the vectors (Church, *supra*). The membranes were washed to rinse off non-specifically bound probe, and exposed to X-ray film to visualize individual sequence ladders. After autoradiography, the
25 hybridized probe was removed by incubation at 65° C, and the hybridization cycle repeated with another tag sequence until the membrane had been probed 38 times for chemical sequencing membranes and 10 times for the dideoxy sequencing membranes. Thus, each gel produced a large number of films, each containing new sequencing information. Whenever a new blot was processed, it was initially probed for an internal
30 standard sequence added to each of the pools.

Digital images of the films were generated using a laser-scanning densitometer (Molecular Dynamics, Sunnyvale, CA). The digitized images were processed on computer workstations (VaxStation 4000's) using the program REPLICA™ (Church et al., *Automated DNA Sequencing and Analysis* (J.C. Venter, ed.), Academic Press, 1994).
35 Image processing included lane straightening, contrast adjustment to smooth out intensity differences, and resolution enhancement by iterative gaussian deconvolution. The sequences were then automatically picked in REPLICA™ and displayed for

- 70 -

interactive proofreading before being stored in a project database. The proofreading was accomplished by a quick visual scan of the film image followed by mouse clicks on the bands of the displayed image to modify the base calls. Many of the sequence errors could be detected and corrected because multiple sequence reads covering the same
5 portion of the genomic DNA provide adequate sequence redundancy for editing. Each sequence automatically received an identification number (corresponding to microtiter plate, probe information, and lane set number). This number serves as a permanent identifier of the sequence so it is always possible to identify the original of any particular sequence without recourse to a specialized database.

10 Routine assembly of *H. pylori* sequences was done using the program FALCON (Church, Church et al., *Automated DNA Sequencing and Analysis* (J.C. Venter, ed.), Academic Press, 1994). This program has proven to be fast and reliable for most sequences. The assembled contigs were displayed using a modified version of GelAssemble, developed by the Genetics Computer Group (GCG) (Devereux et al.,
15 *Nucleic Acid Res.* 12:387-95, 1984) that interacts with REPLICA™. This provided for an integrated editor that allows multiple sequence gel images to be instantaneously called up from the REPLICA™ database and displayed to allow rapid scanning of contigs and proofreading of gel traces where discrepancies occurred between different sequence reads in the assembly.

20

II. Identification, cloning and expression of recombinant *H. pylori* DNA sequences

To facilitate the cloning, expression and purification of membrane and secreted proteins from *H. pylori* a powerful gene expression system, the pET System (Novagen), for cloning and expression of recombinant proteins in *E. coli*, was selected. Also, a
25 DNA sequence encoding a peptide tag, the His-Tag, was fused to the 3' end of DNA sequences of interest in order to facilitate purification of the recombinant protein products. The 3' end was selected for fusion in order to avoid alteration of any 5' terminal signal sequence. The exception to the above was ppiB, a gene cloned for use as a control in the expression studies. In this study, the sequence for *H. pylori* ppiB
30 contains a DNA sequence encoding a His-Tag fused to the 5' end of the full length gene, because the protein product of this gene does not contain a signal sequence and is expressed as a cytosolic protein.

PCR Amplification and cloning of DNA sequences containing ORF's for membrane and secreted proteins from the J99 Strain of Helicobacter pylori.

Sequences chosen (from the list of the DNA sequences of the invention) for cloning from the J99 strain of *H. pylori* were prepared for amplification cloning by polymerase chain reaction (PCR). Synthetic oligonucleotide primers (Table 3) specific for the 5' and 3' ends of open reading frames (ORFs) were designed and purchased (GibcoBRL Life Technologies, Gaithersburg, MD, USA). All forward primers (specific for the 5' end of the sequence) were designed to include an NcoI cloning site at the extreme 5' terminus, except for HpSeq. 4821082 where NdeI was used. These primers were designed to permit initiation of protein translation at a methionine residue followed by a valine residue and the coding sequence for the remainder of the native *H. pylori* DNA sequence. An exception is *H. pylori* sequence 4821082 where the initiator methionine is immediately followed by the remainder of the native *H. pylori* DNA sequence. All reverse primers (specific for the 3' end of any *H. pylori* ORF) included a EcoRI site at the extreme 5' terminus to permit cloning of each *H. pylori* sequence into the reading frame of the pET-28b. The pET-28b vector provides sequence encoding an additional 20 carboxy-terminal amino acids (only 19 amino acids in HpSeq. 26380318 and HpSeq. 14640637) including six histidine residues (at the extreme C-terminus), which comprise the His-Tag. An exception to the above, as noted earlier, is the vector construction for the *ppiB* gene. A synthetic oligonucleotide primer specific for the 5' end of *ppiB* gene encoded a BamHI site at its extreme 5' terminus and the primer for the 3' end of the *ppiB* gene encoded a XhoI site at its extreme 5' terminus.

TABLE 3

Oligonucleotide primers used for PCR amplification of *H. pylori* DNA sequences

Outer membrane Proteins	Forward primer 5' to 3'	Reverse Primer 5' to 3'
Protein 16225006	5'-TATACCATGGTGGG CGCTAA-3' (SEQ ID NO:195)	5'-ATGAATTCGAGTA AGGATTTTG-3' (SEQ ID NO:196)
Protein 26054702	5'-TTAACCATGGTGA AAAGCGATA-3' (SEQ ID NO:197)	5'-TAGAATTCGCATA ACGATCAATC-3' (SEQ ID NO:198)
Protein 7116626	5'-ATATCCATGGTGA GTTTGATGA-3' (SEQ ID NO:199)	5'-ATGAATTCAATT TTTATTTTGCCA-3' (SEQ ID NO:200)

- 72 -

Protein 29479681	5'-AATTCCATGGTGG GGGCTATG-3' (SEQ ID NO:201)	5'-ATGAATTCTCGAT AGCCAAAATC-3' (SEQ ID NO:202)
Protein 14640637	5'-AATTCCATGGTG CATAACTTCCATT-3' (SEQ ID NO:203)	5'-AAGAATTCTCTA GCATCCAAATGGA-3' (SEQ ID NO:204)
Periplasmic/ Secreted Proteins		
Protein 30100332	5'-ATTTCCATGGTCATG TCTCATATT-3' (SEQ ID NO:205)	5'-ATGAATTCCATC TTTTATTCCAC-3' (SEQ ID NO:206)
Protein 4721061	5'-AACCATGGTGATT TAAGCATTGAAAG-3' (SEQ ID NO:207)	5'-AAGAATTCCAC TCAAATTTTTTAAC AG-3' (SEQ ID NO:208)
Other Surface Proteins		
Protein 4821082	5'-GATCATCCATATGTT ATCTTCTAAT-3' (SEQ ID NO:209)	5'-TGAATTCAACCA TTTAAACCCTG-3' (SEQ ID NO:210)
Protein 978477	5'-TATACCATGGTGAA ATTTTTTCTTTTA-3' (SEQ ID NO:211)	5'-AGAATTCAATT GCGTCTTGTAAG- 3' (SEQ ID NO:212)
Inner Membrane Protein		
Protein 26380318	5'-TATACCATGGTGAT GGACAAACTC-3' (SEQ ID NO:213)	5'-ATGAATTCCCACTT GGGGCGATA-3' (SEQ ID NO:214)
Cytoplasmic Protein		
ppi	5'-TTATGGATCCAAAC CAATTAAACT-3' (SEQ ID NO:215)	5'-TATCTCGAGTTATA GAGAAGGGC-3' (SEQ ID NO:216)

Genomic DNA prepared from the J99 strain of *H. pylori* (ATCC #55679; deposited by Genome Therapeutics Corporation, 100 Beaver Street, Waltham, MA 02154) was used as the source of template DNA for PCR amplification reactions

5 (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel et al., eds., 1994). To amplify a DNA sequence containing an *H. pylori* ORF, genomic DNA (50 nanograms) was introduced into a reaction vial containing 2 mM MgCl₂, 1 micromolar synthetic oligonucleotide primers (forward and reverse primers) complementary to and flanking a defined *H. pylori* ORF, 0.2 mM of each

10 deoxynucleotide triphosphate; dATP, dGTP, dCTP, dTTP and 2.5 units of heat stable

- 73 -

DNA polymerase (Amplitaq, Roche Molecular Systems, Inc., Branchburg, NJ, USA) in a final volume of 100 microliters. The following thermal cycling conditions were used to obtain amplified DNA products for each ORF using a Perkin Elmer Cetus/ GeneAmp PCR System 9600 thermal cycler:

5

Protein 26054702, Protein 7116626, Protein 29479681, Protein 30100332, and Protein 4821082;

Denaturation at 94°C for 2 min,

2 cycles at 94°C for 15 sec, 30°C for 15 sec and 72°C for 1.5 min

10 23 cycles at 94°C for 15 sec, 55°C for 15 sec and 72°C for 1.5 min

Reactions were concluded at 72°C for 6 minutes.

Protein 16225006;

Denaturation at 94°C for 2 min,

15 25 cycles at 95°C for 15 sec, 55°C for 15 sec and 72°C for 1.5 min

Reaction was concluded at 72°C for 6 minutes.

Protein 4721061;

Denaturation at 94°C for 2 min,

20 2 cycles at 94°C for 15 sec, 36°C for 15 sec and 72°C for 1.5 min

23 cycles at 94°C for 15 sec, 60°C for 15 sec and 72°C for 1.5 min

Reactions were concluded at 72°C for 6 minutes.

Protein 26380318;

25 Denaturation at 94°C for 2 min,

2 cycles at 94°C for 15 sec, 38°C for 15 sec and 72°C for 1.5 min

23 cycles at 94°C for 15 sec, 62°C for 15 sec and 72°C for 1.5 min

Reactions were concluded at 72°C for 6 minutes.

30 **Protein 14640637;**

Denaturation at 94°C for 2 min,

2 cycles at 94°C for 15 sec, 33°C for 15 sec and 72°C for 1.5 min

30 cycles at 94°C for 15 sec, 55°C for 15 sec and 72°C for 1.5 min

Reactions were concluded at 72°C for 6 minutes.

35

- 74 -

Conditions for amplification of *H. pylori* ppiB;

Denaturation at 94°C for 2 min,

2 cycles at 94°C for 15 sec, 32°C for 15 sec and 72°C for 1.5 min

25 cycles at 94°C for 15 sec, 56°C for 15 sec and 72°C for 1.5 min

5 Reactions were concluded at 72°C for 6 minutes

Upon completion of thermal cycling reactions, each sample of amplified DNA was washed and purified using the Qiaquick Spin PCR purification kit (Qiagen, Gaithersburg, MD, USA). All amplified DNA samples were subjected to digestion with
10 the restriction endonucleases, NcoI and EcoRI (New England BioLabs, Beverly, MA, USA), or in the case of HpSeq. 4821082 (SEQ ID NO: 1309), with NdeI and EcoRI (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel et al., eds., 1994). DNA samples were then subjected to electrophoresis on 1.0 % NuSeive (FMC BioProducts, Rockland, ME USA) agarose gels. DNA was visualized by
15 exposure to ethidium bromide and long wave uv irradiation. DNA contained in slices isolated from the agarose gel was purified using the Bio 101 GeneClean Kit protocol (Bio 101 Vista, CA, USA).

Cloning of H. pylori DNA sequences into the pET-28b prokaryotic expression vector.

20 The pET-28b vector was prepared for cloning by digestion with NcoI and EcoRI, or in the case of *H. pylori* protein 4821082 with NdeI and EcoRI (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel et al., eds., 1994). In the case of cloning ppiB, the pET-28a vector, which encodes a His-Tag that can be fused to the 5' end of an inserted gene, was used and the cloning site prepared for cloning with the
25 ppiB gene by digestion with BamHI and XhoI restriction endonucleases.

Following digestion, DNA inserts were cloned (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel et al., eds., 1994) into the previously digested pET-28b expression vector, except for the amplified insert for ppiB, which was cloned into the pET-28a expression vector. Products of the ligation reaction were then
30 used to transform the BL21 strain of *E. coli* (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel et al., eds., 1994) as described below.

Transformation of competent bacteria with recombinant plasmids

Competent bacteria, *E. coli* strain BL21 or *E. coli* strain BL21(DE3), were
35 transformed with recombinant pET expression plasmids carrying the cloned *H. pylori* sequences according to standard methods (Current Protocols in Molecular, John Wiley and Sons, Inc., F. Ausubel et al., eds., 1994). Briefly, 1 microliter of ligation reaction

- 75 -

was mixed with 50 microliters of electrocompetent cells and subjected to a high voltage pulse, after which, samples were incubated in 0.45 milliliters SOC medium (0.5% yeast extract, 2.0 % tryptone, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl₂, 10 mM MgSO₄ and 20, mM glucose) at 37°C with shaking for 1 hour. Samples were then spread on LB agar plates containing 25 microgram/ml kanamycin sulfate for growth overnight. Transformed colonies of BL21 were then picked and analyzed to evaluate cloned inserts as described below.

Identification of recombinant pET expression plasmids carrying H. pylori sequences

Individual BL21 clones transformed with recombinant pET-28b-H.pylori ORFs were analyzed by PCR amplification of the cloned inserts using the same forward and reverse primers, specific for each *H. pylori* sequence, that were used in the original PCR amplification cloning reactions. Successful amplification verified the integration of the *H. pylori* sequences in the expression vector (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel et al., eds., 1994).

Isolation and Preparation of plasmid DNA from BL21 transformants

Individual clones of recombinant pET-28b vectors carrying properly cloned *H. pylori* ORFs were picked and incubated in 5 mls of LB broth plus 25 microgram/ml kanamycin sulfate overnight. The following day plasmid DNA was isolated and purified using the Qiagen plasmid purification protocol (Qiagen Inc., Chatsworth, CA, USA).

Expression of recombinant H. pylori sequences in E. coli

The pET vector can be propagated in any *E. coli* K-12 strain e.g. HMS174, HB101, JM109, DH5, etc. for the purpose of cloning or plasmid preparation. Hosts for expression include *E. coli* strains containing a chromosomal copy of the gene for T7 RNA polymerase. These hosts are lysogens of bacteriophage DE3, a lambda derivative that carries the lacI gene, the lacUV5 promoter and the gene for T7 RNA polymerase. T7 RNA polymerase is induced by addition of isopropyl-B-D-thiogalactoside (IPTG), and the T7 RNA polymerase transcribes any target plasmid, such as pET-28b, carrying a T7 promoter and a gene of interest. Strains used include: BL21(DE3) (Studier, F.W., Rosenberg, A.H., Dunn, J.J., and Dubendorff, J.W. (1990) Meth. Enzymol. 185, 60-89).

To express recombinant *H. pylori* sequences, 50 nanograms of plasmid DNA isolated as described above was used to transform competent BL21(DE3) bacteria as described above (provided by Novagen as part of the pET expression system kit). The lacZ gene (beta-galactosidase) was expressed in the pET-System as described for the *H.*

- 76 -

pylori recombinant constructions. Transformed cells were cultured in SOC medium for 1 hour, and the culture was then plated on LB plates containing 25 micrograms/ml kanamycin sulfate. The following day, bacterial colonies were pooled and grown in LB medium containing kanamycin sulfate (25 micrograms/ml) to an optical density at 600 nm of 0.5 to 1.0 O.D. units, at which point, 1 millimolar IPTG was added to the culture for 3 hours to induce gene expression of the *H. pylori* recombinant DNA constructions.

After induction of gene expression with IPTG, bacteria were pelleted by centrifugation in a Sorvall RC-3B centrifuge at 3500 x g for 15 minutes at 4°C. Pellets were resuspended in 50 milliliters of cold 10 mM Tris-HCl, pH 8.0, 0.1 M NaCl and 0.1 mM EDTA (STE buffer). Cells were then centrifuged at 2000 x g for 20 min at 4°C. Wet pellets were weighed and frozen at -80°C until ready for protein purification.

III. Purification of recombinant proteins from *E. coli*

Analytical Methods

The concentrations of purified protein preparations were quantified spectrophotometrically using absorbance coefficients calculated from amino acid content (Perkins, S.J. 1986 Eur. J. Biochem. 157, 169-180). Protein concentrations were also measured by the method of Bradford, M.M. (1976) Anal. Biochem. 72, 248-254, and Lowry, O.H., Rosebrough, N., Farr, A.L. & Randall, R.J. (1951) J. Biol. Chem. 193, pages 265-275, using bovine serum albumin as a standard.

SDS-polyacrylamide gels (12% or 4.0 to 25 % acrylamide gradient gels) were purchased from BioRad (Hercules, CA, USA), and stained with Coomassie blue. Molecular weight markers included rabbit skeletal muscle myosin (200 kDa), *E. coli* (-galactosidase (116 kDa), rabbit muscle phosphorylase B (97.4 kDa), bovine serum albumin (66.2 kDa), ovalbumin (45 kDa), bovine carbonic anhydrase (31 kDa), soybean trypsin inhibitor (21.5 kDa), egg white lysozyme (14.4 kDa) and bovine aprotinin (6.5 kDa).

1. Purification of soluble proteins

All steps were carried out at 4°C. Frozen cells were thawed, resuspended in 5 volumes of lysis buffer (20 mM Tris, pH 7.9, 0.5 M NaCl, 5 mM imidazole with 10% glycerol, 0.1 % 2-mercaptoethanol, 200 µg/ml lysozyme, 1 mM phenylmethylsulfonyl fluoride (PMSF), and 10 µg/ml each of leupeptin, aprotinin, pepstatin, L-1-chloro-3-[4-tosylamido]-7-amino-2-heptanone (TLCK), L-1-chloro-3-[4-tosylamido]-4-phenyl-2-butanone (TPCK), and soybean trypsin inhibitor, and ruptured by several passages through a small volume microfluidizer (Model M-110S, Microfluidics International

- 77 -

Corporation, Newton, MA). The resultant homogenate was made 0.1 % Brij 35, and centrifuged at 100,000 x g for 1 hour to yield a clear supernatant (crude extract).

Following filtration through a 0.8 μ m Supor filter (Gelman Sciences, FRG) the crude extract was loaded directly onto a Ni^{2+} -nitrilotriacetate-agarose (NTA) with a 5
5 milliliter bed volume (Hochuli, E., Dbeli, H., and Schacheer, A. (1987) J. Chromatography 411, 177-184) pre-equilibrated in lysis buffer containing 10 % glycerol, 0.1 % Brij 35 and 1 mM PMSF. The column was washed with 250 ml (50 bed volumes) of lysis buffer containing 10 % glycerol, 0.1 % Brij 35, and was eluted with sequential steps of lysis buffer containing 10 % glycerol, 0.05 % Brij 35, 1 mM PMSF,
10 and 20, 100, 200, and 500 mM imidazole in succession. Fractions were monitored by absorbance at OD₂₈₀ nm, and peak fractions were analyzed by SDS-PAGE. Fractions containing the recombinant protein eluted at 100 mM imidazole.

Recombinant protein 14640637 and proteins, beta-galactosidase (lacZ) and peptidyl-
15 *prolyl cis-trans isomerase (ppiB)*

Fractions containing the recombinant proteins from the Ni^{2+} -NTA-agarose columns were pooled and then concentrated to approximately 5 ml by centrifugal filtration (Centriprep-10, Amicon, MA), and loaded directly onto a 180-ml column (1.6 X 91 cm) of Sephacryl S-100 HR gel filtration medium equilibrated in Buffer A (10 mM
20 Hepes, pH 7.5, 150 mM NaCl, 0.1 mM EGTA) and run in Buffer A at 18 ml/h. Fractions containing the recombinant protein were identified by absorbance at 280 nm and analyzed by SDS-PAGE. Fractions were pooled and concentrated by centrifugal filtration.

25 *Recombinant protein 7116626*

Fractions containing the recombinant protein from the Ni^{2+} -NTA-agarose column were pooled and dialyzed overnight against 1 liter of dialysis buffer (10 mM MOPS, pH 6.5, 50 mM NaCl, 0.1 mM EGTA, 0.02% Brij 35 and 1 mM PMSF). In the morning, a fine white precipitate was removed by centrifugation and the resulting
30 supernatant was loaded onto an 8 ml (8 x 75 mm) MonoS high performance liquid chromatography column (Pharmacia Biotechnology, Inc., Piscataway, NJ, USA) equilibrated in buffer B (10 mM MOPS, pH 6.5, 0.1 mM EGTA) containing 50 mM NaCl. The column was washed with 10 bed volumes of buffer B containing 50 mM NaCl, and developed with a 50-ml linear gradient of increasing NaCl (50 to 500 mM).
35 Recombinant protein 7116626 eluted as a sharp peak at 300 mM NaCl.

- 78 -

2. Purification of insoluble proteins from inclusion bodies

The following steps were carried out at 4°C. Cell pellets were resuspended in lysis buffer with 10% glycerol 200 µg/ml lysozyme, 5 mM EDTA, 1mM PMSF and 0.1 % -mercaptoethanol. After passage through the cell disrupter, the resulting homogenate
5 was made 0.2 % deoxycholate, stirred 10 minutes, then centrifuged at 20,000 x g, for 30 min. The pellets were washed with lysis buffer containing 10 % glycerol, 10 mM EDTA, 1% Triton X-100, 1 mM PMSF and 0.1% -mercaptoethanol, followed by several washes with lysis buffer containing 1 M urea, 1 mM PMSF and 0.1 % 2-mercaptoethanol. The resulting white pellet was composed primarily of inclusion
10 bodies, free of unbroken cells and membranous materials.

Recombinant proteins 26054702, 16225006, 30100332, 4721061

The following steps were carried out at room temperature. Purified inclusion bodies were dissolved in 20 ml 8.0 M urea in lysis buffer with 1 mM PMSF and 0.1 %
15 2-mercaptoethanol, and incubated at room temperature for 1 hour. Materials that did not dissolve were removed by centrifugation. The clear supernatant was filtered, then loaded onto a Ni²⁺-NTA agarose column pre-equilibrated in 8.0 M urea in Lysis Buffer. The column was washed with 250 ml (50 bed volumes) of lysis buffer containing 8 M urea, 1.0 mM PMSF and 0.1 % 2-mercaptoethanol, and developed with
20 sequential steps of lysis buffer containing 8M urea, 1 mM PMSF, 0.1 % 2-mercaptoethanol and 20, 100, 200, and 500 mM imidazole in succession. Fractions were monitored by absorbance at OD₂₈₀ nm, and peak fractions were analyzed by SDS-PAGE. Fractions containing the recombinant protein eluted at 100 mM imidazole.

Recombinant proteins 29479681, 26380318

The pellet containing the inclusion bodies was solubilized in buffer B containing 8 M urea, 1 mM PMSF and 0.1 % 2-mercaptoethanol, and incubated for 1 hour at room temperature. Insoluble materials were removed by centrifugation at 20,000 x g for 30 min, and the cleared supernatant was loaded onto a 15 ml (1.6 x 7.5 cm) SP-Sepharose
30 column pre-equilibrated in buffer B, 6 M urea, 1 mM PMSF, 0.1 % 2-mercaptoethanol. After washing the column with 10 bed volumes, the column was developed with a linear gradient from 0 to 500 mM NaCl.

Dialysis and concentration of protein samples

35 Urea was removed slowly from the protein samples by dialysis against Tris-buffered saline (TBS; 10 mM Tris pH 8.0, 150 mM NaCl) containing 0.5 % deoxycholate (DOC) with sequential reduction in urea concentration as follows; 6M,

- 79 -

4M, 3M, 2M, 1M, 0.5 M and finally TBS without any urea. Each dialysis step was conducted for a minimum of 4 hours at room temperature.

After dialysis, samples were concentrated by pressure filtration using Amicon stirred-cells. Protein concentrations were measured using the methods of Perkins (1986 Eur. J. Biochem. 157, 169-180), Bradford ((1976) Anal. Biochem. 72, 248-254) and Lowry ((1951) J. Biol. Chem. 193, pages 265-275).

The recombinant proteins purified by the methods described above are summarized in Table 4 below.

10

TABLE 4

J99 Sequence Identifier	Homolog identified by Blast	Gene symbol of Homolog	Bacterial cell fraction used to purify recombinant proteins	Method of purification	Relative MW on SDS-PAGE gel	Final concentration of purified protein	Composition of buffer
Outer Membrane Proteins							
16225006	P28635	YEAC	Inclusion bodies	His-Tag	18 kDa	5 mg/ml	B
26054702	P15929	flgH	Inclusion bodies	His-Tag	37 kDa	1.18 mg/ml	B
						----	as dry pellet
7116626	P26093	e(P4)	Soluble fraction	His-Tag	29 kDa	0.8 mg/ml	A
						1.85 mg/ml	C
29479681	P13036	fecA	Inclusions bodies	SP-Sephadex	23 kDa	2.36 mg/ml	B
						0.5 mg/ml	B
						----	as dry pellet
14640637	P16665	TPF1	Soluble fraction	His-Tag	17 kDa	2.4 mg/ml	A
				gel filtration S100 HR			
Periplasmic/Secreted Protein							
1010032	P23847	dppA	Inclusion bodies	His-Tag	11 kDa	2.88 mg/ml	B
4721061	P36175	GCP	Inclusion bodies	His-Tag	38 kDa	2.8 mg/ml	B

- 80 -

Other Surface Proteins

4821082	P08089	M protein	Inclusion bodies	His-Tag	20 kDa	1.16 mg/ml	B
978477	L28919	FBP54	Inclusion bodies	SP-Sepharose	44 kDa	2.56 mg/ml	B
						0.3 mg/ml	B

Inner Membrane Proteins

26380318	P15933	fliG	Inclusion bodies	SP-Sepharose	11 kDa	22 mg/ml	B

Control Proteins with His-Tag

	P00722	lacZ	Soluble fraction	His-Tag	116 kDa	10 mg/ml	A
				gel filtration S200 HR			
		ppiB	Soluble fraction	His-Tag	21 kDa	4.4 mg/ml	A
				gel filtration S100 HR			
Buffer composition							
S:							
A=10 mM Hepes pH 7.5, 150 mM NaCl, 0.1 mM EGTA							
B= 10 mM Tris pH 8.0, 150 mM NaCl, 0.5 % DOC							
C= 10 mM MOPS pH 6.5, 300 mM NaCl, 0.1 EGTA							

5 **IV. Analysis of *H. pylori* proteins as Vaccine candidates**

To analyze *H. pylori* proteins for use in the vaccine formulations of the invention, several *H. pylori* proteins were expressed, characterized immunologically and tested in animal efficacy studies as outlined below. Specifically, the immunomodulatory effects of *H. pylori* proteins were investigated in a mouse/*H. pylori* model which mimics the human *H. pylori* infection in humans. In these studies, the effect of oral immunization of selected *H. pylori* polypeptides in *H. pylori* infected mice was determined.

Identification, cloning and expression of recombinant Helicobacter pylori sequences.

To facilitate the cloning, expression and purification of membrane and/or secreted proteins from *H. pylori*, the pET gene expression system (Novagen), for cloning and expression of recombinant proteins in *Escherichia coli* was selected. Further, for proteins that have a signal sequence at their amino-terminal end, a DNA sequence encoding a peptide tag (His-tag) was fused to the 5' end of the *H. pylori* DNA sequences of interest in order to facilitate purification of the recombinant protein products.

PCR amplification and cloning of DNA sequences containing ORFs for membrane and secreted proteins from the J99 strain of Helicobacter pylori.

The sequences selected (from the list of the DNA sequences of the invention) for cloning from *H. pylori* strain J99 were prepared for amplification cloning by the polymerase chain reaction (PCR). All of the selected sequences encode for outer membrane *H. pylori* proteins, with vac9 (SEQ ID NO:125), vac10 (SEQ ID NO:147), vac22 (SEQ ID NO:121) and vac41 (SEQ ID NO:176) sequences all sharing a terminal phenylalanine residue. Likewise, the vac32 (SEQ ID NO:108), vac36 (SEQ ID NO:149) and vac37 (SEQ ID NO:139) sequences all share a terminal phenylalanine residue and a tyrosine cluster at the C-terminus. Synthetic oligonucleotide primers for each ORF of interest (Table 5) specific for the predicted mature 5' end of the ORF and downstream (3') of the predicted translational termination codon were designed and purchased (GibcoBRL Life Technologies, Gaithersburg, MD, USA). All forward primers (specific for the 5' terminus of the region of ORF of interest) were designed to include a *Bam*III restriction site followed by a *Nde*I restriction site. These primers were designed to permit the initiation of protein translation at a methionine residue encoding within the *Nde*I restriction site sequence (in the case of producing a non His-tagged recombinant protein) or to fuse in frame with the DNA sequence encoding the His-tag (for producing His-tagged recombinant protein), followed by the coding sequence for the remainder of the native *H. pylori* DNA. All reverse oligonucleotide primers (specific for downstream (3') of the predicted translational termination codon of the ORF) were designed to include an *Eco*RI restriction site at the 5' terminus. This combination of primers would enable each ORF of interest to be cloned into pET28b (to produce a His-tagged recombinant protein) or pET30a (to produce a non His-tagged or native recombinant protein). The pET28b vector provides sequence encoding an additional 20 amino-terminal amino acids (plus the methionine in the *Nde*I restriction site) including a stretch of six histidine residues which makes up the His-tag.

Genomic DNA prepared from *H. pylori* strain J99 (ATCC 55679) was used as the source of template DNA for the PCR amplification reactions (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel *et al.*, eds., 1994). To amplify a DNA sequence containing a specific *H. pylori* ORF, genomic DNA (50 nanograms) was introduced into a reaction tube containing 200 nanograms of both the forward and reverse synthetic oligonucleotide primer specific for the ORF of interest, and 45 microliters of PCR SuperMix purchased (GibcoBRL Life Technologies, Gaithersburg, MD, USA) in a total of 50 microliters. The PCR SuperMix is supplied in 1.1X concentrations and contains 22mM Tris-HCl (pH 8.4), 55mM KCl, 1.65 mM

MgCl₂, 220 micromolar of each dATP, dCTP, dGTP and dTTP, 22 units recombinant *Taq* polymerase/ml and stabilizers. The following thermal cycling conditions were used to obtain amplified DNA products for each ORF using a Perkin Elmer Cetus/Gene Amp PCR System thermal cycler.

5

Table 5: Oligonucleotide primers

Gene	Forward primer	Reverse primer
vac9 (nt SEQ ID NO:28) (aa SEQ ID NO: 125)	CGCGGATCCATATGGCTGAAA AAACGCCTTTTTTTTAAACTAA AAACCAC (SEQ ID NO: 257)	CCGGAATTCATCAGTATTCAA TGGGAATAAAGCC (SEQ ID NO: 258)
vac10 (nt SEQ ID NO:50) (aa SEQ ID NO: 147)	CGCGGATCCATATGAAAGAAG AAGAAAAAGAAGAAAAAAG ACAGAAAGG (SEQ ID NO: 259)	CCGGAATTCGCTTAAAAGAAA ATAGTCCCCCAAACGC (SEQ ID NO: 260)
vac22 (nt SEQ ID NO:24) (aa SEQ ID NO: 121)	CGCCGGATCCATATGAAAGAG GTCATTCCACCCCTTCAACCCC (SEQ ID NO: 261)	CCGGAATTCATATAAATATCA TATAGGCAGAAAAAC (SEQ ID NO: 262)
vac32 (nt SEQ ID NO:11) (aa SEQ ID NO: 108)	CGCGGATCCATATGGAGGCAG AGCTTGATGAAAAATC (SEQ ID NO: 263)	CCGGAATTCGATTGATTTGTC AAATCTAAAATCCC (SEQ ID NO: 264)
vac36 (hop B) (nt SEQ ID NO:52) (aa SEQ ID NO:149)	TATTATACATATGGAAGAAGA TGGG (SEQ ID NO: 265)	TAATCTCGAGTTTAGAAGGCG TA (SEQ ID NO: 266)

- 83 -

vac37 (i-hop) (nt SEQ ID NO:42) (aa SEQ ID NO: 139)	TTATATTCATATGGAAGACGAT GGC (SEQ ID NO: 267)	AATTCTCGAGCCTCTTTATAA GCC (SEQ ID NO: 268)
vac41 (nt SEQ ID NO:79) (aa SEQ ID NO: 176)	CGCGGATCCATATGGTAGAAG CCTTTCAAAAACACCAAAAAG ACGG (SEQ ID NO: 269)	CCGGAATTCGGAGCCAATAGG GAGCTAAAGCC (SEQ ID NO: 270)

Sequences for Vac32, Vac9 and Vac22

Denaturation at 94°C for 30 sec

35 cycles at 94°C for 15 sec, 55°C for 15 sec, and 72°C for 1.5 min

- 5 Reactions were concluded at 72°C for 8 minutes

Sequences for Vac10 and Vac41

Denaturation at 94°C for 30 sec

35 cycles at 94°C for 15 sec, 55°C for 15 sec, and 72°C for 2.5 min

- 10 Reactions were concluded at 72°C for 8 minutes

Sequences for Vac36 and Vac37

Denaturation at

2 cycles at 94°C for 15 sec, 30°C for 15 sec, and 72°C for 1.5 min

- 15 23 cycles at 94°C for 15 sec, 55°C for 15 sec, and 72°C for 1.5 min

Reactions were concluded at 72°C for 6 minutes

- 20 Upon completion of the thermal cycling reactions, each sample of amplified DNA was subjected to electrophoresis on 1.0% agarose gels. The DNA was visualized by exposure to ethidium bromide and long wave UV irradiation, and cut out in gel slices. DNA was purified using the Wizard PCR Preps Kit (Promega Corp., Madison, WI, USA), and then subjected to digestion with *Bam*HI and *Eco*RI (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel *et al.*, eds., 1994). The digested PCR amplicon was then re-electrophoresed and purified as before.

- 84 -

Ligation of H. pylori DNA sequences into cloning vectors

The pOK12 vector (J. Vieira and J. Messing, Gene 100:189-194, 1991) was prepared for cloning for digestion with *Bam*HI and *Eco*RI in the case of Vac9, 10, 22, 31 and 32, whereas the pSU21 vector (B. Bartolome *et al.*, Gene 102:75-78, 1991) was prepared for cloning by digestion with *Bam*HI and *Eco*RI in the case of Vac 41 (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel *et al.*, eds., 1994). The vectors were subjected to electrophoresis on 1.0% agarose gels and purified using the Wizard PCR Preps kit (Promega Corp., Madison, WI, USA). Following ligation of the purified, digested vector and the purified, digested amplified *H. pylori* ORF, the products of the ligation reaction were transformed into *E. coli* JM109 competent cells according to standard methods (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel *et al.*, eds., 1994). Individual bacterial colonies were screened for those containing the correct recombinant plasmids by incubating in LB broth overnight (plus 25ug/ml kanamycin sulfate for the pOK12 based plasmids or 25ug/ml chloramphenicol for the pSU21 based plasmids) followed by plasmid DNA preparation using the Magic Minipreps system (Promega Corp., Madison, WI, USA), and then analyzed by restriction digestion (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel *et al.*, eds., 1994).

Cloning of H. pylori DNA sequences into the pET28b and pET30a prokaryotic expression vectors

Both the pET28b and pET30a expression vectors were prepared for cloning by digestion with *Nde*I and *Eco*RI (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel *et al.*, eds., 1994). The *H. pylori* DNA sequences were removed from pOK12 (Vac9,10,23,31 and 32) or pSU21 (Vac41) plasmid backbones by digestion with *Nde*I and *Eco*RI (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel *et al.*, eds., 1994). The pET28b, pET30a and *H. pylori* DNA sequences were all electrophoresed on a 1% agarose gel and purified using the Wizard PCR Preps kit (Promega Corp., Madison WI, USA). Following ligation of the purified, digested expression vector and the purified, digest *H. pylori* DNA sequences, the products of the ligation reaction were transformed into *E. coli* JM109 competent cells (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel *et al.*, eds., 1994). Individual bacterial colonies were screened for those containing the correct recombinant plasmids by preparing plasmid DNA as described above followed by analysis by restriction digestion profiles and DNA sequencing (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel *et al.*, eds., 1994). These recombinant plasmids were then used to transform specific *E. coli* expression strains.

Transformation of competent bacteria with recombinant expression plasmids

Competent bacterial strains (BL21(DE3), BL21(DE3)pLysS, HMS174(DE3) and HMS174(DE3)pLysS were prepared and transformed with the recombinant pET28b expression plasmids carrying the cloned *H. pylori* sequences according to standard methods (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel *et al.*, eds., 1994). These expression host strains contain a chromosomal copy of the gene for T7 RNA polymerase. These hosts are lysogens of bacteriophage DE3, a lambda derivative that carries the *lacI* gene, the *lacUV5* promoter and the gene for T7 RNA polymerase. T7 RNA polymerase expression is induced by the addition of isopropyl- β -D thiogalactoside (IPTG), and the T7 RNA polymerase then transcribes any target plasmid, such as pET28b, that carries a T7 promoter sequence and a gene of interest.

Expression of recombinant H. pylori sequences in E. coli

Transformants were collected from LB agar plates containing 25ug/ml kanamycin sulfate (ensures maintenance of the pET28b-based recombinant plasmids) and used to inoculate LB broth containing 25ug/ml kanamycin sulfate and grown to an optical density at 600nm of 0.5 to 1.0 OD units, at which point 1mM IPTG was added to the culture for one to three hours to induce gene expression of the *H. pylori* recombinant DNA constructions. After induction of gene expression with IPTG, bacteria were pelleted by centrifugation and resuspended in SDS-PAGE solubilization buffer and subjected to SDS-PAGE (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel *et al.*, eds., 1994). Proteins were visualized by staining with Coomassie Brilliant Blue or detected by western immunoblotting using the specific anti-His tag monoclonal antibody (Clontech, Palo Alto, CA, USA) using standard methods (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel *et al.*, eds., 1994). The host strain that provided the highest level of recombinant protein production was then chosen for use in a large-scale induction in order to purify the recombinant protein. All of the following proteins listed were expressed recombinantly and the strain giving the highest level of expression listed: BL21(DE3) (vac31, vac26, vac37); BL21(DE3) pLysS (vac 9, 32); HMS174(DE3) (vac10,11).

- 86 -

Purification of recombinant proteins and generation of specific antiserum

Large scale cultures were inoculated and grown as above, and induced with 1mM IPTG for 3 hours. After induction, bacteria were pelleted by centrifugation in a Sorvall centrifuge at 3500 x g for 15 min at 4°C. All of the expressed recombinant proteins were present in the insoluble inclusion body fraction. Inclusion bodies were purified according to standard protocols (Antibodies, Cold Spring Harbor Laboratory Press, E. Harlow and D. Lane, eds., 1988). The recombinant protein produced by vac32 was solubilized in 8M urea and partially purified by nickel chromatography (REF here). Denatured recombinant proteins were purified by electrophoresis on SDS-PAGE gels, and after visualization with Coomassie Brilliant Blue, the protein was excised from the gel and the gel slices homogenized. This material was used to raise specific polyclonal antibodies in mice or rabbits according to standard protocols (Antibodies, Cold Spring Harbor Laboratory Press, E. Harlow and D. Lane, eds., 1988).

Immunological characterization of recombinant proteins

In all cases where antibody was attempted to be raised, high titre antisera was generated, confirming the immunogenicity of the recombinant proteins. Further, these specific antisera were used to analyze whether the protein encoded by the cloned gene was expressed in *H. pylori*. Western immunoblot analysis using standard protocols (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel *et al.*, eds., 1994) confirmed that the *H. pylori* strain J99 did express proteins of the expected molecular weight that reacted with the vac10, vac32, vac31, vac36 antiserum. The specific antiserum was also used to determine the level of antigenic conservation between a large number of *H. pylori* isolates that had been obtained from distinct geographical sites around the world, and from all types of clinical manifestations, including gastritis, duodenal ulcer, gastric ulcer and gastric cancer. It was found that every strain produced a protein that reacted specifically with each antiserum.

Further, *H. pylori* cells from strains J99, 17874, AH244 and SS1 were fractionated into different cellular compartments (Doig and Trust 1994 Infect. Immun. 62:4526-4533; O'Toole *et al.* 1995 J. Bacteriol. 177:6049-6057). The specific antiserum was used to probe these fractions by western immunoblot to identify in which fraction the protein was localized. In all cases, the immunoreactive protein was present in the outer membrane as had been predicted by the sequence features and motif searches described herein.

*Demonstration of protein efficacy as a vaccine*Purification of vac36 for efficacy studies

All the following steps were carried out at 4°C. Cell pellets were resuspended in 5 volumes per gram of cell of lysis buffer (50mM Sodium Phosphate pH 8.0, 0.5 M NaCl, 5mM Imidazole) with 10mM EDTA, 1 mM phenylmethylsulfonyl fluoride (PMSF) and 0.1 % β -mercaptoethanol, and ruptured by several passages through a small volume microfluidizer (Model M-110S, Microfluidics International Corporation, Newton, MA). The resulting homogenate was made 0.2% sodium deoxycholate (DOC), stirred 20 minutes, then centrifuged (10,000 g x 30 min). The pellets were washed twice with Lysis Buffer containing 10mM EDTA, 1% Triton X-100, 1 mM PMSF and 0.1% β -mercaptoethanol, then with lysis buffer containing 1M urea, 1mM PMSF and 0.1 % β -mercaptoethanol. The resulting white pellet is composed primarily of inclusion bodies, free of unbroken cells and membranous materials.

The inclusion bodies were dissolved in 20 ml 6M guanidine-HCl in lysis buffer with 1 mM PMSF and 0.1% β -mercaptoethanol, and incubated on ice for 1 hour. Materials that did not dissolve were removed by centrifugation (100,000 g x 30 min.) The clear supernatant was filtered through a 0.8 μ m Supor filter (Gelman Sciences, FRG) and then load directly onto a 10 ml Ni^{2+} - NTA agarose column (Hochuli et al. 1987) pre-equilibrated in 6M guanidine-HCl in Lysis Buffer containing 1 mM PMSF and 0.1% β -Mercaptoethanol. The column was washed with 20ml (2 bed volumes) of Lysis Buffer containing 6M guanidine-HCl, 1mM PMSF and 0.1 % β -mercaptoethanol, then guanidine-HCl was removed slowly with a 100 ml linear gradient (from 6M to 0 M Guanidine-HCl) of lysis buffer containing 0.5% Brij 35, 1 mM PMSF, 0.1% β -mercaptoethanol. Next, the column was developed with a 25 ml linear gradient of increasing imidazole (5 to 500 mM) in Lysis buffer containing 0.5% Brij 35, 1 mM PMSF and 0.1% β -mercaptoethanol. The recombinant proteins elute as a peak centered at 100mM imidazole.

Fractions containing the recombinant proteins were pooled and then concentrated to approximately 8 ml by centrifugal filtration (Centriprep-10, Amicon, MA), and loaded directly onto a 350-ml column (2.2 X 91 cm) of Sephacyl S-100 HR gel filtration medium equilibrated in Buffer A (50mM Sodium Phosphate, pH 8.0, 500 mM NaCl, 0.1 mM EGTA, 1 mM PMSF, 0.1% β -mercaptoethanol, 0.5% Brij 35) and ran in Buffer A at 30 ml/h. Fractions containing the recombinant protein were identified by absorbance at 280 nm and analyzed by SDS-PAGE. Fractions were pooled, concentrated to 1.5 to 2 mg/ml and dialysed overnight against 10 mM Potassium Phosphate pH 7.5, 150 mM NaCl, 0.1 mM EGTA and 0.5% Brij 35. The concentration of protein in the dialysate was quantified, then aliquoted prior to freezing at - 20°C.

Mouse model of *Helicobacter pylori* infection

A murine model of *H. pylori* infection was produced by infection of C57BL/6 mice with *H. pylori* Sydney strain SS1 and was used to assess the efficacy of recombinant *H. pylori* vac36. This mouse-adapted *H. pylori* strain is *cagA*⁺ *vacA*⁺, shows colonization levels in C57BL/6 mice equivalent to those observed in humans, forms adhesion pedestals, colonizes for at least 8 months, and elicits a chronic-active gastritis and mucosal atrophy (Lee *et al.*, Gastroenterology, 112:1386-1397, 1997). Dose-response studies have shown 100% infection rates of inbred C57BL/6 and Balb/C mice at 8 weeks post-challenge with a single inoculation of 10⁶ organisms.

Assessment of gastric *H. pylori* infection

The presence of *H. pylori* organisms in gastric tissue was determined by culture of gastric tissue and by a quantitative urease assay. In the latter method, a longitudinal segment of antrum, representing approximately ¼ of the total antral region was placed in 1 ml of urea broth. After 4 hr, the extent of color change resulting from urea hydrolysis and increased pH was quantitated by spectrophotometric measurement of A₅₅₀ (Fox *et al.*, Immunol. 88:400-406, 1996). The assay sensitivity is ~ 10³ *H. pylori* organisms. A positive (*H. pylori*-infected) gastric tissue was defined as that sample showing 2 standard deviations above the mean A₅₅₀ value derived from a group of unchallenged uninfected age-matched control mice.

Assessment of local immune response to immunization in gastric tissue

Longitudinal sections of gastric tissues from the esophageal to the duodenal junction were embedded in OCT embedding compound, frozen in liquid nitrogen, and cryosections immunostained with monoclonal antibodies recognizing CD4⁺ or CD8⁺T cells or with antisera against mouse IgA for identification of IgA containing (IgACC) plasma cells (Pappo *et al.*, Infect. Immun. 63:1246-1252, 1995). The degree of local gastric immune response was expressed quantitatively as the number of CD4⁺, CD8⁺ or IgACC cells per mm² of gastric region examined.

Protective activity of purified recombinant *H. pylori* vac36 antigen

The ability of purified recombinant vac36 antigen derived from *H. pylori* to interfere with the establishment of an *H. pylori* infection was examined in mice. Groups (n=10) of 6-8 week-old female C57BL/6 mice were immunized orally 4 times at weekly intervals as follows: 1) 100 µg of recombinant vac36 antigen and 10 µg cholera toxin (CT) adjuvant, 2) 1 mg *H. pylori* lysate antigens and 10 µg CT, and 3) 0.2 M

bicarbonate buffer and 10 μ g CT adjuvant. The mice were challenged 2 weeks later on 3 consecutive days by oral administration of 10^8 *H. pylori* organisms. The experiment was terminated 2 weeks post-challenge, and the *H. pylori* infection level assessed by bacterial colony counts and by quantitative urease assays.

- 5 Oral immunization with vac36 antigen interfered with the establishment of *H. pylori* infection upon challenge with live *H. pylori* organisms. Mice immunized with purified recombinant vac36 antigen exhibited a significantly lower level of colonization by *H. pylori*, as assessed by gastric urease activity and bacterial count assays (Table 6). Oral immunization with vac36 antigen also resulted in the generation of a local
- 10 protective gastric immune response. Greater numbers of CD4⁺T cells and of IgACC were recruited in the gastric tissues of vac36-immunized mice when compared with unimmunized *H. pylori*-infected mice (Table 7).

Table 6

- 15 Recombinant vac36 antigen protects mice from challenge with *H. pylori*

Vaccine Treatment Group	Urease Activity ^a	p^b	<i>H. pylori</i> burden ^c	p^b
vac36	0.199 $\pm$ 0.080	0.0022	55,800 $\pm$ 12,599	0.0125
<i>H. pylori</i> lysate	0.057 $\pm$ 0.007	0.0002	2,360 $\pm$ 955	0.0002
buffer	1.655 $\pm$ 0.420	-	131,000 $\pm$ 18,391	-

- a Urease activity is expressed as mean A₅₅₀ $\pm$ SEM of duplicate antral samples from n=10 mice/group.
- b by Wilcoxon Rank Sum Test compared with mice immunized with CT adjuvant alone
- 20 c The level of *H. pylori* in gastric tissue was assessed by bacterial counts, and shown as mean colony forming units $\pm$ SEM

Table 7

vac36-immunized mice generate a local gastric immune response upon challenge with *H. pylori*

Vaccine Treatment Group	CD4+			CD8+			IgACC		
	cardia	corpus	antrum	cardia	corpus	antrum	cardia	corpus	antrum
vac36	33 ± 9 ^a	54 ± 8*	31 ± 8	3 ± 2	0	1 ± 1	24 ± 12	79 ± 16	67 ± 13
<i>H. pylori</i> lysate	31 ± 13	36 ± 19	24 ± 8	4 ± 2	2 ± 1	2 ± 1	31 ± 9	73 ± 13*	79 ± 15
buffer	12 ± 2	27 ± 8	18 ± 4	1 ± 1	0	0	4 ± 2	30 ± 13	46 ± 14

^a Mean number of cells/mm² of gastric region ± SEM

5 * p<0.05 by Wilcoxon Rank Sum Test when compared with unimmunized *H. pylori* infected mice

V. Sequence Variance Analysis of genes in *Helicobacter pylori* strains

10 Four genes were cloned and sequenced from several strains of *H. pylori* to compare the DNA and deduced amino acid sequences. This information was used to determine the sequence variation between the *H. pylori* strain, J99, and other *H. pylori* strains isolated from human patients.

15 Preparation of Chromosomal DNA.

Cultures of *H. pylori* strains (as listed in Table 10) were grown in BLBB (1% Tryptone, 1% Peptamin 0.1% Glucose, 0.2% Yeast Extract 0.5% Sodium Chloride, 5% Fetal Bovine Serum) to an OD₆₀₀ of 0.2. Cells were centrifuged in a Sorvall RC-3B at 3500 x g at 4°C for 15 minutes and the pellet resuspended in 0.95 mls of 10 mM Tris-HCl, 0.1 mM EDTA (TE). Lysozyme was added to a final concentration of 1mg/ml along with, SDS to 1% and RNase A + T1 to 0.5mg/ml and 5 units/ml respectively, and incubated at 37°C for one hour. Proteinase K was then added to a final concentration of 0.4mg/ml and the sample was incubated at 55 C for more than one hour. NaCl was added to the sample to a concentration of 0.65 M, mixed carefully, and 0.15 ml of 10%

- 91 -

CTAB in 0.7M NaCL (final is 1% CTAB/70mM NaCL) was added followed by incubation at 65°C for 20 minutes. At this point, the samples were extracted with chloroform:isoamyl alcohol, extracted with phenol, and extracted again with chloroform:isoamyl alcohol. DNA was precipitated with either EtOH (1.5 x volumes) or isopropanol (0.6 x volumes) at -70°C for 10 minutes, washed in 70% EtOH and resuspended in TE.

PCR Amplification and cloning.

Genomic DNA prepared from twelve strains of *Helicobacter pylori* was used as the source of template DNA for PCR amplification reactions (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel et al., editors, 1994). To amplify a DNA sequence containing an *H. pylori* ORF, genomic DNA (10 nanograms) was introduced into a reaction vial containing 2 mM MgCl₂, 1 micromolar synthetic oligonucleotide primers (forward and reverse primers, see Table 8) complementary to and flanking a defined *H. pylori* ORF, 0.2 mM of each deoxynucleotide triphosphate; dATP, dGTP, dCTP, dTTP and 0.5 units of heat stable DNA polymerase (Amplitaq, Roche Molecular Systems, Inc., Branchburg, NJ, USA) in a final volume of 20 microliters in duplicate reactions.

Table 8

Oligonucleotide primers used for PCR amplification of *H. pylori* DNA sequences.

Outer membrane Proteins	Forward primer 5' to 3'	Reverse Primer 5' to 3'
Protein 26054702 (for strains AH4, AH15, AH61, 5294, 5640, AH18, and AH244)	5'-TTAACCATGGTGAAAA GCGATA-3' (SEQ ID NO:217)	5'-TAGAATTCGCCTCTAA AACTTTAG-3' (SEQ ID NO:218)
Protein 26054702 (for strains AH5, 5155, 7958, AH24, and J99)	5'-TTAACCATGGTGAAAA GCGATA-3' (SEQ ID NO:219)	5'-TAGAATTCGCATAA CGATCAATC-3' (SEQ ID NO:220)
Protein 7116626	5'-ATATCCATGGTGAGTT TGATGA-3' (SEQ ID NO:221)	5'-ATGAATTCAATTTT TTATTTTGCCA-3' (SEQ ID NO:222)
Protein 29479681	5'-AATTCCATGGCTATC CAAATCCG-3' (SEQ ID NO:223)	5'-ATGAATTCGCCAAAA TCGTAGTATT-3' (SEQ ID NO:224)
Protein 346	5'-GATACCATGGAATTT ATGAAAAAG-3' (SEQ ID NO:225)	5'-TGAATTCGAAAAAGTG TAGTTATAC-3' (SEQ ID NO:226)

- 92 -

The following thermal cycling conditions were used to obtain amplified DNA products for each ORF using a Perkin Elmer Cetus/ GeneAmp PCR System 9600 thermal cycler:

- 5 Protein 7116626 and Protein 346;
Denaturation at 94°C for 2 min,
2 cycles at 94°C for 15 sec, 30°C for 15 sec and 72°C for 1.5 min
23 cycles at 94°C for 15 sec, 55°C for 15 sec and 72°C for 1.5 min
Reactions were concluded at 72°C for 6 minutes.
- 10 Protein 26054702 for strains AH5, 5155, 7958, AH24, and J99;
Denaturation at 94°C for 2 min,
2 cycles at 94°C for 15 sec, 30°C for 15 sec and 72°C for 1.5 min
25 cycles at 94°C for 15 sec, 55°C for 15 sec and 72°C for 1.5 min
- 15 Reaction was concluded at 72°C for 6 minutes.

- Protein 26054702 and Protein 294796813 for strains AH4, AH15, AH61, 5294, 5640, AH18, and Hp244 ;
Denaturation at 94°C for 2 min,
- 20 2 cycles at 94°C for 15 sec, 30°C for 20 sec and 72°C for 2 min
25 cycles at 94°C for 15 sec, 55°C for 20 sec and 72°C for 2 min
Reactions were concluded at 72°C for 8 minutes.

- Upon completion of thermal cycling reactions, each pair of samples were
- 25 combined and used directly for cloning into the pCR cloning vector as described below.

Cloning of H. pylori DNA sequences into the pCR TA cloning vector.

- All amplified inserts were cloned into the pCR 2.1 vector by the method described in the Original TA cloning kit (Invitrogen, San Diego, CA). Products of the
- 30 ligation reaction were then used to transform the TOP10F' (INVaF' in the case of *H. pylori* sequence 350) strain of *E. coli* as described below.

Transformation of competent bacteria with recombinant plasmids

- Competent bacteria, *E. coli* strain TOP10F' or *E. coli* strain INVaF' were
- 35 transformed with recombinant pCR expression plasmids carrying the cloned *H. pylori* sequences according to standard methods (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel et al., editors, 1994). Briefly, 2 microliters of 0.5

- 93 -

micromolar BME was added to each vial of 50 microliters of competent cells. Subsequently, 2 microliters of ligation reaction was mixed with the competent cells and incubated on ice for 30 minutes. The cells and ligation mixture were then subjected to a "heat shock" at 42°C for 30 seconds, and were subsequently placed on ice for an additional 2 minutes, after which, samples were incubated in 0.45 milliliters SOC medium (0.5% yeast extract, 2.0 % tryptone, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl₂, 10 mM MgSO₄ and 20, mM glucose) at 37°C with shaking for 1 hour. Samples were then spread on LB agar plates containing 25 microgram/ml kanamycin sulfate or 100 micrograms/ml ampicillin for growth overnight. Transformed colonies of TOP10F' or INVaF' were then picked and analyzed to evaluate cloned inserts as described below.

Identification of recombinant PCR plasmids carrying H. pylori sequences

Individual TOP10F' or INVaF' clones transformed with recombinant pCR-*H.pylori* ORFs were analyzed by PCR amplification of the cloned inserts using the same forward and reverse primers, specific for each *H. pylori* sequence, that were used in the original PCR amplification cloning reactions. Successful amplification verified the integration of the *H. pylori* sequences in the cloning vector (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel et al., editors, 1994).

Individual clones of recombinant pCR vectors carrying properly cloned *H. pylori* ORFs were picked for sequence analysis. Sequence analysis was performed on ABI Sequencers using standard protocols (Perkin Elmer) using vector-specific primers (as found in PCR2.1 or pCR2.1, Invitrogen, San Diego, CA) and sequencing primers specific to the ORF as listed in Table 9 below.

Table 9

Oligonucleotide primers used for sequencing of *H. pylori* DNA sequences.

Outer membrane Proteins	Forward primers 5' to 3'	Reverse Primers 5' to 3'
Protein 26054702	5'CCCTTCATTTTAGAAATC G-3' (SEQ ID NO:227) 5'ATTCAACCAATTCAAT GCG-3' (SEQ ID NO:228) 5'GCCCCTTTTGATTTGAA GCT-3' (SEQ ID NO:229) 5'TCGCTCCAAGATACCAA GAAGT-3' (SEQ ID NO:230) 5'CTTGAATTAGGGGCAAA GATCG-3' (SEQ ID NO:231) 5'ATGCGTTTTTACCCAAA GAAGT-3' (SEQ ID NO:232) 5'ATAACGCCACTTCCTTA TTGGT-3' (SEQ ID NO:233)	5'CTTTGGGTAAAAACGCATC -3' (SEQ ID NO:234) 5'CGATCTTTGATCCTAATTC A-3' (SEQ ID NO:235) 5'ATCAAGTTGCCTATGCTGA -3' (SEQ ID NO:236)
Protein 7116626	5'TTGAACACTTTTGATTAT GCGG-3' (SEQ ID NO:237) 5'GGATTATGCGATTGTTTT ACAAG-3' (SEQ ID NO:238)	5'GTCTTTAGCAAAAATGGCG TC-3' (SEQ ID NO:239) 5'AATGAGCGTAAGAGAGCC TTC-3' (SEQ ID NO:240)
Protein 29479681	5'CTTATGGGGGTATTGTC A-3' (SEQ ID NO:241) 5'AGCATGTGGGTATCCAG C-3' (SEQ ID NO:242)	5'AGGTTGTTGCCTAAAGACT -3' (SEQ ID NO:243) 5'-CTGCCTCCACCTTTGATC- 3' (SEQ ID NO:244)
Protein 346	5'ACCAATATCAATTGGCA CT-3' (SEQ ID NO:245) 5'ACTTGGAAGCTCTGC A-3' (SEQ ID NO:246)	5'CTTGCTTGTCATATCTAGC- 3' (SEQ ID NO:247) 5'-GTTGAAGTGTGGTGCTA- 3' (SEQ ID NO:248)
	5'CAAGCAAGTGGTTTGGT TTTAG-3' (SEQ ID NO:249) 5'TGGAAAGAGCAAATCAT TGAAG-3' (SEQ ID NO:250)	5'GCCCATAATCAAAAAGCC CAT-3' (SEQ ID NO:251) 5'CTAAAACCAAACCACTTGC TTGTC-3' (SEQ ID NO:252)
Vector Primers	5'-GTAAAACGACGGCCAG- 3' (SEQ ID NO:253)	5'-CAGGAAACAGCTATGAC- 3' (SEQ ID NO:254)

Results

To establish the PCR error rate in these experiments, five individual clones of Protein 26054702, prepared from five separate PCR reaction mixtures from *H. pylori* strain J99, were sequenced over a total length of 897 nucleotides for a cumulative total of 4485 bases of DNA sequence. DNA sequence for the five clones was compared to a DNA sequence obtained previously by a different method, i.e., random shotgun cloning and sequencing. The PCR error rate for the experiments described herein was determined to be 2 base changes out of 4485 bases, which is equivalent to an estimated error rate of less than or equal to 0.04%.

DNA sequence analysis was performed on four different open reading frames identified as genes and amplified by PCR methods from a dozen different strains of the bacterium *Helicobacter pylori*. The deduced amino acid sequences of three of the four open reading frames that were selected for this study showed statistically significant BLAST homology to defined proteins present in other bacterial species. Those ORFs included: Protein 26054702, homologous to the val A & B genes encoding an ABC transporter in *F. novicida*; Protein 7116626, homologous to lipoprotein e (P4) present in the outer membrane of *H. influenzae*; Protein 29479681, homologous to fecA, an outer membrane receptor in iron (III) dicitrate transport in *E. coli*. Protein 346 was identified as an unknown open reading frame, because it showed low homology with sequences in the public databases.

To assess the extent of conservation or variance in the ORFs across various strains of *H. pylori*, changes in DNA sequence and the deduced protein sequence were compared to the DNA and deduced protein sequences found in the J99 strain of *H. pylori* (see Table 10 below). Results are presented as percent identity to the J99 strain of *H. pylori* sequenced by random shotgun cloning. To control for any variations in the J99 sequence each of the four open reading frames were cloned and sequenced again from the J99 bacterial strain and that sequence information was compared to the sequence information that had been collected from inserts cloned by random shotgun sequencing of the J99 strain. The data demonstrate that there is variation in the DNA sequence ranging from as little as 0.12 % difference (Protein 346, J99 strain) to approximately 7% change (Protein 26054702, strain AH5). The deduced protein sequences show either no variation (Protein 346, strains AH18 and AH24) or up to as much as 7.66% amino acid changes (Protein 26054702, Strain AH5).

- 96 -

Table 10Multiple Strain DNA Sequence analysis of *H. pylori* Vaccine Candidates

5	<u>J99 Protein #:</u>	26054702	26054702	7116626	7116626	29479681	29479681	346	346
	Length of Region								
	Sequenced:	248 a.a.	746 nt.	232 a.a.	96 nt.	182 a.a.	548 nt.	273 a.a.	819 nt.
10	<u>Strain Tested</u>								
		AA identity	Nuc. identity	AA identity	Nuc. identity	AA identity	Nuc. identity	AA identity	Nuc. identity
	J99	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	99.63%	99.88%
	AH244	95.16%	95.04%	n.d.	n.d.	99.09%	96.71%	98.90%	96.45%
	AH4	95.97%	95.98%	97.84%	95.83%	n.d.	n.d.	97.80%	95.73%
	AH5	92.34%	93.03%	98.28%	96.12%	98.91%	96.90%	98.53%	95.73%
	AH15	95.16%	94.91%	97.41%	95.98%	99.82%	97.99%	99.63%	96.09%
	AH61	n.d.	n.d.	97.84%	95.98%	99.27%	97.44%	n.d.	n.d.
	5155	n.d.	n.d.	n.d.	n.d.	99.45%	97.08%	98.53%	95.60%
	5294	94.35%	94.37%	98.28%	95.40%	99.64%	97.26%	97.07%	95.48%
	7958	94.35%	94.10%	97.84%	95.40%	n.d.	n.d.	99.63%	96.46%
	5640	95.16%	94.37%	97.41%	95.69%	99.09%	97.63%	98.53%	95.48%
	AH18	n.d.	n.d.	98.71%	95.69%	99.64%	97.44%	100.00%	95.97%
	AH24	94.75%	95.04%	97.84%	95.40%	99.27%	96.71%	100.00%	96.46%

n.d.= not done.

VI. Experimental Knock-Out Protocol for the Determination of Essential *H. pylori* Genes as Potential Therapeutic Targets

Therapeutic targets are chosen from genes whose protein products appear to play key roles in essential cell pathways such as cell envelope synthesis, DNA synthesis, transcription, translation, regulation and colonization/virulence.

The protocol for the deletion of portions of *H. pylori* genes/ORFs and the insertional mutagenesis of a kanamycin-resistance cassette in order to identify genes which are essential to the cell is modified from previously published methods (Labigne-Roussel et al., 1988, J. Bacteriology 170, pp. 1704-1708; Cover et al., 1994, J. Biological

Chemistry 269, pp. 10566-10573; Reytrat et al., 1995, Proc. Natl. Acad. Sci. 92, pp 8768-8772). The result is a gene "knock-out."

Identification and Cloning of H. pylori Gene Sequences

5 The sequences of the genes or ORFs (open reading frames) selected as knock-out targets are identified from the *H. pylori* genomic sequence and used to design primers to specifically amplify the genes/ORFs. All synthetic oligonucleotide primers are designed with the aid of the OLIGO program (National Biosciences, Inc., Plymouth, MN 55447, USA), and can be purchased from Gibco/BRL Life Technologies (Gaithersburg, MD, 10 USA). If the ORF is smaller than 800 to 1000 base pairs, flanking primers are chosen outside of the open reading frame.

Genomic DNA prepared from the *Helicobacter pylori* HpJ99 strain (ATCC 55679; deposited by Genome Therapeutics Corporation, 100 Beaver Street, Waltham, MA 02154) is used as the source of template DNA for amplification of the ORFs by 15 PCR (polymerase chain reaction) (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel et al., editors, 1994). For the preparation of genomic DNA from *H. pylori*, see Example I. PCR amplification is carried out by introducing 10 nanograms of genomic HpJ99 DNA into a reaction vial containing 10 mM Tris pH 8.3, 50 mM KCl, 2 mM MgCl₂, 2 microMolar synthetic oligonucleotide primers 20 (forward=F1 and reverse=R1), 0.2 mM of each deoxynucleotide triphosphate (dATP, dGTP, dCTP, dTTP), and 1.25 units of heat stable DNA polymerase (Amplitaq, Roche Molecular Systems, Inc., Branchburg, NJ, USA) in a final volume of 40 microliters. The PCR is carried out with Perkin Elmer Cetus/GeneAmp PCR System 9600 thermal cyclers.

25 Upon completion of thermal cycling reactions, each sample of amplified DNA is visualized on a 2% TAE agarose gel stained with Ethidium Bromide (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel et al., editors, 1994) to determine that a single product of the expected size had resulted from the reaction. Amplified DNA is then washed and purified using the Qiaquick Spin PCR purification 30 kit (Qiagen, Gaithersburg, MD, USA).

PCR products are cloned into the pT7Blue T-Vector (catalog#69820-1, Novagen, Inc., Madison, WI, USA) using the TA cloning strategy (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel et al., editors, 1994). The ligation of the PCR product into the vector is accomplished by mixing a 6 fold molar excess of the 35 PCR product, 10 ng of pT7Blue-T vector (Novagen), 1 microliter of T4 DNA Ligase Buffer (New England Biolabs, Beverly, MA, USA), and 200 units of T4 DNA Ligase

- 98 -

(New England Biolabs) into a final reaction volume of 10 microliters. Ligation is allowed to proceed for 16 hours at 16°C.

Ligation products are electroporated (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel et al., editors, 1994) into electroporation-competent XL-1 Blue or DH5-a *E.coli* cells (Clontech Lab., Inc. Palo Alto, CA, USA). Briefly, 1 microliter of ligation reaction is mixed with 40 microliters of electrocompetent cells and subjected to a high voltage pulse (25 microFarads, 2.5 kV, 200 ohms) after which the samples are incubated in 0.45 ml SOC medium (0.5% yeast extract, 2% tryptone, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl₂, 10 mM MgSO₄ and 20 mM glucose) at 37°C with shaking for 1 hour. Samples are then spread onto LB (10 g/l bacto tryptone, 5 g/l bacto yeast extract, 10 g/l sodium chloride) plates containing 100 microgram/ml of Ampicillin, 0.3% X-gal, and 100 microgram/ml IPTG. These plates are incubated overnight at 37°C. Ampicillin-resistant colonies with white color are selected, grown in 5 ml of liquid LB containing 100 microgram/ml of Ampicillin, and plasmid DNA is isolated using the Qiagen miniprep protocol (Qiagen, Gaithersburg, MD, USA).

To verify that the correct *H.pylori* DNA inserts had been cloned, these pT7Blue plasmid DNAs are used as templates for PCR amplification of the cloned inserts, using the same forward and reverse primers used for the initial amplification of the J99 *H.pylori* sequence. Recognition of the primers and a PCR product of the correct size as visualized on a 2% TAE, ethidium bromide stained agarose gel are confirmation that the correct inserts had been cloned. Two to six such verified clones are obtained for each knock-out target, and frozen at -70°C for storage. To minimize errors due to PCR, plasmid DNA from these verified clones are pooled, and used in subsequent cloning steps.

The sequences of the genes/ORFs are again used to design a second pair of primers which flank the region of *H. pylori* DNA to be either interrupted or deleted (up to 250 basepairs) within the ORFs but are oriented away from each other. The pool of circular plasmid DNAs of the previously isolated clones are used as templates for this round of PCR. Since the orientation of amplification of this pair of deletion primers is away from each other, the portion of the ORF between the primers is not included in the resultant PCR product. The PCR product is a linear piece of DNA with *H. pylori* DNA at each end and the pT7Blue vector backbone between them which, in essence, results in the deletion of a portion of the ORFs. The PCR product is visualized on a 1% TAE, ethidium bromide stained agarose gel to confirm that only a single product of the correct size has been amplified.

- 99 -

A Kanamycin-resistance cassette (Labigne-Roussel et al., 1988 J. Bacteriology 170, 1704-1708) is ligated to this PCR product by the TA cloning method used previously (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel et al., editors, 1994). The Kanamycin cassette containing a *Campylobacter* kanamycin resistance gene is obtained by carrying out an EcoRI digestion of the recombinant plasmid pCTB8:*kan* (Cover et al., 1994, J. Biological Chemistry 269, pp. 10566-10573). The proper fragment (1.4 kb) is isolated on a 1% TAE gel, and isolated using the QIAquick gel extraction kit (Qiagen, Gaithersburg, MD, USA). The fragment is end repaired using the Klenow fill-in protocol, which involved mixing 4ug of the DNA fragment, 1 microliter of dATP, dGTP, dCTP, dTTP at 0.5 mM, 2 microliter of Klenow Buffer (New England Biolabs) and 5 units of Klenow DNA Polymerase I Large (Klenow) Fragment (New England Biolabs) into a 20 microliter reaction, incubating at 30°C for 15 min, and inactivating the enzyme by heating to 75°C for 10 minutes. This blunt-ended Kanamycin cassette is then purified through a Qiaquick column (Qiagen, Gaithersburg, MD, USA) to eliminate nucleotides. The "T" overhang is then generated by mixing 5 micrograms of the blunt-ended kanamycin cassette, 10 mM Tris pH 8.3, 50 mM KCl, 2 mM MgCl₂, 5 units of DNA Polymerase (Amplitaq, Roche Molecular Systems, Inc., Branchburg, NJ, USA), 20 microliters of 5 mM dTTP, in a 100 microliter reaction and incubating the reaction for 2 hours at 37°C. The "Kan-T" cassette is purified using a QIAquick column (Qiagen, Gaithersburg, MD, USA). The PCR product of the deletion primers (F2 and R2) is ligated to the Kan-T cassette by mixing 10 to 25 ng of deletion primer PCR product, 50 - 75 ng Kan-T cassette DNA, 1 microliter 10x T4 DNA Ligase reaction mixture, 0.5 microliter T4 DNA Ligase (New England Biolabs, Beverly, MA, USA) in a 10 microliter reaction and incubating for 16 hours at 16°C.

The ligation products are transformed into XL-1 Blue or DH5-a *E. coli* cells by electroporation as described previously. After recovery in SOC, cells are plated onto LB plates containing 100 microgram/ml Ampicillin and grown overnight at 37°C. These plates are then replica plated onto plates containing 25 microgram/ml Kanamycin and allowed to grow overnight. Resultant colonies have both the Ampicillin resistance gene present in the pT7Blue vector, and the newly introduced Kanamycin resistance gene. Colonies are picked into LB containing 25 microgram/ml Kanamycin and plasmid DNA is isolated from the cultured cells using the Qiagen miniprep protocol (Qiagen, Gaithersburg, MD, USA).

Several tests by PCR amplification are conducted on these plasmids to verify that the Kanamycin is inserted in the *H. pylori* gene/ORF, and to determine the orientation of the insertion of the Kanamycin-resistance gene relative to the *H. pylori* gene/ORF. To

verify that the Kanamycin cassette is inserted into the *H. pylori* sequence, the plasmid DNAs are used as templates for PCR amplification with the set of primers originally used to clone the *H. pylori* gene/ORFs. The correct PCR product is the size of the deleted gene/ORF but increased in size by the addition of a 1.4 kilobase Kanamycin cassette. To avoid potential polar effects of the kanamycin resistance cassette on *H. pylori* gene expression, the orientation of the Kanamycin resistance gene with respect to the knock-out gene/ORF is determined and both orientations are eventually used in *H. pylori* transformations (see below). To determine the orientation of insertion of the kanamycin resistance gene, primers are designed from the ends of the kanamycin resistance gene ("Kan-1" 5'-ATCTTACCTATCACCTCAAAT-3' (SEQ ID NO:255)), and "Kan-2" 5'-AGACAGCAACATCTTTGTGAA-3' (SEQ ID NO:256)). By using each of the cloning primers in conjunction with each of the Kan primers (4 combinations of primers), the orientation of the Kanamycin cassette relative to the *H. pylori* sequence is determined. Positive clones are classified as either in the "A" orientation (the same direction of transcription is present for both the *H. pylori* gene and the Kanamycin resistance gene), or in the "B" orientation (the direction of transcription for the *H. pylori* gene is opposite to that of the Kanamycin resistance gene). Clones which share the same orientation (A or B) are pooled for subsequent experiments and independently transformed into *H. pylori*.

Transformation of Plasmid DNA into H. pylori cells

Two strains of *H. pylori* are used for transformation: ATCC 55679, the clinical isolate which provided the DNA from which the *H. pylori* sequence database is obtained, and AH244, an isolate which had been passaged in, and has the ability to colonize the mouse stomach. Cells for transformation are grown at 37°C, 10% CO₂, 100% humidity, either on Sheep-Blood agar plates or in Brucella Broth liquid. Cells are grown to exponential phase, and examined microscopically to determine that the cells are "healthy" (actively moving cells) and not contaminated. If grown on plates, cells are harvested by scraping cells from the plate with a sterile loop, suspended in 1 ml of Brucella Broth, spun down (1 minute, top speed in eppendorf microfuge) and resuspended in 200 microliters Brucella Broth. If grown in Brucella Broth liquid, cells are centrifuged (15 minutes at 3000 rpm in a Beckman TJ6 centrifuge) and the cell pellet resuspended in 200 microliters of Brucella broth. An aliquot of cells is taken to determine the optical density at 600 nm, in order to calculate the concentration of cells. An aliquot (1 to 5 OD₆₀₀ units/25 microliter) of the resuspended cells is placed onto a prewarmed Sheep-Blood agar plate, and the plate is further incubated at 37°C, 6% CO₂,

- 101 -

100% humidity for 4 hours. After this incubation, 10 microliters of plasmid DNA (100 micrograms per microliter) is spotted onto these cells. A positive control (plasmid DNA with the ribonuclease H gene disrupted by kanamycin resistance gene) and a negative control (no plasmid DNA) are done in parallel. The plates are returned to 37°C, 6% CO₂ for an additional 4 hours of incubation. Cells are then spread onto that plate using a swab wetted in Brucella broth, and grown for 20 hours at 37°C, 6% CO₂. Cells are then transferred to a Sheep-Blood agar plate containing 25 micrograms/ml Kanamycin, and allowed to grow for 3 to 5 days at 37°C, 6% CO₂, 100% humidity. If colonies appear, they are picked and regrown as patches on a fresh Sheep-Blood agar plate containing 25 micrograms/ml Kanamycin.

Three sets of PCR tests are done to verify that the colonies of transformants have arisen from homologous recombination at the proper chromosomal location. The template for PCR (DNA from the colony) is obtained by a rapid boiling DNA preparation method as follows. An aliquot of the colony (stab of the colony with a toothpick) is introduced into 100 microliters of 1% Triton X-100, 20 mM Tris, pH 8.5, and boiled for 6 minutes. An equal volume of phenol : chloroform (1:1) is added and vortexed. The mixture is microfuged for 5 minutes and the supernatant is used as DNA template for PCR with combinations of the following primers to verify homologous recombination at the proper chromosomal location.

TEST 1. PCR with cloning primers originally used to amplify the gene/ORF. A positive result of homologous recombination at the correct chromosomal location should show a single PCR product whose size is expected to be the size of the deleted gene/ORF but increased in size by the addition of a 1.4 kilobase Kanamycin cassette. A PCR product of just the size of the gene/ORF is proof that the gene had not been knocked out and that the transformant is not the result of homologous recombination at the correct chromosome location.

TEST 2. PCR with F3 (primer designed from sequences upstream of the gene/ORF and not present on the plasmid), and either primer Kan-1 or Kan-2 (primers designed from the ends of the kanamycin resistance gene), depending on whether the plasmid DNA used was of "A" or "B" orientation. Homologous recombination at the correct chromosomal location will result in a single PCR product of the expected size (i.e., from the location of F3 to the insertion site of kanamycin resistance gene). No PCR product or PCR product(s) of incorrect size(s) will prove that the plasmid had not integrated at the correct site and that the gene had not been knocked out.

TEST 3. PCR with R3 (primer designed from sequences downstream of the gene/ORF and not present on the plasmid) and either primer Kan-1 or Kan-2, depending on whether the plasmid DNA used was of "A" or "B" orientation. Homologous

- 102 -

recombination at the correct chromosomal location will result in a single PCR product of the expected size (i.e., from the insertion site of kanamycin resistance gene to the downstream location of R3). Again, no PCR product or PCR product(s) of incorrect size(s) will prove that the plasmid had not integrated at the correct site and that the gene had not been knocked out.

Transformants showing positive results for all three tests above indicate that the gene is not essential for survival *in vitro*.

A negative result in any of the three above tests for each transformant indicates that the gene had not been disrupted, and that the gene is essential for survival *in vitro*.

In the event that no colonies result from two independent transformations while the positive control with the disrupted ribonuclease H plasmid DNA produces transformants, the plasmid DNA is further analyzed by PCR on DNA from transformant populations prior to plating for colony formation. This will verify that the plasmid can enter the cells and undergo homologous recombination at the correct site. Briefly, plasmid DNA is incubated according to the transformation protocol described above. DNA is extracted from the *H. pylori* cells immediately after incubation with the plasmid DNAs and the DNA is used as template for the above TEST 2 and TEST 3. Positive results in TEST 2 and TEST 3 would verify that the plasmid DNA could enter the cells and undergo homologous recombination at the correct chromosomal location. If TEST 2 and TEST 3 are positive, then failure to obtain viable transformants indicates that the gene is essential, and cells suffering a disruption in that gene are incapable of colony formation.

VII. High-throughput drug screen assay

Cloning, expression and protein purification

Cloning, transformation, expression and purification of the *H. pylori* target gene and its protein product, e.g., an *H. pylori* enzyme, to be used in a high-throughput drug screen assay, is carried out essentially as described in Examples II and III above. Development and application of a screening assay for a particular *H. pylori* gene product, peptidyl-propyl *cis-trans* isomerase, is described below as a specific example.

Enzymatic Assay

The assay is essentially as described by Fisher (Fischer, G., et.al. (1984) *Biomed. Biochim. Acta* 43:1101-1111). The assay measures the *cis-trans* isomerization of the Ala-Pro bond in the test peptide N-succinyl-Ala-Ala-Pro-Phe-p-nitroanilide (Sigma # S-7388, lot # 84H5805). The assay is coupled with α -chymotrypsin, where the ability of the protease to cleave the test peptide occurs only when the Ala-Pro bond is in *trans*. The conversion of the test peptide to the *trans* isomer in the assay is followed at 390 nm

on a Beckman Model DU-650 spectrophotometer. The data are collected every second with an average scanning of time of 0.5 second. Assays are carried out in 35 mM Hepes, pH 8.0, in a final volume of 400 μ l, with 10 μ M α -chymotrypsin (type 1-5 from bovine Pancreas, Sigma # C-7762, lot 23H7020) and 10 nM PPIase. To initiate the reaction, 10 μ l of the substrate (2 mM N-Succinyl-Ala-Ala-Pro-Phe-p-nitroanilide in DMSO) is added to 390 μ l of reaction mixture at room temperature.

Enzymatic assay in crude bacterial extract.

A 50 ml culture of *Helicobacter pylori* (strain J99) in Brucella broth is harvested at mid-log phase ($OD_{600\text{ nm}} \sim 1$) and resuspended in lysis buffer with the following protease inhibitors: 1 mM PMSF, and 10 μ g/ml of each of aprotinin, leupeptin, pepstatine, TLCK, TPCK, and soybean trypsin inhibitor. The suspension is subjected to 3 cycles of freeze-thaw (15 minutes at -70°C , then 30 minutes at room temperature), followed by sonication (three 20 second bursts). The lysate is centrifuged (12,000 g x 30 minutes) and the supernatant is assayed for enzymatic activity as described above.

Many *H. pylori* enzymes can be expressed at high levels and in an active form in *E. coli*. Such high yields of purified proteins provide for the design of various high throughput drug screening assays.

VIII. Truncated gene expression and protein production

Identification, cloning and expression of recombinant Helicobacter pylori sequences.

To facilitate the cloning, expression and purification of membrane proteins from *H. pylori*, the pET gene expression system (Novagen), for cloning and expression of recombinant proteins in *Escherichia coli* was selected. Further, for proteins that have a signal sequence at their amino-terminal end, a DNA sequence encoding a peptide tag (His-tag) was fused to the 5' end of the *H. pylori* DNA sequences of interest in order to facilitate purification of the recombinant protein products. In some cases, the DNA sequence was cloned in frame with the glutathione-S-transferase protein to produce a GST-fusion protein. The vectors used in this case were the pGEX series from Pharmacia LKB (Uppsala, Sweden).

PCR amplification and cloning of DNA sequences containing ORFs for membrane and secreted proteins from the J99 strain of Helicobacter pylori.

The sequences chosen (from the list of the DNA sequences of the invention) for cloning from *H. pylori* strain J99 were prepared for amplification cloning by the polymerase chain reaction (PCR). Synthetic oligonucleotide primers for the ORF of interest (Table 11) specific for the predicted mature 5' end of the ORF and either

downstream (3') of the predicted translational termination codon or at specific points within the coding region were designed and purchased (GibcoBRL Life Technologies, Gaithersburg, MD, USA). All forward primers (specific for the 5' terminus of the region of ORF of interest) were designed to include either a *Bam*HI or a *Nde*I restriction site. These primers within the *Nde*I restriction site sequence were designed to permit the initiation of protein translation at a methionine residue (encoded within the *Nde*I restriction site sequence, in the case of producing a non His-tagged recombinant protein) or to fuse in frame with the DNA sequence encoding the His-tag (for producing His tagged recombinant protein), followed by the coding sequence for the remainder of the native *H. pylori* DNA. The primer with the *Bam*HI restriction site was produced to fuse the *H. pylori* specific sequence in-frame with the C-terminus of the glutathione-S-transferase gene in the pGEX vectors (Pharmacia LKB, Uppsala, Sweden). All reverse oligonucleotide primers designed to include an *Eco*RI restriction site at the 5' terminus. Several reverse oligonucleotide primers were selected that would cause a truncation of the polypeptide to remove certain portions of the C-terminus, and in these cases the *Eco*RI restriction site at the 5' end was followed by a translational termination codon. This combination of primers would enable the ORF of interest (or parts of the ORF of interest) to be cloned into pET28b (to produce a His-tagged recombinant protein), pET30a (to produce a non His tagged or native recombinant protein) or the pGEX-4T or pGEX-5X series (to produce a GST fusion protein). The pET28b vector provides sequence encoding an additional 20 amino-terminal amino acids (plus the methionine in the *Nde*I restriction site) including a stretch of six histidine residues which makes up the His-tag, whereas the pGEX vectors fuse the *H. pylori* protein to a 26,000Da glutathione-S-transferase protein.

Genomic DNA prepared from *H. pylori* strain J99 (ATCC 55679) was used as the source of template DNA for the PCR amplification reactions (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel *et al.*, eds., 1994). To amplify a DNA sequence containing a specific *H. pylori* ORF, genomic DNA (50 nanograms) was introduced into a reaction tube containing 200 nanograms of both the forward and reverse synthetic oligonucleotide primer specific for the ORF of interest, and 45 microliters of PCR SuperMix purchased (GibcoBRL Life Technologies, Gaithersburg, MD, USA) in a total of 50 microliters. The PCR SuperMix is supplied in 1.1X concentrations and contains 22mM Tris-HCl (pH 8.4), 55mM KCl, 1.65mM MgCl₂, 220 micromolar of each dATP, dCTP, dGTP and dTTP, 22units recombinant *Taq* polymerase/ml and stabilizers. The following thermal cycling conditions were used to obtain amplified DNA products for each ORF using a Perkin Elmer Cetus/GeneAmp PCR System thermal cycler.

Table 11: Oligonucleotide primers

Gene and location	Sequence
Vac38- BamHI post signal sequence	CGGGATCCGAAGGTGATGGTGTGTTATA TAGG (SEQ ID NO: 271)
Vac38- NdeI post signal sequence	CGCATATGGAAGGTGATGGTGTGTTATA TAGGG (SEQ ID NO: 272)
Vac38- EcoRI/stop codon (removes C-terminal third of protein)	GCGAATTCTCACTCTTTCCAATAGTTG CTGCAGAGC (SEQ ID NO: 273)
Vac38- EcoRI/stop codon (removes C-terminal 11 amino acids)	CCGGAATTCTTAATCCCGTTTCAAATG GTAATAAAGG (SEQ ID NO: 274)
Vac38- EcoRI downstream of native stop codon	GCGAATTCCTTTTATTTAAAAAGTGT AGTTATACC (SEQ ID NO: 275)

5 Sequences for Vac38 (full length or truncated)

Denaturation at 94°C for 30 sec

35 cycles at 94°C for 15 sec, 55°C for 15 sec, and 72°C for 1.5 min

Reactions were concluded at 72°C for 8 minutes

- 10 Upon completion of the thermal cycling reactions, each sample of amplified DNA was subjected to electrophoresis on 1.0% agarose gels. The DNA was visualized by exposure to ethidium bromide and long wave UV irradiation, and cut out in gel slices. DNA was purified using the Wizard PCR Preps kit (Promega Corp., Madison WI, USA), and then subjected to digestion with *Bam*HI and *Eco*RI (Current Protocols in
- 15 Molecular Biology, John Wiley and Sons, Inc., F. Ausubel *et al.*, eds., 1994). The digested PCR amplicon was then re-electrophoresed and purified as before.

Ligation of H. pylori DNA sequences into cloning vectors

- The pOK12 vector (J. Vieira and J. Messing, Gene 100:189-194, 1991) was
- 20 prepared for cloning by digestion with *Bam*HI and *Eco*RI or *Nde*I and *Eco*RI in the case of Vac41 (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel *et al.*, eds., 1994). The vectors were subjected to electrophoresis on 1.0% agarose gels and purified using the Wizard PCR Preps kit (Promega Corp., Madison WI, USA). Following ligation of the purified, digested vector and the purified, digested

amplified *H. pylori* ORF, the products of the ligation reaction were transformed into *E. coli* JM109 competent cells according to standard methods (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel *et al.*, eds., 1994). Individual bacterial colonies were screened for those containing the correct recombinant plasmids by incubating in LB broth overnight (plus 25ug/ml kanamycin sulfate) followed by plasmid DNA preparation using the Magic Minipreps system (Promega Corp., Madison WI, USA), and then analyzed by restriction digestion (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel *et al.*, eds., 1994).

10 *Cloning of H. pylori DNA sequences into the pET28b, pET30a and pGEX4T-3 prokaryotic expression vectors*

Both the pET28b and pET30a expression vectors were prepared for cloning by digestion with *Nde*I and *Eco*RI, and the pGEX4T-3 vector was prepared for cloning by digestion with *Bam*HI and *Eco*RI (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel *et al.*, eds., 1994). The *H. pylori* DNA sequences were removed from pOK12 plasmid backbones by digestion with *Nde*I and *Eco*RI or *Bam*HI and *Eco*RI (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel *et al.*, eds., 1994). The pET28b, pET30a, pGEX4T-3 and *H. pylori* DNA sequences were all electrophoresed on a 1% agarose gel and purified using the Wizard PCR Preps kit (Promega Corp., Madison WI, USA). Following ligation of the purified, digested expression vector and the purified, digested *H. pylori* DNA sequences, the products of the ligation reaction were transformed into *E. coli* JM109 competent cells (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel *et al.*, eds., 1994). Individual bacterial colonies were screened for those containing the correct recombinant plasmids by preparing plasmid DNA as described above followed by analysis by restriction digestion profiles and DNA sequencing (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel *et al.*, eds., 1994). These recombinant plasmids were then used to transform specific *E. coli* expression strains.

30 *Transformation of competent bacteria with recombinant expression plasmids*

Competent bacterial strains BL21(T)E3, BL21(DE3)pLysS, HMS174(DE3) and HMS174(DE3)pLysS were prepared and transformed with the recombinant pET28b expression plasmids carrying the cloned *H. pylori* sequences according to standard methods (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel *et al.*, eds., 1994). These expression host strains contain a chromosomal copy of the gene for T7 RNA polymerase. These hosts are lysogens of bacteriophage DE3, a lambda derivative that carries the *lacI* gene, the *lacUV5* promoter and the gene for T7

RNA polymerase. T7 RNA polymerase expression is induced by the addition of isopropyl- β -D-thiogalactoside (IPTG), and the T7 RNA polymerase then transcribes any target plasmid, such as pET28b, that carries a T7 promoter sequence and a gene of interest.

- 5 Competent bacterial strains JM109 and DH5 α were prepared and transformed with the recombinant pGEX4T-3 expression plasmid carrying the cloned *H. pylori* sequences according to standard methods (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel *et al.*, eds., 1994).

10 *Expression of recombinant H. pylori sequences in E. coli*

- Transformants were collected from LB agar plates containing 25ug/ml kanamycin sulfate (ensures maintenance of the pET28b-based recombinant plasmids) or 100ug/ml ampicillin (ensures maintenance of the pGEX4T-3-based recombinant plasmids) and used to inoculate LB broth containing 25ug/ml kanamycin sulfate or
15 100ug/ml ampicillin and grown to an optical density at 600nm of 0.5 to 1.0 OD units, at which point 1mM IPTG was added to the culture for one to three hours to induce gene expression of the *H. pylori* recombinant DNA constructions. After induction of gene expression with IPTG, bacteria were pelleted by centrifugation and resuspended in SDS-PAGE solubilization buffer and subjected to SDS-PAGE (Current Protocols in
20 Molecular Biology, John Wiley and Sons, Inc., F. Ausubel *et al.*, eds., 1994). Proteins were visualized by staining with Coomassie Brilliant Blue or detected by western immunoblotting using the specific anti-His tag monoclonal antibody (Clontech, Palo Alto, CA, USA) or the anti-GST tag antibody (Pharmacia LKB) using standard methods (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel *et al.*,
25 eds., 1994). The host strain that provided the highest level of recombinant protein production was then chosen for use in a large-scale induction in order to purify the recombinant protein. The strains used were HMS174(DE3) (pET28b-based constructs) and DH5 α (pGEX4T-3-based constructs).

- Removal of the C-terminal regions appeared in both systems to improve the level
30 of expression, although this increase was far more prominent in the GST-fusion system. All recombinant proteins produced were of the predicted molecular weight based on the DNA sequence plus, if necessary, the size of the fusion tag. The truncated portion of the *H. pylori* protein contains some extremely hydrophobic stretches, and removal of these may be the reason for the increased expression.

EQUIVALENTS

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments and methods described herein. Such equivalents are intended to be encompassed by the scope of the following claims.

- 109 -

SEQUENCE LISTING

1) GENERAL INFORMATION:

- 5 (i) APPLICANT:
 (A) NAME: Astra Aktiebolag
 (B) STREET: S-151 85
 (C) CITY: Sodertalje
 (D) STATE:
10 (E) COUNTRY: Sweden
 (F) POSTAL CODE (ZIP)
- (ii) TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
RELATING TO HELICOBACTER PYLORI AND
15 VACCINE COMPOSITIONS THEREOF
- (iii) NUMBER OF SEQUENCES: 275
- (iv) COMPUTER READABLE FORM:
20 (A) MEDIUM TYPE: CD/ROM ISO9660
 (B) COMPUTER:
 (C) OPERATING SYSTEM:
 (D) SOFTWARE:
- 25 (v) CURRENT APPLICATION DATA:
 (A) APPLICATION NUMBER
 (B) FILING DATE:
- 30 (vi) PRIOR APPLICATION DATA:
 (A) APPLICATION NUMBER: US 08/759,625
 (B) FILING DATE: 05-DEC-1996
- (vii) PRIOR APPLICATION DATA:
35 (A) APPLICATION NUMBER: US 08/823,745
 (B) FILING DATE: 25-MAR-1997
- (viii) PRIOR APPLICATION DATA:
40 (A) APPLICATION NUMBER: US 08/891,928
 (B) FILING DATE: 14-JULY-1997
- (ix) CORRESPONDENCE ADDRESS:
 (A) ADDRESSEE: LAHIVE & COCKFIELD
 (B) STREET: 28 State Street
 (C) CITY: Boston
45 (D) STATE: Massachusetts
 (E) COUNTRY: USA
 (F) ZIP: 02109-1875
- 50 (x) ATTORNEY/AGENT INFORMATION:
 (A) NAME: Mandragouras, Amy E.
 (B) REGISTRATION NUMBER: 36,207
 (C) REFERENCE/DOCKET NUMBER: GTN-011CP2PC
- (xi) TELECOMMUNICATION INFORMATION:

- 110 -

(A) TELEPHONE: (617)227-7400
 (B) TELEFAX: (617)227-5941

(2) INFORMATION FOR SEQ ID NO:1:

5

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 687 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: circular

10

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

15

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

20

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc_feature
 (B) LOCATION 1...687

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

30

35

ATGAGATTTA	AGGGTTCAAG	AGTGGAAGCG	TTTTTAGGAG	CGTTAGAATT	TCAAGAGAAT	60
GAATATGAAG	AGTTTAAAGA	GCTTTATGAG	AGCTTAAAAA	CCAAGCAAAA	GCCCCACACT	120
TTGTTTCATTT	CTTGCGTGGA	TTCACGAGTC	GTGCCTAATT	TAATCACAGG	CACCCAACCG	180
GGCGAATTGT	ATGTGATCCG	CAACATGGGC	AATGTGATCC	CCCCTAAAAC	AAGCTATAAA	240
GAATCCCTTT	CTACCATTGC	GAGCGTTGAA	TACGCTATCG	CGCATGTGGG	CGTTCAAAAC	300
TTAATCATTT	GCGGGCATAG	CGATTGTGGG	GCTTGCGGGA	GCATTCATTT	AATCCATGAT	360
GAAACCACCA	AAGCTAAAAC	CCCTTACATT	GCAAACTGGA	TACAATTTTT	AGAGCCTATT	420
AAAGAAGAAT	TAAAAAACCA	CCCGCAATTC	AGCAACCATT	TCGCCAAGCG	TTCATGGCTT	480
ACAGAGCGTT	TGAATGCGCG	CTTGCAACTC	AACAACCTCT	TAAGCTATGA	TTTCATTCAA	540
GAAAGAGTAA	TAAATAACGA	ATTAAAAATT	TTTGGTTGGC	ACTATATCAT	AGAAACAGGC	600
AGGATTTATA	ATTATAATTT	TGAAAGCCAT	TTTTTTGAGC	CGATTGAAGA	AACCATTAAA	660
CAAAGGATAA	GTCATGAAAA	CTTCTAA				687

40

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

45

(A) LENGTH: 666 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

50

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- 111 -

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...666

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```

10 GTGGAAGCGT TTTTAGGAGC GTTAGAATTT CAAGAGAATG AATATGAAGA GTTTAAAGAG 60
   CTTTATGAGA GCTTAAAAAC CAAGCAAAAG CCCACACACTT TGTCATTTC TTGCGTGGAT 120
   TCACGAGTCG TGCCTAATTT AATCACAGGC ACCCAACCGG GCGAATTGTA TGTGATCCGC 180
   AACATGGGCA ATGTGATCCC CCCTAAAACA AGCTATAAAG AATCCCTTTC TACCATTGCG 240
   AGCGTTGAAT ACGCTATCGC GCATGTGGGC GTTCAAACT TAATCATTTG CGGGCATAGC 300
   GATTGTGGGG CTGCGGGAG CATTCAATTA ATCCATGATG AAACCACCAA AGCTAAAACC 360
15 CCTTACATTG CAACTGGAT ACAATTTTGA GAGCCTATTA AAGAAGAATT AAAAAACCAC 420
   CCGCAATTCA GCAACCATT CGCCAAGCGT TCATGGCTTA CAGAGCGTTT GAATGCGCGC 480
   TTGCAACTCA ACAACCTCTT AAGCTATGAT TTCATTCAAG AAAGAGTAAT AAATAACGAA 540
   TTAAAAATTT TTGGTTGGCA CTATATCATA GAAACAGGCA GGATTTATAA TTATAATTTT 600
   GAAAGCCATT TTTTGTAGCC GATTGAAGAA ACCATTAAAC AAAGGATAAG TCATGAAAAC 660
20 TTCTAA 666

```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1008 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...1008

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```

45 ATGTTAGTTA CTCGTTTTAA AAAAGCCTTC ATTTCTTATT CTTTAGGCGT GCTTGTTGTT 60
   TATTATTAT TGAATGTGTG CAACGCTTCA GCACAAGAAG TCAAAGTCAA GGATTATTTT 120
   GGGGAGCAAA CCATAAAGCT TCCTGTTTCC AAAATAGCCT ATATAGGGAG TTATGTAGAA 180
   GTGCCTGCCA TGCTTAATGT TTGGGATAGG GTTGTAGGCG TTTCTGATTA TGCCTTTAAG 240
   GATGACATTG TCAAAGCCAC TCTCAAAGGC GAGGATCTTA AACGAGTCAA ACACATGAGC 300
50 ACCGATCATA CAGCCGCGTT GAATGTGGAA TTATTAAAAA AGCTTAGCCC TGATCTTGTTG 360
   GTAACCTTTG TGGGTAACCC TAAAGCGGTA GAGCATGCGA AAAAATTGTTG GATTTCATTC 420
   CTTTCTTTCC AAGAGACAAC GATTGCAGAG GCCATGCAAG CTATGCAAGC TCAAGCCACG 480
   GTCTTAGAAA TTGACGCTTC CAAAAAATTC GCCAAAATGC AAGAACTTT GGACTTTATT 540
   GCTGAGCGTT TGAAGGGCGT TAAAAGAAA AAGGGGGTGG AGCTTTTCCA TAAAGCCAAT 600

```

- 112 -

5 AAAATCAGCG GCCATCAAGC CATTAGCTCA GACATTTTATG AAAAAGGGGG TATAGATAAT 660
 TTTGGCTTGA AATACGTAA GTTTGGACGC GCTGACATTA GTGTGGAAAA AATCGTTAAA 720
 GAAAACCCTG AAATCATTTT CATTGGGTGG GTAAGCCCAC TCACTCCTGA AGACGTGTTG 780
 AACAACCCTA AATTTTCCAC TATCAAAGCC ATTAAAAATA AGCAAGTCTA TAAGCTCCCC 840
 ACGATGGATA TTGGCGGTCC TAGAGCCCCA CTCATTAGTC TTTTATCGC TTTAAAAGCC 900
 CACCCTGAAG CCTTTAAAGG CGTGGATATT AATGCGATAG TCAAAGATTA TTATAAAGTG 960
 GTCTTTGATT TGAATGATGC GGAAATTGAG CCATTCTTAT GGCCTGA 1008

(2) INFORMATION FOR SEQ ID NO:4:

10

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 825 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: circular

15

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

20

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

25

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
 (B) LOCATION 1...825

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

ATGTTAGTTA CTCGTTTTAA AAAAGCCTTC ATTTCTTATT CTTTAGGCGT GCTTGTTGTT 60
 TCATTATTAT TGAATGTGTG CAACGCTTCA GCACAAGAAG TCAAAGTCAA GGATTATTTT 120
 GGGGAGCAAA CCATAAAGCT TCCTGTTTCC AAAATAGCCT ATATAGGGAG TTATGTAGAA 180
 35 GTGCCTGCCA TGCTTAATGT TTGGGATAGG GTTGTAGGCG TTTCTGATTA TGCCTTTAAG 240
 GATGACATTG TCAAAGCCAC TCTCAAAGGC GAGGATCTTA AACGAGTCAA ACACATGAGC 300
 ACCGATCATA CAGCCGCGTT GAATGTGGAA TTATTAAAAA AGCTTAGCCC TGATCTTGTG 360
 GTAACCTTTG TGGGTAACCC TAAAGCGGTA GAGCATGCGA AAAAATTTGG GATTTTCATT 420
 CTTTCTTTCC AAGAGACAAC GATTGCAGAG GCCATGCAAG CTATGCAAGC TCAAGCCACG 480
 40 GTCTTAGAAA TTGACGCTTC CAAAAAATTC GCCAAAATGC AAGAACTTTT GGACTTTATT 540
 GCTGATCGTT TGAAGGGCGT TAAAAAGAAA AAGGGGGTGG AGCTTTTCCA TAAAGCCAAT 600
 AAAATCAGCG GCCATCAAGC CATTAACCTCA GACATTTTAC AACAAGGGGG TATTGATAAT 660
 TTTGGCTTGA AATACGTCAA GTTTGGACGC GCTGACATTA GTGTGGAAAA AATCGTTAAA 720
 GAAAACCCTG AAATCATTTT CATTAGGTGG GTAACCCAC TCACTCCTGA TTACGTGTTG 780
 45 AACAACCCAA AATTTTCTAC TATCAATGCC ATTAAAAACA TATAA 825

(2) INFORMATION FOR SEQ ID NO:5:

50

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1287 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: circular

- 113 -

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

5 (iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

10 (ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...1287

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

15

ATGAAGAAAA	AATTTCTGTC	ATTAACCTTA	GGTTCGCTTT	TAGTTTCCGC	TTTAAGCGCT	60
GAAGACAACG	GCTTTTTTGT	GAGCGCCGGC	TATCAAATCG	GTGAATCCGC	TCAAATGGTG	120
AAAAACACCA	AAGGCATTCA	AGATCTTTCA	GACAGCTATG	AAAGATTGAA	CAACCTTTTA	180
ACGAATTATA	GCGTCCTAAA	CGCTCTCATC	AGGCAGTCCG	CCGACCCCAA	CGCCATCAAT	240
20 AACGCAAGGG	GCAATTTGAA	CGCGAGCGCG	AAGAATTTGA	TCAATGATAA	AAAGAATTCC	300
CCGGCGTATC	AAGCCGTGCT	TTTAGCCTTG	AATGCGGCAG	CGGGGTTGTG	GCAAGTCATG	360
AGCTATGCGA	TCAGCCCTTG	TGGTCCCGGT	AAAGACACAA	GCAAAAATGG	GGGCGTTCAA	420
ACTTTCCACA	ACACGCCTTC	AAATCAATGG	GGAGGCACTA	CCATTACTTG	TGGCACTACT	480
GGTTATGAAC	CAGGACCATA	CAGCATTTTA	TCCACTGAAA	ATTACGCGAA	AATCAATAAA	540
25 GCTTATCAAA	TCATCCAAAA	GGCTTTTGGG	AGCAGCGGAA	AAGATATTCC	TGCCTTAAGC	600
GACACCAACA	CAGAACTCAA	ATTCAACAATC	AATAAAAAATA	ATGGAAACAC	GAATACGAAT	660
AATAATGGAG	AAGAAATTGT	TACAAAAAAT	AACGCTCAAG	TTCTTTTAGA	ACAGGCTAGC	720
ACCATTATAA	CTACCCTTAA	TAGCGCATGC	CCATGGATCA	ACAATGGTGG	TGCAGGTGGT	780
GCGAGTAGTG	GTAGTTTATG	GGAAGGAATA	TATTTGAAAAG	GCGATGGGAG	CGCTTGCGGG	840
30 ATTTTTTAAAA	ATGAAATCAG	CGCGATTCAA	GACATGATCA	AAAACGCTGC	AATAGCCGTA	900
GAGCAATCCA	AGATCGTTGC	TGCAAACGCG	CAAAACCAGC	GCAACCTAGA	CACCGGGAAG	960
ACATTCAACC	CCTATAAAGA	CGCCAACCTC	GCCCAAAGCA	TGTTGCTAA	CGCCAAAGCG	1020
CAAGCGGAGA	TTTTAAACCG	CGCCCAAGCA	GTGGTGAAAAG	ACTTTGAAAAG	AATCCCTGCA	1080
GAGTTCGTAA	AAGACTCTTT	AGGGGTGTGC	CATGAAGTGC	AAAACGGCCA	TCTCCGTGGC	1140
35 ACGCCATCCG	GCACGGTAAC	TGATAACACT	TGGGGAGCCG	GTTGCGCGTA	TGTGGGAGAG	1200
ACCGTAACGA	ATCTAAAAGA	CAGCATCGCT	CATTTTGGCG	ACCAAGCCGA	GCGAATCCAT	1260
AACGCGCGCA	ACCTCGCTAC	ACTTTAG				1287

(2) INFORMATION FOR SEQ ID NO:6:

40

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 537 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

45

(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

50

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(B) LOCATION 1...537

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

10	ATGAACCCCT	TATTGCAAGA	TTATGCGCGC	ATCCTTTTGTAG	AATGGAATCA	AACGCACAAC	60
	TTGAGCGGCG	CGAGAAATTT	AAGCGAATTA	GAACCCCGAGA	TCACAGACGC	TC'TAAAGCCC	120
	TTAGAATTGT	TCAAAGATTT	TAAAAGCTGC	TTGGATATTG	GGAGCGGGGC	GGGACTTCCT	180
	GCTATCCCTT	TAGCCCTTGA	AAAACCTGAA	GCGCAATTCA	TTCTTTTGTAGA	GCCAAGGGTA	240
	AAAAGAGCGG	CTTTTTTAAA	CTACCTTAAA	AGCGTTTTTGC	CTTTAAACAA	CATTGAAATC	300
ATTAAAAAGC	GTTTAGAAGA	TTATCAAAAT	CTTTTACAAG	TGGATTTAAT	CAC'TTCTAGA	360	
15	GCGGTCGCTA	GCTCTTCTTT	TTTGATAGAA	AAAAGCCAAC	GCTTCCTAAA	AGATAAGGGG	420
	TATTTTTTAT	TCTATAAAGG	CGAGCAGTTA	AAGAATGAAA	TCGCTTATAA	AACCACTGAA	480
	TGCTTTATGC	ATCAAAAGCG	CGTTTATTTT	TACAAATCAA	AGGAAAGTTT	ATGTTAA	537

(2) INFORMATION FOR SEQ ID NO:7:

20

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 723 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: circular

25

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

30

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

35

(ix) FEATURE:

(A) NAME/KEY: misc feature

(B) LOCATION 1...723

40

(xi) SEQUENCE DESCRIPTION: SEO ID NO:7:

45

50

TTGGGTCTTA	AAAAACGAGC	TATTTTATGG	TCTTTAATGG	GATTTTGTGC	AGGATTGAGC	60
GCGCTTGATT	ATGACACCCT	AGACCCAAA	TATTACAAAT	ATATCAAGTA	TTATAAGGCT	120
TATGAAGATA	AAGAA GTTGA	AGAATTGATC	AGAGACTTGA	AAAGGGCGAA	CGCTAAAAGC	180
GGGCTTATTT	TAGGGATCAA	TACCGGTTTT	TTTTATAACC	ATGAAATCAT	GGTCAAAACC	240
AATAGCTCCA	GTATCACC3G	GAATATTTTA	AATTATTTGT	TCGCCTATGG	CTTGCGTTTT	300
GGCTATCAA	CTTTC.3,CC	GTCGTTTTTT	GCGCGCTTGG	TTAAGCCCAA	TATCATTGGC	360
AGGCGCATCT	ATATTCAATA	TTATGGAGGA	GCTCCTAAGA	AAGCGGGCTT	TGGGAGCGTG	420
GGGTTTCAAT	CGGTCATGTT	GAATGGGGAT	TTTTTATTAG	ACTTTCCTTT	GCCCTTTGTG	480
GGGAAATACC	TTTATATGGG	GGGGTATATG	GGTTTAGGCT	TGGGGGTGTG	GGCGCATGGG	540
GTGAATTATA	CGGCGGAATG	GGGGATGTCT	TTTAACGCAG	GATTGGCTCT	AACGGTATTA	600
GAAAAAACC	GCATTGAATT	TGAATTTAAA	ATTTTGAATA	ATTTCCCTTT	TTTGCAATCT	660
AATTCTTCAA	AAGAGACTTG	GTGGGGAGCT	ATAGCAAGCA	TTGGGTATCA	ATATGTGTTC	720
TAA						723

- 115 -

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 942 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
 (B) LOCATION 1...942

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

25	TTGAAACTCA AATACTGGTT AGTTTATCTG GCGTTCATTA TAGGACTTCA AGCGACAGAT	60
	TATGACAATT TAGAAGAAGA AAACCAACAA TTAGACGAAA AAATAAACAA TTTAAAGCGA	120
	CAGCTCACCG AAAAAGGGGT TTCACCCAAA GAGATGGATA AGGATAAGTT TGAAGAAGAA	180
	TATTTAGAGC GAACTTACCC AAAGATTTCT TCAAAGAAAA GAAAAAAATT GCTCAAATCT	240
	TTTTCCATAG CCGATGATAA GAGTGGGGTG TTTTATGGGG GCGGGTATGC TTATGGGGAA	300
	CTTAAGTTGT CTTATCAAGG GGAGATGTTA GACAGGTATG GCGCAAATGC CCCTAGCGCG	360
30	TTTAAAAACA ATATCAATAT TAACGCTCCT GTTCTATGA TTAGCGTTAA ATTTGGGTAT	420
	CAAAAATACT TCGTGCCTTA TTTTGGGACA CGATTTTATG GGGATTTGTT GCTTGGGGGA	480
	GGGGCGTTAA AAGAGAACGC GCTCAAGCAG CCTGTAGGCT CGTTTTTTTA TGTTTTAGGG	540
	GCTATGAATA CCGATTTATT GTTTGACATG CCTTTAGATT TTAAGACTAA AAAGCATTTT	600
	TTAGGCGTTT ATGCGGGTTT TGGGATAGGG CTTATGCTTT ATCAAGACAA GCCTAATCAA	660
35	AACGGGAGGA ATTTGATAGT AGGGGGTTAT TCAAGCCCTA ATTTTTTATG GAAATCTTTG	720
	ATTGAAGTGG ATTACACTTT TAATGTGGGC GTGAGTTTAA CGCTTTATAG GAAACACCGC	780
	TTAGAGATTG GCACAAAATT ACCGATTAGC TATTTGAGGA TGGGAGTAGA AGAGGGAGCG	840
	ATTTATCACA ATAAAGAAAA TGATGAACGA TTGTTGATTT CGGCTAACAA CCAGTTCAAA	900
40	CGATCCAGTT TTTTATTAGT GAATTATGCG TTCATTTTTT GA	942

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1182 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

- 116 -

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

5 (A) NAME/KEY: misc_feature

(B) LOCATION 1...1182

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

```

10  ATGACTTCAG CTTCAAGCCA TTCTTTTAAA GAACAAGATT TTCATATTCC TATCGCTTTC      60
    GCTTTTGATA AGAATTATCT CATTCCGTGC GGCGCATGCA TTTATTCCTT GCTAGAAAAGC      120
    ATCGCTAAAG CCAATAAAAA AATCCGTTAC ACCTTACACG CTTTAGTGGT AGGCTTGAAT      180
    GAAGAAGATA AAACAAAAC TAAACAAATC ACAGAGCCTT TTAAAGAATT TGCTGTTTTA      240
    GAAGTAAAAG ATATTGAACC TTTTGTAGAC ACTATCCCTA ACCCTTTTGA TGAGGATTTT      300
15  ACCAAGCGTT TTTCTAAAAT GGTGTTAGTG AAGTATTTTC TAGCGGATTT ATTCCCCAAA      360
    TATTCTAAAA TGGTGTGGAG CGATGTGGAT GTTATCTTTT GTAATGAATT TAGCGCTGAT      420
    TTCTTAAACA TTAAAGAAGA TGATGAGAAT TATTTTATATG GGGTTTATGA CAAAATATAC      480
    CCGTATGAAG GCTTTTTTTT TTGCAACTTA ACTTACCAGC GAAAAAATCA ATTTTGTAAT      540
    AAAATATTAG AAATCATACG CGCACAAAAA ATAGATAAAG AACCAGCAAT GACAGAATTT      600
20  TGTCGTTCAA AGATCGCGCC ATTAATAATA GAGTATTGTA TTTTCCCACA CTATTATAGC      660
    CTTTCTGAAG AGCATTTTAA GGGCGTGGCC AATGCAATTT ATCATAACAC CATTAAACAA      720
    GCCCTAAGAG AACCTATCGT TATACAATAT GACTCTCATC CTTATTTTCA AATCAAGCCT      780
    TGGACATATC CTTTGTGTTT GAAAGCGGAT TTATGGCTGA ACGCTTTGGC TAAAACCCCA      840
    TTTATGAGCG ATTGGTCTTA TTTGATCACA GGGGGTGGGG GGATAGGTGG AGAAAAATGG      900
25  CATTACTACC ATGGCATTGC CGCTTATCAT TACTACTTTC CTTTATGGAA AGCAGAAGAA      960
    CAGATTGCCC ATGACGCTCT TAAGACATTT TTAAACATT ATTTTTTGCA CATTATGAG      1020
    ATTCCCCAAA ACGCAAGGCG AAGACTATTC AAATACTGCA TTTCAATACC GCTTAAGAGC      1080
    TTTATTAGTA AAACCCTTAA ATTTCTAAAA CTCCATGCAT TGGTGAAAAA AATCCTAATC      1140
    CAACTCAAGC TCTTAAAAAA GAACCAGAGC CAAACTTTTT AA      1182

```

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

```

35  (A) LENGTH: 1308 base pairs
    (B) TYPE: nucleic acid
    (C) STRANDEDNESS: double
    (D) TOPOLOGY: circular

```

40 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

45 (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

50 (A) NAME/KEY: misc_feature

(B) LOCATION 1...1308

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

```

TTGATTTTCT TAAAAAATC TCTTTGCGCG TTGTTAATTT CAGGTTTTTT CATACCACCC      60

```


- 117 -

	TTAATGAAAG	CGGCTAGTTT	TGTCTATGAC	TTGAAGTTTA	TGAGCTTTAA	TTTCAATCTG	120
	GCTTCCCCTC	CAAATAACCC	CTATTGGAAT	AGCCTAACCA	AAATGCAAGG	TCGTCTCATG	180
	CCTCAAATTG	GCGTCCAATT	AGACAAAAGA	CAGGCCTTGA	TGTTTGGGGC	GTGGTTCATT	240
	CAAAATTTGC	ACACGCATTA	TAGCTATTTT	CCTTATTCGT	GGGGGGTTAC	CATGTATTAC	300
5	CAATACATAG	GGAAAAATTT	GAGATTTTTT	TTAGGCATTG	TGCCACGAAG	CTATCAAATA	360
	GGGCATTACC	CTTTAAGCGC	TTTTAAAAAA	CTTTTCTGGT	TTATAGACCC	TACTTTTAGG	420
	GGAGGAGCGT	TCCAATTCAA	ACCGGCTTAT	GATCCCAATC	GTTGGTGGAA	TGGGTGGTTT	480
	GAGGGCGTTG	TGGATTGGTA	TGGGGGGCGT	AATTGGAACA	ACCAGCCCAA	AAAGAAAAAT	540
	TACGATTTTG	ATCAATTCTT	GTATTTTGTT	TCTTCAGAAT	TTCAGTTTCT	TAAAGGGTAT	600
10	TTAGGTTTGG	GGGGACAGCT	TGTCATTTTT	CATAACGCCA	ACTCTCATAG	TATGGGGGAT	660
	AACTACCCTT	ATGGCGGGAA	TTCCTACTTA	AAACCAGGCG	ATGCAACCCC	ACAATGGCCT	720
	AATGGCTACC	CTTATTTTCA	CCAAAAAGAT	AACCCACAAG	GCGGAGAAAT	AGGGAAATAC	780
	TCTAACCCCTA	CCATTTTATA	CAGGGTTTAT	TACCATGCTT	ATTTAAAAGC	AGATTTTAAA	840
	AATCTCATGC	CTTATATGGA	CAATATTTTC	ATGACCTTTG	GCACGCAGTC	GTCTCAAACC	900
15	CATTATTGCG	TGCGTTATGC	TAGCGAGTGT	AAAAACGCC	GATTTTATAA	CAGCTTTGGG	960
	GGGGAATTTT	ACGCTCAAGC	GCAATACAAA	GGCTTTGGGA	TCTTTAACAG	ATACTATTTT	1020
	TCCAACAAAC	CCCAAATGCA	TTTTTATGCC	ACTTATGGCC	AATCCCTTTA	TACCGGATTG	1080
	CCATGGTATA	GAGCCCCTAA	TTTTTGACATG	ATAGGGCTTT	ATTATCTTTA	TAAAAACAAA	1140
	TGGTTAAGCG	TGCGAGCGGA	TGCGTTTTTT	AGCTTTGTGG	GTGGGGGCGA	TGGGTACCAT	1200
20	TTGTATGGCA	AGGGGGGTAA	GTGGTTTGTG	ATGTATCAGC	AATTTTAAAC	CCTAACCATA	1260
	GACACAAGAG	AGTTGATTGA	TTTTGTCAAA	TCTAAATCC	CTAAATAA		1308

(2) INFORMATION FOR SEQ ID NO:11:

- 25 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 663 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: circular
- 30 (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- 35 (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: *Helicobacter pylori*
- 40 (ix) FEATURE:
- (A) NAME/KEY: misc_feature
 - (B) LOCATION 1...663
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

45	ATGAATAAAA	CAACAATTAA	AATATTAATG	GJC	ATGGCGT	TATTATCATC	GCTTCAAGCC	60
	GCAGAGGCAG	AGCTTGATGA	AAAATCAAAA	AAACCTAAAT	TTGCGGATAG	GAATACGTTT		120
	TATTTAGGGG	TTGGGTATCA	GCTTAGCGCG	ATCAACACGT	CTTTTAGCAC	CAGTTCTATA		180
	GATAAATCGT	ATTTTATGAC	CGGCAATGGT	TTTGGCGTGG	TGTTGGGGGG	GAAATTTGTG		240
50	GCTAAACGC	AAGCTGTAGA	GCATGTGGGT	TTTCGTTACG	GGTTGTTTTA	TGATCAGACC		300
	TTTTCTTCTC	ACAAATCCTA	TATTTCTACC	TATGGTTTAG	AATTTAGCGG	TTTGTGGGAC		360
	GCTTTCAATT	CGCCAAAGAT	GTTTTTGGGG	TTGGAGTTTG	GCTTAGGCAT	CGCTGGGGCG		420
	ACTTACATGC	CAGGAGGGGC	CATGCATGGG	ATTATCGCTC	AATATTTAGG	CAAAGAAAAT		480
	TCGCTTTTCC	AATTGCTTGT	GAAAGTGGGT	TTTCGTTTTG	GCTTTTTTCCA	CAATGAAATC		540

- 118 -

ACCTTTGGGT TGAAATTCCC TGTCATTCCT AACAAAAAA CGGAAATCGT TGATGGCTTG 600
AGCGCGACCA CTTTATGGCA ACGCTTGCCG GTAGCCTATT TCAATTATAT CTATAATTTT 660
TAG 663

5 (2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 351 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

15 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- 20 (A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

- 25 (A) NAME/KEY: misc_feature
(B) LOCATION 1...351

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

30 TTGAATCTCC ATTTTATGAA AGGATTTGTT ATGAGTGGAT TAAGAACATT TAGTTGTGTA 60
GTGGTTTTAT GCGGTGCAAT GGTTAATGTA GCTGTAGCTG GTCCTAAAAT AGAGGCAAGG 120
GGTGAATTAG GCAAAATTGT AGGGGGAGCT GTTGGAAATT TTGTTGGTGA TAAAATGGGC 180
GGATTTGTTG GTGGTGCAAT AGGAGGATAT ATTGGGTCTG AAGTAGGCGA TAGGGTAGAA 240
GATTATATCC GTGGCGTTGA TAGAGAGCCA CAAAACAAAG AACCACAAAC CCCAAGAGAA 300
CCTATCCGTG ATTTTATGA TTACGGCTAT AGTTTTGGGC ATGCTTGGTG A 351

35 (2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- 40 (A) LENGTH: 1311 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

45 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- 50 (A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
(B) LOCATION 1...1311

- 119 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

```

5  ATGTCAAGGG ATTTTAAATT TGATTCTAAC TATTTAAATG TCAATACCAA TCCTAAATTA 60
   GGCCCCGTTT ATACCAATCA AAATTATCCA GGATTTTTTA TCTTTGATCA TTTAAGGCGT 120
   TATGTGATGA ACGCTTTTGA GCCTAATTTG AACTTAGTTG TCAATACCAA TAAAGTTAAG 180
   CAAACTTTTA ATGTGGGCAT GCGTTTTATG ACAATGGATA TGTTCATTAG ATCCGATCAA 240
   AGCACATGCG AAAAAACAGA TATTATCAAT GGGGTGTGCC ACATGCCTCC TTATGTCCTT 300
   TCTAAAACGC CTAACAATAA TCAAGAAATG TTTAATAACT ATACAGCGGT ATGGTTGAGC 360
10 GATAAAATAG AGTTTTTTTGA TTCTAAATTG GTGATAACTC CAGGGCTTAG ATACACTTTT 420
   TTGAACTATA ACAACAAAGA GCCAGAAAAG CATGATTTTT CCGTATGGAC CAGTAAAAAA 480
   CAGCGTCAAA ACGAATGGAG TCCTGCCCTT AATATTGGCT ATAAACCTAT GGAAAATTGG 540
   ATATGGTATG CGAACTACCG CCGCAGTTTT ATCCCCCAC AACACACAAT GGTAGGCATT 600
   ACTAGGACTA ATTACAACCA AATTTTTAAT GAAATTGAAG TGGGGCAGCG CTATAGTTAT 660
15 AAAAATCTAT TGAGTTTTAA CACCAATTAT TTTGTGATTT TTGCCAAGCG TTACTATGCG 720
   GGAGGCTATA GCCCACAGCC TGTGGATGCC AGAAGTCAAG GGGTGGAATT GGAATTGTAT 780
   TACGCGCCGA TTAGGGGTTT GCAATTCCAT GTGGCTTACA CTTATATTGA TCGCGGCATC 840
   ACTTCTAACG CTGATGATAT TGCTTATTAT TTTACAGGCA TTGTCAATAA ACCCTTTGAC 900
   ATTAAAGGGA AGCGCTTGCC CTATGTGAGT CCTAACCAAT TCATATTGA CATGATGTAT 960
20 ACTTACAAGC ACACGACTTT TGGTATCAGC AGCTATTTTT ATAGCCGCGC TTATAGTTCC 1020
   ATGCTCAATC AAGCCAAAGA TCAAACCGTA TGCCTGCCCT TAAACCCAGA ATACACAGGG 1080
   GGGTTAAAGT ATGGTTGTAA TTCAGTGGGG TTATTGCCCT TGTATTTTGT GTTGAATGTC 1140
   CAAGTAAGCT CAATCTTATG GCAAAGCGGT AGGCATAAAA TCACAGGGAG TTTGCAAATC 1200
   AATAACCTTT TTAACATGAA GTATTATTTT AGGGGGATTG GCACAAGCCC TACAGGGAGA 1260
25 GAACCCGCGC CAGGGAGATC CATTACAGCG TATTTGAATT ATGAGTTTTA A 1311

```

(2) INFORMATION FOR SEQ ID NO:14:

```

30 (i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 2304 base pairs
    (B) TYPE: nucleic acid
    (C) STRANDEDNESS: double
    (D) TOPOLOGY: circular

35 (ii) MOLECULE TYPE: DNA (genomic)

    (iii) HYPOTHETICAL: NO

    (iv) ANTI-SENSE: NO

40 (vi) ORIGINAL SOURCE:
    (A) ORGANISM: Helicobacter pylori

    (ix) FEATURE:
45 (A) NAME/KEY: misc_feature
    (B) LOCATION 1...2304

```

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

```

50 ATGAAAAGAA TTTTAGTTTC TTTGGCTGTT TTGAGTCATA GCGCGCATGC TGTCAAAACT 60
   CATAATTTGG AAAGGGTGGA AGCTTCAGGG GTGGCTAACG ATAAAGAAGC GCCTTTAAGC 120
   TGGAGGAGCA AGGAAGTTAG AAATTATATG GGTTCCTGCA CCGTGATTTT TAACAAGCAA 180
   CTCATAAAA GCGCCAATCA AAGCATTGAA GAAGCTTTGC AAAATGTGCC AGGCGTGCAT 240
   ATTAGAAACT CTACCGGTAT TGGAGCTGTG CCTAGCATTT CCATTAGGGG GTTTGGTGCT 300

```

- 120 -

GGAGGCCCCAG GGCATTCTAA TACGGGAATG ATTCTAGTCA ATGGGATTCC TATTTATGTC 360
GCGCCCTATG TTGAAATTGG CACGGTTATT TTTCTGTAA CCTTTCAGTC TGTGGATAGA 420
ATCAGCGTAA CTAAGGGTGG GGAGAGCGTG CGTTATGGCC CTAACGCTTT TGGCGGTGTG 480
ATCAACATCA TCACCAAAGG CATTCCCTACC AATTGGGAAA GTCAGGTGAG CGAGAGGACC 540
5 ACTTTTTTGGG GCAAGTCTGA AAACGGGGGC TTTTCAATC AAAATTCTAA AAACATTGAT 600
AAAAGCTTAG TTAATAACAT GCTTTTTTAAC ACCTATTTAA GAACGGGGGG TATGATGAAT 660
AAGCATTTTG GAATCCAAGC TCAAGTCAAT TGGCTCAAAG GGCAAGGGTT TAGATACAAC 720
AGCCCTACGG ATATTCAAAA TTACATGTTA GATTCAATTG ATCAAATCAA TGATAGCAAT 780
AAAATCACCG CTTTTTTTCA ATATTATAGT TATTTCTTGA CAGACCCTGG ATCTTTAGGC 840
10 ATAGCCGCTT ACAATCAAAA TCGTTTTTCAA AACAACCGCC CCAATAACGA TAAAAGCGGG 900
AGAGCGAAGC GATGGGGAGC TGTGTATCAA AACTTTTTTG GGGACACGGA TAGGGTAGGG 960
GGGGATTTC CTTTTAGCTA CTATGGGCAT GACATGTCAA GGGATTTTAA ATTTGATTCT 1020
AACTATTTAA ATGTCAATAC CAATCCTAAA TTAGGCCCG TTTATACCAA TCAAAATTAT 1080
CCAGGATTTT TTATCTTTGA TCATTTAAGG CGTTATGTGA TGAACGCTTT TGAGCCTAAT 1140
15 TTGAACCTAG TTGTCAATAC CAATAAAGTT AAGCAAACCT TTAATGTGGG CATGCGTTTT 1200
ATGACAATGG ATATGTTTAT TAGATCCGAT CAAAGCACAT GCGAAAAAAC AGATATTATC 1260
AATGGGGTGT GCCACATGCC TCCTTATGTC CTTTCTAAAA CGCCTAACAA TAATCAAGAA 1320
ATGTTTAATA ACTATACAGC GGTATGGTTG AGCGATAAAA TAGAGTTTTT TGATTCTAAA 1380
TTGGTGATAA CTCCAGGGCT TAGATACACT TTTTGAACCT ATAACAACAA AGAGCCAGAA 1440
20 AAGCATGATT TTTCCGTATG GACCAGTAAA AAACAGCGTC AAAACGAATG GAGTCCTGCC 1500
CTTAATATTG GCTATAAACC TATGGAAAAT TGGATATGGT ATGCGAACTA CCGCCGCAGT 1560
TTTATCCCCC CACAACACAC AATGGTAGGC ATTACTAGGA CTAATTACAA CCAAATTTTT 1620
AATGAAATTG AAGTGGGGCA GCGCTATAGT TATAAAAATC TATTGAGTTT TAACACCAAT 1680
TATTTTGTGA TTTTGTCCAA GCGTTACTAT GCGGGAGGCT ATAGCCCACA GCCTGTGGAT 1740
25 GCCAGAAGTC AAGGGGTGGA ATTGGAATTG TATTACGCGC CGATTAGGGG TTTGCAATTC 1800
CATGTGGCTT ACACCTATAT TGATGCGCGC ATCACTTCTA ACGCTGATGA TATTGCTTAT 1860
TATTTTACAG GCATTGTCAA TAAACCCTTT GACATTAAAG GGAAGCGCTT GCCCTATGTG 1920
AGTCCTAACC AATTCATATT TGACATGATG TATACTTACA AGCACACGAC TTTTGGTATC 1980
AGCAGCTATT TTTATAGCCG CGCTTATAGT TCCATGCTCA ATCAAGCCAA AGATCAAACC 2040
30 GTATGCCCTG CTTTAAACCC AGAATACACA GGGGGGTTAA AGTATGGTTG TAATTCAGTG 2100
GGGTATTG CTTGTATTT TGTGTTGAAT GTCCAAGTAA GCTCAATCTT ATGGCAAAGC 2160
GGTAGGCATA AAATCACAGG GAGTTTGCAA ATCAATAACC TTTTAAACAT GAAGTATTAT 2220
TTTAGGGGGA TTGGCACAA G CCTACAGG AGAGAACCCG CGCCAGGGAG ATCCATTACA 2280
35 GCGTATTTGA ATTATGAGTT TTA 2304

(2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:
40 (A) LENGTH: 348 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
45 (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- 50 (vi) ORIGINAL SOURCE:
(A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
(A) NAME/KEY: misc_feature

- 121 -

(B) LOCATION 1...348

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

5	TTGCACCCTC TATGCGCACA CGGCCAATGT GGAAGCGAAG CGATTGCGTG TTTAGAAGCC	60
	ATTAGCGTGG GGATTGTGCC TGTTATCGCT AATAGCCCTT TAAGCGCGAC CAGGCAATTC	120
	GCGCTAGATG AACGATCGTT ATTTGAGCCT AATAACGCTA AAGATTTGAG CGCTAAAATA	180
	GACTGGTGGT TAGAAAACAA ACTTGAAAGA GAAAGAATGC AAAACGAATA CGCTAAAAGC	240
	GCTTTAAACT ACACTTTAGA AAATTCAGTC ATTCAAATTG AAAAAGTTTA TGAAGAAGCG	300
10	ATCAAAGATT TTAAAAACAA CCCCAACCTC TTAAAACCT TATCGTAA	348

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

15	(A) LENGTH: 1170 base pairs
	(B) TYPE: nucleic acid
	(C) STRANDEDNESS: double
	(D) TOPOLOGY: circular

20	(ii) MOLECULE TYPE: DNA (genomic)
----	-----------------------------------

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

25

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

30

(A) NAME/KEY: misc_feature

(B) LOCATION 1...1170

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

35	ATGGTTATTG TTTTAGTCGT GGATAGCTTT AAAGACACCA GTAATGGCAC TTCTATGACA	60
	GCGTTTCGTT TTTTGAAGC GCTGAAAAA AGAGGGCATG CGATGAGAGT GGTGCCCCCT	120
	CATGTGGATA ATTTAGGGAG TGAAGAAGAG GGGTATTACA ACCTTAAAGA GCGCTATATC	180
	CCCCTAGTTA CAGAAATTTT ACACAAGCAA CACATTCCTT TTGCCAAACC GGATGAAAAA	240
	ATTCTACGAA AGGCTTTTAA GGGAGCGGAT ATGATCCATA CTTACTTGCC TTTTGTGCTA	300
40	GAAAAAACAG CCGTAAAAAT CGCGCGAGAA ATGCGAGTGC CTTATATTGG CTCTTCCAT	360
	TTACAGCCAG AGCATATTTT TTATAACATG AAATTGGGGC AATTTTCTTG GCTAAATACC	420
	ATGCTTTTTT CATGGTTTAA ATCTTCGCAT TACCGCTATA TCCACCATAT CCATTGCCCA	480
	TCAAAATTCA TTGTAGAAGA ATTGGAAAAA TACAACATATG GAGGAAAAAA ATACGCTATC	540
	TCTAACGGCT TTGATCCCAT GTTTAAGTTT GAGCACCCGC AAAAAAGCCT TTTTGACACC	600
45	ACGCCCTTTA AAATCGCTAT GGTAGGCGCG TATTCTAATG AAAAAATCA AAGCGTTCTC	660
	ATTAAAGCGG TTGCTTTAAG CCGA.A.AAA CAAGACATTG TATTATTACT CAAAGGCAAG	720
	GGGCTGATG AGAAAAAAT CAAACTTCTA GCCCAAAAAC TAGGCGTAAA AACGGAGTTT	780
	GGGTTGTCA ATTCCCATGA ATTGTTAGAG ATTTTAAAAA CTTGCACCCT CTATGCGCAC	840
	ACGGCCAATG TGGAAAGCGA AGCGATTGCG TGTTTAGAAG CCATTAGCGT GGGGATTGTG	900
50	CCTGTTATCG CTAATAGCCC TTAAAGCGCG ACCAGGCAAT TCGCGCTAGA TGAACGATCG	960
	TTATTTGAGC CTAATAACGC TAAAGATTG AGCGCTAAAA TAGACTGGTG GTTAGAAAAC	1020
	AAACTTGAAA GAGAAAGAAT GCAAAACGAA TACGCTAAAA GCGCTTTAAA CTACACTTTA	1080
	GAAATTCAG TCATTCAAAT TGAAAAAGTT TATGAAGAAG CGATCAAAGA TTTTAAAAAC	1140
	AACCCCAACC TCTTTAAAC CTTATCGTAA	1170

- 122 -

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 939 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: circular

10 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

15 (iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

- 20 (A) NAME/KEY: misc_feature
 (B) LOCATION 1...939

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

25 TTGGCTTCTT ACGGGTTTTT TTTAGGAGCG TTGTTTATTT TAGCGAGCGG GATCGTGTGC 60
 TTACAGACTG CCGGTAATCC CTTTGTAACC TTGCTTTCTA AAGGTAAAGA AGCCAGAAAC 120
 TTGGTTTTAG TCCAGGCGTT CAATTCGCTT GGCACGACTT TAGGGCCTAT TTTTGGGAGC 180
 TTGTTGATTT TTAGCGCGAC CAAAACGAGC GATAATTTAA GCCTGATAGA CAAGTTAGCG 240
 GACGCTAAAA GCGTTCAAAT GCCTTATTTG GGTTTAGCGG TGTTTTCGCT TCTTTTAGCG 300
 30 CTTGTGATGT ATCTTTTAAA ATTGCCTGAT GTGGAAAAG AAATGCCCAA AGAAACGACG 360
 CAAAAAGCC TGTTTTTCGCA CAAACACTTT GTTTTTGGGG CTTTAGGGAT CTTTTTCTAT 420
 GTGGGGGGAG AAGTGGCGAT TGGATCATTG TTGGTGCTAA GCTTTGAAAA GCTTTTGAAT 480
 TTAGACGCTC AATCAAGCGC GCATTACTTG GTGTATTATT GGGGCGGCGC GATGGTAGGG 540
 CGTTTCTTAG GCAGCGCTTT GATGAATAAA ATCGCTCCTA ATAAATACCT GGCTTTCAAC 600
 35 GCCTTAAGCT CTATCATTCT TATCGCTTTG GCTATTCTTA TTGGAGGCAA GATCGCTTTA 660
 TTCGCTCTGA CTTTTGTGGG CTTTTTCAAC TCTATCATGT TCCCTACAAT CTTTTCTTTG 720
 GCTACGCTCA ATTTAGGGCA TCTCACTTCT AAGGCTTCTG GAGTGATTAG CATGGCGATT 780
 GTGGGAGGGG CGTTAATCCC CCCCAATCAA GCGTGTTA CAGACATGCT CACAGCAACC 840
 40 GAATCGAATC TGCTCTACGC TTATAGCGTG CCGTTGTTGT GCTATTTTTT TATCCTCTTC 900
 TTTGCACTTA AGGGGTATAA ACAAGAAGAA AACTCCTAA 939

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- 45 (A) LENGTH: 1224 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: circular

50 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

- 123 -

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

5

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...1224

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

10

ATGCAAAAAA	CTTCTAACAC	TTTAGCGCTG	GGGAGTTTGA	CGGCGCTATT	CTTTCTAATG	60
GGTTTTATCA	CGGTTTTAAA	CGACATTTTG	ATCCCGCATT	TAAAGCCCAT	TTTTGACTTG	120
ACCTATTTTG	AAGCTTCGCT	CATTCAATTT	TGCTTTTTTG	GGGCGTATTT	CATCATGGGG	180
GGAGTCTTTG	GGAACGTGAT	CAGTAAAATC	GGCTACCCTT	TTGGCGTGGT	GCTTGGTTTT	240
15 GTGATCACAG	CGAGCGGGTG	CGCGTTGTTT	TATCCGGCGG	CGCATTTTGG	CTCTTACGGG	300
TTTTTTTTAG	GAGCGTTGTT	TATTTTAGCG	AGCGGGATCG	TGTGCTTACA	GACTGCCCGT	360
AATCCCTTTG	TAACCTTGCT	TTCTAAAGGT	AAAGAAGCCA	GAAACTTGGT	TTTAGTCCAG	420
GCGTTCAATT	CGCTTGGCAC	GACTTTAGGG	CCTATTTTTG	GGAGCTTGTT	GATTTTTAGC	480
GCGACCAAAA	CGAGCGATAA	TTTAAGCCTG	ATAGACAAGT	TAGCGGACGC	TAAAAGCGTT	540
20 CAAATGCCTT	ATTTGGGTTT	AGCGGTGTTT	TCGCTTCTTT	TAGCGCTTGT	GATGTATCTT	600
TTAAAATTGC	CTGATGTGGA	AAAAGAAATG	CCCAAAGAAA	CGACGCAAAA	AAGCCTGTTT	660
TCGCACAAAC	ACTTTGTTTT	TGGGGCTTTA	GGGATCTTTT	TCTATGTGGG	GGGAGAAGTG	720
GCGATTGGAT	CATTCTTGGT	GCTAAGCTTT	GAAAAGCTTT	TGAATTTAGA	CGCTCAATCA	780
AGCGCGCATT	ACTTGGTGTA	TTATTGGGGC	GGCGCGATGG	TAGGGCGTTT	CTTAGGCAGC	840
25 GCTTTGATGA	ATAAAATCGC	TCCTAATAAA	TACCTGGCTT	TCAACGCCTT	AAGCTCTATC	900
ATTCTTATCG	CTTTGGCTAT	TCTTATTGGA	GGCAAGATCG	CTTTATTTCG	TCTGACTTTT	960
GTGGGGCTTTT	TCAACTCTAT	CATGTTCCCT	ACAATCTTTT	CTTTGGCTAC	GCTCAATTTA	1020
GGGCATCTCA	CTTCTAAGGC	TTCTGGAGTG	ATTAGCATGG	CGATTGTGGG	AGGGGCGTTA	1080
ATCCCCCCCC	TTCAAGGCGT	GGTTACAGAC	ATGCTCACAG	CAACCGAATC	GAATCTGCTC	1140
30 TACGCTTATA	GCGTGCCGTT	GTTGTGCTAT	TTTTATATCC	TCTTCTTTGC	ACTTAAGGGG	1200
TATAACAAG	AAGAAACTC	CTAA				1224

(2) INFORMATION FOR SEQ ID NO:19:

35

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 378 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: circular

40

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

45

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

50

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...378

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

- 124 -

```

ATGAATAAAA TCGCTCCTAA TAAATACCTG GCTTTCGGCG CCTTAAGCTC TATCATTCCT 60
ATCGCTTTGG CTATTCCTAT TGGAGGCAAG ATCGCTTTAT TCGCTCTGAC TTTTGTGGGC 120
TTTTTCAACT CTATCATGTT CCCTACAATC TTTTCTTTGG CTACGCTCAA TTTAGGCATC 180
5 TCACTTCTAA TGGCTTCTGG AGTGATTAGC ATGGCGATTG TGGGAGGGGC GTTAATCCCC 240
CCCATTCAAG GCGTGGTTAC AGACATGCTC ACAGCAACCG AATCGAATCT GCTCTACGCT 300
TATAGCGTGC CGTTGTTGTG CTATTTTAT ATCCTCTTCT TTGCACTTAA GGGGTATAAA 360
CAAGAAGAAA ACTCCTAA 378

```

10 (2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 993 base pairs
 (B) TYPE: nucleic acid
 15 (C) STRANDEDNESS: double
 (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

20 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

25 (A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
 (B) LOCATION 1...993

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

```

TTGAAAAAAA TATTACCGGC TTTGTTAATG GGGTTTGTGG GATTGAATGC TAGTGATCGT 60
TTGTTAGAAA TCATGCGCCT TTATCAAAAA CAAGGCTTGG AAGTGGTGGG TCAAAAATTG 120
35 GATTCTTATT TAGCGGATAA GTCTTTTGG GCAGAAGAGC TTCAAAACAA GGACACGGAT 180
TTTGGCTATT ATCAAAACAA GCAGTTTTTA TTTGTGGCGG ATAAATCCAA GCCCAGTTTG 240
GAGTTTTATG AAATAGAAAA TAACATGCTT AAAAAATCA ACAGCTCTAA AGCCCTTGTA 300
GGCTCTAAAA AGGGCGATAA AACTTTAGAG GGCGATTGG CCACGCCTAT TGGAGTGTAT 360
CGTATCACGC AGAAATTAGA GCGTTTGGAT CAATATTATG GCGTTTTGGC TTTTGTAACG 420
40 AATTACCCTA ATTTGTATGA CACTTTGAAA AAACGCACCG GGCATGGCAT TTGGGTGCAT 480
GGAATGCCTT TAAATGGCGA TAGGAATGAA TTGAACACTA AGGGTTGCAT TGCGATTGAA 540
AACCCTATTC TAAGCTCTTA TGACAAAGTG TTTAAAGGCG AAAAAGCGTT CCTTATCACT 600
TATGAAGACA AGTTTTCCCC TAGCACTAAA GAAGAATTGA GCATGATTTT AAGCTCCCTT 660
TTCCAATGGA AAGAAGCTTG GGCTAGGGGC GATTTTGAAC GCTACATGCG TTTTATAAC 720
45 CCCAATTTCA CTCGCTATGA CGGCATGAGT TTTAACGCTT TTAAGAGTA TAAAAAAGG 780
GTGTTTGCAA AAAATGAAAA AAAGAATATC GCTTTTTTCTT TATCAATGT GATCCCTTAC 840
CCCAACTCTC AAAACAAACG CTTGTTTAT GTGGTATTGG ACCAAGATTA CAAAGCCTAC 900
CAGCAAAACA AGCTCTCTTA TAGCTCCAAT TCTCAAAAAG AACTCTATGT AGAGATTGAA 960
50 AACAATCAAG CGTCTATTAT AATGGAAAAA TAA 993

```

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 510 base pairs

- 125 -

(B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: circular

5 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

10

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

15

(A) NAME/KEY: misc_feature

(B) LOCATION 1...510

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

```

20 TTGTTTGAGA AATGGATTGG TCTGACCTTA CTCCTTAGTT CCTTAGGCTA TCCATGCCAA    60
   AAGGTAAGTA TTAGTTTCAA GCAATACGAA AATCTTATCC ATATCCATCA AAAAGGTTGC    120
   AACAAATGAAG TGGTGTGCAG AACGCTCATC TCTATCGCTT TACTAGAAAG CTCTCTAGGG    180
   TTGAACAACA AGCGAGAAAA ATCCCTTAAA GACACTTCTT ACTCCATGTT CCATATCACC    240
   TTAAACACCG CTAAAAAGTT CTACCCTACC TATTCTAAAA CGCTCCTCAA AACCAAATTG    300
25 TTAAATGATG TGGGTTTTGC GATCCAATTA GCCAAACAAA TTTTAAAAGA AAATTTTGAT    360
   TATTACCACC AAAAACACCC CAACAAAAGC GTGTATCAAT TAGTACAAAT GGCCATAGGC    420
   GCTTACAATG GGGGAATGAA ACACAACCCT AATGGCGCTT ACATGAAGAA GTTTCGTTGC    480
   ATTTATTCTC AAGTGCATA CAACGAATAA                                510

```

30 (2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 648 base pairs

(B) TYPE: nucleic acid

35

(C) STRANDEDNESS: double

(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

40 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

45

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...648

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

```

ATGAAAAAAC CCTACAGAAA GATTTCTGAT TATGCGATCG TGGGTGGTTT GAGCGCGTTA    60
GTGATGGTAA GCATTGTGGG GTGTAAGAGC AATGCCGATG ACAAACCAAA AGAGCAAAGC    120

```

- 126 -

5 TCTTTAAGTC AAAGCGTTCA AAAAGGCGCG TTTGTGATTT TAGAAGAGCA AAAGGATAAA 180
 TCTTACAAGG TTGTTGAAGA ATACCCCAGC TCAAGAACCC ACATTGTAGT GCGCGATTG 240
 CAAGGCAATG AACGCGTGTT GAGCAATGAA GAGATTCAA AGCTCATCAA AGAAGAAGAA 300
 GCCAAAATTG ATAACGGCAC GAGCAAGCTT GTCCAGCCTA ATAATGGAGG GAGTAATGAA 360
 GGATCAGGCT TTGGCTTGGG AAGCGCGATT TTAGGGAGCG CGGCGGGGGC GATTTTAGGG 420
 AGTTATATTG GCAATAAGCT TTTTAATAAC CCTAATTATC AGCAAAACGC CCAACGGACC 480
 TACAAATCCC CACAAGCTTA CCAACGCTCT CAAAATTCTT TTTCTAAAAG CGCACCCAGC 540
 GCTTCAAGCA TGGGCACAGC GAGTAAGGGA CAGAGCGGGT TTTTGGCTC TAGTAGGCCT 600
 ACTAGTTCGC CTGCAATAAG CTCTGGGACA AGGGGCTTTA ACGCATAA 648

10

(2) INFORMATION FOR SEQ ID NO:23:

15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 762 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: circular
 20 (ii) MOLECULE TYPE: DNA (genomic)
 (iii) HYPOTHETICAL: NO
 (iv) ANTI-SENSE: NO
 25 (vi) ORIGINAL SOURCE:
 (A) ORGANISM: *Helicobacter pylori*
 (ix) FEATURE:
 (A) NAME/KEY: misc_feature
 30 (B) LOCATION 1...762

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

35 TTGAAAAC TC TATTTAGTGT TTATCTCTTT TTGTCGTTGA ATCCACTCTT TTTAGAAGCT 60
 AAAGAAATCA CTTGGTCTCA ATTCTTGGA AATTTTAAAA ACAAGAATGA AGACGACAAA 120
 CCTAAACCCC TAACCATTGA CAAAAACAAT GAAAAACAGC AAATCCTAGA CAAAAACCAG 180
 CAAATCTTAA AAAGGGCTTT AGAAAAAAGC CTTAAATTTT TCTTTATTTT TGGATACAAC 240
 TATTCGCAAG CCGCTTATTC AACCCTAAT CAAAACCTGA CTCTTACGGC GAATAGCATA 300
 GGGTTTAACA CCGCTACAGG CTTGGAGCAT TTTTAAAGAA ACCACCCTAA AGTCGGTTTT 360
 40 AGAATCTTTA GCGTCTATAA CTATTTCCAT TCCGTTTCGC TCTCCAGCC TCAAATCCTA 420
 ATGGTGCAAA ATTACGGAGG CGCGTTAGAT TTTCTTGGA TTTTGTGGA TAAAAAACC 480
 TATCGCTTTA GGAGTTATTT AGGAATCGCT TTAGAGCAAG GGGTGTGTT AGTGATACG 540
 ATTAAAACCG GCTCTTTTAC AACCATCATC CCAAGAACCA AGAAAACCTT TTTTCAAGCC 600
 CCTTGCGTT TTGGTTTTAT CGTGGATTTT ATCGGCTATT TGTCTTTGCA ATTAGGGATT 660
 45 GAAATGCCCT TAGTGAGGAA TGTTTTTTAC ACCTACAATA ACCATCAAGA AAGATTCAA 720
 CCACGATTTA ACGCTAAAT TTCTTTAATC GTTTCGTTTT AG 762

(2) INFORMATION FOR SEQ ID NO:24:

50 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1011 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: circular

- 127 -

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...1011

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

TTGTTTTTCA	AATTTATTTT	ATGTTTATCA	TTAGGAATAT	TTGCATGGGC	AAAAGAGGTC	60
ATTCCCACCC	CTTCAACCCC	ATTAACGCCC	TCTAAACGCT	ATTCTATCAA	TTTGATGACT	120
GAAAATGATG	GTTATATCAA	TCCTTACATT	GATGAGTATT	ACACGGCAGG	CAATCAAATA	180
GGCTTTTCTA	CTAAAGAGTT	TGATTTTCT	AAAAATAAAG	CGATGAAATG	GTCTTCGTAT	240
TTAGGGTTTT	TCAATAAAAG	CCCTAGGGTT	ACTCGTTTTG	GCATTTCTCT	CGCCCAAGAC	300
ATGTATACCC	CCTCACTTGC	AAACAGAAAA	CTGGTGCATT	TGCATGACAA	CCACCCCTTAT	360
GGGGGGTATT	TGAGGGTGAA	TTTGAACGTG	TATAACCGCC	ATCAAACCTT	CATGGAGTTA	420
TTCACGATTT	CTTTAGGCAC	GACAGGGCAA	GATTCCTTGG	CCGCTCAAAC	GCAGCGTCTC	480
ATTCATAAAT	GGGGTCATGA	TCCCCAATTT	TATGGCTGGA	ACACGCAGCT	CAAAAACGAA	540
TTTATCTTTG	AACTGCACTA	CCAATTGCTT	AAAAAAGTCC	CCCTTTTAAA	GACTCGTTTT	600
TTTTCTATGG	AGTTGATGCC	TGGGTTTAAT	GTGGAACCTG	GTAATGCGAG	GGATTATTTC	660
CAACTCGGCT	CGCTCTTTAG	GGCTGGGTAT	AACTTGGACG	CTGATTATGG	GGTCAATAAG	720
GTCAATACCG	CTTTTGATGG	GGGCATGCCT	TATAGCGATA	AGTTTTCCAT	CTATTTTTTT	780
GCAGGGGGCT	TTGGGCGCTT	CCAACCCCTT	AACATCTTCA	TTCAAGGCAA	TAGCCCTGAA	840
ACTAGGGGCA	TTGCCAATTT	GGAATACTTT	GTTTATGCCA	GTGAAATAGG	AGCGGCTATG	900
ATGTGGCGTA	GCCTCAGGGT	GGCTTTTACA	ATCACTGATA	TTAGTAAAC	CTTTCAGTCC	960
CAGCCTAAGC	ACCATCAGAT	CGGCACCTTA	GAATTGAATT	TCGCCTTTTG	A	1011

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 327 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...327

- 128 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

5 ATGAAACCAA TCTTTAGCCT CTTTTTCCTC CTTATTGTTT TAAAAGCGCA CCCCATAAAC 60
 CCCTTATTAG AGCCGTTATA TTCCCCCAGT TACACGCAAT TTTTAGATTT AGAACCTCAT 120
 TTTGTCATTA AAAAAAAGCG CGCTTACAGG CCTTTTCAAT GGGGGAACAC TATTATTATC 180
 AAACGCCATG ATTTAGAAGA GCGCCAGAGC AACCAACCAA GCGATATTTT CCGCCAGAAC 240
 GCTGAAATCA ATGTGTCTTC TCAAACTTTT TTAAGAGGAA TCAGCAGCGC TTCTTCACGC 300
 ATAGTGATCG ATTCGGTCGC TCAGTAA 327

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

15 (A) LENGTH: 588 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

20

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

25

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

30

(A) NAME/KEY: misc_feature

(B) LOCATION 1...588

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

35 ATGAGCAATA ACCCCTTTAA AAAAGTGGGC ATGATCAGCT CTCAAAACAA TAACGGCGCT 60
 TTGAACGGGC TTGGCGTGCA AGTGGGTTAT AAACAATTCT TTGGCGAAAG CAAAAGATGG 120
 GGGTTAAGGT ATTATGGTTT CTTTGATTAC AACCACGGCT ATATCAAATC CAGCTTTTTT 180
 AATTCTTCTT CTGATATATG GACTTATGGC GGTGGGAGCG ATTTGTTAGT GAATTTTATC 240
 AACGATAGCA TCACAAGAAA GAACAACAAG CTTTCTGTGG GTCTTTTTTG TGGTATCCAA 300
 CTAGCAGGGA CTACATGGCT TAATTCTCAA TACATGAATT TAACAGCGTT CAATAACCCT 360
 40 TACAGCGCGA AAGTCAATGC TTCCAATTC CAATTTTTGT TCAATCTCGG CTTGAGGACG 420
 AATCTCGCTA CAGCTAAGAA AAAAGACAGC GAACGTTCCG CGCAACATGG CGTTGAACTG 480
 GGCATTAAAA TCCCTACCAT TAACACCAAT TATTATTCTT TTCTAGGCAC TAAGCTAGAA 540
 TACAGAAGGC TTTATAGCGT GTATCTCAAT TATGTGTTTG CTTATTAA 588

45

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

50 (A) LENGTH: 684 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

- 129 -

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

5

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

10

(A) NAME/KEY: misc_feature

(B) LOCATION 1...684

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

15	GTGCGTTTTG	GTAAAATTGA	TTATTTGAAC	ATGCTCCCTT	TTGATGTGTT	TATCAAATCC	60
	TACCCACCC	CTTGTTATTT	CAAACAATTC	TTACGGCTTA	AAAAACCTA	CCCCCCTCAA	120
	CTCAATGAGA	GTTTTTTATT	CAGGCGCATT	GATGCGGGGT	TTATTTCTTC	TATCGCTGGC	180
	TATCCATTCG	CTCTTTGTTC	TTATTCTCTA	GGCATTGTCG	CTTATAAGGA	AGTTTAAAGC	240
	GTGTTGGTTG	TAAATAGAGA	AAACGCTTTT	GACAAAGAAA	GCGCTTCTTC	AAACGCCCTC	300
	TCTAAAGTGT	TAGGGTTAAA	AGGCGAGGTC	TTAATCGGCA	ATAAAGCGCT	GCAATTTTAT	360
20	TATCCAACCC	CTAAAAAAGA	TTTTATAGAT	TTAGCCGCTC	TGTGGTATGA	AAAAAACCGC	420
	TTGCCGTTTG	TTTTTGGGCG	TCTGTGCTAT	TATCAAAACA	AGGATTTTTA	CAAACGCTTG	480
	TCTTTAGCCT	TCAAACATCA	AAAAACAAAA	ATCCCTCACT	ACATCCTTAA	AGAAGCCGCT	540
	TTGAAAACCA	ACTTGAAACG	CCAAGATATT	CTAAACTACT	TGCAAAAAAT	TTACTACACT	600
	TTAGGCAAAA	AGGAACAATC	AGGCCTTAAA	GCGTTCTATC	GTGAATTGTT	GTTCAAACGC	660
25	ATCCAAAAAC	CCAAGCGGTT	TTAG				684

(2) INFORMATION FOR SEQ ID NO:28:

30

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 918 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: circular

35

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

40

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

45

(A) NAME/KEY: misc_feature

(B) LOCATION 1...918

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

50	ATGGGTAGAA	TTGAATCAAA	AAAGCGTTTG	AAAGCACTCA	TTTTTTTAGC	GAGTTTGGGG	60
	GTGTTGTGGG	GCAATGCGGC	TGAAAAAACG	CCTTTTTTTA	AAACTAAAAA	CCACATTTAT	120
	TTGGGTTTTA	GGCTAGGCAC	AGGGGCTACT	ACGCGCACAA	GCATGTGGCA	ACAAGCCTAT	180
	AAAGACAACC	CCACTTGCCC	TAGCAGCGTG	TGTTATGGCG	AGAAATTAGA	AGCCCATTAT	240
	AAGGGGGGTA	AAACTTATC	TTATACCGGG	CAAATAGGCG	ATGAAATAGC	TTTTGATAAA	300

- 130 -

	TACCATATTT	TAGGCTTAAG	GGTGTGGGGG	GATGTAGAAT	ACGCTAAGGC	TCAATTAGGT	360
	CAAAAAGTGG	GGGGTAACAC	CCTTTTATCC	CAAGCGAATT	ATAACCCAAG	CGCGATTAAA	420
	ACCTACGATC	CTACTTCAAA	CGCTCAAGGC	TCTTTAGTTT	TGCAAAAAAC	CCCAAGCCCC	480
	CAAGATTTC	TTTTCAATAA	CGGGCATTTC	ATGGCGTTTG	GTTTGAACGT	GAACATGTTT	540
5	GTCAATCTCC	CTATAGACAC	CCTTTTAAAA	CTCGCTTTAA	AAACGGAAAA	AATGCTGTTT	600
	TTTAAAATAG	GCGTGTTTGG	TGGGGGTGGG	GTGGAATACG	CAATCTTGTG	GAGTCCTCAA	660
	TATAAAAATC	AAAATACCCA	TCAAGACGAT	AAATTTTTTG	CCGCAGGTGG	GGGGTTTTTT	720
	GTGAATTTTG	GAGGCTCTTT	GTATATAGGC	AAGCGCAACC	GCTTCAATGT	GGGGCTAAAA	780
	ATCCCTTATT	ATAGCTTGAG	CGCGCAAAGT	TGGAAAAATT	TTGGCTCTAG	CAATGTGTGG	840
10	CAGCAACAAA	CGATCCGACA	AACTTCAGC	GTTTTTAGGA	ATAAAGAAGT	TTTTGTCAGC	900
	TACGCGTTCT	TGTTTTAG					918

(2) INFORMATION FOR SEQ ID NO:29:

- 15 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 777 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: circular
- 20 (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- 25 (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: *Helicobacter pylori*
- 30 (ix) FEATURE:
- (A) NAME/KEY: misc_feature
 - (B) LOCATION 1...777

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

35	ATGTTTTTTAA	GATCATACCC	AAAGCTTAGA	TACGCTTTAT	GTTTACCCCT	ACTCACTGAG	60
	ACTTGCTATA	GCGAGGAGCG	CACTTTAAAT	AAGGTTACCA	CCCAAGCTAA	AAGGATTTTC	120
	ACTTACAATA	ATGAGTTTAA	GGTTACTTCT	AAAGAATTGG	ATCAACGCCA	AAGCAATGAA	180
	GTCAAAGACC	TGTTTTAGGAC	TAACCCTGAT	GTGAATGTGG	GCGGAGGGAG	CGTGATGGGG	240
40	CAGAAAATCT	ACGTGAGAGG	CATTGAAGAC	AGGCTTTTAA	GGGTTACGGT	GGATGGGGCT	300
	GCGCAAAATG	GCAACATCTA	CCACCACCAA	GGCAACACCG	TGATTGACCC	TGGCATGCTC	360
	AAAAGCGTGG	AAGTTACTAA	AGGCGCGGCG	AATGCGAGCG	CGGGGCCAGG	AGCGATCGCG	420
	GGAGTGATTA	AAATGGAGAC	TAAAGGAGCG	GCTGATTTTA	TCCCTAGGGG	GAAAAATTAT	480
	GCAGCGAGTG	GGGCGGTGAG	TTTTTATACC	AATTTTGGGG	ACAGAGAGAC	TTTTAGATCG	540
45	GCCTATCAAA	GCGCGCATTT	TGATATTATC	GCTTACTACA	CGCACCAAAA	TATTTTCTAT	600
	TATAGGAGCG	JC JCCACAGT	GATGAAAAAC	CTTTTCAAAC	CCACACAAGC	CGATAAAGAG	660
	CCAGGAACTC	CCAGCGAGCA	AAACAACGCT	TTGATTAAAA	TGAATGGCTA	TTTGAGCGAC	720
	AGAGACACGC	TCACTTTCAG	CTGGAACATG	ACACGAGATA	ACGCCACACG	CCTTTAA	777

50 (2) INFORMATION FOR SEQ ID NO:30:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 579 base pairs
 - (B) TYPE: nucleic acid

- 131 -

(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

5 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

10 (vi) ORIGINAL SOURCE:

(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:

15 (A) NAME/KEY: misc_feature

(B) LOCATION 1...579

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

20	ATGTTTTTTAA GATCATACCC AAAGCTTAGA TACGCTTTAT GTTTACCCCT ACTCACTGAG	60
	ACTTGCTATA GCGAGGAGCG CACTTTAAAT AAGGTTACCA CCAAGCTAA AAGGATTTTC	120
	ACTTACAATA ATGAGTTTAA GGTTACTTCT AAAGAATTGG ATCAACGCCA AAGCAATGAA	180
	GTCAAAGACC TGTTTAGGAC TAACCCTGAT GTGAATGTGG GCGGAGGGAG CGTGATGGGG	240
	CAGAAAATCT ACGTGAGAGG CATTGAAGAC AGGCTTTTAA GGGTTACGGT GGATGGGGCT	300
	GCGCAAAATG GCAACATTTA CCACCACCAA GGCAACACCG TGATTGACCC TGGCATGCTC	360
25	AAAAGCGTGG AAGTTACTAA AGGCGCGGCG AATGCGAGCG CGGGGCCAGG AGCGATCGCG	420
	GGAGTGATTA AAATGGAGAC TAAAGGAGCG GCTGATTTTA TCCCTAGGGG GAAAAATTAT	480
	GCAGCGAGTG GGGCGGTGAG TTTTATAACC AATTTTGGGG ACAGAGAGAC TTTTAGATCG	540
	GCCTATCAAA GCGCGCATTT TGATATTATC GCTTACTAG	579

30 (2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 381 base pairs

(B) TYPE: nucleic acid

35 (C) STRANDEDNESS: double

(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

40 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

45 (A) ORGANISM: Helicobacter pylori

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...381

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

GTGCCCTTGA GTTTGGGAGG CAACCTCTTA AACCTAACA ACAGTAGCGT GCTGAATTTA	60
AAAAACAGCC AGCTTGTTTT TAGCGATCAA GGGAGCTTGA ATATCGCTAA CATTGATTTA	120

- 132 -

CTAAGCGATC TGAATGGTAA TAAAAATCGT GTGTATAACA TCATTCAAGC GGACATGAAT 180
 GGTAATTGGT ATGAGCGTAT CAACTTCTTT GGCATGCGCA TTAATGATGG GATTTATGAC 240
 GCTAAAAACC AAACCTTATAG TTTCACTAAC CCTCTCAATA ACGCCGTAAA ATTCACCGAG 300
 AGCTTTTTTCA TACACCGCCT GTGCGGTTTCG CTCTCTCAAA TACAAAAAAA AAAAAACACA 360
 5 ATAGTCTCAC CTCGGCTCTG A 381

(2) INFORMATION FOR SEQ ID NO:32:

- (i) SEQUENCE CHARACTERISTICS:
 10 (A) LENGTH: 1698 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
 15
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
 20
- (vi) ORIGINAL SOURCE:
 (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
 25 (A) NAME/KEY: misc_feature
 (B) LOCATION 1...1698

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

30 GTGTATTCTT ATAGCGATGA CGCACAAAGG GTGTTTTATC TCACGAGCAG CGTGAAAGGC 60
 TATTACAACC CCAACCAATC CTATCAAGCC AGCGGCAGCA ATAACACCAC GAAAAATAAC 120
 AATCTAACCT CTGAATCTTC TGTCATTTTCG CAAACCTATA ACGCGCAAGG CAACCCTATC 180
 AGCGCGTTAC ACGTCTATAA CAAGGGCTAT AATTTCAAGTA ATATCAAAGC GTTAGGGCAA 240
 ATGGCGCTCA AACTCTACCC TGAAATCAAA AAGATATTAG GGAATGATTT TTCGCTTTCA 300
 35 AGTTTGAGCA ATTTAAAGG CGATGCGCTA AACCAGCTTA CCAAGCTCAT CACGCCTAGC 360
 GATTGGAAAA ACATTAACGA GTTGATTGAT AACGCAACA ATTCGGTCGT GCAAAATTTT 420
 AATAACGGCA CTTTGATTAT AGGAGCGACT AAAATAGGGC AAACAGACAC CAATAGTGCG 480
 GTGGTTTTTG GGGGCTTGGG CTATCAAAAG CCTTGCAGAT ACACTGATAT TGTGTGCCAA 540
 AAATTTAGAG GCACTTATTT GGGGCAGCTT TTGGAGTCCA ACTCCGCTGA TTTGGGCTAT 600
 40 ATTGACACGA CTTTTAACGC TAAAGAAATT TATCTTACCG GCACTTTAGG GAGCGGGAAC 660
 GCATGGGGGA CTGGGGGGGAG TCGGAGCGTA ACTTTTAACA GCCAACTTC GCTCATTCCTC 720
 AACCAAGCGA ATATCGTAAG CTCGCAAACC GATGGGATTT TTAGCATGCT GGGTCAAGAG 780
 GGCATCAATA AGGTTTTCAA TCAAGCCGGG CTCGCTAATA TTTTGGGCGA AGTGGCAATG 840
 CAATCCATTA ACAAAGCCGG GGGATTAGGG AATTTGATAG TAAATACGCT AGGGAGTGAT 900
 45 AGCGTGATTG GGGGGTATTT AACGCCTGAG CAAAAAATC AAACCCTAAG CCAGCTTTTG 960
 GGGCAGAATA ATTTTGATAA CCTCATGAAC GATAGCGGTT TGAACACG C GATTAAGGAT 1020
 TTGATCAGAC AAAAATTAGG CTTTGGACC GGGCTAGTGG GGGGATTAGC CGGACTGGGG 1080
 GGCATTGATT TGCAAAACCC TGAAAAGCTT ATAGGCAGCA TGTCCATCAA TGATTTATTG 1140
 AGTAAAAAGG GGTGTGTTCAA TCAGATCACC GGCTTTATTT CCGCTAACGA TATAGGGCAA 1200
 50 GTCATAAGCG TGATGCTGCA AGATATTGTC AAGCCGAGCG ACGCTTTAAA AAACGATGTA 1260
 GCCGCTTTGG GCAAGCAAAT GATTGGCGAA TTTTTAGGCC AAGACACGCT CAATTCTTTA 1320
 GAAAGCTTGC TGCAAAACCA GCAGATTAAA AGCGTTTTAG ACAAAGTCTT AGCGGCTAAA 1380
 GGATTAGGGT CTATTTATGA ACAAGGTTTG GGGGATTTGA TCCCTAATCT TGGTAAAAAG 1440
 GGGATTTTCG CTCCCTATGG CTTGAGTCAA GTGTGGCAAA AAGGGGATTT TAGTTTCAAC 1500

- 133 -

GCGCAAGGCA ATGTTTTTGT GCAAAATTQC ACTTTCTCTA ACGCTAATGG AGGCACGCTC 1560
 AGTTTTTAACG CAGGAAATTC GCTCATTTTT GCCGGAAACA ACCACATCGC TTCTACTAAC 1620
 CATTCTGGAA CGCTCAATTT GTTGTCTAAT CAAGTTTCTA ACATTAACGT CACCATGCTT 1680
 AACGCAGCAA CGGCCTAA 1698

5

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 519 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: circular

15 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

20 (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

- 25 (A) NAME/KEY: misc_feature
 (B) LOCATION 1...519

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

30 GTGTTTGGAT TGAGTTTGGC GGATATGATT TTAGAGCGTT TTAAAGATTT TATGAGAGAA 60
 TACCTTGAGC CTTACAAGTT TTTACAGGTT TTTTACGCGC AAGAAAAAGA ACGCTTCTTA 120
 AATCATAAAA TGAACGATTA TATCAAGCAA AATAAGAGCA AGGAAGAGGC TAGTATTTTG 180
 GCCAGACAAG GCTTTGTCAG CGTAATTGGA AGAGCGTTAG AAAAAATCAT AGAACTTTTA 240
 TTAAAAGATT TTTGTATTAA AAACAATGTA AAAATGACGA ACGATAAAAC CTTAAGGGCT 300
 AAGCGCATTA ATGGCGAATT AGATAAGGTC AAACGGGCTT TATTGGTGCA TTTTGGAGGA 360
 35 TATAGCGTTT TACCCGATAT TATTCTTTAT CAAACCAACA AAGATAATAT CAAAATCCTA 420
 GCGATTTTAT CGGTAAAAAA TTCGTTTAGA GAGCGTTTCA CAAAAGACGC CTTATTGGAA 480
 ATTAATAACTT TTGCAATCGC CTGTAACCTC TCACATTAA 519

40

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:

- 45 (A) LENGTH: 996 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

50

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

- 134 -

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...996

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

10 ATGAAAAGAT TTGTTTATT CTTGTTATTC ATATGTGTTT GCGTTTTCGT TCAAGCTTAC 60
 GCTGAGCAAG ATTACTTTTT TAGGGATTTT AAATCTATAG ATTTGCCCCA AAAACTCCAC 120
 CTTGATAAAA AGCTCTCCCA AACAATACAG CCATGCGCGC AACTTAACGC ATCAAAACAC 180
 TACACTGCTA CTGGGGTTAG AGAGCCTGAT GCCTGCACCA AGAGTTTAA AAAATCCGCT 240
 ATGGTTTCCT ATGATTTAGC GCTAGGCTAT TTAGTGAGCC AAAACAAACC ATACGGCTTA 300
 AAAGCTATAG AGATTTTAA CGCTTGGGCT AATGAGCTTC AAAGCGTGGA TACTTATCAA 360
 AGCGAGGACA ATATCAATTT TTACATGCCT TATATGAACA TGGCTTATTG GTTTGTCAA 420
 15 AAAGAATTTT CTAGCCCAGA ATATGAAGAT TTCATTAGGC GGATGCGTCA GTATTCTCAA 480
 TCAGCTCTTA AACTAACCA TGGGGCGTGG GGGATTCTCT TTGATGTGAG CTCTGCACTA 540
 GCGCTAGATG ATCATGCCCT TTTGCAAAGT AGCGCTAATC GGTGGCAGGA GTGGGTGTTT 600
 AAAGCCATAG ATGAGAACGG GGTATTGCT AGCGCGATCA CTAGGAGCGA TACGAGCGAT 660
 TATCATGGCG GCCCTACAA GGGCATTAA GGGATAGCTT ATACCAATTT TCGCTTCTT 720
 20 GCGATAACTA TATCAGGCGA ATTGCTTTTT GAGAACGGGT ATGATTGTG GGGTAGTGGA 780
 GCCGGGCAAA GGCTCTCTGT GCGGTATAAC AAAGCCGCAA CATGGATTCT AAACCCTGAA 840
 ACTTTCCCTT ATTTTCAGCC TAACCTCATT GGGGTGCATA ACAACGCTTA TTTCATTATT 900
 TTAGCCAAAC ATTATTCTAG CCCTAGCGCG GATGAGCTTT TAGAGCAAGG CGATTTCAT 960
 GAAGATGGCT TCAGGCTGAA ACTCCGATCG CCATGA 996

25

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:

30

(A) LENGTH: 384 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

35

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

40

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:

45

(A) NAME/KEY: misc_feature

(B) LOCATION 1...384

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

50 ATGCGTCAGT ATTCTCAATC AGCTCTTAAC ACTAACCATG GGGCGTGGGG GATTCTCTTT 60
 GATGTGAGCT CTGCACTAGC GCTAGATGAT CATGCCCTTT TGCAAAGTAG CGCTAATCGG 120
 TGGCAGGAGT GGGTGTTTAA AGCCATAGAT GAGAACGGGG TTATTGCTAG CGCGATCACT 180
 AGGAGCGATA CGAGCGATTA TCATGGCGGC CCTACAAAGG GCATTAAGGG GATAGCTTAT 240
 ACCAATTTTG CGCTTCTTGC GATAACTATA TCAGGCGAAT TGCTTTTGA GAACGGGTAT 300
 GATTTGTGGG GTAGTGAGC CGGGCAAAGG CTCTCTGTGG CGTATAACAA AGCCGCAACA 360

- 135 -

TGGATTCTAA ACCCTGAAAC TTTC

384

(2) INFORMATION FOR SEQ ID NO:36:

5 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 738 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

10

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

15

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

20

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION 1...738

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

TTGAGAACCT	TGTTAAAAAT	GTTGGTTGGT	GTGAGCTTAC	TAACACACGC	TTTAATGGCT	60
ACAGAAGAAA	GCGCTGCCCC	TTCTTGGACA	AAAAATTTGT	ATATGGGATT	CAATTACCAA	120
ACAGGTTCTA	TCAATTTAAT	GACTAATATT	CATGAAGTTA	GAGAAGTTAC	TAGCTATCAA	180
ACCGGTTACA	CCAATGTAAT	GACTAGCATT	AATAGCGTTA	AAAAACTCAC	TAACATGGGT	240
30 TCTAATGGGA	TTGGCTTAGT	CATGGGCTAT	AACCACTTTT	TCCATCCGGA	TAAAGTCTTG	300
GGTTTGCGCT	ATTTTGCTTT	TTTAGATTGG	CAAGGCTATG	GCATGAGATA	CCCTAAAGGC	360
TATTATGGGG	GCAATAACAT	GATCACTTAT	GGCGTGGGCG	TGGATGCGAT	ATGGAATTTT	420
TTCCAAGGGA	GTTTTTATCA	AGATGATATT	GGCGTGGATA	TTGGCGTTTT	TGGGGGGATT	480
GCGATTGCTG	GGAATAGCTG	GTATATTGGC	AATAAAGGGC	AGGAATTATT	AGGCATCACC	540
35 AATAGTAGTG	CGGTTGATAA	CACCTCTTTT	CAATTCCTCT	TTAACTTTGG	TTTCAAAGCT	600
TTATTTGTAG	ATGAACATGA	ATTTGAAATT	GGGTTTAAAT	TCCCCACTCT	TAACAACAAA	660
TACTACACCA	CCGACGCGCT	CAAGGTTCAA	ATGCGTAGGG	TCTTTGCCTT	TTATGTGGGG	720
TATAATTACC	ACTTCTAA					738

40

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 873 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

45

(ii) MOLECULE TYPE: DNA (genomic)

50

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- 136 -

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc_feature

5 (B) LOCATION 1...873

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

```

10 ATGTTTGAAG AAATTACCTT AGCGCATAAG GACTTGTTTT CAAGGTTTTT ACAAACTCAA      60
   AAAATCGTTT TATCGGATGT GAGTTTACC AATTGCTTTT TATGGCAGCA CGCAAGGCTC      120
   ATTCAAGTGG CTGTGATTAG GGATTGTTG GTGATTCAAA CCACTTATGA AAACCAAAAA      180
   CCCTTTTATT TCTATCTAT CCGTAAGAGG CCGCATGAAT GCGTGAAAGA GCTTTTGGAA      240
   TTAGAAAAAA ATTTAAGATT CCACTCCCTG ACTTTAGAGC AAAAAGACGA TTTGAAAGAC      300
   AATTTTGTAG GGGTGTTTGA TTCACTTAC AACCAGAGACA GGAGCGATTA TGTTTATTCT      360
15 ATTGAAGAAC TAATCGCGCT CAAAGGGAAA AAATACCATA AGAAAAAAA CCACTTAAAC      420
   CAGTTTTTAA CCAATCATGC GAATTTTGTT TATGAAAAAA TTTCTCCTCA AAACAGAAAG      480
   GAAGTTTTAG AAGCCTCTAA AGCGTGTTT TTAGAAAGCC AGACCGATGA TATAGGGTTA      540
   ATCAACGAAA ATAAGGGCAT TCAAAGCGTT TTAGAAAATT ATGAAAGCTT GGATTTAAAG      600
   GGGGGGCTTA TTAGGGTTAA TGGGGAAATA GTCTCGTTA GTTTTGGGGA AGTTTAAAC      660
20 GAAGAGAGCG CGCTCATCCA CATTGAAAAA GCCCGCACAG ATATTGCAGG CGCGTATCAA      720
   ATCATCAACC AACAATTGCT TTTGAATGAA TTAGCCATT TAACTTACGC TAACAGAGAA      780
   GAAGATCTAG GATTAGAGGG CTTAAGAAGG TCTAAAATGA GCTATAACCC GGTGTTTTTG      840
   ATAGACAAAT ACGAAGCGGT TGCTAGAAAT TAA                                     873

```

25 (2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 333 base pairs

30 (B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

35 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

40 (A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc_feature

45 (B) LOCATION 1...333

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

```

50 ATGATGTTCA TTGTAGCGGT TTTGATGCTG GCGTTTTTGA TCTTTGTCCA TGAGTTAGGG      60
   CATTTTATTA TCGCTAGGAT TTGTGGGGTG AAAGTGGAAG TGTTTAGCAT TGGTTTTGGT      120
   AAAAAACTCT GGTTTTTCAT GCTTTTGGC ACGCAATTCG CTCTGTCTTT GATCCCGCTT      180
   GGGGGCTATG TGAAATTAAA GGGCATGGAT AAAGAAGAAA ATGAAGAAAA TAAAATTAAT      240
   CAAGCGAATG ATAGCTACGC CAAAAAGCC CTTTCCAAA GCTATGGATA TTGTTTGGTG      300
   GGGCGTTTTT TAATTTTCTT TTTGCGGTTT TAG                                     333

```

- 137 -

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1056 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
 (B) LOCATION 1...1056

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

25	ATGATGTTCA TTGTAGCGGT TTTGATGCTG GCGTTTTTGA TCTTTGTCCA TGAGTTAGGG	60
	CATTTTCATTA TCGCTAGGAT TTGTGGGGTG AAAGTGGAAG TGTTTAGCAT TGGTTTTGGT	120
	AAAAAACTCT GGTTTTTCAA GCTTTTTGGC ACGCAATTCT CTCTGTCTTT GATCCCGCTT	180
	GGGGGCTATG TGAAATTAAA GGGCATGGAT AAAGAAGAAA ATGAAGAAAA TAAAATTAAT	240
	CAAGCGAATG ATAGCTACGC GCAAAAAAGC CCTTTCCAAA AGCTATGGAT ATTGTTTGGT	300
	GGGGCGTTTT TTAATTTTCT TTTTGC GGTT TTAGTGTATT TTTTCTGGC ATTGAGCGGG	360
30	GAAAAAGTCT TACTGCCCGT CATGGCGGT TTAGAAAAAA ACGCGCTAGA AGCCGGGCTG	420
	TTAAAGGGGG ATAGAATCCT TTCTATCAAC CATCAAAAAA TAGCGAGTTT TAGAGAGATT	480
	AGAGAGATAG TGGCGCGTTC TCAAGGCGAG TTAATTTTAG AAATAGAGCG AAACAATCAG	540
	ATTTTAGAAA AACGACTGAC CCCCAAATC GTGGCGGTGA TAAGCGAGTC TAATGATCCT	600
	AATGAAATCA TCAAGTATAA AATAATAGGC ATTAAACCGG ACATGCAAAA AATGGGCGTT	660
35	GTCCTCTATT CCGTGTTCAT AGCGTTTGAA AAGGCTTTGA GTCGGTTTAA AGAGGGCGTT	720
	GTTTTGATTG TGGATTCTTT AAGGCGTTTG ATTATGGGGA GCGCTTCAGT TAAAGAATTG	780
	AGTGGGGTAA TAGGCATTGT GGGGGCGTTA AGCCATGCCA ATAGCGTGAG CATGCTTTTG	840
	TTGTTTGGGG CGTTTTTATC TATCAATCTA GGGATTTTAA ATTTATTACC CATTCCAGCC	900
	TTAGATGGGG CGCAAATGCT AGGGGTCGTT TTTAAAAATA TTTTTCATAT CGCTTTGCCA	960
40	ACGCCCATAC AAAATGCGTT GTGGCTAGTG GGGGTGGGGT TTTTGGTTTT TGTCATGTTT	1020
	TTAGGGCCTT TTAATGACAT TACTCGTTTG CTATAA	1056

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 103 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

- 138 -

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...303

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

```
ATGCAAAAGA ATTTGGATAG TCTTTTAGAA AATTTAAGGG CTGAAATTGA TCGGTTGGAT 60
AATGAATTGA GCGATCTTTT AGACAAACGC TTAGGAATCG CTTTAAAAAT CGCTCTCATC 120
AAACAAGAAA GCCCCAAGA AAACCCCAT TATTGCCCTA AAAGAGAGCA AGAGATTTTA 180
AAACGACTCA GCCAAAGGGG TTTCAAGCAT TTGAATGGAG AAATCCTTGC AAGTTTTTAT 240
GCAGAGGTTT TTAAGATTTT TAGAAATTTT CAAGAAAACG CCCTAAAAGA GTTAAAAAAA 300
TAA 303
```

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 525 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...525

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

```
GTGAAAATGC GTTTTTTTAG TGGTTTTGGG TTTGTTAATG AAAGCGTTTT GTTTGAAGAG 60
TGGCTTTTAA AAGGGGCTTA TGATGTGTCA GGCTTTTCTA TGGGGGCGAT TAAGGCGATA 120
GAATACGCCT ATAATGAAGT CTTGCAACAA CGGCGCATCC ATTCCTTATT GTTGTTTTCG 180
CCTTGCAATGC TAGCGCATAA GAGTTTGGCG TTCAAACGCT TGCAACTTTT CTTGTTTCAA 240
AAAGATCCGC AAAGCTACAT GGATAACTTT TATAAGGAAG TGGGATTGGA CGCTCAATTG 300
GAGCGTTTTA AAAAAGAGGG TTCTTTAGAA GAATTGGAAT TTTTATTGGA TTACAAGTAT 360
AGTGATTCTA TAATTAGATT TTTATTGGAA AAGGGCGTGA AGATTGAAGT GTTTATCGGT 420
TTAAAAGATA GAATCACTGA CATTCAAGCC CTTTTAGAAT TTTTATGCC CTTAGTTCAA 480
GTGTGGCAGT TTAAGGATTG TAACCATTG TTGCAAAAAT CTAA 525
```

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:

- 139 -

- (A) LENGTH: 1416 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

5

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

10

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

15

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...1416

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

20

ATGAAAAATA	CCAATACAAA	AGAGATAAAG	AATACAAGGA	TGAAAAAAGG	TTATAGTCAA	60
TACCACACGC	TCAAAAAAAGG	GCTTTTAA	ACCGCTCTGC	TTTTTAGCCT	TCCTTTAAGC	120
GTGGCGTTAG	CTGAAGACGA	TGGCTTTTAT	ATGGGAGTGG	GCTATCAAAT	CGGCGGCGCG	180
CAACAAAACA	TCAACAACAA	AGGCAGCACC	CTAAGGAATA	ATGTCATTGA	TGATTTCCGC	240
CAAGTGGGCG	TGGGTATGGC	AGGGGGTAAT	GGGCTTTTAG	CTTTAGCGAC	AAACACGACC	300
ATGGACGCTC	TTTTAGGGAT	AGGCAACCAA	ATTGTCAATA	CTAATACAAC	TGTTGGCAAC	360
AACAACGCAG	AGTTAACCCA	GTTTAAAAAA	ATACTCCCCC	AAATTGAACA	ACGCTTTGAG	420
ACGAATAAAA	ACGCTTATAG	CGTTCAAGCC	TTGCAAGTGT	ATTTGAGTAA	TGTGCTTTAT	480
AACTTGTTTA	ATAATAGTAA	TAATGGTAGC	AATAATGGAG	TCGTTCCCTGA	ATATGTAGGG	540
ATTATAAAAG	TTCTCTATGG	TTCTCAAAAT	GAATTCAGTC	TCTTAGCCAC	GGAGAGTGTG	600
GCGCTTTTAA	ACGCGCTCAC	GAGAGTGAAT	CTGGATAGTA	ATTCGGTGTT	TTTAAAAGGG	660
CTATTAGCCC	AAATGCAGCT	TTTTAATGAC	ACTTCTTCAG	CAAAGCTAGG	TCAGATCGCA	720
GAAAACTTGA	AGAACGGTGG	TGCAGGGGCC	ATGCTTCAAA	AGGATGTGAA	AACCATCTCG	780
GATCGAATCG	CTACTTACCA	AGAGAATCTA	AAACAGCTAG	GAGGGATGTT	AAAGAATTAC	840
GATGAGCCAT	ACCTACCCCA	ATTTGGGCCA	GGCACAAGCT	CTCAGCATGG	GGTTATTAAT	900
GGCTTTGGCA	TTCAAGTGGG	CTATAAGCAA	TTTTTTGGGA	GCAAGAAGAA	TATAGGCTTA	960
CGATATTACG	CTTTCTTTGA	TTATGGCTTT	ACGCAATTGG	GCACTCTTAA	CAGTGCTGTT	1020
AAAGCGAACA	TCTTTACTTA	TGGTGCTGGC	ACGGACTTTT	TATGGAATAT	CTTTAGAAGG	1080
GTTTTTAGCG	ATCAGTCCTT	GAATGTGGGG	GTGTTTGGGG	GCATTCAAAT	AGCGGGTAAC	1140
ACTTGGGATA	GCTCTTTAAG	AGGTCAAATT	GAAAACTCGT	TTAAAGAATA	CCCCACTCCC	1200
ACGAATTTCC	AATTTTGTGTT	TAATTTGGGC	TTAAGGGCTC	ATTTTGCCAG	CACCATGCAC	1260
CGCCGGTTTT	TGAGCGCGTC	TCAAAGCATT	CAGCATGGTA	TGGAATTTGG	CGTGAAAATC	1320
CCAGCTATCA	ATCAAAGGTA	TTTGAAAGCG	AATGGGGCTG	ATGTGGATTA	CAGGCGTTTG	1380
TATGCGTTCT	ATATCAATTA	CACGATAGGT	TTTTAA			1416

45

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 390 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

50

(ii) MOLECULE TYPE: DNA (genomic)

- 140 -

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...390

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

```

15  ATGAAAAGCA TCAGAAGAGG CGATGGGCTG AATGTTGTCC CTTTCATTGA TATTATGCTC      60
    GTCTTACTAG CGATTGTGTT GAGTATTTCT ACTTTTATCG CGCAAGGTAA GATTAAAGTC      120
    AGTCTCCCTA ACGCTAAAAA TCGGGA AAAA TCCCAGCCAA ACGATCAAAA AGTGGTGGTC      180
    ATCTCTGTGG ATGAGCATGA CAATATTTTC GTAGATGACA AACCGACGAA TTTAGAAGCT      240
    TTGAGCGCTG TAGTCAAGCA AACAGACCCT AAAACCCTTA TAGATTTAAA AAGCGACAAG      300
20  AGCTCTCGTT TTGAAACTTT TATCAGCATT ATGGATATTT TAAAAGAGCA TAATCATGAA      360
    AATTTCTCCA TCTCCACGCA AGCTCAGTAA                                     390

```

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 225 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...225

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

```

45  ATGCTCGTCT TACTAGCGAT TGTGTTGAGT ATTTCTTCTT TTATCGCGCA AGGTAAGATT      60
    AAAGTCAGTC TCCCTAACGC TAAAAATGCG GAAAAATCCC GACCAAACGA TCAAAAAGTG      120
    GTGGTCATCT CTGTGGATGA GCATGACAAT ATTTTCGTAG ATGACAAACC GACGAATTTA      180
    GAAGCTTTGA GCGCTGTAGT CAAGCAAACA GACCCTAAAA CCCTT                                     225

```

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 672 base pairs

- 141 -

(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

5 (ii) MOLECULE TYPE: DNA (genomic)
(iii) HYPOTHETICAL: NO
(iv) ANTI-SENSE: NO
10 (vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori
(ix) FEATURE:
15 (A) NAME/KEY: misc_feature
(B) LOCATION 1...672

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

20	ATGTTTTTCAC TTTCTTATGT TTCCAAGAAA TTTTAAAGCG TGTTGCTATT GATTTCGCTG	60
	TTTTTAAGCG CTTGCAAATC CAACAATAAA GACAAATTGG ATGAAAATCT TTAAAGCTCC	120
	GGCACTCAAA GCTCCAAAGA ATTAAACGAC AAGCGAGACA ATATAGACAA AAAGAGCTAC	180
	GCTGGTTTAG AAGATGTTTT TTTAGACAAC AAGTCCATTA GCCCTAATGA TAAATACATG	240
	CTTTTAGTTT TTGGCCGTAA TGGTTGCTCC TATTGTGAAA GGCTTAAAAA AGATCTCAAA	300
25	AATGTCAAAG AATTGCGCAA CTATATTAAA GAGCATTTTA GTGCTTACTA TGTCAATATC	360
	AGCTATTCTA AAGAGCATAA TTTTAAAGTC GCGGATAAGG ATAAAAATGA TGAAAAAGAA	420
	ATCAAAATGT CCACAGAAGA ATTAGCGCAA ATTTATGCCG TCCAATCCAC CCCTACGATT	480
	GTTTTATCCG ATAAAACCGG CAAAACCATC TATGAATTGC CGGGCTATAT GCCTTCTGTG	540
	CAATTTTITAG CCGTGTTAGA ATTTATCGGC GATGGGAAGT ATCAAGACAC GAAAAACGAT	600
30	GAGGATCTCA CTAATAAATT AAAGGCTTAC ATCAAGTATA AAACCAACCT TTCTAAGAGC	660
	AAGTCCAGCT AG	672

(2) INFORMATION FOR SEQ ID NO:46:

35 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 351 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
40 (D) TOPOLOGY: circular
(ii) MOLECULE TYPE: DNA (genomic)
(iii) HYPOTHETICAL: NO
45 (iv) ANTI-SENSE: NO
(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori
50 (ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...351
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

- 142 -

TTGATGAAAT CTAAAATCAC TCATTTTATC GTTATCTCTT TTGTTTAAAG CGTGTGAGC 60
 GCCTGCAAAG ATGAGCCTAA AAAATCGTCC CAATCGCACC AAAACAACAC TAAAACCACT 120
 CAAAACAATC AAATCAATCA ACCTAATAAG GATATAAAAA AGATTGAGCA TGAAGAAGAA 180
 5 GATGAAAAAG TCACCAAAGA AGTGAATGAT CTGATCAATA ACGAAAATAA AATTGATGAA 240
 ATCAATAATG AAGAAAACGC TGATCCTTCG CAAAAAAGAA CGAACAATGT TTTGCAACGA 300
 GCCACTAACC ACCAAGACAA TCTCAGTTCC CCACTCAACA GGAAGTATTA A 351

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 240 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- 15 (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION 1...240

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

ATGTTTGAAA AAATACGCAA GATTTTAGCG GATATTGAAG ATTTCGCAAAA TGAAATTGAA 60
 ATGCTTTTAA AATTAGCGAA TTTGAGTTTG GGGGATTTTA TTGAGATTAA AAGAGGGGAGC 120
 ATGGACATGC CAAAGGGCGT GAATGAAGCG TTTTTTACGC AATTAAGCGA AGAAGTGGAG 180
 35 CGCCTAAAGG AGCTTATCAA CGCTTTGAAT AAAATCAAAA AAGGGTTATT GGTGTTTTAA 240

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 156 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- 40 (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

- 143 -

(A) NAME/KEY: misc_feature
(B) LOCATION 1...156

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

5

```

ATGTCTATGT TCATTTCTAA TCTGGCTTTC ACGAGCGAAC ATAAGGACGC TATGGAAGTG      60
GCAAAAATTG CGATTTTACT CGGATCTTTG ATTTCTGGGA TCATAGGGGC TTTATATTTA      120
TTCGCACTAG ATAAAAGAGC GGCTTTAAAG AAATAG                                156

```

10 (2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1350 base pairs
(B) TYPE: nucleic acid
15 (C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

20 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

25 (A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc_feature
30 (B) LOCATION 1...1350

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

```

ATGGGTTTGA AAATAAAAAT TTTAAGGTTG TCTATGAATC TCAAAAAAAC AGAAAACGCG      60
CTCAGTTTGA CGCTTAAAAA CTTCAATTAA AGCGAGTCTT TTGGAGGGAT TTCCTCTTT      120
35 TTGAACGCCG TTTTAGCGAT GGTGGTGGCT AATTCGTTTT TAAAAGAAAG TTATTTTGCG      180
CTATGGCACA CCCCTTTTGG GTTCAAGTA GGGGATTTTT TTATCGGCTT TAGTTTGCAC      240
AACTGGATTG ATGATGTCTT AATGGCGTTA TTCTTTTAA TGATAGGCTT AGAGATCAAG      300
CGAGAATTGT TGTTTGGGGA ATTATCCAGT TTTAAAAAG CTTCTTTCCC TGTGATCGCA      360
GCCATAGGGG GCATGATAGC TCCAGGATTG ATTTATTTTT TTCTTAACGC CAACACGCCC      420
40 TCTCAGCATG GTTTTGGGAT CCCTATGGCA ACGGATATTG CGTTCGCTTT AGGCGTGATC      480
ATGCTTTTAG GCAAGAGGGT GCCAACCGCC TTAAAGGTTT TTTTAATCAC TCTAGCGGTG      540
GCTGATGACT TAGGGGCTAT TGTGGTGATC GCGCTCTTTT ATACCACGAA TTTAAAATTC      600
GCATGGCTTT TAGGGGCTTT AGGGGTGGTT CTTGTTTTAG CCATATTGAA CCGCCTGAAT      660
ATCCGATCGC TCATCCCTTA CTTGCTTTTA GGGGTGTTGC TTTGGTTTTG CGTGATCAA      720
45 AGCGGTATCC ATGCGACGAT CGCTGCGGTG GTTCTAGCTT TTATGATACC GGTGAAAATC      780
CCTAAAGATT CTAAAAATGT AGAGCTTTTG GAATTAGGCA AACGATACGC AAGACGAGT      840
TCAGGAGTGC TTTTAACCAA AGAGCAGCAA GAAATCTTGC ATTCTATTGA AGAAAAAGCG      900
AGTGCTTTAC AAAGCCCCTT AGAAAGATTG GAGCATTTTC TAGCCCCCAT CAGCGGGTAT      960
TTCATCATGC CCTTATTGCG GTTTGCAAAC GCTGGGGTGA GCGTTGATTC TAGCATCAAT      1020
50 TTAGAAGTGG ATAAGGTGCT TTTAGGGGTT ATTTTAGGGC TTTGTTTGGG CAAGCCTTTA      1080
GGGATTTTCT TAATCACTTT CATAAGCGAA AAGCTTAAAA TCACTGCGCG CCCTAAAGGC      1140
ATCGGCTGGT GGCATATTTT AGGGGCTGGG CTTTTAGCAG GGATTGGCTT TACCATGTCT      1200
ATGTTCAATT CTAATCTGGC TTTCACGAGC GAACATAAGG ACGCTATGGA AGTGGCAAAA      1260
ATTGCGATTT TACTCGGATC TTTGATTCTT GGGATCATAG GGGCTTTATA TTTATTGCA      1320

```

- 144 -

CTAGATAAAA GAGCGGCTTT AAAGAAATAG

1350

(2) INFORMATION FOR SEQ ID NO:50:

- 5 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 2448 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: circular

- 10 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

- 15 (iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

- 20 (ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...2448

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

25 ATGAATGACA AGCGTTTTAG AAAATATTGT AGTTTTTCTA TTTTTTTGTC CTTATTAGGA 60
 ACGTTTGAAT TAGAGGCTAA AGAAGAAGAA AAAGAAGAAA AAAAGACAGA AAGGAACAAA 120
 GATAAAGAAA AGAACGCCCA ACACACTTTG GGTAAAGTTA CCACTCAAGC GGCTAAAATC 180
 30 TTTAATTACA ACAACCAGAC AACCATTTC AGTAAAGAAT TAGAAAGAAG GCAAGCCAAC 240
 CAAATCAGCG ACATGTTTAG AAGAAACCCC AATATCAATG TGGGCGGTGG TGCGGTGATA 300
 GCGCAAAAAA TTTACGTGCG CGGTATTGAA GACAGATTGG CTAGGGTTAC GGTGGATGGC 360
 GTGGCGCAAA TGGGCGCAAG CTATGGGCAT CAAGGCAATA CAATCATTGA CCCTGGAATG 420
 CTCAAAAGCG TGGTGGTTAC CAAGGGGGCG GCTCAAGCGA GCGCGGGGCC TATGGCTTTA 480
 ATTGGCGCGA TCAAAATGGA GACTAGGAGC GCGAGCGATT TTATCCCTAA AGGCAAAGAC 540
 35 TACGCCATAA GTGGGGCTGC CACTTTTTTA ACCAACTTTG GGGATAGGGA AACCATTATG 600
 GGCGCTTATC GTAACCATCA TTTTGATGCG CTTTGTATT ACACGCACCA AAATATTTTT 660
 TATTATCGTG ATGGGGATAA CGCGATGAAA AATCTTTTGG ACCCTAAAGC GGATAATAAA 720
 GTTACAGCAA GCCCTAGCGA ACAAACAAT GTGATGGCTA AGATCAATGG TTATTTGAGC 780
 GAAAGGGATA CCTTAACGCT CAGTTATAAC ATGACTAGAG ATAACGCCAA TCGCCCTTTA 840
 40 AGAGCGAATT TTACCGGCAC TTTTTTACCC TATTCTTG TGATTTTCAA CGCTTCCCT 900
 AACGAGAAA ACCCTAGCGA TTGTTTGTTT GAAAATGACG CCAGTTTGTT TAAACTTAT 960
 AGCGTCAATT TAGTGATAA CGTGAGCTTG AATTATGAAA GGAAGGGGG GAGTCGCTTT 1020
 GGCGATCCTA AATTAAAAAT CAATGGCTAC ACGAGCATT GGAATGTCCA AATTGATCCG 1080
 CTTTTAGAC CTAGCGATAT AGCGACTACC ATTCCTTTCA CCCCACCC GCAGCTCTCT 1140
 45 CAAGGCGAAG AAAATCAATG CGTGGCGCAA GGGGGCATT ATGACGCTCT TAAACAACT 1200
 TGCTCCATCA CTTTAAAAAG CCTTGGAGCG GGTCTGTTG TCGCTAATAA AAATTTATTC 1260
 ATCATCAATT CTGGGTTTAA TGCGAACGTG ATCCACACCA TAGACCACAA GAATGACAAT 1320
 CTTTTGGAAT ACGGGTTGAA TTACCAGAAT TTAACCACT TTGATAAAGC GATCCCTGAT 1380
 AGCGAATTAG TCAAGCCCCG CGATGCCCT GATGCGTGCT TAAGAGTTAC AGGACCTGAT 1440
 50 GATCCTAACA TGAACGGGCG CTGCCAACGG AATGGCGCTA CGGCGAATGT GGTGGGGTG 1500
 TATGCGCAAG CGAATTACAC CTTGCACCCT ATGGTAACT TAGGGGCAGG GACTCGTTAT 1560
 GACGTTTATA CTTTAGTGGA TAAAGACTGG CAATTGCACG TAACTCAAGG GTTTAGCCCT 1620
 AGCGCGGCTT TAAACGTCTC GCCTTTAGAA AATTTGAATT TCAGGCTTTC TTACGCGTAT 1680
 GTAAC TAGAG GTGTGGATGC GTCAAGACAA TTTGCGCTAT 1740

- 145 -

AACCGCAATT TAAAGCCAGA AATTGGGCAA AATGCGGAAT TTAACACCGA ATACAGCAGT 1800
 CAGTATTTTG ATTTTCAGAGC CGCCGGTTTT GTCCAATTGA TTTCTAATTA CATCAATCAA 1860
 TTTTCTTCAA CGCTTTTTGT CACCAACTTG CCCGCACAAG ATATTATTTA TGTGCCTGGC 1920
 TATGAAGTTT CAGGGACGGC TAAATACAAG GGTTTTTCTT TAGGCTTGAG CGTGGCGCGA 1980
 5 TCATGGCCTT CTTTAAAGG GCGTTTGATC GCTGACGTGT ATGAATTGGC GGCTACGACA 2040
 GGCAATGTGT TTATTTTAAC GGCAAGCTAT ACAATCCAC GCACCGGCCT TAGCATCACT 2100
 TGGCTTTCAC GCTTTGTTAC TAATTTGAGT TATGCTCTT ATAGCCCTTA TCGTAACGGC 2160
 CCTACGGATA TTGACAGAAG GCCTAGTAAT TGCCCTAAAA CGCCCGGGAT TTTTCATGTG 2220
 CATAAACCCG GCTATGGGGT GAGCAGTTTC TTTATCACTT ACAAGCCTAC TTATAAGAAA 2280
 10 CTCAAAGGGT TGAGCCTGAA CGCGGTGTTT AATAATGTTT TTAACCAACA ATATATTGAT 2340
 CAAGCAAGCC CGGTGATGAG CCCTGATGAA CCAATCAAG ACAATACGC AAGGGGCATG 2400
 GCAGAGCCTG GCTTTAACGC TAGGTTTGAA ATTTCTTATA AGTTTTAA 2448

(2) INFORMATION FOR SEQ ID NO:51:

15

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2445 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

20

(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

25

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

30

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...2445

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

ATGACAAGCG TTTTAGAAAA ATATTGTAGT TTTTCTATTT TTTTGTCTTT ATTAGGAACG 60
 TTTGAATTAG AGGCTAAAGA AGAAGAAAAA GAAGAAAAAA AGACAGAAAG GAACAAAGAT 120
 AAAGAAAAGA ACGCCCAACA CACTTTGGGT AAAGTTACCA CTCAAGCGGC TAAAATCTTT 180
 40 AATTACAACA ACCAGACAAC CATTTCAGT AAAGAATTAG AAAGAAGGCA AGCCAACCAA 240
 ATCAGCGACA TGTTTAGAAG AAACCCCAAT ATCAATGTGG GCGGTGGTGC GGTGATAGCG 300
 CAAAAAATTT ACGTGCGCGG TATTGAAGAC AGATTGGCTA GGGTTACGGT GGATGGCGTG 360
 GCGCAAATGG GCGCAAGCTA TGGGCATCAA GGCAATACAA TCATTGACCC TGGAAATGCTC 420
 AAAAGCGTGG TGGTTACCAA GGGGGCGGCT CAAGCGAGCG CGGGGCCTAT GGCTTTAATT 480
 45 GGC GCGATCA AAATGGAGAC TAGGAGCGCG AGCGATTTTA TCCCTAAAGG CAAAGACTAC 540
 GCCATTAAG GGGCTGCCAC TTTTTTAACC AACTTTGGGG ATAGGGAAAC CATTATGGGC 600
 GCTTATCGTA ACCATCATTT TGATGCGCTT TTGTATTACA CGCACCAAAA TATTTTAT 660
 TATCGTGATG GGGATAACGC GATGAAAAAT CTTTTTGACC CTAAAGCGGA TAATAAAGTT 720
 ACAGCAAGCC CTAGCGAACA AAACAATGTG ATGGCTAAGA TCAATGGTTA TTTGAGCGAA 780
 50 AGGGATACCT TAACGCTCAG TTATAACATG ACTAGAGATA ACGCCAATCG CCCTTTAAGA 840
 GCGAATTTTA CCGGCACTTT TTTACCCTAT TCTTGTGGTG ATTTCAACGC TTTCCCTAAC 900
 GAGAAAAACC CTAGCGATTG TTTGTTTGAA AATGACGCCA GTTTGTTTAA AACTTATAGC 960
 GTCAATTTAG TGCATAACGT GAGCTTGAAT TATGAAAGGG AAGGGGGGAG TCGCTTTGGC 1020
 GATCCTAAAT TAAAAATCAA TGGCTACACG AGCATTAGGA ATGTCCAAT TGATCCGCTT 1080

- 146 -

```

TTCAGACCTA GCGATATAGC GACTACCATT CCTTTCACCC CAAACCCGCA GCTCTCTCAA 1140
GGCGAAGAAA ATCAATGCGT GGCGCAAGGG GGCATTTATG ACGCTCTTAA ACAAACTTGC 1200
TCCATCACTT TTAAAAGCCT TGGAGGGGGT TCTGTTGTCTG CTAATAAAAA TTTATTTCATC 1260
ATCAATTCTG GGTTTAATGC GAACGTGATC CACACCATAG ACCACAAGAA TGACAATCTT 1320
5 TTGGAATACG GGTTGAATTA CCAGAATTTA ACCACTTTTG ATAAAGCGAT CCCTGATAGC 1380
GAATTAGTCA AGCCCGGCGA TGCCCCCTGAT GCGTGCTTAA GAGTTACAGG ACCTGATGAT 1440
CCTAACATGA ACGGGCGCTG CCAACGGAAT GGCGCTACGG CGAATGTGGT TGGGGTGTAT 1500
GCGCAAGCGA ATTACACCTT GCACCCTATG GTAACCTTTAG GGGCAGGGAC TCGTTATGAC 1560
GTTTATACTT TAGTGGATAA AGACTGGCAA TTGCACGTAA CTCAAGGGTT TAGCCCTAGC 1620
10 GCGGCTTTTAA ACGTCTCGCC TTTAGAAAAT TTGAATTTCA GGCTTTCTTA CGCGTATGTA 1680
ACTAGAGGCC CTATGCCTGG AGGTTTGGTG TGGATGCGTC AAGACAATTT GCGCTATAAC 1740
CGCAATTTAA AGCCAGAAAT TGGGCAAAAT GCGGAATTTA ACACCGAATA CAGCAGTCAG 1800
TATTTTGATT TCAGAGCCGC CGGTTTTGTC CAATTGATTT CTAATTACAT CAATCAATTT 1860
TCTTCAACGC TTTTGTGTC CAACTTGCCC GCACAAGATA TTATTTATGT GCCTGGCTAT 1920
15 GAAGTTTCAG GGACGGCTAA ATACAAGGGT TTTTCTTTAG GCTTGAGCGT GGCGCGATCA 1980
TGGCCTTCTT TAAAAGGGCG TTTGATCGCT GACGTGTATG AATTGGCGGC TACGACAGGC 2040
AATGTGTTTA TTTTAACGGC AAGCTATACA ATCCACGCA CCGGCCTTAG CATCACTTGG 2100
CTTTCACGCT TTGTTACTAA TTTGAGTTAT TGCTCTTATA GCCCTTATCG TAACGGCCCT 2160
ACGGATATTG ACAGAAGGCC TAGTAATTGC CCTAAAACGC CCGGGATTTT TCATGTGCAT 2220
20 AAACCCGGCT ATGGGGTGAG CAGTTTCTTT ATCACTTACA AGCCTACTTA TAAGAACTC 2280
AAAGGGTTGA GCCTGAACGC GGTGTTTAAAT AATGTTTTTA ACCACAATA TATTGATCAA 2340
GCAAGCCCCG TGATGAGCCC TGATGAACCC AATCAAGACA AATACGCAAG GGCATGGCA 2400
GAGCCTGGCT TTAACGCTAG GTTTGAAATT TCTTATAAGT TTAA 2445

```

25 (2) INFORMATION FOR SEQ ID NO:52:

(i) SEQUENCE CHARACTERISTICS:

30 (A) LENGTH: 1584 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

35 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

40 (A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

45 (A) NAME/KEY: misc_feature
 (B) LOCATION 1...1584

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

```

ATGAAACAAA ATTTAAAGCC ATTCAAAATG ATTAAGGAAA ATTTAATGAC ACAATCTCAA 60
AAAGTAAGAT TCTTAGCCCC TTTGAGCCTA GCGTTAAGCT TGAGCTTCAA TCCAGTGGGC 120
50 GCTGAAGAAG ATGGGGGCTT TATGACCTTT GGGTATGAAT TAGGTCAGGT GGTCCAGCAA 180
GTGAAAACC CGGGTAAAT CAAAGCCGAA GAATTAGCGG GCCTGTTAAA CTCTACCACG 240
ACAAACAACA CCAATATCAA TATTGCAGGC ACAGGAGGGA ATGTCGCCGG GACTTTGGGC 300
AACCCTTTTA TGAACCAATT GGGCAATTTG ATTGATTTGT ATCCTACTTT GAAAACATAAT 360
AATCTTCACC AATGCGGTAG CACTAATAGC GGTAATGGCG CTACTGCTGC CGCTGCTACT 420

```

- 147 -

	AACAATAGCC	CTTGTTTCCA	AGGTAACCTG	GCTCTTTATA	ACGAAATGGT	TGACTCTATC	480
	AAAACTTTGA	GTCAAAACAT	CAGCAAGAAC	ATCTTTCAAG	GCGACAACAA	CACCACGAGC	540
	GCTAATCTCT	CCAACCAGCT	CAGTGAGTTG	AACACCGCTA	GCGTTTATTT	GACTTACATG	600
	AACCTCGTTCT	TAAACGCCAA	CAACCAAGCG	GGTGGGATTT	TTCAAAACAA	CACCAATCAA	660
5	GCTTACGAGA	ATGGTGTTAC	CGCTCAACAA	ATCGCTTATG	TCCTAAAGCA	AGCTTCAATC	720
	ACTATGGGGC	CAAGCGGTGA	TAGTGGGGCT	GCGGGAGCGT	TTTTAGACGC	CGCTTTAGCC	780
	CAACATGTTT	TCAACTCGGC	TAACGCTGGG	AACGATTTGA	GCGCTAAGGA	ATTCACTAGC	840
	TTGGTGCAAA	ACATCGTCAA	TAATTCTCAA	AACGCTTTAA	CGCTAGCCAA	CAACGCTAAC	900
	ATCAGCAATT	CAACAGGCTA	TCAAGTGAGC	TATGGTGGGA	ATATTGATCA	AGCGCGCTCT	960
10	ACCCAAGTGT	TAAACAACAC	CACAAACACT	TTGGCTAAAG	TTACCGCTCT	AAACAACGAG	1020
	CTTAAAGCTA	ACCCATGGCT	TGGGAATTTT	GCTGCTGGTA	ACAGCTCTCA	AGTGAATGCG	1080
	TTTAACGGGT	TTATCACTAA	AATCGGTTAT	AAGCAATTCT	TCGGGGAAAA	CAAGAATGTG	1140
	GGCTTACGCT	ACTACGGGTT	CTTCAGCTAT	AACGGCGCGG	GCGTGGGTAA	TGGCCCCACT	1200
	TACAATCAAG	TCAATCTGCT	CACTTATGGG	GTGGGGACTG	ATGTGCTTTA	CAATGTGTTT	1260
15	AGCCGCTCTT	TTGGCAGTAG	GAGTCTTAAT	GCGGGCTTCT	TTGGGGGGAT	CCAACCTCGCA	1320
	GGGGACACTT	ACATCAGCAC	GCTAAGAAAC	AGCCCTCAGC	TTGCGAGCAG	ACCTACAGCG	1380
	ACAAAATTCC	AATTCTTGTT	TGATGTGGGC	TTACGCATGA	ACTTTGGTAT	CTTGAAAAAA	1440
	GACCTAAAAA	GCCATAACCA	GCATTCTATA	GAAATCGGTG	TGCAAATCCC	TACGATTTAC	1500
	AACACTTACT	ATAAAGCTGG	TGGCGCTGAA	GTGAAATACT	TCCGCCCTTA	TAGCGTGTAT	1560
20	TGGGTCTATG	GCTACGCCTT	CTAA				1584

(2) INFORMATION FOR SEQ ID NO:53:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1380 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
- (A) NAME/KEY: misc_feature
- (B) LOCATION 1...1380

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

45	GTGGTGTTAT	TAACAATGAC	AAAACGACTT	TTTAAAGGGT	TGTTAGCGAT	TTCTCTTGCG	60
	GTGAGTTTGC	ATGGTGGTGA	AGTTAAGGAA	AAAAAGCCGG	TCA...CGGT	CAAAGAAGAT	120
	CCGCAAGAAT	TAGCGGCTAA	AAGGGTGGAA	GCGTTCAGTC	GTTTCTCTAA	TGTGGTTACA	180
	GAAATTGAAA	AAAAGTATGT	GGATAAGATC	AGTATTTCTG	AGATCATGAC	TAAAGCGATT	240
	GAAGGCTTAC	TCTCTAATTT	GGACGCGCAT	TCAGCGTATT	TGAATGAAAA	GAAGTTTAAG	300
50	GAATTTTCAGG	CCCAAACCGA	GGGCGAATTT	GGGGGGCTTG	GGATCACGGT	GGGCATGCGC	360
	GATGGCGTTT	TGACCGTTAT	TGCACCTTTA	GAGGGCACTC	CAGCTTACAA	GGCTGGGGTT	420
	AAATCAGGCG	ATAGCATTTT	AAAAATCAAT	AACGAAAGCA	CGCTGAGCAT	GAGCATTGAT	480
	GATGCGGTTA	ATCTCATGCG	CGGCAAGCCA	AAAACCTCTA	TTCAGATCAC	TGTTGTTAGG	540
	AAAAATGAGC	CAAACCCCTT	GGTATTTAAT	ATCGTTAGGG	ATATTATCAA	GATCCCCTCT	600

- 148 -

```

5  GTCTATGTGA AAAAGATTAA AGACACACCT TATTTGTACG TGAGAGTCAA TTCTTTTGAT 660
   AAAAATGTTA CCAAATCGGT TTTAGACGGC TTGAAGGCTA ACCCTAACAT TAAGGGCGTT 720
   GTGTTGGATT TGAGGGGGAA TCCTGGAGGG CTATTAAACC AGGCGGTAGG CTTGTCTAAC 780
   CTTTTCATTA AAGAGGGGGT TTTAGTCTCT CAAAGAGGCA AAAATAAGGA GGAAACTTA 840
10 GAATACAAGG CTAATGGCAG AGCCCCCTAT ACCAATTTAC CTGTTGTGGT GTTAGTCAAT 900
   GCGGTTTCAG CGAGCGCGAG CGAGATCGTC GCAGGGGCAC TGCAAGATCA CAAGCGAGCC 960
   ATCATTATCG GTGAAAAAAC CTTTGGTAAG GGAAGCGTGC AAGTGTGCT CCCTGTCAAT 1020
   AAAGACGAAG CCATTAAAT CACGACCGCG CGCTATTATT TGCCGAGCGG GCGCACCATT 1080
   CAAGCTAAGG GGATCACGCC TGATATTGTG ATTTATCCGG GTAAAGTGCC AGAAAATGAA 1140
10 AATAAATTCA GTTTGAAAGA AGCGGATTTA AAACACCATT TAGAGCAAGA GCTTAAAAA 1200
   CTTGATGATA AAACCCCTAT TTCCAAAGAG GCGGATAAAG ACAAGAAAAG CGAAGAGGAA 1260
   AAAGAGTTA CTCCTAAAT GATCAATGAT GATATTCAGC TAAAAACCGC TATTGACAGC 1320
   TTGAAAACCT GGTCTATCGT AGATGAGAAA ATGGATGAAA AAGTGCCTAA GAAGAAATAA 1380

```

15 (2) INFORMATION FOR SEQ ID NO:54:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 315 base pairs

(B) TYPE: nucleic acid

20 (C) STRANDEDNESS: double

(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

25 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

30 (A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...315

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

```

40 TTGCTTTTGC ACCCCTTGCA TGCTCATGCA CAAGTGCTTG GCTTCACAAA CCACGATCAC 60
   GCCCCTTGGC TCTATGATTT CATCAAAAGT TTCTGCAATT TGAGTGGTCA GCCTTTCTTG 120
   GATTTGCAGG CGTTTGCTAT AAATTTCAAT GAGTTTAGCG ATCGCGCTAA TGCCTACAAT 180
   CTTTTCCTTA GGGATATATC CCACGCTAAT ATTCCCAAAA AAAGGGAGCA AATGGTGCTC 240
   GCAAGTGGAG TAAATTCAA TGTTTTGAGC CACTATCATT TCATCGCAAA CGCCTTGAAA 300
   ATACGCGCTT TTAA 315

```

45 (2) INFORMATION FOR SEQ ID NO:55:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 498 base pairs

(B) TYPE: nucleic acid

50 (C) STRANDEDNESS: double

(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

- 149 -

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

5 (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc_feature

10 (B) LOCATION 1...498

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

15	ATGATTGAAC	TAATCTTACA	CAATAAGTCC	ATACAAATTG	ATGAAACATT	GCTGAATGTA	60
	AAAGAGCATT	TAGAAAAGTT	TTATTCAAAC	AAAGAACAAG	AGACAATCGC	AAAAACCTTA	120
	GAGAGCCAAA	CAGAGCTTAC	TTGCAGTTAT	TTATTGGATA	AAGATTTTTC	ATTGCTAGAA	180
	AAGCATTTAG	AAAATAGCTT	AGGGCATTTC	ACTTTTGAGA	GTGAGTTTGC	CCTACTAAAA	240
	GACAAAGAGC	CTTTGAATTT	AGCTCAAATC	AAACAAATCG	GTGTTTTAAA	GGTTATTACC	300
	TATGAAATGA	CACAAGCCTT	AAAAAATCAA	ATCATTTCATT	TAACGCAAAAT	TGTCAATGAA	360
20	GAAAATTTAG	AGTTTGATGA	AGAACTTGTT	ATTTATCACT	TAAATTTTAA	GCTCAATCAA	420
	AATACTTACA	AAGTGTTAGC	GAAATTTTGC	GTATTAAAAA	AGAAAGGAAC	ATTGCATGAA	480
	AAATTTAAGG	CATTTTAG					498

(2) INFORMATION FOR SEQ ID NO:56:

25

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 642 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

30

(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

35

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

40

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...642

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

50	ATGGATACCG	AAACACAAGA	AAAGTTTTTA	GCGTATTTGT	TTGAAAAAGC	TTTACAAAAA	60
	AATCTACAAG	CTTATTGGAT	AACAACAAC	GAACTAAGA	ATGAATTAAC	AAGAGAAGAG	120
	TTTTCAAATT	TAATAAGAAA	AACAATGATT	GACTAATCT	TACACAATAA	GTCCATACAA	180
	ATTGATGAAA	CATTGCTGAA	TGTAAAAGAG	CATTTAGAAA	AGTTTTATTC	AAACAAAGAA	240
	CAAGAGACAA	TCGCAAAAAC	CTTAGAGAGC	CAAACAGAGC	TTACTTGCAG	TTATTTATTG	300
	GATAAAGATT	TTTCATTGCT	AGAAAAGCAT	TTAGAAAATA	GCTTAGGGCA	TTTTACTTTT	360
	GAGAGTGAGT	TTGCCCTACT	AAAAGACAAA	GAGCCTTTGA	ATTTAGCTCA	AATCAAACAA	420
	ATCGGTGTTT	TAAAGGTTAT	TACCTATGAA	ATGACACAAG	CCTTAAAAAA	TCAAATCATT	480

- 150 -

CATTTAACGC AAATTGTCAA TGAAGAAAAT TTAGAGTTTG ATGAAGAACT TGTTATTTAT 540
CACTTAAATT TTAAGCTCAA TCAAAATACT TACAAAGTGT TAGCGAAATT TTGCGTATTA 600
AAAAAGAAAG GAACATTGCA TGAAAAATTT AAGGCATTTT AG 642

5 (2) INFORMATION FOR SEQ ID NO:57:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 762 base pairs
(B) TYPE: nucleic acid
10 (C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

15 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- 20 (A) ORGANISM:
- Helicobacter pylori*

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
25 (B) LOCATION 1...762

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

ATGGCGATCT CTATTAAAAG CCCAAAAGAA ATCAAAGCCC TAAGAAAAGC CGGGGAATTA 60
ACCGCTCAAG CGTTAGCCCT TTTAGAGCGA GAAGTAAGGC CTGGGGTTTC ACTTTTAGAG 120
30 CTGGATAAAA TGGCTGAAGA TTTTATCAAA TCCTCGCATG CTAGGCCTGC TTTTAAGGGG 180
CTCTATGGTT TCCCTAAGTC TGTGTGCATG TCCTTAAATG AGGTGGTTAT TCATGGTATT 240
CCTACGGATT ATGTTTTTACA AGAAGGGGAT ATTATAGGCT TGGATTGGG GGTGGAGGTG 300
GATGGCTATT ATGGCGATTC AGCCCTCAGC CTTCCCATAG GCGCGATAAG CCCGCAAGAT 360
GAAAAATTGC TCGCTTGCTC TAAAGAGAGC TTGATGCATG CCATTAGCTC AATTAGAGTG 420
35 GGCATGCATT TTAAAGAGTT GAGTCAGATT TTAGAGGGCG CTATTACAGA AAGGGGCTTT 480
GTGCCTTTGA AGGGATTTTG CGGGCATGGC ATTGGTAAAA AGCCCCATGA AGAGCCAGAA 540
ATCCCCAACT ACCTAGAAAA AGGCGTCAAA GCTAATAGCG GCCCTAAAAT CAAAGAGGGC 600
ATGGTGTTTT GTTTAGAGCC TATGGTGTGT CAAAAACAAG GCGAGCCTAA AATACTAGCG 660
GATAAGTGGA GCGTGGTTTC AGTGGATGGA CTTAACACAA GCCACCATGA GCATACTATC 720
40 GCCATAGTTG GCAATAAAGC AGTGATTCTT ACGGAGCGTT AA 762

(2) INFORMATION FOR SEQ ID NO:58:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 744 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
45 (D) TOPOLOGY: circular

50 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

- 151 -

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

5 (ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...744

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

10

AAGCCCAAAA	GAAATCAAAG	CCCTAAGAAA	AGCCGGGAAT	TAACCGCTCA	AGCGTTAGCC	60
CTTTTAGAGC	GAGAAGTAAG	GCCTGGGGTT	TCACTTTTAG	AGCTGGATAA	AATGGCTGAA	120
GATTTTATCA	AATCCTCGCA	TGCTAGGCCT	GCTTTTAAGG	GGCTCTATGG	TTCCCTAAC	180
TCTGTGTGCA	TGTCCTTAAA	TGAGGTGGTT	ATTCATGGTA	TTCTACGGA	TTATGTTTTA	240
15 CAAGAAGGGG	ATATTATAGG	CTTGGATTG	GGGGTGGAGG	TGGATGGCTA	TTATGGCGAT	300
TCAGCCCTCA	CGCTTCCCAT	AGGCGCGATA	AGCCCGCAAG	ATGAAAAATT	GCTCGCTTGC	360
TCTAAAGAGA	GCTTGATGCA	TGCCATTAGC	TCAATTAGAG	TGGGCATGCA	TTTTAAAGAG	420
TTGAGTCAGA	TTTTAGAGGG	CGCTATTACA	GAAAGGGGCT	TTGTGCCTTT	GAAGGGATTT	480
TGCGGGCATG	GCATTGGTAA	AAAGCCCCAT	GAAGAGCCAG	AAATCCCCAA	CTACCTAGAA	540
20 AAAGGCGTCA	AAGCTAATAG	CGGCCCTAAA	ATCAAAGAGG	GCATGGTGTT	TTGTTTAGAG	600
CCTATGGTGT	GTCAAAAACA	AGGCGAGCCT	AAAATACTAG	CGGATAAGTG	GAGCGTGGTT	660
TCACTGGATG	GACTTAACAC	AAGCCACCAT	GAGCATACTA	TCGCCATAGT	TGGCAATAAA	720
GCACTGATTC	TTACGGAGCG	TTAA				744

25 (2) INFORMATION FOR SEQ ID NO:59:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1023 base pairs

(B) TYPE: nucleic acid

30 (C) STRANDEDNESS: double

(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

35 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

40 (A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...1023

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

ATGTATCGTA	AAGATTGGA	TAATTACTTA	AAACAGCGCC	TCCCTAAAGC	GGTGTTTTTG	60
TATGGGGAGT	TTGATTTTTT	CATCCATTAT	TATATTCAAA	CGATTAGCGC	GCTTTTTTAA	120
50 GGCAATAACC	CTGACACAGA	AACTTCGCTT	TTTTATGCGA	GCGATTATGA	AAAAAGCCAG	180
ATTGCGACCC	TTTTAGAGCA	GGATTCTTTA	TTTGGAGGGA	GCAGTTTAGT	TATTTTAAAA	240
CTGGATTTTG	CATTGCATAA	GAAATTTAAG	GAAAATGATA	TCAATCCTTT	TTTAAAAGCT	300
TTAGAGCGGC	CTAGCCATAA	TAGGCTTATC	ATAGGGCTTT	ATAATGCTAA	AAGCGACACC	360
ACAAAATACA	AATACACTAG	CGAAATTATC	GTAAATTTT	TCCAAAAAAG	CCCCTTGAAA	420

- 152 -

```

GATGAAGCCA TTTGCGTGCG CTTTTTTTACC CCTAAAGCGT GGGAGAGTTT GAAATTCTTG 480
CAAGAAAGGG CTAATTTTTT GCATTTAGAC ATCAGCGGCC ATCTTTTAAA CGCTCTTTTT 540
GAAATTAATA ACGAAGATT AAGCGTTTCG TTTAACGATT TAGACAAGCT AGCGGTTTTA 600
AACGCGCCCA TCACTTTAGA AGACATTCAA GAATTAAGCT CCAATGCGGG GGATATGGAT 660
5 TTGCAAAAGC TCATTTTAGG GCTTTTTTTG AAAAAAAGCG TCCTTGATAT TTATGATTAT 720
TTGTTAAAAG AGGGCAAAAA GGATGCGGAT ATTTTAAGGG GGTTAGAGCG CTATTTTTTAC 780
CAGCTTTTTT TATTTTTCGC CCACATTAAA ACGACCGGTT TAATGGACGC TAAAGAGGTC 840
TTAGGCTACG CTCCTCCTAA AGAGATTGTA GAAAATTACG CTAAAAACGC CCTGCGTTTG 900
AAAGAAGCCG GCTATAAGAG GGTTTTGTAA ATTTTATAGG TATGGCACCT TCAAAGCATG 960
10 CAAGGGCAAA AGGAATTGGG CTTTTTGTAT TTGACCCCCA TTCAAAAAAT CATTAACCCT 1020
TGA 1023

```

(2) INFORMATION FOR SEQ ID NO:60:

- 15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 603 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: circular
- 20 (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- 25 (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
 (A) ORGANISM: *Helicobacter pylori*
- 30 (ix) FEATURE:
 (A) NAME/KEY: misc_feature
 (B) LOCATION 1...603
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

```

35 GTGTTTATGA CAAGCGCTCT GTTAGGCTTA CAAATTGTTT TAGCGGTATT GATTGTGGTG 60
GTGGTTTTGT TGCAAAAAG TTCTAGCATC GGCTTAGGGG CTTATAGCGG AAGCAACGAT 120
TCTTTATTTG GCGCTAAAGG GCCCGCAAGC TTTATGGCGA AATTGACCAT GTTTTTAGGT 180
TTATTGTTTG TCATCAACAC CATCGCTTTG GGCTATTTTT ACAACAAAGA ATACGGCAAG 240
40 AGCGTTTTAG ATGAAACTAA AACCAATAAA GAGCTTTCGC CCTTAGTCCC TGCCACCGGC 300
ACGCTCAACC CTACGCTTAA TCCCACATTA AACCCAACGC TCAACCCTTT AGAGCAAGCC 360
CCCACTAATC CTTTAATGCC TACACAAACG CCTAAAGAGC TTCCTAAAGA GCCAGCCAAA 420
ACGCCTTTTG TTGAAAGCCC CAAACAGAAT GAAAAGAATG AAAAGAATGA TGCCAAAGAA 480
AATGGTATAA AGGGTGTGTA AAAAAACAAA GAGAACGCCA AAACGCCCCC AACCACCCAC 540
45 CAAAAGCCTA AAACGCATGC GACAACCAAC GCCCATACCA ACCAAAAAAA GGATGAAAAA 600
TAA 603

```

(2) INFORMATION FOR SEQ ID NO:61:

- 50 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 480 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: circular

- 153 -

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

5 (iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

10 (A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...480

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

ATGCGTTCTC	CAAATTTAGA	AAAAGAAGAA	ACTGAAATCA	TAGAAACGCT	TCTTGTGCGT	60
GAAAAAATGC	GTTTATGCCC	CTTGTATTGG	CGCATCTTAG	CGTTTTTAAT	CGATAGTTTA	120
TTGGTGCGCT	TTTTATTGAG	CGATCTTTTA	AGGGCATGCG	CTTTTTTACA	TTCTTTATAT	180
20 TGGCTGACTA	ACCCCATTTA	TTACAGCGCG	TTTGTGTGTA	TGGGTTTTAT	CATCTTGTAT	240
GGCGTTTATG	AAATCTTTTT	TGTGTGTTTG	TGCAAGATGA	GTTTGGCTAA	ACTGGTTTTT	300
AGGATTAAGA	TCATTGATAT	TTATTTAGCG	GATTGCCCCA	GTAGGGCTAT	TTTATTGAAG	360
CGTTTAGGGT	TAAAAATCGT	GGTTTTTCTA	TGCCCCTTTT	TATGGTTTGT	GGTGTTTTAA	420
25 AACCCCTATC	ATAGGGCATG	GCATGAAGAA	AAAAGCAAAA	GTCTTTTGGT	GTTGTTTTAA	480

(2) INFORMATION FOR SEQ ID NO:62:

(i) SEQUENCE CHARACTERISTICS:

30 (A) LENGTH: 705 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

35 (iii) HYPOTHETICAL: NO.

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

40 (A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc_feature

45 (B) LOCATION 1...705

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

TTGAATACGG	ACTTTAGCCA	TATCACCGAT	ATTGAGGGCA	TGCGTTTTGT	TAATGAAGAA	60
50 GACGCTTTAA	ACAAATTGAT	TAATGAAATC	CACACGCGCC	ACATTGATTT	AAAAGATTCC	120
ATCATGCTCG	CTTTGAGTTT	TAACGCCTTG	TATTTAGCTA	ACGCTTTAGC	GCAAAAATTT	180
GGGGCGACTT	ATGATATACT	TTTTTTAGAA	CCTATCTTAG	CCCCTTTAAA	CTCAAAGTGT	240
GAAATCGCTT	TAGTGAGTGA	AAGCATGGAT	ATAGTGATGA	ATGAAAGTTT	AATCAATTCC	300
TTTGACATCG	CTTTAGACTA	TGTTTATGGG	GAAGCCAAGC	GGGCTTATGA	AGAAGACATT	360

- 154 -

CTGTCTCACA TCTATCAGTA TCGCAAAGGC AATGCGATCA AAAGCCTAAA AGATAAAAAAT 420
 ATTTTATCG TAGATAGGGG GATTGAGACC GGGTTTAGAG CAGGGTTAGG CGTGCAAACT 480
 TGTTTGAAAA AAGAATGCCA AGACATTTAT ATTTTAACCC CCATTCTCGC GCAAAATGTC 540
 GCTCAAGGCT TAGAAAGCTT GTGCGATGGG GTGATTAGCG TGTATCGCCC TGAATGTTTT 600
 5 GTCTCTGTGG AACACCATTA TAAAGAACTC AAGCGATTAA GCAATGAAGA AATTGAAAAA 660
 TACTTGGGCG CTAACAACGC GCCCAATCTC AAAAAGGAAC ATTAA 705

(2) INFORMATION FOR SEQ ID NO:63:

- 10 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 864 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: circular
 15
 (ii) MOLECULE TYPE: DNA (genomic)
 (iii) HYPOTHETICAL: NO
 20 (iv) ANTI-SENSE: NO
 (vi) ORIGINAL SOURCE:
 (A) ORGANISM: *Helicobacter pylori*
 25 (ix) FEATURE:
 (A) NAME/KEY: misc_feature
 (B) LOCATION 1...864

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:

30 TTGAAACAAA GCGAAATGGC CATGGAATTT AATGATCCTA GGATGCGTTT TTTTATTGGC 60
 GATGTCAGGG ATTTAGAACG CTTGAATTAC GCTTTAGAGG GCGTGGATAT TTGTATCCAT 120
 GCGGCCGCGC TCAAGCATGT GCCTATCGCT GAATACAACC CCCTAGAATG CATTAAAACT 180
 AACATCATGG GAGCGAGCAA TGTGATTAAC GCATGCTTAA AAAATGAAAT CAGCCAGGTT 240
 35 ATTGCCCTAA GCACCGATAA AGCCGCTAAC CCCATTACC TCTACGGCGC AACCAAATTG 300
 TGCAGCGACA AGCTCTTTGT GAGCGCGAAC AACTTTAAAG GCCCTTCTCA AACGCAATTT 360
 GCGTGGTGC GTTATGGTAA TGTGGTGGGG AGTCGTGGGA GCGTGGTGCC GTTTTTTAAA 420
 AAATTAGTCC AAAACAAAGC GAGTGAAATC CCCATTACCG ATATTCGCAT GACACGATTT 480
 TGGATCACCT TAGATGAGGG GGTTCCTTTT GTGCTTAAAA GCTTGAAAAG AATGCATGGG 540
 40 GGGGAAATTT TTGTGCCTAA AATCCCCAGC ATGAAAATGA TTGATCTCGC CAAAGCCCTA 600
 GCCCCCAATA TCCCTACTAA AATCATAGGG ATTCGCCCCG GCGAAAACT CCATGAAGTG 660
 ATGATCCCTA AAGATGAAAG CCATTTAGCC CTAGAATTTG AAGACTTTTT TATTATTTCAG 720
 CCCACTATAA GCTTCCAAAC GCCTAAAGAT TACACGCTCA CCAAACTCCA TGAAAAAGGC 780
 CAAAAAGTCG CCCCTGATTT TGAATACAGC AGCCATACTA ATAACCAATG GCTAGAGCCT 840
 45 GATGATTTGT TAAATTATT ATGA 864

(2) INFORMATION FOR SEQ ID NO:64:

- 50 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 606 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: circular

- 155 -

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

5 (iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: *Helicobacter pylori*

10 (ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...606

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:

15 ATGCGTTTGC AACTGCGCTT TTTTGGTATT AATTCGTTGC TTGTCGCCAC TCTTTTGATA 60
 AGCGGTTGCA GTCTCTTTAA AAAGCGTAAC ACTAACGCTC AGCTAATCCC CCCTTCAGCT 120
 AACGGGTTGC AAGCCCCCAT TTATCCCCCA ACCAATTTC AAGCCAGAAA GAGCATTTCAG 180
 CCTCTCCCAA GCCCTCGCCT TGAGAATAAC GATCAGCCCA TCATTAGCTC TAATCCCACT 240
 20 AACGCTATCC CTAACACCCC CATTCTCACG CCCAATAATG TCATTGAGTT GAATGCGGTG 300
 GGCATGGGTG TGGCTCCAGA ATCCACCATT TCGCCCTCTC AAGCTCTAGC TTTAGCTAAG 360
 CGAGCGGCTA TTGTTGATGG CTACCGCCAG TTGGGTGAAA AAATGTATGG CATCAGAGTG 420
 AACGCTCAAG ACACCGTCAA AGACATGGTT TTACAAAATT CCGTGATTAA AACGAGAGTG 480
 AATGCCCTCA TTCGTAACGC TGAAATCACT GAGACTATCT ATAAAGACGG CTTGTGCCAG 540
 25 GTAAGCATGG AGCTTAAATT AGACGGCAGG ATTTGGTATC GTATTTTGAG CGGATCGAGA 600
 GGATAA 606

(2) INFORMATION FOR SEQ ID NO:65:

30 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1068 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

35 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

40 (iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: *Helicobacter pylori*

45 (ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...1068

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:

50 ATGAGTTATA CTATTAATAA ACGCTTTTCT GTGGGTGTGG GTTTAAGGGG GCTTTATGCG 60
 ACCGGGAGCT TTAATAACAC CGTTTATGTG CCTTTAGAGG GCGCTTCAGT TTTGAGCGCG 120
 GAGCAAATCT TAAACTTACC CAACAATGTT TTTGCCGATC AAGTGCCAAG TAACATGATG 180
 ACTTTATTAG GCAATATTGG CTACCAACCA GCGCTTAATT GCCAAAAAGC CCGTGGGGAC 240

- 156 -

```

5  ATGAGTGATC AGAGCTGTCA AGAGTTTAC AACGGCTTGA AAAAAATCAT GGGTTATAGC 300
   GGTTTAATCA AAGCGAGCGC GAATCTTTAT GGCACGACTC AAGTCGTGCA AAAATCTAAC 360
   GGACAAGGCG TATCGGGGGG GTATAGAGTG GGTTCGAGTT TCGGTGTGTT TGATCATGGC 420
   ATGTTTTCTG TGGTGATATA TTCTTCAGTT ACCTTTAACA TGAAAGGCGG TTTGGTGGCT 480
10 ATCACAGAGC TTGGCCCTTC TTTAGGGAGC GTTTTACTA AAGGCAGCTT GAATATCAAT 540
   GTTTCACTCC CCCAAACTTT AAGCTTAGCC TACGCCACC AATTTTTTAA AGATCGCCTA 600
   AGGGTTGAAG GGGTGTTTGA GCGCACTTTT TGGAGTCAAG GGAATAAATT TTTAGTCACC 660
   CCTGATTTTG CGAACGCCAC TTACAAGGGC TTGAGCGGGA CGGTGGCTTC CTTGGACTCT 720
   GAAACGCTTA AAAAAATGGT AGGCCTAGCG AATTTTAAAA GCGTGATGAA CATGGGGGCT 780
10 GGCTGGAGGG ACACCAACAC CTTTAGATTA GGGGTAACCTT ACATGGGTAA AAGCTTGCGT 840
   TTAATGGGCG CTATTGATTA TGATCAAGCC CCAAGCCCCC AAGACGCGAT AGGCATTCCG 900
   GACTCTAATG GCTATACCGT GGCTTTTGGG ACTAAATACA ATTTTAGGGG CTTTGATTG 960
   GCGGTAGCGG GGAGTTTCAC TTTTAAGAGC AACCCTCCA GTTTGTATCA ATCCCCAACT 1020
   ATTGGGCAAT TGAGAATCTT TAGCGCCTCT TTAGGCTATC GCTGGTAA 1068

```

(2) INFORMATION FOR SEQ ID NO:66:

(i) SEQUENCE CHARACTERISTICS:

- 20 (A) LENGTH: 1764 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

25

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

30

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

- 35 (A) NAME/KEY: misc_feature
 (B) LOCATION 1...1764

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:

```

40 ATGAAAAACT TTTCCCCACT CTATTGTCTT AAAAAGCTCA AAAACGCCA TTTAATCGCT 60
   CTGAGTCTGC CCTTGCTTTC TTATGCGAAT GGCTTTAAAA TCCAAGAGCA AAGCTTGAAT 120
   GGCACGGCTT TAGGCTCGGC GTATGTCGCT GGGGCTAGGG GTGCTGACGC TTCTTTTTTAC 180
   AACCCGGCTA ACATGGGCTT TACTAACGAT TGGGGCGAAA ACAGAAGCGA ATTTGAAATG 240
   ACCACCACCG TGATCAATAT CCCGGCCTTT AGCTTTAAAG TCCCTACGAC CAATCAAGGC 300
   TTATATTCGG TAACAAGTTT AGAAATTGAT AAAAGCCAAC AAAATATTTT AGGCATCATC 360
45 AACACTATAG GGTTAGGCAA TATCCTTAAA GCGCTTGGCA ATACGGCCGC TACCAATGGC 420
   TTA.C.CAAG CTATCAATCG TGTTCAGGG CTTATGAACT TAACCAATCA AAAAGTCGTA 480
   ACCCTCGCTT CAAAACCTGA CACTCAAATC GTGAATGGCT GGACAGGCAC GACTAATTTT 540
   GTTTTACCTA AATTCTTTTA TAAAACGCGC ACGCATAACG GCTTCACTTT TGGGGGGAGT 600
   TTTACCGCTC CTAGTGGGTT GGGTATGAAA TGGAATGGTA AGGGGGGGGA ATTTTTCAT 660
50 GACGTGTTTA TCATGATGGT AGAGCTTGCC CCTAGCATGA GTTATACTAT TAATAAACGC 720
   TTTTCTGTGG GTGTGGGTTT AAGGGGGCTT TATGCGACCG GGAGCTTTAA TAACACCGTT 780
   TATGTGCCTT TAGAGGGCGC TTCAGTTTTG AGCGCGGAGC AAATCTTAAA CTTACCCAAC 840
   AATGTTTTTG CCGATCAAGT GCCAAGTAAC ATGATGACTT TATTAGGCAA TATTGGCTAC 900
   CAACCAGCGC TTAATTGCCA AAAAGCCGGT GGGGACATGA GTGATCAGAG CTGTCAAGAG 960

```


- 157 -

5 TTTTACAACG GCTTGAAAAA AATCATGGGT TATAGCGGTT TAATCAAAGC GAGCGCGAAT 1020
 CTTTATGGCA CGACTCAAGT CGTGCAAAAA TCTAACGGAC AAGGCGTATC GGGGGGGTAT 1080
 AGAGTGGGTT CGAGTTTGCG TGTGTTTGAT CATGGCATGT TTTCTGTGGT GTATAATTCT 1140
 TCAGTTACCT TTAACATGAA AGGCGGTTTG GTGGCTATCA CAGAGCTTGG CCCTTCTTTA 1200
 GGGAGCGTTT TGACTAAAGG CAGCTTGAAT ATCAATGTTT CACTCCCCCA AACTTTAAGC 1260
 TTAGCCTACG CCCACCAATT TTTTAAAGAT CGCCTAAGGG TTGAAGGGGT GTTTGAGCGC 1320
 ACTTTTGGGA GTCAAGGGAA TAAATTTTAA GTCACCCCTG ATTTTGCGAA CGCCACTTAC 1380
 AAGGGCTTGA GCGGGACGGT GGCTTCCTTG GACTCTGAAA CGCTTAAAAA AATGGTAGGC 1440
 CTAGCGAATT ~~TTAAAAGCGT~~ GATGAACATG GGGGCTGGCT GGAGGGACAC CAACACCTTT 1500
 10 AGATTAGGGG TAACCTACAT GGGTAAAAGC TTGCGTTTAA TGGGCGCTAT TGATTATGAT 1560
 CAAGCCCCAA GCCCCAAGA CGCGATAGGC ATTCCGGACT CTAATGGCTA TACCGTGGCT 1620
 TTTGGGACTA AATACAATTT TAGGGGCTTT GATTTGGGCG TAGCGGGGAG TTTCACCTTT 1680
 AAGAGCAACC GCTCCAGTTT GTATCAATCC CCAACTATTG GGCAATTGAG AATCTTTAGC 1740
 GCCTCTTTAG GCTATCGCTG GTAA 1764

(2) INFORMATION FOR SEQ ID NO:67:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 618 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
 (B) LOCATION 1...618

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:

40 TTGATTTTTTA GATTTTTTCTT AATCTTAAGC CTTTTTAAAAG GGGTTTTACT GGCCAAAAAG 60
 GATTGGAATT TTTTCAAACC TTTAGAGCCT ACTAAAAAAT ATTTTGGCTC TTTTAAAATC 120
 GGCTATCTTT ACCAACATGC AGAAACGACT AAAAGATTCC CCATCCGCCC TAAAAACCGC 180
 CCGCCTATTT TAATGGATAA AATTTACCAT GACGCTTCTT TGGGTTTTGA CGCAGGGTAT 240
 GTTTTGAAAA AGAAAGCTTT ATTGGGGGGG TATTTGGATG CAGGAATGGG CGATTCGTAT 300
 TTCATGAGCG CTGGGCTAGT CGCTGGGGTG AGGCTTTTTA AGGGGTGGGT TATCCCTAAA 360
 45 ATCGCCTTAG GCTATCAGCT TCAAATTTTA GGGGCTAAGA TTGATAAGTA TCAATTCAAT 420
 ATCCAATCAG CGGTGGGGAG TGTGGGCTTG TTTTTCATG CGGCTAAAAA TTTTGGCTTG 480
 AGTATAGAAG CAAGGGGCGG TATCCCTTTT TATTTCAATC AGAGCAGGTT TTCTAAGGCT 540
 TTCGGCACGC CACGATTGAA TATCTATTCT GTTGGTATCA CATTCACTTT TTATGACTTT 600
 50 ACGAGATTTT TAGGGTAA 618

(2) INFORMATION FOR SEQ ID NO:68:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 762 base pairs

- 158 -

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

5 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

10

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

15

(A) NAME/KEY: misc_feature

(B) LOCATION 1...762

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:

20	TTGTGGCATG CTGCCTTTAG CGTTGGCGAG TGGGGATGGA ACGGCGATGA AATCCCCTAT	60
	AGGGATTGCG ATGAGTGGGG GCTTGATGAT TTCTATGGTG TTAAGCCTAC TGATTGTGCC	120
	GGTGTTTTAT CGTTTGCTCG CTCCCATAGA CGACAAAATC AAGCGGTTTT ATCAAAACCA	180
	AAAAGCTTTA GAATGAAAAA AATTGCTTTC ATTTTGGCTT TATGGGTGGG CTTGTTAGGG	240
	GCGTTTGAGC CTAAAAAAG TCATATTTAT TTTGGGGCTA TGGTGGGTTT AGCCCCCTGT	300
25	AAAATAACCC CAAAACCGGC TAGTGATTCT TCTTATACGG CTTTTTTATG GGGGGCTAAA	360
	GGGGGGTATC AATTCCGCTT TTTTAAAGCT CTAGCGTTAA GGGGTGAATT TTCCTACCTT	420
	ATGGCGATCA AACCACCGC ACTGCACACG ATTAACACTT CTTTATTGAG TTAAATATG	480
	GATGTGTTGA GCGATTTTTA CACTTATAAA AAATACAGCT TTGGGGTGTA TGGGGGGCTT	540
	GGGATAGGGT ATTTTATCA AAGCAACCAT TTAGGCATGA AAAATAGTTC GTTTATGGGT	600
30	TATAACGGCT TGTTTAATGT GGGGCTTGGC AGCAGATCG ATCGCCACCA CCGCGTAGAG	660
	CTTGGGGCTA AGATCCCTTT TTCAAAGACT AGAAATTCTT TTAAAAATTC TTATTTTTTA	720
	GAGAGCGTTT TTATCCATGC GGCTTATAGT TATATGTTTT AA	762

(2) INFORMATION FOR SEQ ID NO:69:

35

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1239 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

40

(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

45

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

50

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...1239

- 159 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:

```

5  ATGGAATCAG TAAAAACAGT AAAAACAAAT AAAGTTGGCA AAAACACAGA GACAGCTAAC 60
   ACAGAGGCAA GTAAAGAGAC TCATTTTAAA CAAGCGAGTG CCATTACAAA TACGCTCCGA 120
   TCAATTGGTG GGATTTTAC AAAAATTGCA AAGAAAGTTA GAGAACTTGT GAAAAAACAT 180
   CCCAAGAAAA GCAGTGTGGC ATTAGTAGTA TTGACCCATA TTGCGTGCAA GAGGGCAAAA 240
   GAATTGGACG ATAAAGTCCA AGATAAATCC AAACAAGCTG AAAAAGAAAA TCAAATCAAT 300
   TGGTGGAAAT ATTCAGGATT AACAATAGCG GCAAGTTTAT TATTAGCCGC TTGTAGCACT 360
   GGTGATATTG ATAAACAAAT AGAACTAGAA CAAGAAAAAA AGGAAGCAAA TAAGAGTGGG 420
10 ATAAAGTTAG AACAAGAAAG ACAGAAAACA GAACAAGAAA GACAGAAGAC AAATAAGAGT 480
   GAGATAGAGT TAGAACAAGA AAGACAAAAA ACAAACAAGA GTGGGATAGA ACTCGCTAAT 540
   AGTCAAATAA AAGCAGAACA AGAAAGACAA AAGACAGAAC AAGAAAAACA AAAAGCAAAT 600
   AAGAGTGAGA TAGAGTTAGA ACAGCAAAAA CAAAAGACAA TTAATACACA AAGAGATTTG 660
   ATTAAGAAGC AGAAAGATTT CATTAAAGAA ACAGAACAAA ATTGCCAAGA AAAACATGGC 720
15 CAATTGTTTA TTAAAAAGC AAGAATTAAG ACCGGTATTA CTACTGGTAT TGCCATAGAA 780
   ATAGAAGCTG AATGCAAAAC CCCTAAACCT GCAAAAACCA ATCAAACCCC TATCCAGCCA 840
   AAACACCTCC CAAACTCTAA ACAACCCCGC TCTCAAAGAG GATCAAAAGC GCAAGAGCTT 900
   ATCGCTTATT TGCAAAAAGA GCTAGAATCT CTGCCCTATT CGCAAAAAGC TATCGCTAAA 960
   CAAGTGGATT TTTATAAACC AAGTTCTATC GCTTATTTAG AACTAGACCC TAGAGATTTT 1020
20 AAGGTTACAG AAGAATGGCA AAAAGAAAAT TTAAAAATAC GCTCTAAAGC TCAAGCTAAA 1080
   ATGCTTGAAA TGAGAAACCC ACAAGCCAC CTTCCAACCT CTCAAAGCCT TTTGTTTCGT 1140
   CAAAAAATAT TTGCTGATAT TAAATAAGAA ATAGAAGCAG TTGCTAATAC TGAAAAGAAA 1200
   ACAGAAAAAG CGGGTTATGG TTATAGTAAA AGGATGTAG 1239

```

25 (2) INFORMATION FOR SEQ ID NO:70:

(i) SEQUENCE CHARACTERISTICS:

- 30 (A) LENGTH: 450 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

35 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- 40 (A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

- 45 (A) NAME/KEY: misc_feature
 (B) LOCATION 1...450

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:

```

50 TTGAATTGGG AGCATTTGAT GAAAAAATTA GCGTTTCTT TATTATTTAC AGGGACTTTT 60
   TTGGGGCTTT TTTTGAATGC GAGTGATTTT AAGAGCATGG ATAACAAGCA ACTATTAGAG 120
   CAAGCAGGGA AAGTCGCTCC TAGCGAAGTT CCAGAGTTTC GCACAGAAGT CAATAAACGA 180
   TTAGAAGCGA TGAAAGAAGA AGAGCGTCAA AAATATAAAG CGGATTTTAA GAAAGCGATG 240
   GATAAGAATT TGGCTTCTTT AAGCCAAGAA GATCGCAACA AGCGTAAAAA AGAAATCCTT 300
   GAAGTCATTG CTAACAAAAA GAAAACAATG ACCATGAAAG AGTATCGTGA AGAGGGGTTG 360
   GATTTCATG ATTGCGCATG CGAAGGCCCT TTTCATGATC ATGAAAAAAA GGGGCAAAAA 420

```

- 160 -

GGGAAAAAAC CAAGCCATCA TAAGCATTAG

450

(2) INFORMATION FOR SEQ ID NO:71:

5 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 615 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: circular

10 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

15 (iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
 (A) ORGANISM: *Helicobacter pylori*

20 (ix) FEATURE:
 (A) NAME/KEY: misc_feature
 (B) LOCATION 1...615

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:71:

25 ATGCAAGCAG TGATTTTAGC GAATGGGGAG TTTCCTAAAT CTAAAAAATG CTTAGACATT 60
 TTACAAAACG CTCCTTTTTT AATCGCATGC GATGGGGCTG TTATATCATT GCATGCGCTT 120
 CAATTCAAAC CCAGCGTTGT TATAGGCGAT TTGGATAGCA TTGATTCGCA TTTGAAAGCC 180
 TTGTATAACC CTATACGCGT GAGCGAACAA GACAGCAACG ATTTGTCCAA AGCCTTTTTT 240
 30 TATGCTTTGA ATAGGGGTTG TGATGATTTT ATTTTTTTAG GGTGAATGG CAAGCGAGAA 300
 GACCACGCTT TAGCGAACAC TTTTTTATTG TTGGAGTATT TTAAATTTTG CAAAAAATC 360
 CAATCCGTAA GCGATTATGG CCTTTTTTAGG GTGTTAGAAA CCCCTTTTAC TTTGCCCAGT 420
 TTTAAGGGGG AGCAAATCTC GCTTTTTTAGC TTGGATCTTA AAGCCCGATT CACTTCTAAA 480
 AACCTCAAAT ACCCCTTAAA AGACTTGCGT CTAAAAACGC TCTTTTCCGG CTCGCTCAAT 540
 35 GAAGCCACTA ATCATTGTTT TAGCCTTAGC TCTGAACCTA AATCGGTGGT GCTAGTGTAT 600
 CAAAAATTCT CATGA 615

(2) INFORMATION FOR SEQ ID NO:72:

40 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 843 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: circular

45 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

50 (iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
 (A) ORGANISM: *Helicobacter pylori*

- 161 -

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...843

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:72:

```

GTGTTTGACT CATTGGGCGG ATTTTGGGG TATAAACTT TTAAGCCGAT AGTGGATAAG    60
GTTAAAAATA TAAACGCTTG GATAAAAAAT TACGATAATA AAAAAGCTCA AGAGATTATG    120
GGTTTTATAG AAAATCCTAC GCCTGATTTC CAAAATAATA AGTTTTTG TGTTTTAAAC    180
10 CGACAAGGAA CAAGGCACAA CAATTATCTT GGTTTAACCT CTACAAACCT TCTAATCGGC    240
GCGATCTATT TCTCCATCCG CCATTGCATC AAAGCCACAT GGCAAAACGA TAGGGATCAA    300
TTCTACGCC CTTATGATGA CGCTTTCCAA GACGACAGCG AGTTTAAAA CAATTGTTTG    360
GCGTTCATGC TTTTTCACAC CCAAACCGC ATCACTGCCA CTCAAGGGAC TAACCATTTT    420
ATCCCTTTTA GCGAAGATGA AGTTGATTCT AAAGAAAGGT ATTTGAGCCA TGCTTTATTA    480
15 GACTTTTTTA AAGGCGAAAT CAAAGAACCT AAAAAGAGCG ATAGCCTCTT TTTAAACGCC    540
AAAAAGAAA ACAAGCCCCT AAAATTCAGC TCGAGCGCTT CAAAGGTGTT TGACGCTGGC    600
AGAGAGATTT ATCGCTATTA CCACACACAA GATTTCATCC ACACCCCTA TAACGCTAAC    660
GCAAGCCTTT ATGACATCAA AGAATTTTTT CAAGGCCGTA ACAAGCAAGG CAGATTAAAC    720
TCACCCACCA AAGCCAAAGA TGAATATTAC AAACAGCTTT ACGCTAACTT GCAATACGCC    780
20 CTAAAGATC TCGCCAAAGA AATACAGCCT AAAGTCTATG AATACGGATT TTTAAGGGAG    840
TAG

```

(2) INFORMATION FOR SEQ ID NO:73:

25 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 930 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: circular

30

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

35

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

40

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...930

45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:73:

```

TGTGACAGGG CAATTCCCCA TTGGCTTTTT AGTCTGGGAT ACCGCTACCC CCCCCC.P.A    60
AAACCAACCA ACGCGTTCAA TTTAGAAGTG TTTGACTCAT TGGGCGGATT TTTGGGGTAT    120
AAAACCTTTA AGCCGATAGT GGATAAGGTT AAAAATATAA ACGCTTGGAT AAAAATTAC    180
GATAATAAAA AAGCTCAAGA GATTATGGGT TTTATAGAAA ATCCTACGCC TGATTTCCAA    240
50 AATAATAAGT TTTGTGTGT TTTAAACCGA CAAGGAACAA GGCACAACAA TTATCTTGGT    300
TTAACCTCTA CAAACCTTCT AATCGGCGCG ATCTATTTCT CCATCCGCCA TTGCATCAAA    360
GCCACATGGC AAAACGATAG GGATCAATTC TACGCCCTT ATGATGACGC TTTCCAAGAC    420
GACAGCGAGT TTA AAAACAA TTTGTTGGCG TTCATGCTTT TTCACACCCA AAACCGCATC    480
ACTGCCACTC AAGGGACTAA CCATTTTATC CCCTTTAGCG AAGATGAAGT TGATTCTAAA    540

```

- 162 -

5 GAAAGGTATT TGAGCCATGC TTTATTAGAC TTTTAAAAG GCGAAATCAA AGAACCTAAA 600
 AAGAGCGATA GCCTCTTTT AAACGCCAAA AAAGAAAACA AGCCCCTAAA ATTCAGCTCG 660
 AGCGCTTCAA AGGTGTTTGA CGCTGGCAGA GAGATTTATC GCTATTACCA CACACAAGAT 720
 TTCATCCACA CCCCCTATAA CGCTAACGCA AGCCTTTATG ACATCAAAGA ATTTTTCAT 780
 GGCCGTAACA AGCAAGGCAG ATTAACTCA CCCACCAAAG CCAAAGATGA ATATTACAAA 840
 CAGCTTTACG CTAACCTGCA ATACGCCCTA AAAGATCTCG CCAAAGAAAT ACAGCCTAAA 900
 GTCTATGAAT ACGGATTTT AAGGGAGTAG 930

(2) INFORMATION FOR SEQ ID NO:74:

10

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 564 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

15

(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

20

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

25

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...564

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:74:

35 TTGGAACTT ATATCATTGA TGCAGATAAT ATAGATGGGG ATTTATTTTT CTATAATCTT 60
 ACTAGAAACA GCAATGATT TTCCATGTTG CCCGTTTTTG AACTCGATCG CATTGCCCAA 120
 AAAATTAGAA ATATTCTTAA AAAACATGGC AGTAGAAAAG ACATTATTTT AAAACACAAT 180
 GAAATTAAAG AAGCCTTTTT TAGCCCGTTC AAACCGCAGC TAAAACCGT TCAAGTGTTT 240
 CTCTCGCACT CGCATGCGGA TAAAAATAAG GCTTTAGGGG TTAAGGACTA TTTGGAAAGC 300
 AAAACAAAAC GCAAAGTGTT TATCGATTCT CTTTTTTGGG ATTATAAAGA CGATGTTTTA 360
 AACAAATTGG CAAACACGA TGATATAAGC AAGATTGAAG ACGCTTTCAC GCTCATTCTC 420
 AGAAAATCTT TACAAGATAT GATTGAAAAA TGCCCTTATT TTGTGTTTTT ACAAAGCAAG 480
 40 AACAGCGTTT CTAATCAAGG GCTATCACGC ATCACTTATT CCGCATGGAT TTATGAAGAA 540
 TTAAAAATCG CTTCACTCTA TTAG 564

(2) INFORMATION FOR SEQ ID NO:75:

45

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 597 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

50

(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

- 163 -

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...597

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:75:

```

15  TTGGAAACTT ATATCATTGA TGCAGATAAT ATAGATGGGG ATTTATTTTT CTATAATCTT    60
    ACTAGAAACA GCAATGATTT TTCCATGTTG CCCGTTTTTG AACTCGATCG CATTGCCCAA    120
    AAAATTAGAA ATATTCTTAA AAAACATGGC AGTAGAAAAG ACATTATTTT AAAACACAAT    180
    GAAATTAAAG AAGCCTTTTT TAGCCCGTTC AAACCGCAGC TAAAAACCGT TCAAGTGTTT    240
    CTCTCGCACT CGCATGCGCT TAAAAATAAG GCTTTAGGGG TTAAGGACTA TTTGGAAAGC    300
    AAAACAAAAC GCAAAGTGTT TATCGATTCT CTTTTTGGG ATTATAAAGA CGATGTTTAA    360
    AACAAATTGG CAAAACACGA TGATATAAGC AAGATTGAAG ACGCTTTCAC GCTCATTCTC    420
    AGAAAATCTT TACAAGATAT GATTGAAAAA TGCCCTTATT TTGTGTTTTT ACAAAGCAAG    480
20  AACAGCGTTT CTAATCAAGG GCTATCACGC ATCACTTATT CCGCATGGAT TTATGAAGAA    540
    TTAAAAATCG CTTCAATTCT ATTAGCGCTA TTAACGAGAG TCGCCCAATT CCAATGA    597

```

(2) INFORMATION FOR SEQ ID NO:76:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 570 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...570

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:76:

```

45  ATGATGACTA A.AA.CGCTTA TGCGTTTGTC GTGATTGAAA AAAGTATTAT GGTGTTTAAA    60
    TGCGCCAAAG ACAAGGGGCT AATCCCTATC ACTGAAGGCT TTGTGCCGTT AAAAGAGGGC    120
    TTTTGTAGAA GTTTTAAAGA GCGTTGCAAT CTGGATTTTT TAGAAAATTT AGACCTTTTG    180
    TTTTGTATG ACTACCAATT TCCAAGCGAG GTTTTTTCAT TGTGTAAGGA TTTGAAAAAT    240
50  TCCATTGGG ACAGAAAGCT TGTGGTAGTG CTAGTGGAGG CTTTGGAGGG TTTTAAGGGT    300
    TTGAATTTGT CTCTTAAGAT AGAAGATAGG CATTCTAATA GCTTGGGTAA TGGCGTTCAA    360
    AAATGCTCA CCAACGCTGA TTTGGGGAGC AACCACAAAC CAATCGTAAT AGACAGCATG    420
    AAAACATACC ACCAAAGCCA GCAAGAAAAA TACAAAAGAG AAAGAGGCGA AACGCTAGAG    480
    GTTCGCCCCA CAACACCCC TAGCTATGGG GGTGGGAGCA TTAGAATCAG CGGCGATAAA    540

```

- 164 -

AAGCCTGATT CCAATGAAGA AAATTTTAA

570

(2) INFORMATION FOR SEQ ID NO:77:

5 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1773 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: circular

10 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

15 (iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
 (A) ORGANISM: *Helicobacter pylori*

20 (ix) FEATURE:
 (A) NAME/KEY: misc_feature
 (B) LOCATION 1...1773

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:77:

25 ATGAAAGCGA TAAAAATACT TCTTATAATG ACACTCAGTT TAAACGCTAT CAGCGTGAAT 60
AGGGCGTTGT TTGATTTAAA AGATTTCGCA TTAAGAGGGG AATTAACGCC AAAAATAGTG 120
GATTTTGGGG GTTATAAAAAG CAACACCACA GAGTGGGGAG CTACGGCTTT AAACATATATC 180
AATGCGGCTA ATGGCGATGC GAAAAAATTC AGCGCGTTAG TGGAAAAAAT GCGTTTAAAC 240
30 TCTGGTATCT TGGGGAATTT TAGAGCGCAT GCACATTTGA GGCAAGCCCT AAAATTGCAA 300
AAGAATTTGA AATATTGCCT TAAATCATC GCTAGGGATT CTTTTTATAG TTACCGCACC 360
GGTATTTATA TCCCCTTAGG CATTCTTTA AAAGATCAAA AAACGGCTCA AAAAATGCTC 420
GCTGATTTGA GCGTGGTAGG GGCGTATCTT AAAAAGCAAC AGGAGAATGA AAAGGCTCAA 480
AGCCCTTATT ACAGGAGCAA CAACTATTAC AACTCCTACT ATAGCCCTTA TTATGGCATG 540
35 TATGGCATGT ATGGAATGGG CATGTATGGA ATGTATGGCA TGGGCATGTA TGATTTTAT 600
GACTTTTATG ATGGCATGTA TGGGTTCTAC CCTAACATGT TTTTCATGAT GCAAGTTCAA 660
GACTACTTGA TGTTAGAAAA TTACATGTAT GCACTCGATC AAGAAGAGAT TTTAGACCAT 720
GACGCTTCCA TCAACCAACT TGATACGCCT ACTGATGATG ACAGAGACGA TAAAGACGAT 780
AAATCTTCGC AACCAGCGAA TCTCATGAGC TTTTATCGTG ATCCCAAATT CAGCAAAGAC 840
40 ATTCAAACCA ACCGCTTGAA TAGCGCCTTA GTCAATTTAG ACAACAGCCA CATGCTCAA 900
GACAATTCGC TCTTCCACAC TAAAGCCATG CCCACTAAAA GCGTGGATGC GATCACTTCT 960
CAAGCTAAAG AGCTTAACCA TTTGGTGGGG CAAATCAAAG AGATGAAGCA AGACGGGGCG 1020
AGTCCTAATA AGATTGATTC AGTGGTCAAT AAAGCTATGG AGGTAGGGA CAAATTAGAC 1080
AACAACCTCA ACCAACTAGA CAATGACTTA AAAGATCAAA AAGGGCTTTC AAGCGAGCAG 1140
45 CAAGCCCAAG TGGATAAAGC CTTAGACAGC GTGCAACAAT TAAGCCATAG CAGCGATGTG 1200
GTAGGGAATT ATTTAGACGG GAGTTTGAAA ATTGATGGCG ATGACAGAGA CGATTTGAAT 1260
GATGCGATCA ATAACCCTAT GCAACAACCT GCGCAACAAA CGCCTATTAA CAACATGGAC 1320
AACACCCATG CAAATGACAG CAAAGATCAA GGGGGTAACG CGCTCATAAA CCCTAACAAAC 1380
GCCACCAACG ATGATCACAA CGATGATCAC ATGGACACTA ACACCACTGA CACTAGCAAC 1440
50 GCAAACGACA CCCCCACTGA TGATAAAGAT GCTAGCGGCA ACAATACCGG CGATATGAAT 1500
AACACCGACA CCGGCAATAC GGACACTGGC AACACCGACA CCGGTAACAC TGATGATATG 1560
AGCAACATGA ACAACGGCAA CGATGATACG GGTAACACTA ACGACGACAT GGGTAATAGC 1620
AACGACATGG GCGATGACAT GAATAACGCG AACGACATGA ACGACGACAT GGGTAACAGC 1680
AACGATGACA TGGGCGATAT GGGGGACATG AACGATGACA TGGGTGGCGA TATGGGAGAC 1740

- 165 -

ATGGGGGATA TGGGTGGCGA TATGGGGAAT TGA

1773

(2) INFORMATION FOR SEQ ID NO:78:

5 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 588 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: circular

10 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

15 (iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
 (A) ORGANISM: *Helicobacter pylori*

20 (ix) FEATURE:
 (A) NAME/KEY: misc_feature
 (B) LOCATION 1...588

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:78:

25 TTGAATTTAC GATTGGCTGG AGCAAGCGTT TTAACGGCTT GTGTCTTTTC GGGGTGTTTT 60
 TTTTAAAAA TGTTTGACAA AAAACTTTCT AGCAACGATT GGCATATCCA AAAAGTAGAA 120
 ATGAACCATC AAGTGTATGA CATTGAAACC ATGCTCGCTG ATAGCGCTTT TAGAGAGCAT 180
 GAAGAAGAGC AAGACTCCTC TTAAATACC GCTTTGCCTG AAGATAAAAC AGCGATTGAA 240
 30 GCCAAAGAGC AAGAGCAAAA AGAAAAAAGG AAACACTGGT ATGAGCTTTT TAAAAAGAAG 300
 CCAAAGCCCA AAAGCTCTAT GGGAGAGTTT GTGTTTGATC AAAAAGAAAA TCGTATTTAT 360
 GGGAAAGGCT ATTGCAACCG GTATTTTGCT AGCTACACAT GGCAGGGCGA TAGGCACATC 420
 GCAATTGAAG ATAGCGGGAT TTCAAGAAAA GTGTGTAGAG ATGAGCATTT GATGGCGTTT 480
 GAATTGGAAT TTATGGAGAA TTTTAAGGGT AATTTTGC GG TAAC TAAGGG CAAGGACACG 540
 35 CTCATTTTAG ACAACCAAAA AATGAAAATT TATTTGAAAA CGCCATGA 588

(2) INFORMATION FOR SEQ ID NO:79:

40 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 2235 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: circular

45 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

50 (vi) ORIGINAL SOURCE:
 (A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

- 166 -

(A) NAME/KEY: misc_feature
(B) LOCATION 1...2235

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:79:

```

5  ATGTTAAAC TCGCCAGTAA AACGATTTGT TTGTCCCTAA TCAGCTCATT CACGGCTGTA      60
   GAAGCCTTTC AAAAACACCA AAAAGACGGC TTTTTCATAG AAGCCGGCTT TGAAACCGGG      120
   CTATTACAAG GCACACAAAC CCAAGAACAA ACCATAGCCA CCACTCAAGA AAAACCCAAA      180
   CCCAAACCCA AACCAAAACC CATTACCCCT CAAAGCACCT ATGGGAAATA CTACATCTCC      240
10  CAAAGCACCA TTTTAAAGAA TGCGACTGAG TTGTTTGAG AGGATAATAT CACCAACTTA      300
   ACCTTTTACT CTCAAAACCC TGTGTATGTA ACCGCTTATA ACCAAGAAAG CGCTGAAGAA      360
   GCTGGCTATG GTAATAACAG CTTGATTATG ATACAAACT TCTTGCCTTA TAACCTGAAC      420
   AACATTGAGC TGAGTTACAC GGACGATCAA GGCAATGTGG TCAGTTTGGG CGTGATAGAG      480
   ACTATCCCTA AACAATCTCA AATCATTCTG CCCGCAAGCT TGTTTAACGA CCCACAGCTT      540
15  AACGCCGATG GCTTCCAACA ACTCCAAACC AACACCACAC GATTTTCTGA TGCCAGCAGC      600
   CAGAATCTGT TTAACAAGCT CAGCAAGGTT ACAACCAATC TTCAAATGAC TTATATCAAT      660
   TACAACCAAT TTTCTAGCGG TAACGGCAGT GGCTCTAAAC CCCCATGCCC CCCATACGAA      720
   AACCAAGCAA ATTGTGTGGC TAAAGTGCCG CCTTTCACCT CTCAAGACGC TAAAAATTTG      780
   ACCAATTTAA TGCTGAACAT GATGGCGGTG TTTGATTCTA AATCTTGGGA AGACGCCGTC      840
20  TTAAACGCTC CTTTCCAATT CAGCGACAAC AACCTGTCAG CGCCATGTTA TTCTGATTAC      900
   CTTACATGCG TGAATCCTTA CAACGATGGG CTTGTTGATC CTAAATTGAT CGCCAAAAAT      960
   AAAGGAGATG AATACAATAT AGAAAACGGG CAAACAGGCT CAGTGATATT AACGCCGCAA      1020
   GATGTTATCT ATAGCTATAG AGTCGCTAAT AATATTTATG TGAATCTCTT GCCACAAGA      1080
   GGAGGGGATT TAGGGTTAGG GTCTCAATAT GGTGGCCCGA ATGGCCAGG CGATGATGGC      1140
25  ACCAATTTTG GCGCTTTAGG GATATTGTCC CCTTTCCTAG ACCCTGAAAT ATTGTTTGGC      1200
   AAAGAATTGA ATAAAGTCGC CATCATGCAA TTAAGAGACA TCATCCATGA ATACGGCCAT      1260
   ACTTTAGGCT ATACGCATAA CGGGAACATG ACTTATCAA GAGTGCGCAT GTGCGAAGAA      1320
   AACAATGGGC CAGAAGAGCG CTGTCAGGGC GGAAGGATAG AGCAAGTGGG TGGGAAAGAA      1380
   GTGCAAGTGT TTGACAACGG GCATGAAGTG CGAGACACCG ATGGCTCTAC CTATGATGTG      1440
30  TGTTCTCGTT TTAAAGATAA GCCCTATACA GCGGGCAGCT ATCCTAATTC CATCTATACC      1500
   GATTGCTCTC AAGTCCCCGC TGGGCTTATA GCGCTTACCA GCGCTGTTTG GCAACAACCTC      1560
   ATTGATCAAA ACGCCCTACC GGTGGATTTT ACTAATTTGA GCAGCCAAAC CAACTATTTG      1620
   AACGCCAGCT TGAACACGCA AGACTTTGCG ACCACCATGC TTAGCGCGAT CAGTCAAAGC      1680
   CTTTCATCTT CTAAATCTAG CGCCACTACT TATCGCACTT CAAAACCTC ACGGCCCTTT      1740
35  GGAGCCCCC TATTAGGCGT TAATCTTAAA ATGGGCTATC AAAAAATATT TAATGATTAT      1800
   CTAGGGTTGT CTTCTTATGG CATTATCAA TACAACACG CTCAAGCCAA CAACGAAAAA      1860
   ATCCAGCAAT TAAGCTATGG CGTGGGAATG GATGTGCTGT TTGATTTTCAT CACCAATTAC      1920
   ACTAACGAAA AGAACCCCAA AAGCAATCTA ACCAAGAAAG TTTTCACTTC CTCTCTTGGG      1980
   GTGTTTGGGG GGTAAAGGGG CTTATACAAC AGCTATTATT TGTTGAACCA ATACAAAGGG      2040
40  AGCGGTAATT TAAATGTGAC CGGTGGGTTG AATTACCGCT ACAAGCATTC CAAATATTCT      2100
   ATAGGCATTA GCGTTCCTTT GGTCCAGTTG AAATCTAGGA TCGTTTCTAG CGATGCTGCT      2160
   TATACCAATT CTATACCCT CAATGAAGGG GGCAGTCATT TTAAAGTGTT TTTTAATTAC      2220
   GGGTGGATTT TCTAA
                                     2235

```

45 (2) INFORMATION FOR SEQ ID NO:80:

(i) SEQUENCE CHARACTERISTICS:

50 (A) LENGTH: 1590 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

- 167 -

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

5

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

10

(A) NAME/KEY: misc_feature

(B) LOCATION 1...1590

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:80:

15	ATGACTTATA	TCAATTACAA	CCAATTTTCT	AGCGGTAACG	GCAGTGGCTC	TAAACCCCCA	60
	TGCCCCCAT	ACGAAAACCA	AGCAAATTGT	GTGGCTAAAG	TGCCGCCTTT	CACCTCTCAA	120
	GACGCTAAAA	ATTTGACCAA	TTTAATGCTG	AACATGATGG	CGGTGTTTGA	TTCTAAATCT	180
	TGGGAAGACG	CCGTCTTAAA	CGCTCCTTTC	CAATTCAGCG	ACAACAACCT	GTCAGCGCCA	240
	TGTTATTCTG	ATTACCTTAC	ATGCGTGAAT	CCTTACAACG	ATGGGCTTGT	TGATCCTAAA	300
	TTGATCGCCA	AAAATAAAGG	AGATGAATAC	AATATAGAAA	ACGGGCAAAC	AGGCTCAGTG	360
20	ATATTAACGC	CGCAAGATGT	TATCTATAGC	TATAGAGTCG	CTAATAATAT	TTATGTGAAT	420
	CTCTTGCCCA	CAAGAGGAGG	GGATTTAGGG	TTAGGGTCTC	AATATGGTGG	CCCGAATGGC	480
	CCAGGCGATG	ATGGCACCAA	TTTTGGCGCT	TTAGGGATAT	TGTCCCCTTT	CTTAGACCCT	540
	GAAATATTGT	TTGGCAAAGA	ATTGAATAAA	GTCGCCATCA	TGCAATTAAG	AGACATCATC	600
	CATGAATACG	GCCATACTTT	AGGCTATACG	CATAACGGGA	ACATGACTTA	TCAAAGAGTG	660
25	CGCATGTGCG	AAGAAAACAA	TGGGCCAGAA	GAGCGCTGTC	AGGGCGGAAG	GATAGAGCAA	720
	GTGGATGGGA	AAGAAGTGCA	AGTGTGTTGAC	AACGGGCATG	AAGTGCGAGA	CACCGATGGC	780
	TCTACCTATG	ATGTGTGTTT	TCGTTTTTAA	GATAAGCCCT	ATACAGCGGG	CAGCTATCCT	840
	AATTCCATCT	ATACCGATTG	CTCTCAAGTC	CCCGCTGGGC	TTATAGGCGT	TACCAGCGCT	900
	GTTTGGCAAC	AACTCATTGA	TCAAAACGCC	CTACCGGTGG	ATTTTACTAA	TTTGAGCAGC	960
30	CAAACCAACT	ATTTGAACGC	CAGCTTGAAC	ACGCAAGACT	TTGCGACCAC	CATGCTTAGC	1020
	GCGATCAGTC	AAAGCCTTTC	ATCTTCTAAA	TCTAGCGCCA	CTACTTATCG	CACCTCAAAA	1080
	ACCTCACGGC	CCTTTGGAGC	CCCCCTATTA	GGCGTTAATC	TTAAAATGGG	CTATCAAAAA	1140
	TATTTTAATG	ATTATCTAGG	GTTGTCTTCT	TATGGCATTA	TCAAATACAA	CTACGCTCAA	1200
	GCCAACAACG	AAAAAATCCA	GCAATTAAGC	TATGGCGTGG	GAATGGATGT	GCTGTTTGAT	1260
35	TTCATCACCA	ATTACACTAA	CGAAAAGAAC	CCCAAAAGCA	ATCTAACCBA	GAAAGTTTTT	1320
	ACTTCCTCTC	TTGGGGTGTT	TGGGGGGTTA	AGGGGCTTAT	ACAACAGCTA	TTATTTGTTG	1380
	AACCAATACA	AAGGGAGCGG	TAATTTAAAT	GTGACCGGTG	GGTTGAATTA	CCGCTACAAG	1440
	CATTCCAAAT	ATTCTATAGG	CATTAGCGTT	CCTTTGGTCC	AGTTGAAATC	TAGGATCGTT	1500
	TCTAGCGATG	GTGCTTATAC	CAATTCTATC	ACCCTCAATG	AAGGGGGCAG	TCATTTTAAA	1560
40	GTGTTTTTTT	ATTACGGGTG	GATTTTCTAA				1590

(2) INFORMATION FOR SEQ ID NO:81:

(i) SEQUENCE CHARACTERISTICS:

45

(A) LENGTH: 564 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: circular

50

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

- 168 -

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

5 (ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...564

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:81:

10 TTGGGTTGCG TATCAATGAC TCTAGGTATT GATGAAGCGG GGAGGGGGTG TTTGGCCGGT 60
TCGCTTTTGT TGGCGGGGGT GGTGTGTAAT GAAAAAATAG CCTTAGAATT TCTAAAAATG 120
GGTCTTAAGG ATAGCAAGAA GCTCAGCCCC AAAAAGCGCT TTTTCTTAGA AGATAAAATC 180
AAAACGCATG GTGAGGTGGG GTTTTTCGTG GTTAAAAAAA GCGCGAATGA AATTGATCAT 240
15 TTGGGCTTAG GGGCGTGTTC GAACTCGCT ATTGAAGAAA TTGTAGAAAA TGTTTGCTCT 300
TTAGCCAATG AAATAAAAAAT AGATGGCAAC ACGGCGTTTG GCTTGAACAA ACGCTACCCC 360
AACATACAAA CCATCATCAA GGGCGATGAA ACAATCGCTC AAATCGCTAT GGCGTCTGTT 420
TTGGCGAAAG CTTCTAAGGA TAGGGAAATG TTAGAACTGC ACGCTTTGTT TAAGGAATAC 480
GGCTGGGATA AGAATTGCGG GTATGGGACT AAACAACATA TAGAAGCGAT CAATAAGCTA 540
20 GGGGCTACGC TTTCATCGGC ATAG 564

(2) INFORMATION FOR SEQ ID NO:82:

(i) SEQUENCE CHARACTERISTICS:

25 (A) LENGTH: 615 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

30 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

35

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

40 (A) NAME/KEY: misc_feature
(B) LOCATION 1...615

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:82:

45 ATGACTCTAG GTATTGATGA AGCGGGGAGG GGGTGTTTGG CCGGTTTCGCT TTTTGTGGCG 60
GGGGTGGTGT GTAATGAAAA AATAGCCTTA GAATTTCTAA AAATGGGTCT TAAGGATAGC 120
AAGAAGCTCA GCCCCAAAAA GCGCTTTTTC TTAGAAGATA AAATCAAAAC GCATGGTGAG 180
GTGGGGTTTT TCGTGTTTAA AAAAAGCGCG AATGAAATTG ATCATTTGGG CTTAGGGGCG 240
TGTTTGAAAC TCGCTATTGA AGAAATTGTA GAAAATGGTT GCTCTTTAGC CAATGAAATA 300
50 AAAATAGATG GCAACACGGC GTTTGGCTTG AACAAACGCT ACCCAACAT ACAAAACATC 360
ATCAAGGGCG ATGAAACAAT CGCTCAAATC GCTATGGCGT CTGTTTTGGC GAAAGCTTCT 420
AAGGATAGGG AAATGTTAGA ACTGCACGCT TTGTTTAAAG AATACGGCTG GGATAAGAAT 480
TGCGGGTATG GGAATAACA ACATATAGAA GCGATCAATA AGCTAGGGGC TACGCCTTTT 540
CATCGGCATA GCTTCACGCT TAAAAACCGC ATCTTAAATC CCAAACCTCT AGAGGTGGAA 600

- 169 -

CAACGCCTTG TTTAA

615

(2) INFORMATION FOR SEQ ID NO:83:

5 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 579 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: circular

10 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

15 (iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
 (A) ORGANISM: *Helicobacter pylori*

20 (ix) FEATURE:
 (A) NAME/KEY: misc_feature
 (B) LOCATION 1...579

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:83:

25	ATGAATGCAT	TGAAAAAATT	AAGTTTTTGC	GCCTTGTTAT	CCCTAGGCCT	CTTCGCTCAA	60
	ACAGTGCATG	CTCAGCATTT	AAAGGACACG	ATTAACATC	CTGATTGGCT	TAAAATCAAT	120
	CTTTTTGATA	AAAAGAACCC	GCCCAATCAA	TATGTCGGAT	CGGCTTCAAT	TTCTGGTAAA	180
	AGGAACGATT	TTTATTCCAA	TTACATCCCC	TATGATGACA	AATTGCCCCC	TGAAAAGAAC	240
30	GCTGAAGAAA	TCGCTCTTTT	AAGGGCCAGA	ATGAACGCTT	ACAGCACTTT	AGAAAGCGCT	300
	TTACTCACTA	AAATGTGCAA	TCGCATTGTT	AAAGCGCTTC	AAGTTAAAAA	TAATGTTATC	360
	AGCCATTTAT	TCGGGTTTGT	TGATTTTTTA	ACGTCTAAAT	CCATTTTGGC	TAAAAGGTTT	420
	GTGGATACCA	CCAACCATCG	TGTGTATGTC	ATGGTGCAAT	TCCCTTTCAT	TCAGCCTGAA	480
	GACTTAATCG	CTTACTTTAA	AGCCAAACGC	ATCGACCTTT	CTTTAGCGAG	CGCTACCAAT	540
35	CTCAGCGCCA	TTTTAAACAA	GCGTTGTTT	CACCTCTAA			579

(2) INFORMATION FOR SEQ ID NO:84:

40 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 261 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: circular

45 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

50 (vi) ORIGINAL SOURCE:
 (A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

- 170 -

- (A) NAME/KEY: misc_feature
(B) LOCATION 1...261

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:84:

5
ATGAATGCAT TGAAAAAATT AAGTTTTTGC GCCTTGTTAT CCCTAGGCCT CTTCGCTCAA 60
ACAGTGCATG CTCAGCATTT AAAGGACACG ATTAACATATC CTGATTGGCT TAAAATCAAT 120
CTTTTTGATA AAAAGAACCC GCCCAATCAA TATGTCGGAT CGGCTTCAAT TTCTGGTAAA 180
AGGAACGATT TTTATTCCAA TTACATCCCC TATGATGACA AATTGCCCCC TGAAAGAACG 240
10 CTGAAGAAAT CGCTCTTTTA A 261

(2) INFORMATION FOR SEQ ID NO:85:

- 15 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 228 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular
- 20 (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- 25 (vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori
- (ix) FEATURE:
30 (A) NAME/KEY: misc_feature
(B) LOCATION 1...228

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:85:

35 TTGAAAATTT TAACCCTTTT TTTGATAGGT TTAAACGCAT TGTTGCCCCT AGATTTGAAC 60
GCGCTTAAAA CAGAAATCAA AGAAACCTAT CTCAAAGAAT ACAAAGACTT AAAATTGGAA 120
ATTGAAACAA TTAATTTAGA AATCCCAGAG CGTTTTTCTC ACGCTTCCAT TTTAAGCTAT 180
GAATTGAACG CTTCTAACAA GCTTAAAAAA GATGGGTCGT GTTTTTTA 228

40 (2) INFORMATION FOR SEQ ID NO:86:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 636 base pairs
(B) TYPE: nucleic acid
45 (C) STRANDEDNESS: double
(D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- 50 (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:

- 171 -

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...636

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:86:

```
10 ATGTTTTCAA TAATTCTGGG GGGGGGGGGG GGTAATACCC CATGCGGCTT GACATGGCAA 60
   CACTTCAAAT TAGGGGATTT GTTTGAAATT GAAAAAACCT TAAGCTTTAA TAAAGACGCT 120
   TTAACGCAAG GACAAGATTA CGATTATATT ACAAGAACTT CGCAAAATCA AGGCGTTTTG 180
   CAAACTACAG GATTTGTCAA TGCAGAAAAT TTAAACCCAC CATTTACTTG GAGTTTAGGG 240
   CTTTTGCAAA TGGATTTTTT CTATCGTAAA AAGTCATGGT ATGCGGGACA ATTCATGCGA 300
   AAAATCACAC CAAAAACTGA AATTAAAAAT AAAATTAATT CACGCATAGC CCACTATTTT 360
15 ACAACGCTTT TAAACGCCTT AAAACGCCCT TTATTGAGTG TATTAGTTAG GGATATTGAT 420
   AAAACTTTTA GGGAGCAAAA AATCCAATA CCCCTAAAC CCACCGCTAA AACTCAAAGC 480
   CTTGATGGTA TTGATTTTGA TTTGATGCAC ACCCTAATCA ACGCCCTGAT GAAGCAAACC 540
   ATTCAGGGCG TGGTTCAATA CTGCGACGCT AAAATACAGG CTACAAAAGA AGTTATCAGC 600
   CAAGAAACGC CTATTCAAAA AGACTCGTTA TTTTGA 636
```

(2) INFORMATION FOR SEQ ID NO:87:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1221 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...1221

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:87:

```
45 GTGATTGGCC CCCTTAGCAG CCAACTCAAC GCTATTAAGT GGGGCGAGTT CAAATTAGGG 60
   GATTGTGTTG AAGCGAGTAA CGGCGATTTT GACATTCAAA AACGCCACAT CAATCATAAG 120
   GCGGAATTTG TCATCACCGC AGGGCTTAGC AATAATGGCG TTTTAGGGCA AAGCGATATA 180
   AAAGCAAAAG TTTTTGAAAG CCATACCATT ACTATTGACA TGTTTGTTG CGCGTTTTAT 240
   CGCAGTTTTG CTTATAAAAT GGTAACACAT GCTAGGGTAT TTTCTCTCAA ACCTAAATTT 300
   GAAATCAACC ATAAATCGG CTTGTTTTTA TCCACGCTAT TTTTGGTTA CCATAAAAAA 360
50 TTCGGCTATG AAAACATGTG TTCATGGGCA AAAATTAAAA ACGATAAAGT CATTCTACCC 420
   CTAAACCCCA CCGCTAACAC TCAAACCTT GAGGGTATTG ATTTTGATTT CATGGAAAAA 480
   TTCATAGCCG AACTTGAGCA GTGTCGGCTC GCCGAACTTC AGGCTTATTT AAAAGCTACA 540
   GGGCTAGAAA ACACCACCCT TTCTAACGAT GAAGAAAATG CCCTTAATGT TTTCAATAAT 600
   TCTGGGGGGG GGGGGGGTAA TACCCCATGC GGCTTGACAT GGCAACACTT CAAATTAGGG 660
```

- 172 -

```

GATTTGTTTG AAATTGAAAA AACCTTAAGC TTTAATAAAG ACGCTTTAAC GCAAGGACAA 720
GATTACGATT ATATTACAAG AACTTCGCAA AATCAAGGCG TTTTGCAAAC TACAGGATTT 780
GTCAATGCAG AAAATTTAAA CCCACCATTT ACTTGGAGTT TAGGGCTTTT GCAAATGGAT 840
TTTTTCTATC GTAAAAAGTC ATGGTATGCG GGACAATTCA TCGGAAAAAT CACACCAAAA 900
5 ACTGAAATTA AAAATAAAAT TAATTCACGC ATAGCCCACT ATTTCAACAAC GCTTTTAAAC 960
GCCTTAAAC GCCCTTTATT GAGTGTATTA GTTAGGGATA TTGATAAAAC TTTTAGGGAG 1020
CAAAAATCC AACTACCCCT AAAACCCACC GCTAAACTC AAAGCCTTGA TGGTATTGAT 1080
TTTGATTTCA TGCACACCCT AATCAACGCC CTGATGAAGC AAACCATTCA AGGCGTGGTT 1140
CAATACTGCG ACGCTAAAAT ACAGGCTACA AAAGAAGTTA TCAGCCAAGA AACGCCTATT 1200
10 CAAAAGACT CGTTATTTTG A 1221

```

(2) INFORMATION FOR SEQ ID NO:88:

```

15 (i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 828 base pairs
    (B) TYPE: nucleic acid
    (C) STRANDEDNESS: double
    (D) TOPOLOGY: circular

20 (ii) MOLECULE TYPE: DNA (genomic)

    (iii) HYPOTHETICAL: NO

    (iv) ANTI-SENSE: NO

25 (vi) ORIGINAL SOURCE:
    (A) ORGANISM: Helicobacter pylori

    (ix) FEATURE:
30 (A) NAME/KEY: misc_feature
    (B) LOCATION 1...828

```

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:88:

```

35 ATGAGTAAGA GTTTATACCA AACTTTAAAC GTGAGCGAAA ACGCCAGCCA AGATGAAATC 60
AAAAATCCT ACCGCCGTTT AGCCAGGCAA TACCACCCGG ATTTGAATAA AACCAGAGAA 120
GCCGAAGAGA AATTCAAAGA AATCAACGCC GCTTATGAAA TTTTGAGCGA TGAAGAAAAA 180
CGCCGCCAAT ACGATCAATT TGGCGACAAC ATGTTTGGCG GGCAGAATTT CAGCGATTTT 240
GCCAGAAGCC GTGGTCCTAG TGAAGATTTA GATGATATTT TAAGCTCTAT TTTTGGGAAA 300
40 GGAGGCTTTT CGCAAAGATT TTCTCAAAAT TCGCAAGGCT TTTCTGGCTT TAATTTTTC 360
AATTTTCGCC CTGAAAATTT AGATGTAACC GCTATTTTAA ATGTCTCTGT TTTAGACACC 420
CTTTTAGGCA ATAAAAACA AGTGAGCGTC AATAATGAGA CTTTTCAGCT TAAAATCCCT 480
ATCGGCGTGG AAGAGGGCGA AAAGATTAGG GTTCGCAACA AAGGGAAAAT GGGGCGAACG 540
GGTAGGGGCG ATTTGCTCTT ACAGATCCAT ATTGAAGAAG ATGAAATGTA TAGGCGCGAA 600
45 AAAGACGATA TTATCCAAAT CTTTGATTTA CCCTTAAAAA CGGCTCTTTT TGGAGGGAAA 660
ATTGAAATCG CTAATTGGCA TAAACCTTA ACCCTAACCA TTCCCCCTAA CACCAAAG 720
ATGCAAAAAT TCCGCATCAA AGACAAAGGG ATCAAAAGCA GAAAACTTC GCATGTGGGG 780
GATTGTATTG CAAGCTCGTT TGATCTGCTA AAATTGAAAC GCTTCTAA 828

```

50 (2) INFORMATION FOR SEQ ID NO:89:

```

(i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 837 base pairs
    (B) TYPE: nucleic acid

```


- 173 -

(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

5 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

10 (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

15 (A) NAME/KEY: misc_feature

(B) LOCATION 1...837

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:89:

20	ATGAGTAAGA	GTTTATACCA	AACTTTAAAC	GTGAGCGAAA	ACGCCAGCCA	AGATGAAATC	60
	AAAAAATCCT	ACCGCCGTTT	AGCCAGGCAA	TACCACCCGG	ATTGGAATAA	AACCAAAGAA	120
	GCCGAAGAGA	AATTCAAAGA	AATCAACGCC	GCTTATGAAA	TTTGAGCGA	TGAAGAAAAA	180
	CGCCGCCAAT	ACGATCAATT	TGGCGACAAC	ATGTTTGGCG	GGCAGAATTT	CAGCGATTTT	240
	GCCAGAAGCC	GTGGTCCTAG	TGAAGATTTA	GATGATATTT	TAAGCTCTAT	TTTTGGGAAA	300
	GGAGGCTTTT	CGCAAAGATT	TTCTCAAAAT	TCGCAAGGCT	TTTCTGGCTT	TAATTTTTC	360
25	AATTTTCGCCC	CTGAAAATTT	AGATGTAACC	GCTATTTTAA	ATGTCTCTGT	TTTAGACACC	420
	CTTTTAGGCA	ATAAAAAACA	AGTGAGCGTC	AATAATGAGA	CTTTTAGCCT	TAAAATCCCT	480
	ATCGGCGTGG	AAGAGGGCGA	AAAGATTAGG	GTTTCGCAACA	AAGGGAAAAT	GGGGCGAACG	540
	GGTAGGGGCG	ATTTGCTCTT	ACAGATCCAT	ATTGAAGAAG	ATGAAATGTA	TAGGCGCGAA	600
	AAAGACGATA	TTATCCAAAT	CTTTGATTTA	CCCTTAAAAA	CGGCTCTTTT	TGGAGGGAAA	660
30	ATTGAAATCG	CTACTTGGCA	TAAAACCTTA	ACCCTAACCA	TTCCCCCTAA	CACCAAAGCC	720
	ATGCAAAAAT	TCCGCATCAA	AGACAAAGGG	ATCAAAAAGCA	GAAAAACTTC	GCATGTGGGG	780
	GATTGTATTG	CAAGCTCGTT	TGATCTGCCT	AAAATTGAAA	CGTTTCTAAT	GAGTTGA	837

(2) INFORMATION FOR SEQ ID NO:90:

35

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 699 base pairs

(B) TYPE: nucleic acid

40 (C) STRANDEDNESS: double

(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

45 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

50

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...699

- 174 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:90:

```

5  GTGGTTCAAA AATTAAATTT TTATAAGACA GGTGGCATGC GTTTAAAACA TTTTAAGACA      60
   TTCCTTTTTA TCACAATGGC GGTGATTGTG ATAGGCACTG GTTGTGCGAA TAAAAAGAAA      120
   AAAAAAGATG AATACAACAA ACCGGCGATC TTTTGGTATC AAGGGATTTT GAGAGAAATT      180
   CTTTTTGCTA ATTTAGAAAC AGCGGACAAT TACTATTCTT CCTTACAGAG CGAACACATC      240
   AATTCCCCCC TTGTCCCAGA AGCTATGCTA GCTTTAGGGC AAGCGCACAT GAAAAAGAAA      300
   GAGTATGTTT TAGCGTCTTT TTACTTTGAT GAATACATCA AGCGCTTTGG GACGAAGGAC      360
   AATGTGGATT ATTTGACCTT TTTGAAACTG CAATCGCATT ATTACGCTTT CAAAAACCAT      420
10 TCTAAAGACC AGGAATTTAT CTCTAATTCT ATTGTGAGTT TAGGCGAATT TATAGAAAAA      480
   TACCCTAACA GCCGTTACCG CCCCTATGTA GAATACATGC AAATCAAATT CATTTTAGGG      540
   CAAAATGAGC TCAATCGCGC GATCGCGAAT GTCTATAAAA AACGCCACAA GCCCGAGGGC      600
   GTGAAACGCT ATTTAGAAAAG GATAGATGAG ACTTTAGAAA AAGAGACTAA ACCCAAACCA      660
15 TCGCACATGC CTTGGTATGT GTTAATTTTT GATTGGTAG                                699

```

(2) INFORMATION FOR SEQ ID NO:91:

(i) SEQUENCE CHARACTERISTICS:

```

20 (A) LENGTH: 345 base pairs
   (B) TYPE: nucleic acid
   (C) STRANDEDNESS: double
   (D) TOPOLOGY: circular

```

(ii) MOLECULE TYPE: DNA (genomic)

25

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

30

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

```

35 (A) NAME/KEY: misc_feature
   (B) LOCATION 1...345

```

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:91:

```

40 ATGCGTTTTT TGAATAACAA ACATAGAGAA AAGGGCTTAA AGGCTGAAGA AGAAGCTTGC      60
   GGGTTTTTAA AAACGCTGGG TTTTGAAATG ATAGAGAGGA ACTTTTTTTC ACAATTTGGT      120
   GAAATTGATA TTATCGCTTT GAAAAAAGGG GTTTTGCATT TCATTGAAGT CAAAAGCGGG      180
   GAAAATTTTG ATCCCATTTA TGCGATCACG CCGAGCAAAT TAAAAAGAT GATTAAACG      240
   ATCCGCTGTT ATTTGTCTCA AAAAGATCCC AATAGCGATT TTTGCATTGA CGCTCTTATT      300
45 GTGAAAAATG GTAAATTTGA GCTTTTAGAA AATATCACTT TTTAG                                345

```

(2) INFORMATION FOR SEQ ID NO:92:

(i) SEQUENCE CHARACTERISTICS:

```

50 (A) LENGTH: 306 base pairs
   (B) TYPE: nucleic acid
   (C) STRANDEDNESS: double
   (D) TOPOLOGY: circular

```

(ii) MOLECULE TYPE: DNA (genomic)

- 175 -

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

5 (vi) ORIGINAL SOURCE:
(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

10 (A) NAME/KEY: misc_feature
(B) LOCATION 1...306

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:92:

15 ATGGGCAGCA TTGGGGCTAT GACTAAAGGG AGCTCTGATA GGTATTTTCA AGAGGGCGTG 60
GCGAGTGAAA AATTAGTCCC AGAAGGCATT GAGGGGCGTG TGCCTTATCG TGGTAAGGTT 120
TCGGATATGA TTTTCCAATT AGTAGGGGGC GTGCGTTCTT CTATGGGGTA TCAGGGGGCG 180
AAGAATATTT TGGGAATTGTA TCAAAACGCT GAATTTGTAG AAATCACTAG CGCGGGGTTA 240
AAAAAAGCC ATGTGCATGG CGTGGATATT ACTAAAGAAG CCCCTAATAT TATGGGTGAA 300
20 TTTTAA 306

(2) INFORMATION FOR SEQ ID NO:93:

(i) SEQUENCE CHARACTERISTICS:

25 (A) LENGTH: 1446 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

30 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

35 (vi) ORIGINAL SOURCE:
(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

40 (A) NAME/KEY: misc_feature
(B) LOCATION 1...1446

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:93:

45 ATGAGAATTT TACAAAGGGC TTTGACTTTT GAAGACGTGT TGATGGTGCC TAGAAAATCC 60
AGCGTTTTAC CTAAAGATGT GAGCTTAAAG TCTCGCCTAA CCAAAAACAT TGGTTTGAAT 120
ATCCCTTTTA TTAGTGCGGC TATGGATACG GTTACAGAGC ATAAAACCGC TATCGCTATG 180
GCGCGCCTTG GGGGTATTGG CATCGTGCAT AAAAACATGG ATATTCAAAC GCAAGTCAAA 240
GAAATCACTA AAGTTAAAAA AAGCGAGAGC GGGGTGATTA ATGATCCTAT TTTTATCCAT 300
50 GCGCACAGGA CGCTAGCGGA CGCTAAAGTC ATAACGGATA ATTATAAGAT TTCAGGCGTG 360
CCTGTGGTAG ATGATAAGGG GTTGTGTGATT GGGATTTTAA CCAACAGAGA CGTGCGTTTT 420
GAAACCGATT TGAGTAAAAA AGTGGGCGAT GTGATGACTA AAATGCCTTT AGTTACCGCT 480
CATGTGGGCA TTAGCTTAGA TGAAGCGAGC GATTTGATGC ACAAGCATAA GATTGAAAAA 540
TTGCCCATTG TGGATAAAGA TAATGTTTTA AAAGGCTTGA TCACGATCAA AGACATTCAA 600

- 176 -

```

5  AAACGCATTG AATACCCTGA GGCCAATAAA GATGATTTTG GGAGGTTGAG AGTGGGGGCG 660
   GCTATTGGAG TGGGGCAGTT GGATAGGGCT GAAATGTTAG TTAAAGCGGG GGTGGATGCG 720
   TTGGTGTTAG ACAGCGCGCA TGGGCATTCA GCCAATATTT TACACACTTT AGAAGAGATT 780
   AAAAAAAGCT TGGTAGTGGA TGTGATTGTG GGAATGTGG TTAATAAAGA AGCCACAAGC 840
10 GATTTGATTA GCGCGGGAGC GGACGCTGTT AAAGTGGGTA TTGGGCCAGG AAGCATTTCG 900
   ACCACTAGGA TTGTGGCCGG GGTGGGAATG CCCCAGTGA GCGCAATTGA TAATTGCGTG 960
   GAAGTGGCGT CTAAATTTGA TATTCCTGTG ATTGCCGATG GAGGGATCCG CTATTCAGGC 1020
   GATGTGGCTA AGGCTCTAGC TTTAGGAGCA TCAAGCGTGA TGATAGGCTC TTTACTCGCT 1080
   GGCACAGAAG AATCTCCAGG GGATTTTATG ATTTACCAAG GGAGGCAATA TAAAAGCTAT 1140
15 AGGGGCATGG GCAGCATTGG GGCTATGACT AAAGGGAGCT CTGATAGGTA TTTTCAAGAG 1200
   GCGGTGGCGA GTGAAAATT AGTCCCAGAA GGCATTGAGG GCGGTGTGCC TTATCGTGGT 1260
   AAGGTTTCGG ATATGATTTT CCAATTAGTA GGGGGCGTGC GTTCTTCTAT GGGGTATCAG 1320
   GGGGCGAAGA ATATTTTGGA ATTGTATCAA AACGCTGAAT TTGTAGAAAT CACTAGCGCG 1380
   GGGTTAAAAG AAAGCCATGT GCATGGCGTG GATATTACTA AAGAAGCCCC TAATTATTAT 1440
   GGGTGA 1446

```

(2) INFORMATION FOR SEQ ID NO:94:

```

20 (i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 615 base pairs
    (B) TYPE: nucleic acid
    (C) STRANDEDNESS: double
    (D) TOPOLOGY: circular

25 (ii) MOLECULE TYPE: DNA (genomic)

    (iii) HYPOTHETICAL: NO

    (iv) ANTI-SENSE: NO

30 (vi) ORIGINAL SOURCE:
    (A) ORGANISM: Helicobacter pylori

    (ix) FEATURE:
35 (A) NAME/KEY: misc_feature
    (B) LOCATION 1...615

```

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:94:

```

40 ATGCAAGGGT TTCTTTTACA AACACAAAGC ATAAGAGATG AAGATTGAT CGTGACGTT 60
   TTAACCAAAA ACCAGCTCAA AACCTCTAT CGTTTCTATG GCAAACGCCA CAGCGTGCTG 120
   AATGTGGGTC GTAAAATTGA TTTTGAAGAA GAAACGATG ATAAATTTT ACCCAAGTTA 180
   AGGAATATTT TGCATTTAGG CTATATTTGG GAAAGAGAAA TGGAGCGCTT GTTTTTTTGG 240
   CAACGCTTTT GCGCTCTTTT GTTCAAGCAT TTAGAGGGCG TGCATTCTTT AGATAGCATC 300
45 TATTTTGACA CTTTAGATGA TGGGGCTAGC AAATCTCTCA AACAGCACCC CTTAAGAGTG 360
   ATTTTAGAAA TGTATGCAGT CCTTTTGAAT TTTGAAGGGC GCTTGCAAAG .T.CAATTCT 420
   TGTTTTTTAT GCGATGCAAA ATTAGAGCGT TCTGTCGCTT TAGCGCAAGG GTTTATTTTA 480
   GCGACCCCTT CTTGCTTGAA AGCTAAAAGC TTGGATTAG AAAAAATCCA AGCTTTTTTC 540
   CGCACTCAAA GCACGATTGA TCTAGAAACA GAAGAAGTGG AAGAATTATG GCGCACGCTG 600
50 AATTTAGGGT TTTGA 615

```

(2) INFORMATION FOR SEQ ID NO:95:

(i) SEQUENCE CHARACTERISTICS:

- 177 -

- (A) LENGTH: 249 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

5

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

10

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

15

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...249

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:95:

20

ATGGGCGTCG	GACGGGTCGG	CAATATGGCA	CTGTTGGCGT	GTGCAGGTCC	GATGGGCATC	60
GGCGCTATTG	CTATCGCCAT	TAACGGCGGC	AGACAACGGT	CGCGGATGTT	GGTGGTCGAT	120
ATAGACGACA	AACGTCTGGA	GCAGGTACAG	AAGATGCTGC	CGGGGAATTG	GCGGCCAGTA	180
ACGGCATTGA	GCTGGTGTCT	GTGCATACCA	AAGCGAGGAG	CGATCCGTGC	CAGATGCTGC	240
GAGCGCTGA						249

25

(2) INFORMATION FOR SEQ ID NO:96:

(i) SEQUENCE CHARACTERISTICS:

30

- (A) LENGTH: 204 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

35

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

40

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

45

(A) NAME/KEY: misc_feature

(B) LOCATION 1...204

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:96:

50

TTGTCCGGTA	CAGCCGTGAG	TTGCCGGTGC	ACATGCCGCA	TACAGTTGGT	ATTGGTGCGC	60
ACCAGCATCC	CGGTTGTTAT	CGGGTGCTCA	TGCCATTCC	TTTCCAGTAT	TGGGTTTACA	120
ACGGGAACCC	ACCAATCACC	CGTTAAACGC	TGCGGGGTTA	ACGCCGGA	AACACCGTCA	180
AAAAAACATT	TGCATTTAAA	CTAA				204

- 178 -

(2) INFORMATION FOR SEQ ID NO:97:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 345 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
 (B) LOCATION 1...345

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:97:

```

GTGTGGCTGG CGGCGCTGGG CTTCTGATC ACCGCGGTGG GGCTGCCGGT GATCACCGTG      60
ATCGCCCTGG CCAAGGTCGG CGGTTCTGTC ACGCCCTCAG CCATCCGATC GGCAGGTATG     120
CCGGCGGCCT GCTGGCGGCG GTCTGTCTACC TGGCGGTTCG CCCGCTGTTC GCCATTCCGC     180
GCACCGCCAC GGTGTCCTTC GAAGGTCAGC GTGGTGCCGC TGCTCGGCGA AGAAGCGGCA     240
CGGCGCTGTT CGTCTACAGC CTGGCGTACT TCCTCCTCGC CCTGGCCATC TCCCTCTACC     300
CCGGTCGCCT GCTGGACACC GTCGGACGCT TCCTCGCCCC GCTGA                      345

```

(2) INFORMATION FOR SEQ ID NO:98:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 228 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
 (B) LOCATION 1...228

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:98:

```

Met Arg Phe Lys Gly Ser Arg Val Glu Ala Phe Leu Gly Ala Leu Glu
1           5           10           15
Phe Gln Glu Asn Glu Tyr Glu Glu Phe Lys Glu Leu Tyr Glu Ser Leu
          20           25           30

```

- 179 -

Lys Thr Lys Gln Lys Pro His Thr Leu Phe Ile Ser Cys Val Asp Ser
 35 40 45
 Arg Val Val Pro Asn Leu Ile Thr Gly Thr Gln Pro Gly Glu Leu Tyr
 50 55 60
 5 Val Ile Arg Asn Met Gly Asn Val Ile Pro Pro Lys Thr Ser Tyr Lys
 65 70 75 80
 Glu Ser Leu Ser Thr Ile Ala Ser Val Glu Tyr Ala Ile Ala His Val
 85 90 95
 10 Gly Val Gln Asn Leu Ile Ile Cys Gly His Ser Asp Cys Gly Ala Cys
 100 105 110
 Gly Ser Ile His Leu Ile His Asp Glu Thr Thr Lys Ala Lys Thr Pro
 115 120 125
 Tyr Ile Ala Asn Trp Ile Gln Phe Leu Glu Pro Ile Lys Glu Glu Leu
 130 135 140
 15 Lys Asn His Pro Gln Phe Ser Asn His Phe Ala Lys Arg Ser Trp Leu
 145 150 155 160
 Thr Glu Arg Leu Asn Ala Arg Leu Gln Leu Asn Asn Leu Leu Ser Tyr
 165 170 175
 20 Asp Phe Ile Gln Glu Arg Val Ile Asn Asn Glu Leu Lys Ile Phe Gly
 180 185 190
 Trp His Tyr Ile Ile Glu Thr Gly Arg Ile Tyr Asn Tyr Asn Phe Glu
 195 200 205
 Ser His Phe Phe Glu Pro Ile Glu Glu Thr Ile Lys Gln Arg Ile Ser
 210 215 220
 25 His Glu Asn Phe
 225

(2) INFORMATION FOR SEQ ID NO:99:

- 30 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 221 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 35 (ii) MOLECULE TYPE: protein
 (iii) HYPOTHETICAL: YES
 (vi) ORIGINAL SOURCE:
 40 (A) ORGANISM: *Helicobacter pylori*
 (ix) FEATURE:
 (A) NAME/KEY: misc_feature
 (B) LOCATION 1...221
 45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:99:

Val Glu Ala Phe Leu Gly Ala Leu Glu Phe Gln Glu Asn Glu Tyr Glu
 1 5 10 15
 50 Glu Phe Lys Glu Leu Tyr Glu Ser Leu Lys Thr Lys Gln Lys Pro His
 20 25 30
 Thr Leu Phe Ile Ser Cys Val Asp Ser Arg Val Val Pro Asn Leu Ile
 35 40 45
 Thr Gly Thr Gln Pro Gly Glu Leu Tyr Val Ile Arg Asn Met Gly Asn

- 180 -

```

      50      55      60
Val Ile Pro Pro Lys Thr Ser Tyr Lys Glu Ser Leu Ser Thr Ile Ala
65      70      75      80
Ser Val Glu Tyr Ala Ile Ala His Val Gly Val Gln Asn Leu Ile Ile
5      85      90      95
Cys Gly His Ser Asp Cys Gly Ala Cys Gly Ser Ile His Leu Ile His
      100      105      110
Asp Glu Thr Thr Lys Ala Lys Thr Pro Tyr Ile Ala Asn Trp Ile Gln
      115      120      125
10  Phe Leu Glu Pro Ile Lys Glu Glu Leu Lys Asn His Pro Gln Phe Ser
      130      135      140
Asn His Phe Ala Lys Arg Ser Trp Leu Thr Glu Arg Leu Asn Ala Arg
145      150      155      160
Leu Gln Leu Asn Asn Leu Leu Ser Tyr Asp Phe Ile Gln Glu Arg Val
15      165      170      175
Ile Asn Asn Glu Leu Lys Ile Phe Gly Trp His Tyr Ile Ile Glu Thr
      180      185      190
Gly Arg Ile Tyr Asn Tyr Asn Phe Glu Ser His Phe Phe Glu Pro Ile
      195      200      205
20  Glu Glu Thr Ile Lys Gln Arg Ile Ser His Glu Asn Phe
      210      215      220

```

(2) INFORMATION FOR SEQ ID NO:100:

```

25      (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 335 amino acids
          (B) TYPE: amino acid
          (D) TOPOLOGY: linear

```

```

30      (ii) MOLECULE TYPE: protein

```

```

      (iii) HYPOTHETICAL: YES

```

```

      (vi) ORIGINAL SOURCE:
35      (A) ORGANISM: Helicobacter pylori

```

```

      (ix) FEATURE:
          (A) NAME/KEY: misc_feature
          (B) LOCATION 1...335

```

```

40      (xi) SEQUENCE DESCRIPTION: SEQ ID NO:100:

```

```

Met Leu Val Thr Arg Phe Lys Lys Ala Phe Ile Ser Tyr Ser Leu Gly
1      5      10      15
45 Val Leu Val Val Ser Leu Leu Leu Asn Val Cys Asn Ala Ser Ala Gln
      20      25      30
Glu Val Lys Val Lys Asp Tyr Phe Gly Glu Gln Thr Ile Lys Leu Pro
      35      40      45
Val Ser Lys Ile Ala Tyr Ile Gly Ser Tyr Val Glu Val Pro Ala Met
50      50      55      60
Leu Asn Val Trp Asp Arg Val Val Gly Val Ser Asp Tyr Ala Phe Lys
65      70      75      80
Asp Asp Ile Val Lys Ala Thr Leu Lys Gly Glu Asp Leu Lys Arg Val
      85      90      95

```


- 181 -

```

      Lys His Met Ser Thr Asp His Thr Ala Ala Leu Asn Val Glu Leu Leu
      100                      105                      110
      Lys Lys Leu Ser Pro Asp Leu Val Val Thr Phe Val Gly Asn Pro Lys
      115                      120                      125
5     Ala Val Glu His Ala Lys Lys Phe Gly Ile Ser Phe Leu Ser Phe Gln
      130                      135                      140
      Glu Thr Thr Ile Ala Glu Ala Met Gln Ala Met Gln Ala Gln Ala Thr
      145                      150                      155                      160
      Val Leu Glu Ile Asp Ala Ser Lys Lys Phe Ala Lys Met Gln Glu Thr
10    165                      170                      175
      Leu Asp Phe Ile Ala Glu Arg Leu Lys Gly Val Lys Lys Lys Lys Gly
      180                      185                      190
      Val Glu Leu Phe His Lys Ala Asn Lys Ile Ser Gly His Gln Ala Ile
      195                      200                      205
15    Ser Ser Asp Ile Leu Glu Lys Gly Gly Ile Asp Asn Phe Gly Leu Lys
      210                      215                      220
      Tyr Val Lys Phe Gly Arg Ala Asp Ile Ser Val Glu Lys Ile Val Lys
      225                      230                      235                      240
      Glu Asn Pro Glu Ile Ile Phe Ile Trp Trp Val Ser Pro Leu Thr Pro
20    245                      250                      255
      Glu Asp Val Leu Asn Asn Pro Lys Phe Ser Thr Ile Lys Ala Ile Lys
      260                      265                      270
      Asn Lys Gln Val Tyr Lys Leu Pro Thr Met Asp Ile Gly Gly Pro Arg
      275                      280                      285
25    Ala Pro Leu Ile Ser Leu Phe Ile Ala Leu Lys Ala His Pro Glu Ala
      290                      295                      300
      Phe Lys Gly Val Asp Ile Asn Ala Ile Val Lys Asp Tyr Tyr Lys Val
      305                      310                      315                      320
      Val Phe Asp Leu Asn Asp Ala Glu Ile Glu Pro Phe Leu Trp His
30    325                      330                      335

```

(2) INFORMATION FOR SEQ ID NO:101:

- (i) SEQUENCE CHARACTERISTICS:
- 35 (A) LENGTH: 274 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- 40 (iii) HYPOTHETICAL: YES
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: *Helicobacter pylori*
- 45 (ix) FEATURE:
- (A) NAME/KEY: misc_feature
- (B) LOCATION 1...274
- 50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:101:

```

Met Leu Val Thr Arg Phe Lys Lys Ala Phe Ile Ser Tyr Ser Leu Gly
1      5      10      15
Val Leu Val Val Ser Leu Leu Leu Asn Val Cys Asn Ala Ser Ala Gln

```

- 182 -

```

      20      25      30
Glu Val Lys Val Lys Asp Tyr Phe Gly Glu Gln Thr Ile Lys Leu Pro
      35      40      45
Val Ser Lys Ile Ala Tyr Ile Gly Ser Tyr Val Glu Val Pro Ala Met
5      50      55      60
Leu Asn Val Trp Asp Arg Val Val Gly Val Ser Asp Tyr Ala Phe Lys
      65      70      75      80
Asp Asp Ile Val Lys Ala Thr Leu Lys Gly Glu Asp Leu Lys Arg Val
      85      90      95
10 Lys His Met Ser Thr Asp His Thr Ala Ala Leu Asn Val Glu Leu Leu
      100      105      110
Lys Lys Leu Ser Pro Asp Leu Val Val Thr Phe Val Gly Asn Pro Lys
      115      120      125
15 Ala Val Glu His Ala Lys Lys Phe Gly Ile Ser Phe Leu Ser Phe Gln
      130      135      140
Glu Thr Thr Ile Ala Glu Ala Met Gln Ala Met Gln Ala Gln Ala Thr
      145      150      155      160
Val Leu Glu Ile Asp Ala Ser Lys Lys Phe Ala Lys Met Gln Glu Thr
      165      170      175
20 Leu Asp Phe Ile Ala Asp Arg Leu Lys Gly Val Lys Lys Lys Lys Gly
      180      185      190
Val Glu Leu Phe His Lys Ala Asn Lys Ile Ser Gly His Gln Ala Ile
      195      200      205
Asn Ser Asp Ile Leu Gln Gln Gly Gly Ile Asp Asn Phe Gly Leu Lys
25      210      215      220
Tyr Val Lys Phe Gly Arg Ala Asp Ile Ser Val Glu Lys Ile Val Lys
      225      230      235      240
Glu Asn Pro Glu Ile Ile Phe Ile Arg Trp Val Thr Pro Leu Thr Pro
      245      250      255
30 Asp Tyr Val Leu Asn Asn Pro Lys Phe Ser Thr Ile Asn Ala Ile Lys
      260      265      270
Asn Ile

```

35 (2) INFORMATION FOR SEQ ID NO:102:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 428 amino acids

(B) TYPE: amino acid

40 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

45

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

50

(A) NAME/KEY: misc_feature

(B) LOCATION 1...428

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:102:

- 183 -

Met Lys Lys Lys Phe Leu Ser Leu Thr Leu Gly Ser Leu Leu Val Ser
 1 5 10 15
 Ala Leu Ser Ala Glu Asp Asn Gly Phe Phe Val Ser Ala Gly Tyr Gln
 20 25 30
 5 Ile Gly Glu Ser Ala Gln Met Val Lys Asn Thr Lys Gly Ile Gln Asp
 35 40 45
 Leu Ser Asp Ser Tyr Glu Arg Leu Asn Asn Leu Leu Thr Asn Tyr Ser
 50 55 60
 Val Leu Asn Ala Leu Ile Arg Gln Ser Ala Asp Pro Asn Ala Ile Asn
 10 65 70 75 80
 Asn Ala Arg Gly Asn Leu Asn Ala Ser Ala Lys Asn Leu Ile Asn Asp
 85 90 95
 Lys Lys Asn Ser Pro Ala Tyr Gln Ala Val Leu Leu Ala Leu Asn Ala
 100 105 110
 15 Ala Ala Gly Leu Trp Gln Val Met Ser Tyr Ala Ile Ser Pro Cys Gly
 115 120 125
 Pro Gly Lys Asp Thr Ser Lys Asn Gly Gly Val Gln Thr Phe His Asn
 130 135 140
 Thr Pro Ser Asn Gln Trp Gly Gly Thr Thr Ile Thr Cys Gly Thr Thr
 20 145 150 155 160
 Gly Tyr Glu Pro Gly Pro Tyr Ser Ile Leu Ser Thr Glu Asn Tyr Ala
 165 170 175
 Lys Ile Asn Lys Ala Tyr Gln Ile Ile Gln Lys Ala Phe Gly Ser Ser
 180 185 190
 25 Gly Lys Asp Ile Pro Ala Leu Ser Asp Thr Asn Thr Glu Leu Lys Phe
 195 200 205
 Thr Ile Asn Lys Asn Asn Gly Asn Thr Asn Thr Asn Asn Asn Gly Glu
 210 215 220
 Glu Ile Val Thr Lys Asn Asn Ala Gln Val Leu Leu Glu Gln Ala Ser
 30 225 230 235 240
 Thr Ile Ile Thr Thr Leu Asn Ser Ala Cys Pro Trp Ile Asn Asn Gly
 245 250 255
 Gly Ala Gly Gly Ala Ser Ser Gly Ser Leu Trp Glu Gly Ile Tyr Leu
 260 265 270
 35 Lys Gly Asp Gly Ser Ala Cys Gly Ile Phe Lys Asn Glu Ile Ser Ala
 275 280 285
 Ile Gln Asp Met Ile Lys Asn Ala Ala Ile Ala Val Glu Gln Ser Lys
 290 295 300
 Ile Val Ala Ala Asn Ala Gln Asn Gln Arg Asn Leu Asp Thr Gly Lys
 40 305 310 315 320
 Thr Phe Asn Pro Tyr Lys Asp Ala Asn Phe Ala Gln Ser Met Phe Ala
 325 330 335
 Asn Ala Lys Ala Gln Ala Glu Ile Leu Asn Arg Ala Gln Ala Val Val
 340 345 350
 45 Lys Asp Phe Glu Arg Ile Pro Ala Glu Phe Val Lys Asp Ser Leu Gly
 355 360 365
 Val Cys His Glu Val Gln Asn Gly His Leu Arg Gly Thr Pro Ser Gly
 370 375 380
 Thr Val Thr Asp Asn Thr Trp Gly Ala Gly Cys Ala Tyr Val Gly Glu
 50 385 390 395 400
 Thr Val Thr Asn Leu Lys Asp Ser Ile Ala His Phe Gly Asp Gln Ala
 405 410 415
 Glu Arg Ile His Asn Ala Arg Asn Leu Ala Thr Leu
 420 425

- 184 -

(2) INFORMATION FOR SEQ ID NO:103:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 178 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
 (B) LOCATION 1...178

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:103:

```

Met Asn Pro Leu Leu Gln Asp Tyr Ala Arg Ile Leu Leu Glu Trp Asn
1           5           10           15
Gln Thr His Asn Leu Ser Gly Ala Arg Asn Leu Ser Glu Leu Glu Pro
25          20          25          30
Gln Ile Thr Asp Ala Leu Lys Pro Leu Glu Phe Val Lys Asp Phe Lys
35          40          45
Ser Cys Leu Asp Ile Gly Ser Gly Ala Gly Leu Pro Ala Ile Pro Leu
50          55          60
Ala Leu Glu Lys Pro Glu Ala Gln Phe Ile Leu Leu Glu Pro Arg Val
30          65          70          75          80
Lys Arg Ala Ala Phe Leu Asn Tyr Leu Lys Ser Val Leu Pro Leu Asn
85          90          95
Asn Ile Glu Ile Ile Lys Lys Arg Leu Glu Asp Tyr Gln Asn Leu Leu
35          100         105         110
Gln Val Asp Leu Ile Thr Ser Arg Ala Val Ala Ser Ser Ser Phe Leu
115         120         125
Ile Glu Lys Ser Gln Arg Phe Leu Lys Asp Lys Gly Tyr Phe Leu Phe
130         135         140
40 Tyr Lys Gly Glu Gln Leu Lys Asn Glu Ile Ala Tyr Lys Thr Thr Glu
145         150         155         160
Cys Phe Met His Gln Lys Arg Val Tyr Phe Tyr Lys Ser Lys Glu Ser
165         170         175
Leu Cys

```

(2) INFORMATION FOR SEQ ID NO:104:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 240 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

- 185 -

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

5 (A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...240

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:104:

```

Leu Gly Leu Lys Lys Arg Ala Ile Leu Trp Ser Leu Met Gly Phe Cys
1           5           10           15
15 Ala Gly Leu Ser Ala Leu Asp Tyr Asp Thr Leu Asp Pro Lys Tyr Tyr
    20           25           30
    Lys Tyr Ile Lys Tyr Tyr Lys Ala Tyr Glu Asp Lys Glu Val Glu Glu
    35           40           45
    Leu Ile Arg Asp Leu Lys Arg Ala Asn Ala Lys Ser Gly Leu Ile Leu
    50           55           60
    Gly Ile Asn Thr Gly Phe Phe Tyr Asn His Glu Ile Met Val Lys Thr
    65           70           75           80
    Asn Ser Ser Ser Ile Thr Gly Asn Ile Leu Asn Tyr Leu Phe Ala Tyr
    85           90           95
25 Gly Leu Arg Phe Gly Tyr Gln Thr Phe Arg Pro Ser Phe Phe Ala Arg
    100          105          110
    Leu Val Lys Pro Asn Ile Ile Gly Arg Arg Ile Tyr Ile Gln Tyr Tyr
    115          120          125
    Gly Gly Ala Pro Lys Lys Ala Gly Phe Gly Ser Val Gly Phe Gln Ser
    130          135          140
30 Val Met Leu Asn Gly Asp Phe Leu Leu Asp Phe Pro Leu Pro Phe Val
    145          150          155          160
    Gly Lys Tyr Leu Tyr Met Gly Gly Tyr Met Gly Leu Gly Leu Gly Val
    165          170          175
35 Val Ala His Gly Val Asn Tyr Thr Ala Glu Trp Gly Met Ser Phe Asn
    180          185          190
    Ala Gly Leu Ala Leu Thr Val Leu Glu Lys Asn Arg Ile Glu Phe Glu
    195          200          205
    Phe Lys Ile Leu Asn Asn Phe Pro Phe Leu Gln Ser Asn Ser Ser Lys
    210          215          220
40 Glu Thr Trp Trp Gly Ala Ile Ala Ser Ile Gly Tyr Gln Tyr Val Phe
    225          230          235          240

```

(2) INFORMATION FOR SEQ ID NO:105:

45

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 313 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

50

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

- 186 -

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

5 (A) NAME/KEY: misc_feature

(B) LOCATION 1...313

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:105:

```

10  Leu Lys Leu Lys Tyr Trp Leu Val Tyr Leu Ala Phe Ile Ile Gly Leu
    1          5          10          15
    Gln Ala Thr Asp Tyr Asp Asn Leu Glu Glu Glu Asn Gln Gln Leu Asp
        20          25          30
15  Glu Lys Ile Asn Asn Leu Lys Arg Gln Leu Thr Glu Lys Gly Val Ser
    35          40          45
    Pro Lys Glu Met Asp Lys Asp Lys Phe Glu Glu Glu Tyr Leu Glu Arg
    50          55          60
    Thr Tyr Pro Lys Ile Ser Ser Lys Lys Arg Lys Lys Leu Leu Lys Ser
    65          70          75          80
20  Phe Ser Ile Ala Asp Asp Lys Ser Gly Val Phe Leu Gly Gly Gly Tyr
    85          90          95
    Ala Tyr Gly Glu Leu Asn Leu Ser Tyr Gln Gly Glu Met Leu Asp Arg
    100          105          110
25  Tyr Gly Ala Asn Ala Pro Ser Ala Phe Lys Asn Asn Ile Asn Ile Asn
    115          120          125
    Ala Pro Val Ser Met Ile Ser Val Lys Phe Gly Tyr Gln Lys Tyr Phe
    130          135          140
    Val Pro Tyr Phe Gly Thr Arg Phe Tyr Gly Asp Leu Leu Leu Gly Gly
    145          150          155          160
30  Gly Ala Leu Lys Glu Asn Ala Leu Lys Gln Pro Val Gly Ser Phe Phe
    165          170          175
    Tyr Val Leu Gly Ala Met Asn Thr Asp Leu Leu Phe Asp Met Pro Leu
    180          185          190
35  Asp Phe Lys Thr Lys Lys His Phe Leu Gly Val Tyr Ala Gly Phe Gly
    195          200          205
    Ile Gly Leu Met Leu Tyr Gln Asp Lys Pro Asn Gln Asn Gly Arg Asn
    210          215          220
    Leu Ile Val Gly Gly Tyr Ser Ser Pro Asn Phe Leu Trp Lys Ser Leu
    225          230          235          240
40  Ile Glu Val Asp Tyr Thr Phe Asn Val Gly Val Ser Leu Thr Leu Tyr
    245          250          255
    Arg Lys His Arg Leu Glu Ile Gly Thr Lys Leu Pro Ile Ser Tyr Leu
    260          265          270
45  Arg Met Gly Val Glu Glu Gly Ala Ile Tyr His Asn Lys Glu Asn Asp
    275          280          285
    Glu Arg Leu Leu Ile Ser Ala Asn Asn Gln Phe Lys Arg Ser Ser Phe
    290          295          300
    Leu Leu Val Asn Tyr Ala Phe Ile Phe
    305          310
50

```

(2) INFORMATION FOR SEQ ID NO:106:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 393 amino acids

- 187 -

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...393

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:106:

```

Met Thr Ser Ala Ser Ser His Ser Phe Lys Glu Gln Asp Phe His Ile
1          5          10          15
Pro Ile Ala Phe Ala Phe Asp Lys Asn Tyr Leu Ile Pro Ala Gly Ala
20          20          25          30
Cys Ile Tyr Ser Leu Leu Glu Ser Ile Ala Lys Ala Asn Lys Lys Ile
35          40          45
Arg Tyr Thr Leu His Ala Leu Val Val Gly Leu Asn Glu Glu Asp Lys
50          55          60
Thr Lys Leu Asn Gln Ile Thr Glu Pro Phe Lys Glu Phe Ala Val Leu
25 65          70          75          80
Glu Val Lys Asp Ile Glu Pro Phe Leu Asp Thr Ile Pro Asn Pro Phe
85          90          95
Asp Glu Asp Phe Thr Lys Arg Phe Ser Lys Met Val Leu Val Lys Tyr
30 100          105          110
Phe Leu Ala Asp Leu Phe Pro Lys Tyr Ser Lys Met Val Trp Ser Asp
115          120          125
Val Asp Val Ile Phe Cys Asn Glu Phe Ser Ala Asp Phe Leu Asn Ile
130          135          140
Lys Glu Asp Asp Glu Asn Tyr Phe Tyr Gly Val Tyr Asp Lys Ile Tyr
35 145          150          155          160
Pro Tyr Glu Gly Phe Phe Tyr Cys Asn Leu Thr Tyr Gln Arg Lys Asn
165          170          175
Gln Phe Cys Lys Lys Ile Leu Glu Ile Ile Arg Ala Gln Lys Ile Asp
40 180          185          190
Lys Glu Pro Gln Leu Thr Glu Phe Cys Arg Ser Lys Ile Ala Pro Leu
195          200          205
Lys Ile Glu Tyr Cys Ile Phe Pro His Tyr Tyr Ser Leu Ser Glu Glu
210          215          220
His Leu Lys Gly Val Ala Asn Ala Ile Tyr His Asn Thr Ile Lys Gln
45 225          230          235          240
Ala Leu Arg Glu Pro Ile Val Ile Gln Tyr Asp Ser His Pro Tyr Phe
245          250          255
Gln Ile Lys Pro Trp Thr Tyr Pro Phe Gly Leu Lys Ala Asp Leu Trp
50 260          265          270
Leu Asn Ala Leu Ala Lys Thr Pro Phe Met Ser Asp Trp Ser Tyr Leu
275          280          285
Ile Thr Gly Gly Gly Gly Ile Gly Gly Glu Lys Trp His Tyr Tyr His
290          295          300

```

- 188 -

Gly Ile Ala Ala Tyr His Tyr Tyr Phe Pro Leu Trp Lys Ala Glu Glu
 305 310 315 320
 Gln Ile Ala His Asp Ala Leu Lys Thr Phe Leu Lys His Tyr Phe Leu
 325 330 335
 5 His Ile His Glu Ile Pro Gln Asn Ala Arg Arg Arg Leu Phe Lys Tyr
 340 345 350
 Cys Ile Ser Ile Pro Leu Lys Ser Phe Ile Ser Lys Thr Leu Lys Phe
 355 360 365
 10 Leu Lys Leu His Ala Leu Val Lys Lys Ile Leu Ile Gln Leu Lys Leu
 370 375 380
 Leu Lys Lys Asn Gln Ser Gln Asn Phe
 385 390

(2) INFORMATION FOR SEQ ID NO:107:

15

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 435 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

20

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

25

- (vi) ORIGINAL SOURCE:
 (A) ORGANISM: *Helicobacter pylori*

30

- (ix) FEATURE:
 (A) NAME/KEY: misc_feature
 (B) LOCATION 1...435

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:107:

35 Leu Ile Phe Leu Lys Lys Ser Leu Cys Ala Leu Leu Ile Ser Gly Phe
 1 5 10 15
 Phe Ile Pro Pro Leu Met Lys Ala Ala Ser Phe Val Tyr Asp Leu Lys
 20 25 30
 Phe Met Ser Phe Asn Phe Asn Leu Ala Ser Pro Pro Asn Asn Pro Tyr
 35 40 45
 40 Trp Asn Ser Leu Thr Lys Met Gln Gly Arg Leu Met Pro Gln Ile Gly
 50 55 60
 Val Gln Leu Asp Lys Arg Gln Ala Leu Met Phe Gly Ala Trp Phe Ile
 65 70 75 80
 Gln Asn Leu His Thr His Tyr Ser Tyr Phe Pro Tyr Ser Trp Gly Val
 85 90 95
 45 Thr Met Tyr Tyr Gln Tyr Ile Gly Lys Asn Leu Arg Phe Phe Leu Gly
 100 105 110
 Ile Val Pro Arg Ser Tyr Gln Ile Gly His Tyr Pro Leu Ser Ala Phe
 115 120 125
 50 Lys Lys Leu Phe Trp Phe Ile Asp Pro Thr Phe Arg Gly Gly Ala Phe
 130 135 140
 Gln Phe Lys Pro Ala Tyr Asp Pro Asn Arg Trp Trp Asn Gly Trp Phe
 145 150 155 160
 Glu Gly Val Val Asp Trp Tyr Gly Gly Arg Asn Trp Asn Asn Gln Pro

- 189 -

```

                    165                    170                    175
Lys Lys Lys Asn Tyr Asp Phe Asp Gln Phe Leu Tyr Phe Val Ser Ser
                    180                    185                    190
5  Glu Phe Gln Phe Leu Lys Gly Tyr Leu Gly Leu Gly Gly Gln Leu Val
                    195                    200                    205
    Ile Phe His Asn Ala Asn Ser His Ser Met Gly Asp Asn Tyr Pro Tyr
                    210                    215                    220
    Gly Gly Asn Ser Tyr Leu Lys Pro Gly Asp Ala Thr Pro Gln Trp Pro
225                    230                    235                    240
10  Asn Gly Tyr Pro Tyr Phe Ser Gln Lys Asp Asn Pro Gln Gly Gly Glu
                    245                    250                    255
    Ile Gly Lys Tyr Ser Asn Pro Thr Ile Leu Asp Arg Val Tyr Tyr His
                    260                    265                    270
    Ala Tyr Leu Lys Ala Asp Phe Lys Asn Leu Met Pro Tyr Met Asp Asn
15  Ile Phe Met Thr Phe Gly Thr Gln Ser Ser Gln Thr His Tyr Cys Val
                    275                    280                    285
    Arg Tyr Ala Ser Glu Cys Lys Asn Ala Arg Phe Tyr Asn Ser Phe Gly
290                    295                    300                    305
    Gly Glu Phe Tyr Ala Gln Ala Gln Tyr Lys Gly Phe Gly Ile Phe Asn
310                    315                    320
20  Arg Tyr Tyr Phe Ser Asn Lys Pro Gln Met His Phe Tyr Ala Thr Tyr
                    325                    330                    335
    Gly Gln Ser Leu Tyr Thr Gly Leu Pro Trp Tyr Arg Ala Pro Asn Phe
25  Asp Met Ile Gly Leu Tyr Tyr Leu Tyr Lys Asn Lys Trp Leu Ser Val
                    340                    345                    350
    Arg Ala Asp Ala Phe Phe Ser Phe Val Gly Gly Gly Asp Gly Tyr His
355                    360                    365                    370
    Leu Tyr Gly Lys Gly Gly Lys Trp Phe Val Met Tyr Gln Gln Phe Leu
375                    380                    385                    390
    Thr Leu Thr Ile Asp Thr Arg Glu Leu Ile Asp Phe Val Lys Ser Lys
405                    410                    415                    420
    Ile Pro Lys
425                    430                    435
35  Ile Pro Lys
435

```

(2) INFORMATION FOR SEQ ID NO:108:

- 40 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 220 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
- 45 (ii) MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: YES
- (vi) ORIGINAL SOURCE:
 (A) ORGANISM: *Helicobacter pylori*
- 50 (ix) FEATURE:
 (A) NAME/KEY: misc_feature
 (B) LOCATION 1...220

- 190 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:108:

```

Met Asn Lys Thr Thr Ile Lys Ile Leu Met Gly Met Ala Leu Leu Ser
1           5           10           15
5 Ser Leu Gln Ala Ala Glu Ala Glu Leu Asp Glu Lys Ser Lys Lys Pro
           20           25           30
Lys Phe Ala Asp Arg Asn Thr Phe Tyr Leu Gly Val Gly Tyr Gln Leu
           35           40           45
10 Ser Ala Ile Asn Thr Ser Phe Ser Thr Ser Ser Ile Asp Lys Ser Tyr
           50           55           60
Phe Met Thr Gly Asn Gly Phe Gly Val Val Leu Gly Gly Lys Phe Val
65           70           75           80
Ala Lys Thr Gln Ala Val Glu His Val Gly Phe Arg Tyr Gly Leu Phe
           85           90           95
15 Tyr Asp Gln Thr Phe Ser Ser His Lys Ser Tyr Ile Ser Thr Tyr Gly
           100          105          110
Leu Glu Phe Ser Gly Leu Trp Asp Ala Phe Asn Ser Pro Lys Met Phe
           115          120          125
20 Leu Gly Leu Glu Phe Gly Leu Gly Ile Ala Gly Ala Thr Tyr Met Pro
           130          135          140
Gly Gly Ala Met His Gly Ile Ile Ala Gln Tyr Leu Gly Lys Glu Asn
145          150          155          160
Ser Leu Phe Gln Leu Leu Val Lys Val Gly Phe Arg Phe Gly Phe Phe
           165          170          175
25 His Asn Glu Ile Thr Phe Gly Leu Lys Phe Pro Val Ile Pro Asn Lys
           180          185          190
Lys Thr Glu Ile Val Asp Gly Leu Ser Ala Thr Thr Leu Trp Gln Arg
           195          200          205
30 Leu Pro Val Ala Tyr Phe Asn Tyr Ile Tyr Asn Phe
           210          215          220

```

(2) INFORMATION FOR SEQ ID NO:109:

```

35 (i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 116 amino acids
    (B) TYPE: amino acid
    (D) TOPOLOGY: linear

40 (ii) MOLECULE TYPE: protein

    (iii) HYPOTHETICAL: YES

    (vi) ORIGINAL SOURCE:
        (A) ORGANISM: Helicobacter pylori

45 (ix) FEATURE:
    (A) NAME/KEY: misc_feature
    (B) LOCATION 1...116

```

50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:109:

```

Leu Asn Leu His Phe Met Lys Gly Phe Val Met Ser Gly Leu Arg Thr
1           5           10           15
Phe Ser Cys Val Val Val Leu Cys Gly Ala Met Val Asn Val Ala Val

```

- 191 -

```

                20                25                30
Ala Gly Pro Lys Ile Glu Ala Arg Gly Glu Leu Gly Lys Phe Val Gly
                35                40                45
Gly Ala Val Gly Asn Phe Val Gly Asp Lys Met Gly Gly Phe Val Gly
5      50                55                60
Gly Ala Ile Gly Gly Tyr Ile Gly Ser Glu Val Gly Asp Arg Val Glu
65                70                75                80
Asp Tyr Ile Arg Gly Val Asp Arg Glu Pro Gln Asn Lys Glu Pro Gln
                85                90                95
10    Thr Pro Arg Glu Pro Ile Arg Asp Phe Tyr Asp Tyr Gly Tyr Ser Phe
                100               105               110
Gly His Ala Trp
                115

```

15 (2) INFORMATION FOR SEQ ID NO:110:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 436 amino acids

(B) TYPE: amino acid

20 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

25

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

30

(A) NAME/KEY: misc_feature

(B) LOCATION 1...436

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:110:

```

35  Met Ser Arg Asp Phe Lys Phe Asp Ser Asn Tyr Leu Asn Val Asn Thr
    1                5                10                15
    Asn Pro Lys Leu Gly Pro Val Tyr Thr Asn Gln Asn Tyr Pro Gly Phe
                20                25                30
    Phe Ile Phe Asp His Leu Arg Arg Tyr Val Met Asn Ala Phe Glu Pro
40    35                40                45
    Asn Leu Asn Leu Val Val Asn Thr Asn Lys Val Lys Gln Thr Phe Asn
    50                55                60
    Val Gly Met Arg Phe Met Thr Met Asp Met Phe Ile Arg Ser Asp Gln
65    70                75                80
45  Ser Thr Cys Glu Lys Thr Asp Ile Ile Asn Gly Val Cys Fis Met Pro
                85                90                95
    Pro Tyr Val Leu Ser Lys Thr Pro Asn Asn Asn Gln Glu Met Phe Asn
                100               105               110
    Asn Tyr Thr Ala Val Trp Leu Ser Asp Lys Ile Glu Phe Phe Asp Ser
50    115               120               125
    Lys Leu Val Ile Thr Pro Gly Leu Arg Tyr Thr Phe Leu Asn Tyr Asn
    130               135               140
    Asn Lys Glu Pro Glu Lys His Asp Phe Ser Val Trp Thr Ser Lys Lys
145               150               155               160

```

- 192 -

```

      Gln Arg Gln Asn Glu Trp Ser Pro Ala Leu Asn Ile Gly Tyr Lys Pro
                                165                                170                                175
Met Glu Asn Trp Ile Trp Tyr Ala Asn Tyr Arg Arg Ser Phe Ile Pro
      180                                185                                190
5  Pro Gln His Thr Met Val Gly Ile Thr Arg Thr Asn Tyr Asn Gln Ile
      195                                200                                205
      Phe Asn Glu Ile Glu Val Gly Gln Arg Tyr Ser Tyr Lys Asn Leu Leu
      210                                215                                220
10  Ser Phe Asn Thr Asn Tyr Phe Val Ile Phe Ala Lys Arg Tyr Tyr Ala
      225                                230                                235                                240
      Gly Gly Tyr Ser Pro Gln Pro Val Asp Ala Arg Ser Gln Gly Val Glu
      245                                250                                255
      Leu Glu Leu Tyr Tyr Ala Pro Ile Arg Gly Leu Gln Phe His Val Ala
      260                                265                                270
15  Tyr Thr Tyr Ile Asp Ala Arg Ile Thr Ser Asn Ala Asp Asp Ile Ala
      275                                280                                285
      Tyr Tyr Phe Thr Gly Ile Val Asn Lys Pro Phe Asp Ile Lys Gly Lys
      290                                295                                300
20  Arg Leu Pro Tyr Val Ser Pro Asn Gln Phe Ile Phe Asp Met Met Tyr
      305                                310                                315                                320
      Thr Tyr Lys His Thr Thr Phe Gly Ile Ser Ser Tyr Phe Tyr Ser Arg
      325                                330                                335
      Ala Tyr Ser Ser Met Leu Asn Gln Ala Lys Asp Gln Thr Val Cys Leu
      340                                345                                350
25  Pro Leu Asn Pro Glu Tyr Thr Gly Gly Leu Lys Tyr Gly Cys Asn Ser
      355                                360                                365
      Val Gly Leu Leu Pro Leu Tyr Phe Val Leu Asn Val Gln Val Ser Ser
      370                                375                                380
30  Ile Leu Trp Gln Ser Gly Arg His Lys Ile Thr Gly Ser Leu Gln Ile
      385                                390                                395                                400
      Asn Asn Leu Phe Asn Met Lys Tyr Tyr Phe Arg Gly Ile Gly Thr Ser
      405                                410                                415
      Pro Thr Gly Arg Glu Pro Ala Pro Gly Arg Ser Ile Thr Ala Tyr Leu
      420                                425                                430
35  Asn Tyr Glu Phe
      435

```

(2) INFORMATION FOR SEQ ID NO:111:

- 40 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 767 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
- 45 (ii) MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: YES
- (vi) ORIGINAL SOURCE:
- 50 (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
- (A) NAME/KEY: misc_feature
 - (B) LOCATION 1...767

- 193 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:111:

5	Met	Lys	Arg	Ile	Leu	Val	Ser	Leu	Ala	Val	Leu	Ser	His	Ser	Ala	His
	1			5					10						15	
	Ala	Val	Lys	Thr	His	Asn	Leu	Glu	Arg	Val	Glu	Ala	Ser	Gly	Val	Ala
				20					25					30		
	Asn	Asp	Lys	Glu	Ala	Pro	Leu	Ser	Trp	Arg	Ser	Lys	Glu	Val	Arg	Asn
			35					40					45			
10	Tyr	Met	Gly	Ser	Arg	Thr	Val	Ile	Ser	Asn	Lys	Gln	Leu	Thr	Lys	Ser
		50					55					60				
	Ala	Asn	Gln	Ser	Ile	Glu	Glu	Ala	Leu	Gln	Asn	Val	Pro	Gly	Val	His
	65					70					75					80
	Ile	Arg	Asn	Ser	Thr	Gly	Ile	Gly	Ala	Val	Pro	Ser	Ile	Ser	Ile	Arg
15					85					90					95	
	Gly	Phe	Gly	Ala	Gly	Gly	Pro	Gly	His	Ser	Asn	Thr	Gly	Met	Ile	Leu
				100					105					110		
	Val	Asn	Gly	Ile	Pro	Ile	Tyr	Val	Ala	Pro	Tyr	Val	Glu	Ile	Gly	Thr
		115					120						125			
20	Val	Ile	Phe	Pro	Val	Thr	Phe	Gln	Ser	Val	Asp	Arg	Ile	Ser	Val	Thr
		130					135					140				
	Lys	Gly	Gly	Glu	Ser	Val	Arg	Tyr	Gly	Pro	Asn	Ala	Phe	Gly	Gly	Val
	145					150					155					160
	Ile	Asn	Ile	Ile	Thr	Lys	Gly	Ile	Pro	Thr	Asn	Trp	Glu	Ser	Gln	Val
25					165					170					175	
	Ser	Glu	Arg	Thr	Thr	Phe	Trp	Gly	Lys	Ser	Glu	Asn	Gly	Gly	Phe	Phe
				180					185					190		
	Asn	Gln	Asn	Ser	Lys	Asn	Ile	Asp	Lys	Ser	Leu	Val	Asn	Asn	Met	Leu
			195					200					205			
30	Phe	Asn	Thr	Tyr	Leu	Arg	Thr	Gly	Gly	Met	Met	Asn	Lys	His	Phe	Gly
		210					215					220				
	Ile	Gln	Ala	Gln	Val	Asn	Trp	Leu	Lys	Gly	Gln	Gly	Phe	Arg	Tyr	Asn
	225					230					235					240
	Ser	Pro	Thr	Asp	Ile	Gln	Asn	Tyr	Met	Leu	Asp	Ser	Leu	Tyr	Gln	Ile
35					245					250					255	
	Asn	Asp	Ser	Asn	Lys	Ile	Thr	Ala	Phe	Phe	Gln	Tyr	Tyr	Ser	Tyr	Phe
				260					265					270		
	Leu	Thr	Asp	Pro	Gly	Ser	Leu	Gly	Ile	Ala	Ala	Tyr	Asn	Gln	Asn	Arg
			275					280					285			
40	Phe	Gln	Asn	Asn	Arg	Pro	Asn	Asn	Asp	Lys	Ser	Gly	Arg	Ala	Lys	Arg
		290					295					300				
	Trp	Gly	Ala	Val	Tyr	Gln	Asn	Phe	Phe	Gly	Asp	Thr	Asp	Arg	Val	Gly
	305					310					315					320
	Gly	Asp	Phe	Thr	Phe	Ser	Tyr	Tyr	Gly	His	Asp	Met	Ser	Arg	Asp	Phe
45					325					330					335	
	Lys	Phe	Asp	Ser	Asn	Tyr	Leu	Asn	Val	Asn	Thr	Asn	Pro	Lys	Leu	Gly
				340					345					350		
	Pro	Val	Tyr	Thr	Asn	Gln	Asn	Tyr	Pro	Gly	Phe	Phe	Ile	Phe	Asp	His
			355				360						365			
50	Leu	Arg	Arg	Tyr	Val	Met	Asn	Ala	Phe	Glu	Pro	Asn	Leu	Asn	Leu	Val
		370					375					380				
	Val	Asn	Thr	Asn	Lys	Val	Lys	Gln	Thr	Phe	Asn	Val	Gly	Met	Arg	Phe
	385					390					395					400
	Met	Thr	Met	Asp	Met	Phe	Ile	Arg	Ser	Asp	Gln	Ser	Thr	Cys	Glu	Lys

- 194 -

				405				410					415
	Thr	Asp	Ile	Ile	Asn	Gly	Val	Cys	His	Met	Pro	Pro	Tyr
				420									430
5	Lys	Thr	Pro	Asn	Asn	Asn	Gln	Glu	Met	Phe	Asn	Asn	Tyr
				435									445
	Trp	Leu	Ser	Asp	Lys	Ile	Glu	Phe	Phe	Asp	Ser	Lys	Leu
				450									460
	Pro	Gly	Leu	Arg	Tyr	Thr	Phe	Leu	Asn	Tyr	Asn	Asn	Lys
													480
10	Lys	His	Asp	Phe	Ser	Val	Trp	Thr	Ser	Lys	Lys	Gln	Arg
													495
	Trp	Ser	Pro	Ala	Leu	Asn	Ile	Gly	Tyr	Lys	Pro	Met	Glu
				500									510
15	Trp	Tyr	Ala	Asn	Tyr	Arg	Arg	Ser	Phe	Ile	Pro	Pro	Gln
				515									525
	Val	Gly	Ile	Thr	Arg	Thr	Asn	Tyr	Asn	Gln	Ile	Phe	Asn
													540
	Val	Gly	Gln	Arg	Tyr	Ser	Tyr	Lys	Asn	Leu	Leu	Ser	Phe
													560
20	Tyr	Phe	Val	Ile	Phe	Ala	Lys	Arg	Tyr	Tyr	Ala	Gly	Gly
													575
	Gln	Pro	Val	Asp	Ala	Arg	Ser	Gln	Gly	Val	Glu	Leu	Glu
													590
25	Ala	Pro	Ile	Arg	Gly	Leu	Gln	Phe	His	Val	Ala	Tyr	Thr
													605
	Ala	Arg	Ile	Thr	Ser	Asn	Ala	Asp	Asp	Ile	Ala	Tyr	Tyr
													620
	Ile	Val	Asn	Lys	Pro	Phe	Asp	Ile	Lys	Gly	Lys	Arg	Leu
													640
30	Ser	Pro	Asn	Gln	Phe	Ile	Phe	Asp	Met	Met	Tyr	Thr	Tyr
													655
	Thr	Phe	Gly	Ile	Ser	Ser	Tyr	Phe	Tyr	Ser	Arg	Ala	Tyr
													670
35	Leu	Asn	Gln	Ala	Lys	Asp	Gln	Thr	Val	Cys	Leu	Pro	Leu
													685
	Tyr	Thr	Gly	Gly	Leu	Lys	Tyr	Gly	Cys	Asn	Ser	Val	Gly
													700
	Leu	Tyr	Phe	Val	Leu	Asn	Val	Gln	Val	Ser	Ser	Ile	Leu
													720
40	Gly	Arg	His	Lys	Ile	Thr	Gly	Ser	Leu	Gln	Ile	Asn	Asn
													735
	Met	Lys	Tyr	Tyr	Phe	Arg	Gly	Ile	Gly	Thr	Ser	Pro	Thr
													750
45	Pro	Ala	Pro	Gly	Arg	Ser	Ile	Thr	Ala	Tyr	Leu	Asn	Tyr
													765

(2) INFORMATION FOR SEQ ID NO:112:

(i) SEQUENCE CHARACTERISTICS:

- 50 (A) LENGTH: 115 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

- 195 -

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

5 (A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc_feature

10 (B) LOCATION 1...115

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:112:

Leu His Pro Leu Cys Ala His Gly Gln Cys Gly Ser Glu Ala Ile Ala
 1 5 10 15
 15 Cys Leu Glu Ala Ile Ser Val Gly Ile Val Pro Val Ile Ala Asn Ser
 20 25 30
 Pro Leu Ser Ala Thr Arg Gln Phe Ala Leu Asp Glu Arg Ser Leu Phe
 35 40 45
 20 Glu Pro Asn Asn Ala Lys Asp Leu Ser Ala Lys Ile Asp Trp Trp Leu
 50 55 60
 Glu Asn Lys Leu Glu Arg Glu Arg Met Gln Asn Glu Tyr Ala Lys Ser
 65 70 75 80
 Ala Leu Asp Tyr Thr Leu Glu Asn Ser Val Ile Gln Ile Glu Lys Val
 85 90 95
 25 Tyr Glu Glu Ala Ile Lys Asp Phe Lys Asn Asn Pro Asn Leu Phe Lys
 100 105 110
 Thr Leu Ser
 115

30 (2) INFORMATION FOR SEQ ID NO:113:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 389 amino acids

35 (B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

40 (iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

45 (ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...389

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:113:

50 Met Val Ile Val Leu Val Val Asp Ser Phe Lys Asp Thr Ser Asn Gly
 1 5 10 15
 Thr Ser Met Thr Ala Phe Arg Phe Phe Glu Ala Leu Lys Lys Arg Gly
 20 25 30
 His Ala Met Arg Val Val Ala Pro His Val Asp Asn Leu Gly Ser Glu

- 196 -

		35		40		45													
		Glu	Glu	Gly	Tyr	Tyr	Asn	Leu	Lys	Glu	Arg	Tyr	Ile	Pro	Leu	Val	Thr		
		50					55						60						
5		Glu	Ile	Ser	His	Lys	Gln	His	Ile	Leu	Phe	Ala	Lys	Pro	Asp	Glu	Lys		
		65					70					75				80			
		Ile	Leu	Arg	Lys	Ala	Phe	Lys	Gly	Ala	Asp	Met	Ile	His	Thr	Tyr	Leu		
					85						90				95				
		Pro	Phe	Leu	Leu	Glu	Lys	Thr	Ala	Val	Lys	Ile	Ala	Arg	Glu	Met	Arg		
				100						105				110					
10		Val	Pro	Tyr	Ile	Gly	Ser	Phe	His	Leu	Gln	Pro	Glu	His	Ile	Ser	Tyr		
				115					120					125					
		Asn	Met	Lys	Leu	Gly	Gln	Phe	Ser	Trp	Leu	Asn	Thr	Met	Leu	Phe	Ser		
				130			135						140						
		Trp	Phe	Lys	Ser	Ser	His	Tyr	Arg	Tyr	Ile	His	His	Ile	His	Cys	Pro		
15		145				150						155				160			
		Ser	Lys	Phe	Ile	Val	Glu	Glu	Leu	Glu	Lys	Tyr	Asn	Tyr	Gly	Gly	Lys		
				165						170				175					
		Lys	Tyr	Ala	Ile	Ser	Asn	Gly	Phe	Asp	Pro	Met	Phe	Lys	Phe	Glu	His		
				180					185				190						
20		Pro	Gln	Lys	Ser	Leu	Phe	Asp	Thr	Thr	Pro	Phe	Lys	Ile	Ala	Met	Val		
				195				200					205						
		Gly	Arg	Tyr	Ser	Asn	Glu	Lys	Asn	Gln	Ser	Val	Leu	Ile	Lys	Ala	Val		
				210			215					220							
		Ala	Leu	Ser	Arg	Tyr	Lys	Gln	Asp	Ile	Val	Leu	Leu	Leu	Lys	Gly	Lys		
25		225				230						235				240			
		Gly	Pro	Asp	Glu	Lys	Lys	Ile	Lys	Leu	Leu	Ala	Gln	Lys	Leu	Gly	Val		
				245						250				255					
		Lys	Thr	Glu	Phe	Gly	Phe	Val	Asn	Ser	His	Glu	Leu	Leu	Glu	Ile	Leu		
				260					265				270						
30		Lys	Thr	Cys	Thr	Leu	Tyr	Ala	His	Thr	Ala	Asn	Val	Glu	Ser	Glu	Ala		
				275				280					285						
		Ile	Ala	Cys	Leu	Glu	Ala	Ile	Ser	Val	Gly	Ile	Val	Pro	Val	Ile	Ala		
				290			295					300							
		Asn	Ser	Pro	Leu	Ser	Ala	Thr	Arg	Gln	Phe	Ala	Leu	Asp	Glu	Arg	Ser		
35		305				310						315				320			
		Leu	Phe	Glu	Pro	Asn	Asn	Ala	Lys	Asp	Leu	Ser	Ala	Lys	Ile	Asp	Trp		
				325						330				335					
		Trp	Leu	Glu	Asn	Lys	Leu	Glu	Arg	Glu	Arg	Met	Gln	Asn	Glu	Tyr	Ala		
				340					345				350						
40		Lys	Ser	Ala	Leu	Asn	Tyr	Thr	Leu	Glu	Asn	Ser	Val	Ile	Gln	Ile	Glu		
				355				360					365						
		Lys	Val	Tyr	Glu	Glu	Ala	Ile	Lys	Asp	Phe	Lys	Asn	Asn	Pro	Asn	Leu		
				370			375					380							
		Phe	Lys	Thr	Leu	Ser													
45		385																	

(2) INFORMATION FOR SEQ ID NO:114:

(i) SEQUENCE CHARACTERISTICS:

- 50 (A) LENGTH: 312 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

- 197 -

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

5 (A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...312

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:114:

```

Leu Ala Ser Tyr Gly Phe Phe Leu Gly Ala Leu Phe Ile Leu Ala Ser
1      5      10      15
15 Gly Ile Val Cys Leu Gln Thr Ala Gly Asn Pro Phe Val Thr Leu Leu
    20      25      30
Ser Lys Gly Lys Glu Ala Arg Asn Leu Val Leu Val Gln Ala Phe Asn
    35      40      45
20 Ser Leu Gly Thr Thr Leu Gly Pro Ile Phe Gly Ser Leu Leu Ile Phe
    50      55      60
Ser Ala Thr Lys Thr Ser Asp Asn Leu Ser Leu Ile Asp Lys Leu Ala
65      70      75      80
Asp Ala Lys Ser Val Gln Met Pro Tyr Leu Gly Leu Ala Val Phe Ser
    85      90      95
25 Leu Leu Leu Ala Leu Val Met Tyr Leu Leu Lys Leu Pro Asp Val Glu
    100      105      110
Lys Glu Met Pro Lys Glu Thr Thr Gln Lys Ser Leu Phe Ser His Lys
    115      120      125
30 His Phe Val Phe Gly Ala Leu Gly Ile Phe Phe Tyr Val Gly Gly Glu
    130      135      140
Val Ala Ile Gly Ser Phe Leu Val Leu Ser Phe Glu Lys Leu Leu Asn
145      150      155      160
Leu Asp Ala Gln Ser Ser Ala His Tyr Leu Val Tyr Tyr Trp Gly Gly
    165      170      175
35 Ala Met Val Gly Arg Phe Leu Gly Ser Ala Leu Met Asn Lys Ile Ala
    180      185      190
Pro Asn Lys Tyr Leu Ala Phe Asn Ala Leu Ser Ser Ile Ile Leu Ile
    195      200      205
40 Ala Leu Ala Ile Leu Ile Gly Gly Lys Ile Ala Leu Phe Ala Leu Thr
    210      215      220
Phe Val Gly Phe Phe Asn Ser Ile Met Phe Pro Thr Ile Phe Ser Leu
225      230      235      240
Ala Thr Leu Asn Leu Gly His Leu Thr Ser Lys Ala Ser Gly Val Ile
    245      250      255
45 Ser Met Ala Ile Val Gly Gly Ala Leu Ile Pro Pro Ile Gln Gly Val
    260      265      270
Val Thr Asp Met Leu Thr Ala Thr Glu Ser Asn Leu Leu Tyr Ala Tyr
    275      280      285
Ser Val Pro Leu Leu Cys Tyr Phe Tyr Ile Leu Phe Phe Ala Leu Lys
50      290      295      300
Gly Tyr Lys Gln Glu Glu Asn Ser
305      310

```

(2) INFORMATION FOR SEQ ID NO:115:

- 198 -

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 407 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...407

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:115:

```

20 Met Gln Lys Thr Ser Asn Thr Leu Ala Leu Gly Ser Leu Thr Ala Leu
   1             5             10             15
Phe Phe Leu Met Gly Phe Ile Thr Val Leu Asn Asp Ile Leu Ile Pro
   20             25             30
His Leu Lys Pro Ile Phe Asp Leu Thr Tyr Phe Glu Ala Ser Leu Ile
25   35             40             45
Gln Phe Cys Phe Phe Gly Ala Tyr Phe Ile Met Gly Gly Val Phe Gly
   50             55             60
Asn Val Ile Ser Lys Ile Gly Tyr Pro Phe Gly Val Val Leu Gly Phe
   65             70             75             80
30 Val Ile Thr Ala Ser Gly Cys Ala Leu Phe Tyr Pro Ala Ala His Phe
   85             90             95
Gly Ser Tyr Gly Phe Phe Leu Gly Ala Leu Phe Ile Leu Ala Ser Gly
   100            105            110
Ile Val Cys Leu Gln Thr Ala Gly Asn Pro Phe Val Thr Leu Leu Ser
35   115            120            125
Lys Gly Lys Glu Ala Arg Asn Leu Val Leu Val Gln Ala Phe Asn Ser
   130            135            140
Leu Gly Thr Thr Leu Gly Pro Ile Phe Gly Ser Leu Leu Ile Phe Ser
   145            150            155            160
40 Ala Thr Lys Thr Ser Asp Asn Leu Ser Leu Ile Asp Lys Leu Ala Asp
   165            170            175
Ala Lys Ser Val Gln Met Pro Tyr Leu Gly Leu Ala Val Phe Ser Leu
   180            185            190
Leu Leu Ala Leu Val Met Tyr Leu Leu Lys Leu Pro Asp Val Glu Lys
45   195            200            205
Glu Met Pro Lys Glu Thr Thr Gln Lys Ser Leu Phe Ser His Lys Lys
   210            215            220
Phe Val Phe Gly Ala Leu Gly Ile Phe Phe Tyr Val Gly Gly Glu Val
   225            230            235            240
50 Ala Ile Gly Ser Phe Leu Val Leu Ser Phe Glu Lys Leu Leu Asn Leu
   245            250            255
Asp Ala Gln Ser Ser Ala His Tyr Leu Val Tyr Tyr Trp Gly Gly Ala
   260            265            270
Met Val Gly Arg Phe Leu Gly Ser Ala Leu Met Asn Lys Ile Ala Pro

```

- 199 -

```

                275                280                285
Asn Lys Tyr Leu Ala Phe Asn Ala Leu Ser Ser Ile Ile Leu Ile Ala
  290                295                300
5  Leu Ala Ile Leu Ile Gly Gly Lys Ile Ala Leu Phe Ala Leu Thr Phe
  305                310                315                320
Val Gly Phe Phe Asn Ser Ile Met Phe Pro Thr Ile Phe Ser Leu Ala
                325                330                335
Thr Leu Asn Leu Gly His Leu Thr Ser Lys Ala Ser Gly Val Ile Ser
                340                345                350
10 Met Ala Ile Val Gly Gly Ala Leu Ile Pro Pro Ile Gln Gly Val Val
  355                360                365
Thr Asp Met Leu Thr Ala Thr Glu Ser Asn Leu Leu Tyr Ala Tyr Ser
  370                375                380
15 Val Pro Leu Leu Cys Tyr Phe Tyr Ile Leu Phe Phe Ala Leu Lys Gly
  385                390                395                400
Tyr Lys Gln Glu Glu Asn Ser
                405

```

20 (2) INFORMATION FOR SEQ ID NO:116:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 125 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

25

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

30

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc_feature

35

(B) LOCATION 1...125

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:116:

```

40 Met Asn Lys Ile Ala Pro Asn Lys Tyr Leu Ala Phe Gly Ala Leu Ser
   1                5                10                15
Ser Ile Ile Leu Ile Ala Leu Ala Ile Leu Ile Gly Gly Lys Ile Ala
   20                25                30
Leu Phe Ala Leu Thr Phe Val Gly Phe Phe Asn Ser Ile Met Phe Pro
   35                40                45
45 Thr Ile Phe Ser Leu Ala Thr Leu Asn Leu Gly Ile Ser Leu Leu Met
   50                55                60
Ala Ser Gly Val Ile Ser Met Ala Ile Val Gly Gly Ala Leu Ile Pro
   65                70                75                80
Pro Ile Gln Gly Val Val Thr Asp Met Leu Thr Ala Thr Glu Ser Asn
50                85                90                95
Leu Leu Tyr Ala Tyr Ser Val Pro Leu Leu Cys Tyr Phe Tyr Ile Leu
   100                105                110
Phe Phe Ala Leu Lys Gly Tyr Lys Gln Glu Glu Asn Ser
   115                120                125

```

- 200 -

(2) INFORMATION FOR SEQ ID NO:117:

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 330 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

- 10 (iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

- 15 (A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
 (B) LOCATION 1...330

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:117:

Leu Lys Lys Ile Leu Pro Ala Leu Leu Met Gly Phe Val Gly Leu Asn
 1 5 10 15
 Ala Ser Asp Arg Leu Leu Glu Ile Met Arg Leu Tyr Gln Lys Gln Gly
 25 20 25 30
 Leu Glu Val Val Gly Gln Lys Leu Asp Ser Tyr Leu Ala Asp Lys Ser
 35 40 45
 Phe Trp Ala Glu Glu Leu Gln Asn Lys Asp Thr Asp Phe Gly Tyr Tyr
 50 55 60
 30 Gln Asn Lys Gln Phe Leu Phe Val Ala Asp Lys Ser Lys Pro Ser Leu
 65 70 75 80
 Glu Phe Tyr Glu Ile Glu Asn Asn Met Leu Lys Lys Ile Asn Ser Ser
 85 90 95
 Lys Ala Leu Val Gly Ser Lys Lys Gly Asp Lys Thr Leu Glu Gly Asp
 35 100 105 110
 Leu Ala Thr Pro Ile Gly Val Tyr Arg Ile Thr Gln Lys Leu Glu Arg
 115 120 125
 Leu Asp Gln Tyr Tyr Gly Val Leu Ala Phe Val Thr Asn Tyr Pro Asn
 130 135 140
 40 Leu Tyr Asp Thr Leu Lys Lys Arg Thr Gly His Gly Ile Trp Val His
 145 150 155 160
 Gly Met Pro Leu Asn Gly Asp Arg Asn Glu Leu Asn Thr Lys Gly Cys
 165 170 175
 Ile Ala Ile Glu Asn Pro Ile Leu Ser Ser Tyr Asp Lys Val Leu Lys
 45 180 185 190
 Gly Glu Lys Ala Phe Leu Ile Thr Tyr Glu Asp Lys Phe Ser Pro Ser
 195 200 205
 Thr Lys Glu Glu Leu Ser Met Ile Leu Ser Ser Leu Phe Gln Trp Lys
 210 215 220
 50 Glu Ala Trp Ala Arg Gly Asp Phe Glu Arg Tyr Met Arg Phe Tyr Asn
 225 230 235 240
 Pro Asn Phe Thr Arg Tyr Asp Gly Met Ser Phe Asn Ala Phe Lys Glu
 245 250 255
 Tyr Lys Lys Arg Val Phe Ala Lys Asn Glu Lys Lys Asn Ile Ala Phe

- 201 -

260 265 270
 Ser Ser Ile Asn Val Ile Pro Tyr Pro Asn Ser Gln Asn Lys Arg Leu
 275 280 285
 Phe Tyr Val Val Phe Asp Gln Asp Tyr Lys Ala Tyr Gln Gln Asn Lys
 5 290 295 300
 Leu Ser Tyr Ser Ser Asn Ser Gln Lys Glu Leu Tyr Val Glu Ile Glu
 305 310 315 320
 Asn Asn Gln Ala Ser Ile Ile Met Glu Lys
 325 330

10

(2) INFORMATION FOR SEQ ID NO:118:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 169 amino acids
 15 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein.

20 (iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
 (A) ORGANISM: *Helicobacter pylori*

25 (ix) FEATURE:
 (A) NAME/KEY: misc_feature
 (B) LOCATION 1...169

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:118:

30 Leu Phe Glu Lys Trp Ile Gly Leu Thr Leu Leu Leu Ser Ser Leu Gly
 1 5 10 15
 Tyr Pro Cys Gln Lys Val Ser Ile Ser Phe Lys Gln Tyr Glu Asn Leu
 20 25 30
 35 Ile His Ile His Gln Lys Gly Cys Asn Asn Glu Val Val Cys Arg Thr
 35 40 45
 Leu Ile Ser Ile Ala Leu Leu Glu Ser Ser Leu Gly Leu Asn Asn Lys
 50 55 60
 Arg Glu Lys Ser Leu Lys Asp Thr Ser Tyr Ser Met Phe His Ile Thr
 40 65 70 75 80
 Leu Asn Thr Ala Lys Lys Phe Tyr Pro Thr Tyr Ser Lys Thr Leu Leu
 85 90 95
 Lys Thr Lys Leu Leu Asn Asp Val Gly Phe Ala Ile Gln Leu Ala Lys
 100 105 110
 45 Gln Ile Leu Lys Glu Asn Phe Asp Tyr Tyr His Gln Lys His Pro Asn
 115 120 125
 Lys Ser Val Tyr Gln Leu Val Gln Met Ala Ile Gly Ala Tyr Asn Gly
 130 135 140
 Gly Met Lys His Asn Pro Asn Gly Ala Tyr Met Lys Lys Phe Arg Cys
 50 145 150 155 160
 Ile Tyr Ser Gln Val Arg Tyr Asn Glu
 165

(2) INFORMATION FOR SEQ ID NO:119:

- 202 -

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 215 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
 (A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:
 (A) NAME/KEY: misc_feature
 (B) LOCATION 1...215

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:119:

20	Met	Lys	Lys	Pro	Tyr	Arg	Lys	Ile	Ser	Asp	Tyr	Ala	Ile	Val	Gly	Gly
	1				5				10					15		
	Leu	Ser	Ala	Leu	Val	Met	Val	Ser	Ile	Val	Gly	Cys	Lys	Ser	Asn	Ala
			20					25					30			
25	Asp	Asp	Lys	Pro	Lys	Glu	Gln	Ser	Ser	Leu	Ser	Gln	Ser	Val	Gln	Lys
			35				40					45				
	Gly	Ala	Phe	Val	Ile	Leu	Glu	Glu	Gln	Lys	Asp	Lys	Ser	Tyr	Lys	Val
		50				55				60						
	Val	Glu	Glu	Tyr	Pro	Ser	Ser	Arg	Thr	His	Ile	Val	Val	Arg	Asp	Leu
	65				70				75					80		
30	Gln	Gly	Asn	Glu	Arg	Val	Leu	Ser	Asn	Glu	Glu	Ile	Gln	Lys	Leu	Ile
					85				90					95		
	Lys	Glu	Glu	Glu	Ala	Lys	Ile	Asp	Asn	Gly	Thr	Ser	Lys	Leu	Val	Gln
				100					105				110			
	Pro	Asn	Asn	Gly	Gly	Ser	Asn	Glu	Gly	Ser	Gly	Phe	Gly	Leu	Gly	Ser
35			115				120					125				
	Ala	Ile	Leu	Gly	Ser	Ala	Ala	Gly	Ala	Ile	Leu	Gly	Ser	Tyr	Ile	Gly
		130				135					140					
	Asn	Lys	Leu	Phe	Asn	Asn	Pro	Asn	Tyr	Gln	Gln	Asn	Ala	Gln	Arg	Thr
	145				150				155					160		
40	Tyr	Lys	Ser	Pro	Gln	Ala	Tyr	Gln	Arg	Ser	Gln	Asn	Ser	Phe	Ser	Lys
					165				170					175		
	Ser	Ala	Pro	Ser	Ala	Ser	Ser	Met	Gly	Thr	Ala	Ser	Lys	Gly	Gln	Ser
				180				185					190			
	Gly	Phe	Phe	Gly	Ser	Ser	Arg	Pro	Thr	Ser	Ser	Pro	Ala	Ile	Ser	Ser
45			195				200						205			
	Gly	Thr	Arg	Gly	Phe	Asn	Ala									
		210				215										

(2) INFORMATION FOR SEQ ID NO:120:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 253 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

- 203 -

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...253

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:120:

```

15  Leu Lys Thr Leu Phe Ser Val Tyr Leu Phe Leu Ser Leu Asn Pro Leu
    1             5             10             15
    Phe Leu Glu Ala Lys Glu Ile Thr Trp Ser Gln Phe Leu Glu Asn Phe
        20             25             30
20  Lys Asn Lys Asn Glu Asp Asp Lys Pro Lys Pro Leu Thr Ile Asp Lys
    35             40             45
    Asn Asn Glu Lys Gln Gln Ile Leu Asp Lys Asn Gln Gln Ile Leu Lys
    50             55             60
    Arg Ala Leu Glu Lys Ser Leu Lys Phe Phe Phe Ile Phe Gly Tyr Asn
    65             70             75             80
25  Tyr Ser Gln Ala Ala Tyr Ser Thr Thr Asn Gln Asn Leu Thr Leu Thr
    85             90             95
    Ala Asn Ser Ile Gly Phe Asn Thr Ala Thr Gly Leu Glu His Phe Leu
    100            105            110
    Arg Asn His Pro Lys Val Gly Phe Arg Ile Phe Ser Val Tyr Asn Tyr
    115            120            125
30  Phe His Ser Val Ser Leu Ser Gln Pro Gln Ile Leu Met Val Gln Asn
    130            135            140
    Tyr Gly Gly Ala Leu Asp Phe Ser Trp Ile Phe Val Asp Lys Lys Thr
    145            150            155            160
35  Tyr Arg Phe Arg Ser Tyr Leu Gly Ile Ala Leu Glu Gln Gly Val Leu
    165            170            175
    Leu Val Asp Thr Ile Lys Thr Gly Ser Phe Thr Thr Ile Ile Pro Arg
    180            185            190
    Thr Lys Lys Thr Phe Phe Gln Ala Pro Leu Arg Phe Gly Phe Ile Val
    195            200            205
40  Asp Phe Ile Gly Tyr Leu Ser Leu Gln Leu Gly Ile Glu Met Pro Leu
    210            215            220
    Val Arg Asn Val Phe Tyr Thr Tyr Asn Asn His Gln Glu Arg Phe Lys
    225            230            235            240
45  Pro Arg Phe Asn Ala Asn Leu Ser Leu Ile Val Ser Phe
    245            250

```

(2) INFORMATION FOR SEQ ID NO:121:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 336 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

- 204 -

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

5

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

10

(A) NAME/KEY: misc_feature

(B) LOCATION 1...336

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:121:

```

15  Leu Phe Phe Lys Phe Ile Leu Cys Leu Ser Leu Gly Ile Phe Ala Trp
    1           5           10           15
    Ala Lys Glu Val Ile Pro Thr Pro Ser Thr Pro Leu Thr Pro Ser Lys
        20           25           30
    Arg Tyr Ser Ile Asn Leu Met Thr Glu Asn Asp Gly Tyr Ile Asn Pro
        35           40           45
20  Tyr Ile Asp Glu Tyr Tyr Thr Ala Gly Asn Gln Ile Gly Phe Ser Thr
    50           55           60
    Lys Glu Phe Asp Phe Ser Lys Asn Lys Ala Met Lys Trp Ser Ser Tyr
    65           70           75           80
    Leu Gly Phe Phe Asn Lys Ser Pro Arg Val Thr Arg Phe Gly Ile Ser
25  85           90           95
    Leu Ala Gln Asp Met Tyr Thr Pro Ser Leu Ala Asn Arg Lys Leu Val
        100           105           110
    His Leu His Asp Asn His Pro Tyr Gly Gly Tyr Leu Arg Val Asn Leu
        115           120           125
30  Asn Val Tyr Asn Arg His Gln Thr Phe Met Glu Leu Phe Thr Ile Ser
    130           135           140
    Leu Gly Thr Thr Gly Gln Asp Ser Leu Ala Ala Gln Thr Gln Arg Leu
    145           150           155           160
    Ile His Lys Trp Gly His Asp Pro Gln Phe Tyr Gly Trp Asn Thr Gln
35  165           170           175
    Leu Lys Asn Glu Phe Ile Phe Glu Leu His Tyr Gln Leu Leu Lys Lys
        180           185           190
    Val Pro Leu Leu Lys Thr Arg Phe Phe Ser Met Glu Leu Met Pro Gly
        195           200           205
40  Phe Asn Val Glu Leu Gly Asn Ala Arg Asp Tyr Phe Gln Leu Gly Ser
    210           215           220
    Leu Phe Arg Ala Gly Tyr Asn Leu Asp Ala Asp Tyr Gly Val Asn Lys
    225           230           235           240
    Val Asn Thr Ala Phe Asp Gly Gly Met Pro Tyr Ser Asp Lys Phe Ser
45  245           250           255
    Ile Tyr Phe Phe Ala Gly Ala Phe Gly Arg Phe Gln Pro Leu Asn Ile
        260           265           270
    Phe Ile Gln Gly Asn Ser Pro Glu Thr Arg Gly Ile Ala Asn Leu Glu
        275           280           285
50  Tyr Phe Val Tyr Ala Ser Glu Ile Gly Ala Ala Met Met Trp Arg Ser
    290           295           300
    Leu Arg Val Ala Phe Thr Ile Thr Asp Ile Ser Lys Thr Phe Gln Ser
    305           310           315           320
    Gln Pro Lys His His Gln Ile Gly Thr Leu Glu Leu Asn Phe Ala Phe

```


325

330

335

(2) INFORMATION FOR SEQ ID NO:122:

- 5 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 108 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
- 10 (ii) MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: YES
- (vi) ORIGINAL SOURCE:
 15 (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
 (A) NAME/KEY: misc_feature
 (B) LOCATION 1...108
- 20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:122:

	Met	Lys	Pro	Ile	Phe	Ser	Leu	Phe	Phe	Leu	Leu	Ile	Val	Leu	Lys	Ala	
	1				5					10					15		
25	His	Pro	Ile	Asn	Pro	Leu	Leu	Glu	Pro	Leu	Tyr	Phe	Pro	Ser	Tyr	Thr	
				20					25					30			
	Gln	Phe	Leu	Asp	Leu	Glu	Pro	His	Phe	Val	Ile	Lys	Lys	Lys	Arg	Ala	
			35				40						45				
30	Tyr	Arg	Pro	Phe	Gln	Trp	Gly	Asn	Thr	Ile	Ile	Ile	Lys	Arg	His	Asp	
		50				55						60					
	Leu	Glu	Glu	Arg	Gln	Ser	Asn	Gln	Pro	Ser	Asp	Ile	Phe	Arg	Gln	Asn	
		65			70					75					80		
	Ala	Glu	Ile	Asn	Val	Ser	Ser	Gln	Thr	Phe	Leu	Arg	Gly	Ile	Ser	Ser	
				85					90					95			
35	Ala	Ser	Ser	Arg	Ile	Val	Ile	Asp	Ser	Val	Ala	Gln					
				100					105								

(2) INFORMATION FOR SEQ ID NO:123:

- 40 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 195 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
- 45 (ii) MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: YES
- (vi) ORIGINAL SOURCE:
 50 (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
 (A) NAME/KEY: misc_feature
 (B) LOCATION 1...195

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:123:

```

5  Met Ser Asn Asn Pro Phe Lys Lys Val Gly Met Ile Ser Ser Gln Asn
   1      5      10      15
   Asn Asn Gly Ala Leu Asn Gly Leu Gly Val Gln Val Gly Tyr Lys Gln
      20      25      30
   Phe Phe Gly Glu Ser Lys Arg Trp Gly Leu Arg Tyr Tyr Gly Phe Phe
      35      40      45
10  Asp Tyr Asn His Gly Tyr Ile Lys Ser Ser Phe Phe Asn Ser Ser Ser
   50      55      60
   Asp Ile Trp Thr Tyr Gly Gly Gly Ser Asp Leu Leu Val Asn Phe Ile
   65      70      75      80
   Asn Asp Ser Ile Thr Arg Lys Asn Asn Lys Leu Ser Val Gly Leu Phe
15  Gly Gly Ile Gln Leu Ala Gly Thr Thr Trp Leu Asn Ser Gln Tyr Met
      85      90      95
   Asn Leu Thr Ala Phe Asn Asn Pro Tyr Ser Ala Lys Val Asn Ala Ser
      100      105      110
      115      120      125
20  Asn Phe Gln Phe Leu Phe Asn Leu Gly Leu Arg Thr Asn Leu Ala Thr
   130      135      140
   Ala Lys Lys Lys Asp Ser Glu Arg Ser Ala Gln His Gly Val Glu Leu
   145      150      155      160
   Gly Ile Lys Ile Pro Thr Ile Asn Thr Asn Tyr Tyr Ser Phe Leu Gly
25  Thr Lys Leu Glu Tyr Arg Arg Leu Tyr Ser Val Tyr Leu Asn Tyr Val
      165      170      175
      180      185      190
   Phe Ala Tyr
      195
30

```

(2) INFORMATION FOR SEQ ID NO:124:

(i) SEQUENCE CHARACTERISTICS:

```

35  (A) LENGTH: 227 amino acids
     (B) TYPE: amino acid
     (D) TOPOLOGY: linear

```

(ii) MOLECULE TYPE: protein

40 (iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

45 (ix) FEATURE:

```

     (A) NAME/KEY: misc_feature
     (B) LOCATION 1...227

```

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:124:

```

50  Val Arg Phe Gly Lys Ile Asp Tyr Leu Asn Met Leu Pro Phe Asp Val
   1      5      10      15
   Phe Ile Lys Ser Tyr Pro Thr Pro Cys Tyr Phe Lys Gln Phe Leu Arg
      20      25      30

```

-207-

```

Leu Lys Lys Thr Tyr Pro Ser Lys Leu Asn Glu Ser Phe Leu Phe Arg
      35              40              45
Arg Ile Asp Ala Gly Phe Ile Ser Ser Ile Ala Gly Tyr Pro Phe Ala
      50              55              60
5  Leu Cys Ser Tyr Ser Leu Gly Ile Val Ala Tyr Lys Glu Val Leu Ser
   65              70              75              80
Val Leu Val Val Asn Arg Glu Asn Ala Phe Asp Lys Glu Ser Ala Ser
      85              90              95
10 Ser Asn Ala Leu Ser Lys Val Leu Gly Leu Lys Gly Glu Val Leu Ile
   100              105              110
Gly Asn Lys Ala Leu Gln Phe Tyr Tyr Ser Asn Pro Lys Lys Asp Phe
   115              120              125
Ile Asp Leu Ala Ala Leu Trp Tyr Glu Lys Lys Arg Leu Pro Phe Val
   130              135              140
15 Phe Gly Arg Leu Cys Tyr Tyr Gln Asn Lys Asp Phe Tyr Lys Arg Leu
   145              150              155              160
Ser Leu Ala Phe Lys His Gln Lys Thr Lys Ile Pro His Tyr Ile Leu
      165              170              175
20 Lys Glu Ala Ala Leu Lys Thr Asn Leu Lys Arg Gln Asp Ile Leu Asn
   180              185              190
Tyr Leu Gln Lys Ile Tyr Tyr Thr Leu Gly Lys Lys Glu Gln Ser Gly
   195              200              205
Leu Lys Ala Phe Tyr Arg Glu Leu Leu Phe Lys Arg Ile Gln Lys Pro
   210              215              220
25 Lys Arg Phe
   225

```

(2) INFORMATION FOR SEQ ID NO:125:

- 30 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 305 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
- 35 (ii) MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: YES
- (vi) ORIGINAL SOURCE:
 40 (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
 (A) NAME/KEY: misc_feature
 (B) LOCATION 1...305
- 45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:125:

```

Met Gly Arg Ile Glu Ser Lys Lys Arg Leu Lys Ala Leu Ile Phe Leu
1      5              10              15
50 Ala Ser Leu Gly Val Leu Trp Gly Asn Ala Ala Glu Lys Thr Pro Phe
   20              25              30
Phe Lys Thr Lys Asn His Ile Tyr Leu Gly Phe Arg Leu Gly Thr Gly
   35              40              45
Ala Thr Thr Arg Thr Ser Met Trp Gln Gln Ala Tyr Lys Asp Asn Pro

```

-208-

```

      50              55              60
Thr Cys Pro Ser Ser Val Cys Tyr Gly Glu Lys Leu Glu Ala His Tyr
65              70              75              80
Lys Gly Gly Lys Asn Leu Ser Tyr Thr Gly Gln Ile Gly Asp Glu Ile
5
Ala Phe Asp Lys Tyr His Ile Leu Gly Leu Arg Val Trp Gly Asp Val
      85              90              95
Glu Tyr Ala Lys Ala Gln Leu Gly Gln Lys Val Gly Gly Asn Thr Leu
      100              105              110
10 Leu Ser Gln Ala Asn Tyr Asn Pro Ser Ala Ile Lys Thr Tyr Asp Pro
      115              120              125
Thr Ser Asn Ala Gln Gly Ser Leu Val Leu Gln Lys Thr Pro Ser Pro
      130              135              140
145 Gln Asp Phe Leu Phe Asn Asn Gly His Phe Met Ala Phe Gly Leu Asn
15
Val Asn Met Phe Val Asn Leu Pro Ile Asp Thr Leu Leu Lys Leu Ala
      150              155              160
      165              170              175
Leu Lys Thr Glu Lys Met Leu Phe Phe Lys Ile Gly Val Phe Gly Gly
      180              185              190
20 Gly Gly Val Glu Tyr Ala Ile Leu Trp Ser Pro Gln Tyr Lys Asn Gln
      195              200              205
Asn Thr His Gln Asp Asp Lys Phe Phe Ala Ala Gly Gly Gly Phe Phe
      210              215              220
225 Val Asn Phe Gly Gly Ser Leu Tyr Ile Gly Lys Arg Asn Arg Phe Asn
25
Val Gly Leu Lys Ile Pro Tyr Tyr Ser Leu Ser Ala Gln Ser Trp Lys
      225              230              235              240
      245              250              255
Asn Phe Gly Ser Ser Asn Val Trp Gln Gln Gln Thr Ile Arg Gln Asn
      260              265              270
275              280              285
30 Phe Ser Val Phe Arg Asn Lys Glu Val Phe Val Ser Tyr Ala Phe Leu
      290              295              300
Phe
305

```

35 (2) INFORMATION FOR SEQ ID NO:126:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 258 amino acids

(B) TYPE: amino acid

40 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

45

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

50

(A) NAME/KEY: misc_feature

(B) LOCATION 1...258

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:126:

-209-

```

Met Phe Leu Arg Ser Tyr Pro Lys Leu Arg Tyr Ala Leu Cys Leu Pro
1      5      10      15
Leu Leu Thr Glu Thr Cys Tyr Ser Glu Glu Arg Thr Leu Asn Lys Val
      20      25      30
5  Thr Thr Gln Ala Lys Arg Ile Phe Thr Tyr Asn Asn Glu Phe Lys Val
      35      40      45
Thr Ser Lys Glu Leu Asp Gln Arg Gln Ser Asn Glu Val Lys Asp Leu
      50      55      60
10 Phe Arg Thr Asn Pro Asp Val Asn Val Gly Gly Gly Ser Val Met Gly
      65      70      75      80
Gln Lys Ile Tyr Val Arg Gly Ile Glu Asp Arg Leu Leu Arg Val Thr
      85      90      95
Val Asp Gly Ala Ala Gln Asn Gly Asn Ile Tyr His His Gln Gly Asn
      100     105     110
15 Thr Val Ile Asp Pro Gly Met Leu Lys Ser Val Glu Val Thr Lys Gly
      115     120     125
Ala Ala Asn Ala Ser Ala Gly Pro Gly Ala Ile Ala Gly Val Ile Lys
      130     135     140
20 Met Glu Thr Lys Gly Ala Ala Asp Phe Ile Pro Arg Gly Lys Asn Tyr
      145     150     155     160
Ala Ala Ser Gly Ala Val Ser Phe Tyr Thr Asn Phe Gly Asp Arg Glu
      165     170     175
Thr Phe Arg Ser Ala Tyr Gln Ser Ala His Phe Asp Ile Ile Ala Tyr
      180     185     190
25 Tyr Thr His Gln Asn Ile Phe Tyr Tyr Arg Ser Gly Ala Thr Val Met
      195     200     205
Lys Asn Leu Phe Lys Pro Thr Gln Ala Asp Lys Glu Pro Gly Thr Pro
      210     215     220
30 Ser Glu Gln Asn Asn Ala Leu Ile Lys Met Asn Gly Tyr Leu Ser Asp
      225     230     235     240
Arg Asp Thr Leu Thr Phe Ser Trp Asn Met Thr Arg Asp Asn Ala Thr
      245     250     255
Arg Leu

```

35

(2) INFORMATION FOR SEQ ID NO:127:

(i) SEQUENCE CHARACTERISTICS:

40

(A) LENGTH: 192 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

45

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

50

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...192

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:127:

-210-

```

Met Phe Leu Arg Ser Tyr Pro Lys Leu Arg Tyr Ala Leu Cys Leu Pro
1           5           10           15
Leu Leu Thr Glu Thr Cys Tyr Ser Glu Glu Arg Thr Leu Asn Lys Val
5           20           25           30
Thr Thr Gln Ala Lys Arg Ile Phe Thr Tyr Asn Asn Glu Phe Lys Val
           35           40           45
Thr Ser Lys Glu Leu Asp Gln Arg Gln Ser Asn Glu Val Lys Asp Leu
           50           55           60
10 Phe Arg Thr Asn Pro Asp Val Asn Val Gly Gly Gly Ser Val Met Gly
65           70           75           80
Gln Lys Ile Tyr Val Arg Gly Ile Glu Asp Arg Leu Leu Arg Val Thr
           85           90           95
Val Asp Gly Ala Ala Gln Asn Gly Asn Ile Tyr His His Gln Gly Asn
15           100          105          110
Thr Val Ile Asp Pro Gly Met Leu Lys Ser Val Glu Val Thr Lys Gly
           115          120          125
Ala Ala Asn Ala Ser Ala Gly Pro Gly Ala Ile Ala Gly Val Ile Lys
           130          135          140
20 Met Glu Thr Lys Gly Ala Ala Asp Phe Ile Pro Arg Gly Lys Asn Tyr
145          150          155          160
Ala Ala Ser Gly Ala Val Ser Phe Tyr Thr Asn Phe Gly Asp Arg Glu
           165          170          175
Thr Phe Arg Ser Ala Tyr Gln Ser Ala His Phe Asp Ile Ile Ala Tyr
25           180          185          190

```

(2) INFORMATION FOR SEQ ID NO:128:

- (i) SEQUENCE CHARACTERISTICS:
- 30 (A) LENGTH: 126 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- 35 (iii) HYPOTHETICAL: YES
- (vi) ORIGINAL SOURCE:
- 40 (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
- (A) NAME/KEY: misc_feature
(B) LOCATION 1...126
- 45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:128:

```

Val Pro Leu Ser Leu Gly Gly Asn Leu Leu Asn Pro Asn Asn Ser Ser
1           5           10           15
Val Leu Asn Leu Lys Asn Ser Gln Leu Val Phe Ser Asp Gln Gly Ser
50           20           25           30
Leu Asn Ile Ala Asn Ile Asp Leu Leu Ser Asp Leu Asn Gly Asn Lys
           35           40           45
Asn Arg Val Tyr Asn Ile Ile Gln Ala Asp Met Asn Gly Asn Trp Tyr
50           55           60

```

-211-

Glu Arg Ile Asn Phe Phe Gly Met Arg Ile Asn Asp Gly Ile Tyr Asp
 65 70 75 80
 Ala Lys Asn Gln Thr Tyr Ser Phe Thr Asn Pro Leu Asn Asn Ala Val
 85 90 95
 5 Lys Phe Thr Glu Ser Phe Phe Ile His Arg Leu Cys Gly Ser Leu Ser
 100 105 110
 Gln Ile Gln Lys Lys Lys Asn Thr Ile Val Ser Pro Arg Leu
 115 120 125

10 (2) INFORMATION FOR SEQ ID NO:129:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 565 amino acids

(B) TYPE: amino acid

15 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

20

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

25 (A) NAME/KEY: misc_feature

(B) LOCATION 1...565

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:129:

30 Val Tyr Ser Tyr Ser Asp Asp Ala Gln Gly Val Phe Tyr Leu Thr Ser
 1 5 10 15
 Ser Val Lys Gly Tyr Tyr Asn Pro Asn Gln Ser Tyr Gln Ala Ser Gly
 20 25 30
 Ser Asn Asn Thr Thr Lys Asn Asn Leu Thr Ser Glu Ser Ser Val
 35 35 40 45
 Ile Ser Gln Thr Tyr Asn Ala Gln Gly Asn Pro Ile Ser Ala Leu His
 50 55 60
 Val Tyr Asn Lys Gly Tyr Asn Phe Ser Asn Ile Lys Ala Leu Gly Gln
 65 70 75 80
 40 Met Ala Leu Lys Leu Tyr Pro Glu Ile Lys Lys Ile Leu Gly Asn Asp
 85 90 95
 Phe Ser Leu Ser Ser Leu Ser Asn Leu Lys Gly Asp Ala Leu Asn Gln
 100 105 110
 Leu Thr Lys Leu Ile Thr Pro Ser Asp Trp Lys Asn Ile Asn Glu Leu
 45 115 120 125
 Ile Asp Asn Ala Asn Asn Ser Val Val Gln Asn Phe Asn Asn Gly Thr
 130 135 140
 Leu Ile Ile Gly Ala Thr Lys Ile Gly Gln Thr Asp Thr Asn Ser Ala
 145 150 155 160
 50 Val Val Phe Gly Gly Leu Gly Tyr Gln Lys Pro Cys Asp Tyr Thr Asp
 165 170 175
 Ile Val Cys Gln Lys Phe Arg Gly Thr Tyr Leu Gly Gln Leu Leu Glu
 180 185 190
 Ser Asn Ser Ala Asp Leu Gly Tyr Ile Asp Thr Thr Phe Asn Ala Lys

-212-

195 200 205
 Glu Ile Tyr Leu Thr Gly Thr Leu Gly Ser Gly Asn Ala Trp Gly Thr
 210 215 220
 5 Gly Gly Ser Ala Ser Val Thr Phe Asn Ser Gln Thr Ser Leu Ile Leu
 225 230 235 240
 Asn Gln Ala Asn Ile Val Ser Ser Gln Thr Asp Gly Ile Phe Ser Met
 245 250 255
 Leu Gly Gln Glu Gly Ile Asn Lys Val Phe Asn Gln Ala Gly Leu Ala
 260 265 270
 10 Asn Ile Leu Gly Glu Val Ala Met Gln Ser Ile Asn Lys Ala Gly Gly
 275 280 285
 Leu Gly Asn Leu Ile Val Asn Thr Leu Gly Ser Asp Ser Val Ile Gly
 290 295 300
 15 Gly Tyr Leu Thr Pro Glu Gln Lys Asn Gln Thr Leu Ser Gln Leu Leu
 305 310 315 320
 Gly Gln Asn Asn Phe Asp Asn Leu Met Asn Asp Ser Gly Leu Asn Thr
 325 330 335
 Ala Ile Lys Asp Leu Ile Arg Gln Lys Leu Gly Phe Trp Thr Gly Leu
 340 345 350
 20 Val Gly Gly Leu Ala Gly Leu Gly Gly Ile Asp Leu Gln Asn Pro Glu
 355 360 365
 Lys Leu Ile Gly Ser Met Ser Ile Asn Asp Leu Leu Ser Lys Lys Gly
 370 375 380
 25 Leu Phe Asn Gln Ile Thr Gly Phe Ile Ser Ala Asn Asp Ile Gly Gln
 385 390 395 400
 Val Ile Ser Val Met Leu Gln Asp Ile Val Lys Pro Ser Asp Ala Leu
 405 410 415
 Lys Asn Asp Val Ala Ala Leu Gly Lys Gln Met Ile Gly Glu Phe Leu
 420 425 430
 30 Gly Gln Asp Thr Leu Asn Ser Leu Glu Ser Leu Leu Gln Asn Gln Gln
 435 440 445
 Ile Lys Ser Val Leu Asp Lys Val Leu Ala Ala Lys Gly Leu Gly Ser
 450 455 460
 35 Ile Tyr Glu Gln Gly Leu Gly Asp Leu Ile Pro Asn Leu Gly Lys Lys
 465 470 475 480
 Gly Ile Phe Ala Pro Tyr Gly Leu Ser Gln Val Trp Gln Lys Gly Asp
 485 490 495
 Phe Ser Phe Asn Ala Gln Gly Asn Val Phe Val Gln Asn Ser Thr Phe
 500 505 510
 40 Ser Asn Ala Asn Gly Gly Thr Leu Ser Phe Asn Ala Gly Asn Ser Leu
 515 520 525
 Ile Phe Ala Gly Asn Asn His Ile Ala Phe Thr Asn His Ser Gly Thr
 530 535 540
 45 Leu Asn Leu Leu Ser Asn Gln Val Ser Asn Ile Asn Val Thr Met Leu
 545 550 555 560
 Ala Ala Ala Thr Ala
 565

(2) INFORMATION FOR SEQ ID NO:130:

50

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 172 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

-213-

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

5 (vi) ORIGINAL SOURCE:
(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:
10 (A) NAME/KEY: misc_feature
(B) LOCATION 1...172

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:130:

15	Val	Phe	Gly	Leu	Ser	Leu	Ala	Asp	Met	Ile	Leu	Glu	Arg	Phe	Lys	Asp
	1				5					10					15	
	Phe	Met	Arg	Glu	Tyr	Pro	Glu	Pro	Tyr	Lys	Phe	Leu	Gln	Val	Phe	Tyr
			20						25					30		
	Ala	Gln	Glu	Lys	Glu	Arg	Phe	Leu	Asn	His	Lys	Met	Asn	Asp	Tyr	Ile
20			35					40					45			
	Lys	Gln	Asn	Lys	Ser	Lys	Glu	Ala	Ser	Ile	Leu	Ala	Arg	Gln	Gly	
		50				55					60					
	Phe	Val	Ser	Val	Ile	Gly	Arg	Ala	Leu	Glu	Lys	Ile	Ile	Glu	Leu	Leu
	65				70					75					80	
25	Leu	Lys	Asp	Phe	Cys	Ile	Lys	Asn	Asn	Val	Lys	Met	Thr	Asn	Asp	Lys
				85					90					95		
	Thr	Leu	Arg	Ala	Lys	Arg	Ile	Asn	Gly	Glu	Leu	Asp	Lys	Val	Lys	Arg
			100						105					110		
	Ala	Leu	Leu	Val	His	Phe	Gly	Gly	Tyr	Ser	Val	Leu	Pro	Asp	Ile	Ile
30			115					120					125			
	Leu	Tyr	Gln	Thr	Asn	Lys	Asp	Asn	Ile	Lys	Ile	Leu	Ala	Ile	Leu	Ser
		130				135					140					
	Val	Lys	Asn	Ser	Phe	Arg	Glu	Arg	Phe	Thr	Lys	Asp	Ala	Leu	Leu	Glu
	145				150					155					160	
35	Ile	Lys	Thr	Phe	Ala	Ile	Ala	Cys	Asn	Phe	Ser	His				
				165					170							

(2) INFORMATION FOR SEQ ID NO:131:

40 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 331 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

45 (ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
50 (A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...331

-214-

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:131:

```

5  Met Lys Arg Phe Val Leu Phe Leu Leu Phe Ile Cys Val Cys Val Cys
   1           5           10           15
Val Gln Ala Tyr Ala Glu Gln Asp Tyr Phe Phe Arg Asp Phe Lys Ser
   20           25           30
Ile Asp Leu Pro Gln Lys Leu His Leu Asp Lys Lys Leu Ser Gln Thr
   35           40           45
10 Ile Gln Pro Cys Ala Gln Leu Asn Ala Ser Lys His Tyr Thr Ala Thr
   50           55           60
Gly Val Arg Glu Pro Asp Ala Cys Thr Lys Ser Phe Lys Lys Ser Ala
   65           70           75           80
Met Val Ser Tyr Asp Leu Ala Leu Gly Tyr Leu Val Ser Gln Asn Lys
   85           90           95
15 Pro Tyr Gly Leu Lys Ala Ile Glu Ile Leu Asn Ala Trp Ala Asn Glu
   100          105          110
Leu Gln Ser Val Asp Thr Tyr Gln Ser Glu Asp Asn Ile Asn Phe Tyr
   115          120          125
20 Met Pro Tyr Met Asn Met Ala Tyr Trp Phe Val Lys Lys Glu Phe Pro
   130          135          140
Ser Pro Glu Tyr Glu Asp Phe Ile Arg Arg Met Arg Gln Tyr Ser Gln
   145          150          155          160
Ser Ala Leu Asn Thr Asn His Gly Ala Trp Gly Ile Leu Phe Asp Val
   165          170          175
25 Ser Ser Ala Leu Ala Leu Asp Asp His Ala Leu Leu Gln Ser Ser Ala
   180          185          190
Asn Arg Trp Gln Glu Trp Val Phe Lys Ala Ile Asp Glu Asn Gly Val
   195          200          205
30 Ile Ala Ser Ala Ile Thr Arg Ser Asp Thr Ser Asp Tyr His Gly Gly
   210          215          220
Pro Thr Lys Gly Ile Lys Gly Ile Ala Tyr Thr Asn Phe Ala Leu Leu
   225          230          235          240
Ala Ile Thr Ile Ser Gly Glu Leu Leu Phe Glu Asn Gly Tyr Asp Leu
   245          250          255
35 Trp Gly Ser Gly Ala Gly Gln Arg Leu Ser Val Ala Tyr Asn Lys Ala
   260          265          270
Ala Thr Trp Ile Leu Asn Pro Glu Thr Phe Pro Tyr Phe Gln Pro Asn
   275          280          285
40 Leu Ile Gly Val His Asn Asn Ala Tyr Phe Ile Ile Leu Ala Lys His
   290          295          300
Tyr Ser Ser Pro Ser Ala Asp Glu Leu Leu Glu Gln Gly Asp Leu His
   305          310          315          320
Glu Asp Gly Phe Arg Leu Lys Leu Arg Ser Pro
   325          330
45

```

(2) INFORMATION FOR SEQ ID NO:132:

(i) SEQUENCE CHARACTERISTICS:

- 50 (A) LENGTH: 128 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

-215-

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

5 (A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...128

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:132:

Met Arg Gln Tyr Ser Gln Ser Ala Leu Asn Thr Asn His Gly Ala Trp
 1 5 10 15
 15 Gly Ile Leu Phe Asp Val Ser Ser Ala Leu Ala Leu Asp Asp His Ala
 20 25 30
 Leu Leu Gln Ser Ser Ala Asn Arg Trp Gln Glu Trp Val Phe Lys Ala
 35 40 45
 20 Ile Asp Glu Asn Gly Val Ile Ala Ser Ala Ile Thr Arg Ser Asp Thr
 50 55 60
 Ser Asp Tyr His Gly Gly Pro Thr Lys Gly Ile Lys Gly Ile Ala Tyr
 65 70 75 80
 Thr Asn Phe Ala Leu Leu Ala Ile Thr Ile Ser Gly Glu Leu Leu Phe
 85 90 95
 25 Glu Asn Gly Tyr Asp Leu Trp Gly Ser Gly Ala Gly Gln Arg Leu Ser
 100 105 110
 Val Ala Tyr Asn Lys Ala Ala Thr Trp Ile Leu Asn Pro Glu Thr Phe
 115 120 125

30 (2) INFORMATION FOR SEQ ID NO:133:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 245 amino acids

(B) TYPE: amino acid

35 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

40

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...245

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:133:

50 Leu Arg Thr Leu Leu Lys Met Leu Val Gly Val Ser Leu Leu Thr His
 1 5 10 15
 Ala Leu Met Ala Thr Glu Glu Ser Ala Ala Pro Ser Trp Thr Lys Asn
 20 25 30
 Leu Tyr Met Gly Phe Asn Tyr Gln Thr Gly Ser Ile Asn Leu Met Thr

-216-

```

      35          40          45
Asn Ile His Glu Val Arg Glu Val Thr Ser Tyr Gln Thr Gly Tyr Thr
  50          55          60
5  Asn Val Met Thr Ser Ile Asn Ser Val Lys Lys Leu Thr Asn Met Gly
  65          70          75          80
Ser Asn Gly Ile Gly Leu Val Met Gly Tyr Asn His Phe Phe His Pro
      85          90          95
Asp Lys Val Leu Gly Leu Arg Tyr Phe Ala Phe Leu Asp Trp Gln Gly
      100          105          110
10 Tyr Gly Met Arg Tyr Pro Lys Gly Tyr Tyr Gly Gly Asn Asn Met Ile
      115          120          125
Thr Tyr Gly Val Gly Val Asp Ala Ile Trp Asn Phe Phe Gln Gly Ser
      130          135          140
15 Phe Tyr Gln Asp Asp Ile Gly Val Asp Ile Gly Val Phe Gly Gly Ile
      145          150          155          160
Ala Ile Ala Gly Asn Ser Trp Tyr Ile Gly Asn Lys Gly Gln Glu Leu
      165          170          175
Leu Gly Ile Thr Asn Ser Ser Ala Val Asp Asn Thr Ser Phe Gln Phe
      180          185          190
20 Leu Phe Asn Phe Gly Phe Lys Ala Leu Phe Val Asp Glu His Glu Phe
      195          200          205
Glu Ile Gly Phe Lys Phe Pro Thr Leu Asn Asn Lys Tyr Tyr Thr Thr
      210          215          220
25 Asp Ala Leu Lys Val Gln Met Arg Arg Val Phe Ala Phe Tyr Val Gly
      225          230          235          240
Tyr Asn Tyr His Phe
      245

```

(2) INFORMATION FOR SEQ ID NO:134:

30

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 290 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

35

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

40

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc_feature

45

(B) LOCATION 1...290

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:134:

```

50 Met Phe Glu Glu Ile Thr Leu Ala His Lys Asp Leu Phe Ser Arg Phe
   1          5          10          15
Leu Gln Thr Gln Lys Ile Val Leu Ser Asp Val Ser Phe Thr Asn Cys
      20          25          30
Phe Leu Trp Gln His Ala Arg Leu Ile Gln Val Ala Val Ile Arg Asp
      35          40          45

```

-217-

Cys Leu Val Ile Gln Thr Thr Tyr Glu Asn Gln Lys Pro Phe Tyr Phe
 50 55 60
 Tyr Pro Ile Gly Lys Arg Pro His Glu Cys Val Lys Glu Leu Leu Glu
 65 70 75 80
 5 Leu Glu Lys Asn Leu Arg Phe His Ser Leu Thr Leu Glu Gln Lys Asp
 85 90 95
 Asp Leu Lys Asp Asn Phe Val Gly Val Phe Asp Phe Thr Tyr Asn Arg
 100 105 110
 10 Asp Arg Ser Asp Tyr Val Tyr Ser Ile Glu Glu Leu Ile Ala Leu Lys
 115 120 125
 Gly Lys Lys Tyr His Lys Lys Lys Asn His Leu Asn Gln Phe Leu Thr
 130 135 140
 Asn His Ala Asn Phe Val Tyr Glu Lys Ile Ser Pro Gln Asn Arg Lys
 145 150 155 160
 15 Glu Val Leu Glu Ala Ser Lys Ala Trp Phe Leu Glu Ser Gln Thr Asp
 165 170 175
 Asp Ile Gly Leu Ile Asn Glu Asn Lys Gly Ile Gln Ser Val Leu Glu
 180 185 190
 20 Asn Tyr Glu Ser Leu Asp Leu Lys Gly Gly Leu Ile Arg Val Asn Gly
 195 200 205
 Glu Ile Val Ser Phe Ser Phe Gly Glu Val Leu Asn Glu Glu Ser Ala
 210 215 220
 Leu Ile His Ile Glu Lys Ala Arg Thr Asp Ile Ala Gly Ala Tyr Gln
 225 230 235 240
 25 Ile Ile Asn Gln Gln Leu Leu Leu Asn Glu Phe Ser His Leu Thr Tyr
 245 250 255
 Ala Asn Arg Glu Glu Asp Leu Gly Leu Glu Gly Leu Arg Arg Ser Lys
 260 265 270
 30 Met Ser Tyr Asn Pro Val Phe Leu Ile Asp Lys Tyr Glu Ala Val Ala
 275 280 285
 Arg Asn
 290

(2) INFORMATION FOR SEQ ID NO:135:

35

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 110 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

40

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

45

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc_feature

50

(B) LOCATION 1...110

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:135:

Met Met Phe Ile Val Ala Val Leu Met Leu Ala Phe Leu Ile Phe Val

-218-

```

1           5           10           15
His Glu Leu Gly His Phe Ile Ile Ala Arg Ile Cys Gly Val Lys Val
                20           25           30
Glu Val Phe Ser Ile Gly Phe Gly Lys Lys Leu Trp Phe Phe Lys Leu
5           35           40           45
Phe Gly Thr Gln Phe Ala Leu Ser Leu Ile Pro Leu Gly Gly Tyr Val
        50           55           60
Lys Leu Lys Gly Met Asp Lys Glu Glu Asn Glu Glu Asn Lys Ile Asn
65           70           75           80
10 Gln Ala Asn Asp Ser Tyr Ala Lys Lys Ala Leu Ser Lys Ser Tyr Gly
        85           90           95
Tyr Cys Leu Val Gly Arg Phe Leu Ile Phe Phe Leu Arg Phe
                100           105           110

```

15 (2) INFORMATION FOR SEQ ID NO:136:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 351 amino acids

(B) TYPE: amino acid

20 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

25

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

30

(A) NAME/KEY: misc_feature

(B) LOCATION 1...351

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:136:

```

35 Met Met Phe Ile Val Ala Val Leu Met Leu Ala Phe Leu Ile Phe Val
1           5           10           15
His Glu Leu Gly His Phe Ile Ile Ala Arg Ile Cys Gly Val Lys Val
                20           25           30
Glu Val Phe Ser Ile Gly Phe Gly Lys Lys Leu Trp Phe Phe Lys Leu
40           35           40           45
Phe Gly Thr Gln Phe Ala Leu Ser Leu Ile Pro Leu Gly Gly Tyr Val
        50           55           60
Lys Leu Lys Gly Met Asp Lys Glu Glu Asn Glu Glu Asn Lys Ile Asn
65           70           75           80
45 Gln Ala Asn Asp Ser Tyr Ala Gln Lys Ser Pro Phe Gln Lys Leu Trp
        85           90           95
Ile Leu Phe Gly Gly Ala Phe Phe Asn Phe Leu Phe Ala Val Leu Val
                100           105           110
Tyr Phe Phe Leu Ala Leu Ser Gly Glu Lys Val Leu Leu Pro Val Ile
50           115           120           125
Gly Gly Leu Glu Lys Asn Ala Leu Glu Ala Gly Leu Leu Lys Gly Asp
        130           135           140
Arg Ile Leu Ser Ile Asn His Gln Lys Ile Ala Ser Phe Arg Glu Ile
145           150           155           160

```

- 219 -

```

      Arg Glu Ile Val Ala Arg Ser Gln Gly Glu Leu Ile Leu Glu Ile Glu
                165                170                175
      Arg Asn Asn Gln Ile Leu Glu Lys Arg Leu Thr Pro Lys Ile Val Ala
                180                185                190
5    Val Ile Ser Glu Ser Asn Asp Pro Asn Glu Ile Ile Lys Tyr Lys Ile
      195                200                205
      Ile Gly Ile Lys Pro Asp Met Gln Lys Met Gly Val Val Ser Tyr Ser
      210                215                220
      Val Phe Gln Ala Phe Glu Lys Ala Leu Ser Arg Phe Lys Glu Gly Val
10   225                230                235                240
      Val Leu Ile Val Asp Ser Leu Arg Arg Leu Ile Met Gly Ser Ala Ser
                245                250                255
      Val Lys Glu Leu Ser Gly Val Ile Gly Ile Val Gly Ala Leu Ser His
                260                265                270
15   Ala Asn Ser Val Ser Met Leu Leu Phe Gly Ala Phe Leu Ser Ile
      275                280                285
      Asn Leu Gly Ile Leu Asn Leu Leu Pro Ile Pro Ala Leu Asp Gly Ala
      290                295                300
      Gln Met Leu Gly Val Val Phe Lys Asn Ile Phe His Ile Ala Leu Pro
20   305                310                315                320
      Thr Pro Ile Gln Asn Ala Leu Trp Leu Val Gly Val Gly Phe Leu Val
                325                330                335
      Phe Val Met Phe Leu Gly Leu Phe Asn Asp Ile Thr Arg Leu Leu
                340                345                350
25

```

(2) INFORMATION FOR SEQ ID NO:137:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 100 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
 (B) LOCATION 1...100

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:137:

```

45   Met Gln Lys Asn Leu Asp Ser Leu Leu Glu Asn Leu Arg Ala Glu Ile
      1          5          10          15
      Asp Ala Leu Asp Asn Glu Leu Ser Asp Leu Leu Asp Lys Arg Leu Gly
                20          25          30
50   Ile Ala Leu Lys Ile Ala Leu Ile Lys Gln Glu Ser Pro Gln Glu Asn
      35          40          45
      Pro Ile Tyr Cys Pro Lys Arg Glu Gln Glu Ile Leu Lys Arg Leu Ser
      50          55          60
      Gln Arg Gly Phe Lys His Leu Asn Gly Glu Ile Leu Ala Ser Phe Tyr

```

```

65          70          75          80
Ala Glu Val Phe Lys Ile Ser Arg Asn Phe Gln Glu Asn Ala Leu Lys
          85          90          95
Glu Leu Lys Lys
5          100

```

(i) SEQUENCE CHARACTERISTICS:

10 (A) LENGTH: 174 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

15

(vi) ORIGINAL SOURCE:

20

(A) NAME/KEY: misc_feature
(B) LOCATION 1...174

25

30

50.

(A) LENGTH: 471 amino acids
(B) TYPE: amino acid

- 221 -

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

5 (iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

10 (ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...471

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:139:

15 Met Lys Asn Thr Asn Thr Lys Glu Ile Lys Asn Thr Arg Met Lys Lys
 1 5 10 15
 Gly Tyr Ser Gln Tyr His Thr Leu Lys Lys Gly Leu Leu Lys Thr Ala
 20 20 25 30
 Leu Leu Phe Ser Leu Pro Leu Ser Val Ala Leu Ala Glu Asp Asp Gly
 35 40 45
 Phe Tyr Met Gly Val Gly Tyr Gln Ile Gly Gly Ala Gln Gln Asn Ile
 50 55 60
 Asn Asn Lys Gly Ser Thr Leu Arg Asn Asn Val Ile Asp Asp Phe Arg
 25 65 70 75 80
 Gln Val Gly Val Gly Met Ala Gly Gly Asn Gly Leu Leu Ala Leu Ala
 85 90 95
 Thr Asn Thr Thr Met Asp Ala Leu Leu Gly Ile Gly Asn Gln Ile Val
 100 105 110
 30 Asn Thr Asn Thr Thr Val Gly Asn Asn Asn Ala Glu Leu Thr Gln Phe
 115 120 125
 Lys Lys Ile Leu Pro Gln Ile Glu Gln Arg Phe Glu Thr Asn Lys Asn
 130 135 140
 Ala Tyr Ser Val Gln Ala Leu Gln Val Tyr Leu Ser Asn Val Leu Tyr
 35 145 150 155 160
 Asn Leu Val Asn Asn Ser Asn Asn Gly Ser Asn Asn Gly Val Val Pro
 165 170 175
 Glu Tyr Val Gly Ile Ile Lys Val Leu Tyr Gly Ser Gln Asn Glu Phe
 180 185 190
 40 Ser Leu Leu Ala Thr Glu Ser Val Ala Leu Leu Asn Ala Leu Thr Arg
 195 200 205
 Val Asn Leu Asp Ser Asn Ser Val Phe Leu Lys Gly Leu Leu Ala Gln
 210 215 220
 Met Gln Leu Phe Asn Asp Thr Ser Ser Ala Lys Leu Gly Gln Ile Ala
 45 225 230 235 240
 Glu Asn Leu Lys Asn Gly Gly Ala Gly Ala Met Leu Gln Lys Asp Val
 245 250 255
 Lys Thr Ile Ser Asp Arg Ile Ala Thr Tyr Gln Glu Asn Leu Lys Gln
 260 265 270
 50 Leu Gly Gly Met Leu Lys Asn Tyr Asp Glu Pro Tyr Leu Pro Gln Phe
 275 280 285
 Gly Pro Gly Thr Ser Ser Gln His Gly Val Ile Asn Gly Phe Gly Ile
 290 295 300
 Gln Val Gly Tyr Lys Gln Phe Phe Gly Ser Lys Lys Asn Ile Gly Leu

- 222 -

```

      305              310              315              320
      Arg Tyr Tyr Ala Phe Phe Asp Tyr Gly Phe Thr Gln Leu Gly Ser Leu
                    325              330              335
5    Asn Ser Ala Val Lys Ala Asn Ile Phe Thr Tyr Gly Ala Gly Thr Asp
      340              345              350
      Phe Leu Trp Asn Ile Phe Arg Arg Val Phe Ser Asp Gln Ser Leu Asn
      355              360              365
      Val Gly Val Phe Gly Gly Ile Gln Ile Ala Gly Asn Thr Trp Asp Ser
      370              375              380
10   Ser Leu Arg Gly Gln Ile Glu Asn Ser Phe Lys Glu Tyr Pro Thr Pro
      385              390              395              400
      Thr Asn Phe Gln Phe Leu Phe Asn Leu Gly Leu Arg Ala His Phe Ala
      405              410              415
15   Ser Thr Met His Arg Arg Phe Leu Ser Ala Ser Gln Ser Ile Gln His
      420              425              430
      Gly Met Glu Phe Gly Val Lys Ile Pro Ala Ile Asn Gln Arg Tyr Leu
      435              440              445
      Lys Ala Asn Gly Ala Asp Val Asp Tyr Arg Arg Leu Tyr Ala Phe Tyr
      450              455              460
20   Ile Asn Tyr Thr Ile Gly Phe
      465              470

```

(2) INFORMATION FOR SEQ ID NO:140:

- 25 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 129 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
- 30 (ii) MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: YES
- (vi) ORIGINAL SOURCE:
 35 (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
 (A) NAME/KEY: misc_feature
 (B) LOCATION 1...129

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:140:

```

Met Lys Ser Ile Arg Arg Gly Asp Gly Leu Asn Val Val Pro Phe Ile
1              5              10              15
45 Asp Ile Met Leu Val Leu Leu Ala Ile Val Leu Ser Ile Ser Thr Phe
      20              25              30
      Ile Ala Gln Gly Lys Ile Lys Val Ser Leu Pro Asn Ala Lys Asn Ala
      35              40              45
50 Glu Lys Ser Gln Pro Asn Asp Gln Lys Val Val Val Ile Ser Val Asp
      50              55              60
      Glu His Asp Asn Ile Phe Val Asp Asp Lys Pro Thr Asn Leu Glu Ala
      65              70              75              80
      Leu Ser Ala Val Val Lys Gln Thr Asp Pro Lys Thr Leu Ile Asp Leu
      85              90              95

```

- 223 -

Lys Ser Asp Lys Ser Ser Arg Phe Glu Thr Phe Ile Ser Ile Met Asp
 100 105 110
 Ile Leu Lys Glu His Asn His Glu Asn Phe Ser Ile Ser Thr Gln Ala
 115 120 125
 5 Gln

(2) INFORMATION FOR SEQ ID NO:141:

- 10 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 75 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
- 15 (ii) MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: YES
- (vi) ORIGINAL SOURCE:
 20 (A) ORGANISM: Helicobacter pylori
- (ix) FEATURE:
 (A) NAME/KEY: misc_feature
 (B) LOCATION 1...75
- 25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:141:

Met Leu Val Leu Leu Ala Ile Val Leu Ser Ile Ser Thr Phe Ile Ala
 1 5 10 15
 30 Gln Gly Lys Ile Lys Val Ser Leu Pro Asn Ala Lys Asn Ala Glu Lys
 20 25 30
 Ser Arg Pro Asn Asp Gln Lys Val Val Val Ile Ser Val Asp Glu His
 35 40 45
 Asp Asn Ile Phe Val Asp Asp Lys Pro Thr Asn Leu Glu Ala Leu Ser
 35 50 55 60
 Ala Val Val Lys Gln Thr Asp Pro Lys Thr Leu
 65 70 75

(2) INFORMATION FOR SEQ ID NO:142:

- 40 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 223 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
- 45 (ii) MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: YES
- (vi) ORIGINAL SOURCE:
 50 (A) ORGANISM: Helicobacter pylori
- (ix) FEATURE:
 (A) NAME/KEY: misc_feature

- 224 -

(B) LOCATION 1...223

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:142:

```

5  Met Phe Ser Leu Ser Tyr Val Ser Lys Lys Phe Leu Ser Val Leu Leu
   1           5           10           15
   Leu Ile Ser Leu Phe Leu Ser Ala Cys Lys Ser Asn Asn Lys Asp Lys
     20           25           30
10  Leu Asp Glu Asn Leu Leu Ser Ser Gly Thr Gln Ser Ser Lys Glu Leu
     35           40           45
   Asn Asp Lys Arg Asp Asn Ile Asp Lys Lys Ser Tyr Ala Gly Leu Glu
     50           55           60
   Asp Val Phe Leu Asp Asn Lys Ser Ile Ser Pro Asn Asp Lys Tyr Met
   65           70           75           80
15  Leu Leu Val Phe Gly Arg Asn Gly Cys Ser Tyr Cys Glu Arg Leu Lys
     85           90           95
   Lys Asp Leu Lys Asn Val Lys Glu Leu Arg Asn Tyr Ile Lys Glu His
     100          105          110
   Phe Ser Ala Tyr Tyr Val Asn Ile Ser Tyr Ser Lys Glu His Asn Phe
   115          120          125
20  Lys Val Gly Asp Lys Asp Lys Asn Asp Glu Lys Glu Ile Lys Met Ser
   130          135          140
   Thr Glu Glu Leu Ala Gln Ile Tyr Ala Val Gln Ser Thr Pro Thr Ile
   145          150          155          160
25  Val Leu Ser Asp Lys Thr Gly Lys Thr Ile Tyr Glu Leu Pro Gly Tyr
     165          170          175
   Met Pro Ser Val Gln Phe Leu Ala Val Leu Glu Phe Ile Gly Asp Gly
     180          185          190
   Lys Tyr Gln Asp Thr Lys Asn Asp Glu Asp Leu Thr Lys Lys Leu Lys
   195          200          205
30  Ala Tyr Ile Lys Tyr Lys Thr Asn Leu Ser Lys Ser Lys Ser Ser
     210          215          220

```

(2) INFORMATION FOR SEQ ID NO:143:

35

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 116 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

40

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

45

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc_feature

50

(B) LOCATION 1...116

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:143:

Leu Met Lys Ser Lys Ile Thr His Phe Ile Val Ile Ser Phe Val Leu

- 225 -

```

1           5           10           15
Ser Val Leu Ser Ala Cys Lys Asp Glu Pro Lys Lys Ser Ser Gln Ser
                20                25                30
5 His Gln Asn Asn Thr Lys Thr Thr Gln Asn Asn Gln Ile Asn Gln Pro
                35                40                45
Asn Lys Asp Ile Lys Lys Ile Glu His Glu Glu Glu Asp Glu Lys Val
                50                55                60
Thr Lys Glu Val Asn Asp Leu Ile Asn Asn Glu Asn Lys Ile Asp Glu
65                70                75                80
10 Ile Asn Asn Glu Glu Asn Ala Asp Pro Ser Gln Lys Arg Thr Asn Asn
                85                90                95
Val Leu Gln Arg Ala Thr Asn His Gln Asp Asn Leu Ser Ser Pro Leu
                100                105                110
15 Asn Arg Lys Tyr
                115

```

(2) INFORMATION FOR SEQ ID NO:144:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 79 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: YES
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
- (A) NAME/KEY: misc_feature
- (B) LOCATION 1...79
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:144:

```

Met Phe Glu Lys Ile Arg Lys Ile Leu Ala Asp Ile Glu Asp Ser Gln
1           5           10           15
40 Asn Glu Ile Glu Met Leu Leu Lys Leu Ala Asn Leu Ser Leu Gly Asp
                20                25                30
Phe Ile Glu Ile Lys Arg Gly Ser Met Asp Met Pro Lys Gly Val Asn
                35                40                45
Glu Ala Phe Phe Thr Gln Leu Ser Glu Glu Val Glu Arg Leu Lys Glu
50                55                60
45 Leu Ile Asn Ala Leu Asn Lys Ile Lys Lys Gly Leu Leu Val Phe
65                70                75

```

(2) INFORMATION FOR SEQ ID NO:145:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 51 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

- 226 -

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

5 (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

10 (A) NAME/KEY: misc_feature

(B) LOCATION 1...51

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:145:

15 Met Ser Met Phe Ile Ser Asn Leu Ala Phe Thr Ser Glu His Lys Asp
 1 5 10 15
 Ala Met Glu Val Ala Lys Ile Ala Ile Leu Leu Gly Ser Leu Ile Ser
 20 25 30
 Gly Ile Ile Gly Ala Leu Tyr Leu Phe Ala Leu Asp Lys Arg Ala Ala
 35 40 45
 20 Leu Lys Lys
 50

(2) INFORMATION FOR SEQ ID NO:146:

25 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 449 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

30 (ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

35 (A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc_feature

40 (B) LOCATION 1...449

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:146:

Met Gly Leu Lys Ile Lys Ile Leu Arg Leu Ser Met Asn Leu Lys Lys
 1 5 10 15
 45 Thr Glu Asn Ala Leu Ser Leu Thr Leu Lys Asn Phe Ile Lys Ser Glu
 20 25 30
 Ser Phe Gly Gly Ile Phe Leu Phe Leu Asn Ala Val Leu Ala Met Val
 35 40 45
 Val Ala Asn Ser Phe Leu Lys Glu Ser Tyr Phe Ala Leu Trp His Thr
 50 55 60
 Pro Phe Gly Phe Gln Val Gly Asp Phe Phe Ile Gly Phe Ser Leu His
 65 70 75 80
 Asn Trp Ile Asp Asp Val Leu Met Ala Leu Phe Phe Leu Met Ile Gly
 85 90 95

- 227 -

Leu Glu Ile Lys Arg Glu Leu Leu Phe Gly Glu Leu Ser Ser Phe Lys
 100 105 110
 Lys Ala Ser Phe Pro Val Ile Ala Ala Ile Gly Gly Met Ile Ala Pro
 115 120 125
 5 Gly Leu Ile Tyr Phe Phe Leu Asn Ala Asn Thr Pro Ser Gln His Gly
 130 135 140
 Phe Gly Ile Pro Met Ala Thr Asp Ile Ala Phe Ala Leu Gly Val Ile
 145 150 155 160
 Met Leu Leu Gly Lys Arg Val Pro Thr Ala Leu Lys Val Phe Leu Ile
 10 165 170 175
 Thr Leu Ala Val Ala Asp Asp Leu Gly Ala Ile Val Val Ile Ala Leu
 180 185 190
 Phe Tyr Thr Thr Asn Leu Lys Phe Ala Trp Leu Leu Gly Ala Leu Gly
 195 200 205
 15 Val Val Leu Val Leu Ala Ile Leu Asn Arg Leu Asn Ile Arg Ser Leu
 210 215 220
 Ile Pro Tyr Leu Leu Leu Gly Val Leu Leu Trp Phe Cys Val His Gln
 225 230 235 240
 Ser Gly Ile His Ala Thr Ile Ala Ala Val Leu Ala Phe Met Ile
 20 245 250 255
 Pro Val Lys Ile Pro Lys Asp Ser Lys Asn Val Glu Leu Leu Glu Leu
 260 265 270
 Gly Lys Arg Tyr Ala Glu Thr Ser Ser Gly Val Leu Leu Thr Lys Glu
 275 280 285
 25 Gln Gln Glu Ile Leu His Ser Ile Glu Glu Lys Ala Ser Ala Leu Gln
 290 295 300
 Ser Pro Leu Glu Arg Leu Glu His Phe Leu Ala Pro Ile Ser Gly Tyr
 305 310 315 320
 Phe Ile Met Pro Leu Phe Ala Phe Ala Asn Ala Gly Val Ser Val Asp
 30 325 330 335
 Ser Ser Ile Asn Leu Glu Val Asp Lys Val Leu Leu Gly Val Ile Leu
 340 345 350
 Gly Leu Cys Leu Gly Lys Pro Leu Gly Ile Phe Leu Ile Thr Phe Ile
 355 360 365
 35 Ser Glu Lys Leu Lys Ile Thr Ala Arg Pro Lys Gly Ile Gly Trp Trp
 370 375 380
 His Ile Leu Gly Ala Gly Leu Leu Ala Gly Ile Gly Phe Thr Met Ser
 385 390 395 400
 Met Phe Ile Ser Asn Leu Ala Phe Thr Ser Glu His Lys Asp Ala Met
 40 405 410 415
 Glu Val Ala Lys Ile Ala Ile Leu Leu Gly Ser Leu Ile Ser Gly Ile
 420 425 430
 Ile Gly Ala Leu Tyr Leu Phe Ala Leu Asp Lys Arg Ala Ala Leu Lys
 435 440 445
 45 Lys

(2) INFORMATION FOR SEQ ID NO:147:

50

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 815 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

- 228 -

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

5 (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc_feature

10 (B) LOCATION 1...815

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:147:

```

15 Met Asn Asp Lys Arg Phe Arg Lys Tyr Cys Ser Phe Ser Ile Phe Leu
   1           5           10           15
   Ser Leu Leu Gly Thr Phe Glu Leu Glu Ala Lys Glu Glu Glu Lys Glu
      20           25           30
   Glu Lys Lys Thr Glu Arg Asn Lys Asp Lys Glu Lys Asn Ala Gln His
      35           40           45
20 Thr Leu Gly Lys Val Thr Thr Gln Ala Ala Lys Ile Phe Asn Tyr Asn
   50           55           60
   Asn Gln Thr Thr Ile Ser Ser Lys Glu Leu Glu Arg Arg Gln Ala Asn
   65           70           75           80
   Gln Ile Ser Asp Met Phe Arg Arg Asn Pro Asn Ile Asn Val Gly Gly
25 Gly Ala Val Ile Ala Gln Lys Ile Tyr Val Arg Gly Ile Glu Asp Arg
   100           105           110           115
   Leu Ala Arg Val Thr Val Asp Gly Val Ala Gln Met Gly Ala Ser Tyr
      115           120           125
30 Gly His Gln Gly Asn Thr Ile Ile Asp Pro Gly Met Leu Lys Ser Val
   130           135           140
   Val Val Thr Lys Gly Ala Ala Gln Ala Ser Ala Gly Pro Met Ala Leu
   145           150           155           160
   Ile Gly Ala Ile Lys Met Glu Thr Arg Ser Ala Ser Asp Phe Ile Pro
35 Lys Gly Lys Asp Tyr Ala Ile Ser Gly Ala Ala Thr Phe Leu Thr Asn
      165           170           175           180
   Phe Gly Asp Arg Glu Thr Ile Met Gly Ala Tyr Arg Asn His His Phe
      180           185           190           195
40 Asp Ala Leu Leu Tyr Tyr Thr His Gln Asn Ile Phe Tyr Tyr Arg Asp
   210           215           220
   Gly Asp Asn Ala Met Lys Asn Leu Phe Asp Pro Lys Ala Asp Asn Lys
   225           230           235           240
   Val Thr Ala Ser Pro Ser Glu Gln Asn Asn Val Met Ala Lys Ile Asn
45 Gly Tyr Leu Ser Glu Arg Asp Thr Leu Thr Leu Ser Tyr Asn Met Thr
      245           250           255           260
   Arg Asp Asn Ala Asn Arg Pro Leu Arg Ala Asn Phe Thr Gly Thr Phe
      260           265           270           275
50 Leu Pro Tyr Ser Cys Gly Asp Phe Asn Ala Phe Pro Asn Glu Lys Asn
   290           295           300
   Pro Ser Asp Cys Leu Phe Glu Asn Asp Ala Ser Leu Phe Lys Thr Tyr
   305           310           315           320
   Ser Val Asn Leu Val His Asn Val Ser Leu Asn Tyr Glu Arg Glu Gly

```


- 230 -

755 760 765
 Val Phe Asn Asn Val Phe Asn Gln Gln Tyr Ile Asp Gln Ala Ser Pro
 770 775 780
 Val Met Ser Pro Asp Glu Pro Asn Gln Asp Lys Tyr Ala Arg Gly Met
 5 785 790 795 800
 Ala Glu Pro Gly Phe Asn Ala Arg Phe Glu Ile Ser Tyr Lys Phe
 805 810 815

(2) INFORMATION FOR SEQ ID NO:148:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 814 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...814

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:148:

30 Met Thr Ser Val Leu Glu Lys Tyr Cys Ser Phe Ser Ile Phe Leu Ser
 1 5 10 15
 Leu Leu Gly Thr Phe Glu Leu Glu Ala Lys Glu Glu Glu Lys Glu Glu
 20 25 30
 Lys Lys Thr Glu Arg Asn Lys Asp Lys Glu Lys Asn Ala Gln His Thr
 35 40 45
 35 Leu Gly Lys Val Thr Thr Gln Ala Ala Lys Ile Phe Asn Tyr Asn Asn
 50 55 60
 Gln Thr Thr Ile Ser Ser Lys Glu Leu Glu Arg Arg Gln Ala Asn Gln
 65 70 75 80
 Ile Ser Asp Met Phe Arg Arg Asn Pro Asn Ile Asn Val Gly Gly Gly
 40 85 90 95
 Ala Val Ile Ala Gln Lys Ile Tyr Val Arg Gly Ile Glu Asp Arg Leu
 100 105 110
 Ala Arg Val Thr Val Asp Gly Val Ala Gln Met Gly Ala Ser Tyr Gly
 115 120 125
 45 His Gln Gly Asn Thr Ile Ile Asp Pro Gly Met Leu Lys Ser Val Val
 130 135 140
 Val Thr Lys Gly Ala Ala Gln Ala Ser Ala Gly Pro Met Ala Leu Ile
 145 150 155 160
 Gly Ala Ile Lys Met Glu Thr Arg Ser Ala Ser Asp Phe Ile Pro Lys
 50 165 170 175
 Gly Lys Asp Tyr Ala Ile Ser Gly Ala Ala Thr Phe Leu Thr Asn Phe
 180 185 190
 Gly Asp Arg Glu Thr Ile Met Gly Ala Tyr Arg Asn His His Phe Asp
 195 200 205

- 231 -

	Ala	Leu	Leu	Tyr	Tyr	Thr	His	Gln	Asn	Ile	Phe	Tyr	Tyr	Arg	Asp	Gly
	210						215					220				
	Asp	Asn	Ala	Met	Lys	Asn	Leu	Phe	Asp	Pro	Lys	Ala	Asp	Asn	Lys	Val
	225					230					235					240
5	Thr	Ala	Ser	Pro	Ser	Glu	Gln	Asn	Asn	Val	Met	Ala	Lys	Ile	Asn	Gly
					245					250					255	
	Tyr	Leu	Ser	Glu	Arg	Asp	Thr	Leu	Thr	Leu	Ser	Tyr	Asn	Met	Thr	Arg
				260					265					270		
10	Asp	Asn	Ala	Asn	Arg	Pro	Leu	Arg	Ala	Asn	Phe	Thr	Gly	Thr	Phe	Leu
			275					280					285			
	Pro	Tyr	Ser	Cys	Gly	Asp	Phe	Asn	Ala	Phe	Pro	Asn	Glu	Lys	Asn	Pro
	290						295					300				
	Ser	Asp	Cys	Leu	Phe	Glu	Asn	Asp	Ala	Ser	Leu	Phe	Lys	Thr	Tyr	Ser
	305					310					315					320
15	Val	Asn	Leu	Val	His	Asn	Val	Ser	Leu	Asn	Tyr	Glu	Arg	Glu	Gly	Gly
					325					330					335	
	Ser	Arg	Phe	Gly	Asp	Pro	Lys	Leu	Lys	Ile	Asn	Gly	Tyr	Thr	Ser	Ile
				340					345					350		
20	Arg	Asn	Val	Gln	Ile	Asp	Pro	Leu	Phe	Arg	Pro	Ser	Asp	Ile	Ala	Thr
			355					360					365			
	Thr	Ile	Pro	Phe	Thr	Pro	Asn	Pro	Gln	Leu	Ser	Gln	Gly	Glu	Glu	Asn
	370						375					380				
	Gln	Cys	Val	Ala	Gln	Gly	Gly	Ile	Tyr	Asp	Ala	Leu	Lys	Gln	Thr	Cys
	385					390					395					400
25	Ser	Ile	Thr	Phe	Lys	Ser	Leu	Gly	Gly	Gly	Ser	Val	Val	Ala	Asn	Lys
					405					410					415	
	Asn	Leu	Phe	Ile	Ile	Asn	Ser	Gly	Phe	Asn	Ala	Asn	Val	Ile	His	Thr
				420					425					430		
30	Ile	Asp	His	Lys	Asn	Asp	Asn	Leu	Leu	Glu	Tyr	Gly	Leu	Asn	Tyr	Gln
			435					440					445			
	Asn	Leu	Thr	Thr	Phe	Asp	Lys	Ala	Ile	Pro	Asp	Ser	Glu	Leu	Val	Lys
	450						455					460				
	Pro	Gly	Asp	Ala	Pro	Asp	Ala	Cys	Leu	Arg	Val	Thr	Gly	Pro	Asp	Asp
	465					470					475					480
35	Pro	Asn	Met	Asn	Gly	Arg	Cys	Gln	Arg	Asn	Gly	Ala	Thr	Ala	Asn	Val
					485					490					495	
	Val	Gly	Val	Tyr	Ala	Gln	Ala	Asn	Tyr	Thr	Leu	His	Pro	Met	Val	Thr
				500					505					510		
40	Leu	Gly	Ala	Gly	Thr	Arg	Tyr	Asp	Val	Tyr	Thr	Leu	Val	Asp	Lys	Asp
			515					520					525			
	Trp	Gln	Leu	His	Val	Thr	Gln	Gly	Phe	Ser	Pro	Ser	Ala	Ala	Leu	Asn
	530						535					540				
	Val	Ser	Pro	Leu	Glu	Asn	Leu	Asn	Phe	Arg	Leu	Ser	Tyr	Ala	Tyr	Val
	545					550					555					560
45	Thr	Arg	Gly	Pro	Met	Pro	Gly	Gly	Leu	Val	Trp	Met	Arg	Gln	Asp	Asn
					565					570					575	
	Leu	Arg	Tyr	Asn	Arg	Asn	Leu	Lys	Pro	Glu	Ile	Gly	Gln	Asn	Ala	Glu
				580					585					590		
50	Phe	Asn	Thr	Glu	Tyr	Ser	Ser	Gln	Tyr	Phe	Asp	Phe	Arg	Ala	Ala	Gly
			595					600					605			
	Phe	Val	Gln	Leu	Ile	Ser	Asn	Tyr	Ile	Asn	Gln	Phe	Ser	Ser	Thr	Leu
	610						615					620				
	Phe	Val	Thr	Asn	Leu	Pro	Ala	Gln	Asp	Ile	Ile	Tyr	Val	Pro	Gly	Tyr
	625					630					635					640

- 232 -

Glu Val Ser Gly Thr Ala Lys Tyr Lys Gly Phe Ser Leu Gly Leu Ser
 645 650 655
 Val Ala Arg Ser Trp Pro Ser Leu Lys Gly Arg Leu Ile Ala Asp Val
 660 665 670
 5 Tyr Glu Leu Ala Ala Thr Thr Gly Asn Val Phe Ile Leu Thr Ala Ser
 675 680 685
 Tyr Thr Ile Pro Arg Thr Gly Leu Ser Ile Thr Trp Leu Ser Arg Phe
 690 695 700
 10 Val Thr Asn Leu Ser Tyr Cys Ser Tyr Ser Pro Tyr Arg Asn Gly Pro
 705 710 715 720
 Thr Asp Ile Asp Arg Arg Pro Ser Asn Cys Pro Lys Thr Pro Gly Ile
 725 730 735
 Phe His Val His Lys Pro Gly Tyr Gly Val Ser Ser Phe Phe Ile Thr
 740 745 750
 15 Tyr Lys Pro Thr Tyr Lys Lys Leu Lys Gly Leu Ser Leu Asn Ala Val
 755 760 765
 Phe Asn Asn Val Phe Asn Gln Gln Tyr Ile Asp Gln Ala Ser Pro Val
 770 775 780
 20 Met Ser Pro Asp Glu Pro Asn Gln Asp Lys Tyr Ala Arg Gly Met Ala
 785 790 795 800
 Glu Pro Gly Phe Asn Ala Arg Phe Glu Ile Ser Tyr Lys Phe
 805 810

(2) INFORMATION FOR SEQ ID NO:149:

25

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 527 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

30

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

35

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

40

- (A) NAME/KEY: misc_feature
 (B) LOCATION 1...527

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:149:

45 Met Lys Gln Asn Leu Lys Pro Phe Lys Met Ile Lys Glu Asn Leu Met
 1 5 10 15
 Thr Gln Ser Gln Lys Val Arg Phe Leu Ala Pro Leu Ser Leu Ala Leu
 20 25 30
 Ser Leu Ser Phe Asn Pro Val Gly Ala Glu Glu Asp Gly Gly Phe Met
 35 40 45
 50 Thr Phe Gly Tyr Glu Leu Gly Gln Val Val Gln Gln Val Lys Asn Pro
 50 55 60
 Gly Lys Ile Lys Ala Glu Glu Leu Ala Gly Leu Leu Asn Ser Thr Thr
 65 70 75 80
 Thr Asn Asn Thr Asn Ile Asn Ile Ala Gly Thr Gly Gly Asn Val Ala

- 233 -

				85				90				95				
	Gly	Thr	Leu	Gly	Asn	Leu	Phe	Met	Asn	Gln	Leu	Gly	Asn	Leu	Ile	Asp
				100					105				110			
5	Leu	Tyr	Pro	Thr	Leu	Lys	Thr	Asn	Asn	Leu	His	Gln	Cys	Gly	Ser	Thr
			115					120					125			
	Asn	Ser	Gly	Asn	Gly	Ala	Thr	Ala	Ala	Ala	Ala	Thr	Asn	Asn	Ser	Pro
			130				135					140				
	Cys	Phe	Gln	Gly	Asn	Leu	Ala	Leu	Tyr	Asn	Glu	Met	Val	Asp	Ser	Ile
	145					150					155				160	
10	Lys	Thr	Leu	Ser	Gln	Asn	Ile	Ser	Lys	Asn	Ile	Phe	Gln	Gly	Asp	Asn
				165						170					175	
	Asn	Thr	Thr	Ser	Ala	Asn	Leu	Ser	Asn	Gln	Leu	Ser	Glu	Leu	Asn	Thr
			180						185					190		
15	Ala	Ser	Val	Tyr	Leu	Thr	Tyr	Met	Asn	Ser	Phe	Leu	Asn	Ala	Asn	Asn
			195					200						205		
	Gln	Ala	Gly	Gly	Ile	Phe	Gln	Asn	Asn	Thr	Asn	Gln	Ala	Tyr	Glu	Asn
	210						215					220				
	Gly	Val	Thr	Ala	Gln	Gln	Ile	Ala	Tyr	Val	Leu	Lys	Gln	Ala	Ser	Ile
	225					230					235				240	
20	Thr	Met	Gly	Pro	Ser	Gly	Asp	Ser	Gly	Ala	Ala	Gly	Ala	Phe	Leu	Asp
				245						250					255	
	Ala	Ala	Leu	Ala	Gln	His	Val	Phe	Asn	Ser	Ala	Asn	Ala	Gly	Asn	Asp
			260						265					270		
25	Leu	Ser	Ala	Lys	Glu	Phe	Thr	Ser	Leu	Val	Gln	Asn	Ile	Val	Asn	Asn
			275					280					285			
	Ser	Gln	Asn	Ala	Leu	Thr	Leu	Ala	Asn	Asn	Ala	Asn	Ile	Ser	Asn	Ser
	290						295					300				
	Thr	Gly	Tyr	Gln	Val	Ser	Tyr	Gly	Gly	Asn	Ile	Asp	Gln	Ala	Arg	Ser
	305				310						315				320	
30	Thr	Gln	Leu	Leu	Asn	Asn	Thr	Thr	Asn	Thr	Leu	Ala	Lys	Val	Thr	Ala
				325						330					335	
	Leu	Asn	Asn	Glu	Leu	Lys	Ala	Asn	Pro	Trp	Leu	Gly	Asn	Phe	Ala	Ala
			340						345					350		
35	Gly	Asn	Ser	Ser	Gln	Val	Asn	Ala	Phe	Asn	Gly	Phe	Ile	Thr	Lys	Ile
		355						360					365			
	Gly	Tyr	Lys	Gln	Phe	Phe	Gly	Glu	Asn	Lys	Asn	Val	Gly	Leu	Arg	Tyr
	370						375					380				
	Tyr	Gly	Phe	Phe	Ser	Tyr	Asn	Gly	Ala	Gly	Val	Gly	Asn	Gly	Pro	Thr
	385				390						395				400	
40	Tyr	Asn	Gln	Val	Asn	Leu	Leu	Thr	Tyr	Gly	Val	Gly	Thr	Asp	Val	Leu
				405						410					415	
	Tyr	Asn	Val	Phe	Ser	Arg	Ser	Phe	Gly	Ser	Arg	Ser	Leu	Asn	Ala	Gly
			420						425					430		
45	Phe	Phe	Gly	Gly	Ile	Gln	Leu	Ala	Gly	Asp	Thr	Tyr	Ile	Ser	Thr	Leu
			435					440					445			
	Arg	Asn	Ser	Pro	Gln	Leu	Ala	Leu	Arg	Pro	Thr	Ala	Thr	Lys	Phe	Gln
	450						455					460				
	Phe	Leu	Phe	Asp	Val	Gly	Leu	Arg	Met	Asn	Phe	Gly	Ile	Leu	Lys	Lys
	465				470						475				480	
50	Asp	Leu	Lys	Ser	His	Asn	Gln	His	Ser	Ile	Glu	Ile	Gly	Val	Gln	Ile
				485						490					495	
	Pro	Thr	Ile	Tyr	Asn	Thr	Tyr	Tyr	Lys	Ala	Gly	Gly	Ala	Glu	Val	Lys
			500						505					510		
	Tyr	Phe	Arg	Pro	Tyr	Ser	Val	Tyr	Trp	Val	Tyr	Gly	Tyr	Ala	Phe	

- 234 -

515

520

525

(2) INFORMATION FOR SEQ ID NO:150:

5

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 459 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

10

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

15

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...459

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:150:

	Val	Val	Leu	Leu	Thr	Met	Thr	Lys	Arg	Leu	Phe	Lys	Gly	Leu	Leu	Ala	
	1				5					10					15		
25	Ile	Ser	Leu	Ala	Val	Ser	Leu	His	Gly	Gly	Glu	Val	Lys	Glu	Lys	Lys	
				20					25					30			
	Pro	Val	Lys	Pro	Val	Lys	Glu	Asp	Pro	Gln	Glu	Leu	Ala	Ala	Lys	Arg	
			35					40					45				
30	Val	Glu	Ala	Phe	Ser	Arg	Phe	Ser	Asn	Val	Val	Thr	Glu	Ile	Glu	Lys	
		50					55					60					
	Lys	Tyr	Val	Asp	Lys	Ile	Ser	Ile	Ser	Glu	Ile	Met	Thr	Lys	Ala	Ile	
	65					70					75				80		
	Glu	Gly	Leu	Leu	Ser	Asn	Leu	Asp	Ala	His	Ser	Ala	Tyr	Leu	Asn	Glu	
					85					90				95			
35	Lys	Lys	Phe	Lys	Glu	Phe	Gln	Ala	Gln	Thr	Glu	Gly	Glu	Phe	Gly	Gly	
				100					105					110			
	Leu	Gly	Ile	Thr	Val	Gly	Met	Arg	Asp	Gly	Val	Leu	Thr	Val	Ile	Ala	
			115				120						125				
40	Pro	Leu	Glu	Gly	Thr	Pro	Ala	Tyr	Lys	Ala	Gly	Val	Lys	Ser	Gly	Asp	
		130					135					140					
	Ser	Ile	Leu	Lys	Ile	Asn	Asn	Glu	Ser	Thr	Leu	Ser	Met	Ser	Ile	Asp	
	145					150					155				160		
	Asp	Ala	Val	Asn	Leu	Met	Arg	Gly	Lys	Pro	Lys	Thr	Ser	Ile	Gln	Ile	
					165					170					175		
45	Thr	Val	Val	Arg	Lys	Asn	Glu	Pro	Lys	Pro	Leu	Val	Phe	Asn	Ile	Val	
				180					185					190			
	Arg	Asp	Ile	Ile	Lys	Ile	Pro	Ser	Val	Tyr	Val	Lys	Lys	Ile	Lys	Asp	
			195				200						205				
50	Thr	Pro	Tyr	Leu	Tyr	Val	Arg	Val	Asn	Ser	Phe	Asp	Lys	Asn	Val	Thr	
		210					215					220					
	Lys	Ser	Val	Leu	Asp	Gly	Leu	Lys	Ala	Asn	Pro	Asn	Ile	Lys	Gly	Val	
	225					230					235				240		
	Val	Leu	Asp	Leu	Arg	Gly	Asn	Pro	Gly	Gly	Leu	Leu	Asn	Gln	Ala	Val	
					245					250					255		

- 235 -

Gly Leu Ser Asn Leu Phe Ile Lys Glu Gly Val Leu Val Ser Gln Arg
 260 265 270
 Gly Lys Asn Lys Glu Glu Asn Leu Glu Tyr Lys Ala Asn Gly Arg Ala
 275 280 285
 5 Pro Tyr Thr Asn Leu Pro Val Val Val Leu Val Asn Gly Gly Ser Ala
 290 295 300
 Ser Ala Ser Glu Ile Val Ala Gly Ala Leu Gln Asp His Lys Arg Ala
 305 310 315 320
 10 Ile Ile Ile Gly Glu Lys Thr Phe Gly Lys Gly Ser Val Gln Val Leu
 325 330 335
 Leu Pro Val Asn Lys Asp Glu Ala Ile Lys Ile Thr Thr Ala Arg Tyr
 340 345 350
 Tyr Leu Pro Ser Gly Arg Thr Ile Gln Ala Lys Gly Ile Thr Pro Asp
 355 360 365
 15 Ile Val Ile Tyr Pro Gly Lys Val Pro Glu Asn Glu Asn Lys Phe Ser
 370 375 380
 Leu Lys Glu Ala Asp Leu Lys His His Leu Glu Gln Glu Leu Lys Lys
 385 390 395 400
 20 Leu Asp Asp Lys Thr Pro Ile Ser Lys Glu Ala Asp Lys Asp Lys Lys
 405 410 415
 Ser Glu Glu Glu Lys Glu Val Thr Pro Lys Met Ile Asn Asp Asp Ile
 420 425 430
 Gln Leu Lys Thr Ala Ile Asp Ser Leu Lys Thr Trp Ser Ile Val Asp
 435 440 445
 25 Glu Lys Met Asp Glu Lys Val Pro Lys Lys Lys
 450 455

(2) INFORMATION FOR SEQ ID NO:151:

- 30 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 104 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 35 (ii) MOLECULE TYPE: protein
 (iii) HYPOTHETICAL: YES
 (vi) ORIGINAL SOURCE:
 40 (A) ORGANISM: *Helicobacter pylori*
 (ix) FEATURE:
 (A) NAME/KEY: misc_feature
 (B) LOCATION 1...104
 45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:151:

Leu Leu Leu His Pro Leu His Ala His Ala Gln Val Leu Gly Phe Thr
 1 5 10 15
 50 Asn His Asp His Ala Pro Trp Leu Tyr Asp Phe Ile Lys Ser Phe Cys
 20 25 30
 Asn Leu Ser Gly Gln Pro Phe Leu Asp Leu Gln Ala Phe Ala Ile Asn
 35 40 45
 Phe Asn Glu Phe Ser Asp Arg Ala Asn Ala Tyr Asn Leu Phe Leu Arg

- 236 -

[illegible]

(2) INFORMATION FOR SEQ ID NO:152:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 165 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc feature

(B) LOCATION 1...165

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:152:

[illegible]

(2) INFORMATION FOR SEQ ID NO:153:

(i) SEQUENCE CHARACTERISTICS:

- 237 -

(A) LENGTH: 213 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

5 (ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

10 (A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...213

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:153:

	Met	Asp	Thr	Glu	Thr	Gln	Glu	Lys	Phe	Leu	Ala	Tyr	Leu	Phe	Glu	Lys	
	1				5					10					15		
20	Ala	Leu	Gln	Lys	Asn	Leu	Gln	Ala	Tyr	Trp	Ile	Thr	Thr	Thr	Glu	Thr	
				20					25					30			
	Lys	Asn	Glu	Leu	Thr	Arg	Glu	Glu	Phe	Ser	Asn	Leu	Ile	Arg	Lys	Thr	
				35				40					45				
25	Met	Ile	Glu	Leu	Ile	Leu	His	Asn	Lys	Ser	Ile	Gln	Ile	Asp	Glu	Thr	
		50				55						60					
	Leu	Leu	Asn	Val	Lys	Glu	His	Leu	Glu	Lys	Phe	Tyr	Ser	Asn	Lys	Glu	
	65					70					75					80	
	Gln	Glu	Thr	Ile	Ala	Lys	Thr	Leu	Glu	Ser	Gln	Thr	Glu	Leu	Thr	Cys	
					85					90					95		
30	Ser	Tyr	Leu	Leu	Asp	Lys	Asp	Phe	Ser	Leu	Leu	Glu	Lys	His	Leu	Glu	
				100					105					110			
	Asn	Ser	Leu	Gly	His	Phe	Thr	Phe	Glu	Ser	Glu	Phe	Ala	Leu	Leu	Lys	
			115				120						125				
	Asp	Lys	Glu	Pro	Leu	Asn	Leu	Ala	Gln	Ile	Lys	Gln	Ile	Gly	Val	Leu	
35		130				135					140						
	Lys	Val	Ile	Thr	Tyr	Glu	Met	Thr	Gln	Ala	Leu	Lys	Asn	Gln	Ile	Ile	
	145					150				155					160		
	His	Leu	Thr	Gln	Ile	Val	Asn	Glu	Glu	Asn	Leu	Glu	Phe	Asp	Glu	Glu	
				165					170					175			
40	Leu	Val	Ile	Tyr	His	Leu	Asn	Phe	Lys	Leu	Asn	Gln	Asn	Thr	Tyr	Lys	
				180					185					190			
	Val	Leu	Ala	Lys	Phe	Cys	Val	Leu	Lys	Lys	Lys	Gly	Thr	Leu	His	Glu	
			195				200					205					
45	Lys	Phe	Lys	Ala	Phe												
		210															

(2) INFORMATION FOR SEQ ID NO:154:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 253 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

- 238 -

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

5 (A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...253

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:154:

Met Ala Ile Ser Ile Lys Ser Pro Lys Glu Ile Lys Ala Leu Arg Lys
 1 5 10 15
 15 Ala Gly Glu Leu Thr Ala Gln Ala Leu Ala Leu Leu Glu Arg Glu Val
 20 25 30
 Arg Pro Gly Val Ser Leu Leu Glu Leu Asp Lys Met Ala Glu Asp Phe
 35 40 45
 Ile Lys Ser Ser His Ala Arg Pro Ala Phe Lys Gly Leu Tyr Gly Phe
 50 55 60
 20 Pro Asn Ser Val Cys Met Ser Leu Asn Glu Val Val Ile His Gly Ile
 65 70 75 80
 Pro Thr Asp Tyr Val Leu Gln Glu Gly Asp Ile Ile Gly Leu Asp Leu
 85 90 95
 25 Gly Val Glu Val Asp Gly Tyr Tyr Gly Asp Ser Ala Leu Thr Leu Pro
 100 105 110
 Ile Gly Ala Ile Ser Pro Gln Asp Glu Lys Leu Leu Ala Cys Ser Lys
 115 120 125
 Glu Ser Leu Met His Ala Ile Ser Ser Ile Arg Val Gly Met His Phe
 130 135 140
 30 Lys Glu Leu Ser Gln Ile Leu Glu Gly Ala Ile Thr Glu Arg Gly Phe
 145 150 155 160
 Val Pro Leu Lys Gly Phe Cys Gly His Gly Ile Gly Lys Lys Pro His
 165 170 175
 35 Glu Glu Pro Glu Ile Pro Asn Tyr Leu Glu Lys Gly Val Lys Ala Asn
 180 185 190
 Ser Gly Pro Lys Ile Lys Glu Gly Met Val Phe Cys Leu Glu Pro Met
 195 200 205
 Val Cys Gln Lys Gln Gly Glu Pro Lys Ile Leu Ala Asp Lys Trp Ser
 210 215 220
 40 Val Val Ser Val Asp Gly Leu Asn Thr Ser His His Glu His Thr Ile
 225 230 235 240
 Ala Ile Val Gly Asn Lys Ala Val Ile Leu Thr Glu Arg
 245 250

45

(2) INFORMATION FOR SEQ ID NO:155:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 247 amino acids

50

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

- 239 -

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...247

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:155:

```

Lys Pro Lys Arg Asn Gln Ser Pro Lys Lys Ser Arg Glu Leu Thr Ala
1          5          10          15
Gln Ala Leu Ala Leu Leu Glu Arg Glu Val Arg Pro Gly Val Ser Leu
15          20          25          30
Leu Glu Leu Asp Lys Met Ala Glu Asp Phe Ile Lys Ser Ser His Ala
35          40          45
Arg Pro Ala Phe Lys Gly Leu Tyr Gly Phe Pro Asn Ser Val Cys Met
50          55          60
20 Ser Leu Asn Glu Val Val Ile His Gly Ile Pro Thr Asp Tyr Val Leu
65          70          75          80
Gln Glu Gly Asp Ile Ile Gly Leu Asp Leu Gly Val Glu Val Asp Gly
85          90          95
Tyr Tyr Gly Asp Ser Ala Leu Thr Leu Pro Ile Gly Ala Ile Ser Pro
25          100         105         110
Gln Asp Glu Lys Leu Leu Ala Cys Ser Lys Glu Ser Leu Met His Ala
115         120         125
Ile Ser Ser Ile Arg Val Gly Met His Phe Lys Glu Leu Ser Gln Ile
130         135         140
30 Leu Glu Gly Ala Ile Thr Glu Arg Gly Phe Val Pro Leu Lys Gly Phe
145         150         155         160
Cys Gly His Gly Ile Gly Lys Lys Pro His Glu Glu Pro Glu Ile Pro
165         170         175
Asn Tyr Leu Glu Lys Gly Val Lys Ala Asn Ser Gly Pro Lys Ile Lys
35          180         185         190
Glu Gly Met Val Phe Cys Leu Glu Pro Met Val Cys Gln Lys Gln Gly
195         200         205
Glu Pro Lys Ile Leu Ala Asp Lys Trp Ser Val Val Ser Val Asp Gly
210         215         220
40 Leu Asn Thr Ser His His Glu His Thr Ile Ala Ile Val Gly Asn Lys
225         230         235         240
Ala Val Ile Leu Thr Glu Arg
245

```

(2) INFORMATION FOR SEQ ID NO:156:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 340 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

- 240 -

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

5

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...340

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:156:

10

```

Met Tyr Arg Lys Asp Leu Asp Asn Tyr Leu Lys Gln Arg Leu Pro Lys
1      5      10      15
Ala Val Phe Leu Tyr Gly Glu Phe Asp Phe Phe Ile His Tyr Tyr Ile
      20      25      30
15 Gln Thr Ile Ser Ala Leu Phe Lys Gly Asn Asn Pro Asp Thr Glu Thr
      35      40      45
Ser Leu Phe Tyr Ala Ser Asp Tyr Glu Lys Ser Gln Ile Ala Thr Leu
      50      55      60
Leu Glu Gln Asp Ser Leu Phe Gly Gly Ser Ser Leu Val Ile Leu Lys
20 65      70      75      80
Leu Asp Phe Ala Leu His Lys Lys Phe Lys Glu Asn Asp Ile Asn Pro
      85      90      95
Phe Leu Lys Ala Leu Glu Arg Pro Ser His Asn Arg Leu Ile Ile Gly
      100     105     110
25 Leu Tyr Asn Ala Lys Ser Asp Thr Thr Lys Tyr Lys Tyr Thr Ser Glu
      115     120     125
Ile Ile Val Lys Phe Phe Gln Lys Ser Pro Leu Lys Asp Glu Ala Ile
      130     135     140
Cys Val Arg Phe Phe Thr Pro Lys Ala Trp Glu Ser Leu Lys Phe Leu
30 145     150     155     160
Gln Glu Arg Ala Asn Phe Leu His Leu Asp Ile Ser Gly His Leu Leu
      165     170     175
Asn Ala Leu Phe Glu Ile Asn Asn Glu Asp Leu Ser Val Ser Phe Asn
      180     185     190
35 Asp Leu Asp Lys Leu Ala Val Leu Asn Ala Pro Ile Thr Leu Glu Asp
      195     200     205
Ile Gln Glu Leu Ser Ser Asn Ala Gly Asp Met Asp Leu Gln Lys Leu
      210     215     220
Ile Leu Gly Leu Phe Leu Lys Lys Ser Val Leu Asp Ile Tyr Asp Tyr
40 225     230     235     240
Leu Leu Lys Glu Gly Lys Lys Asp Ala Asp Ile Leu Arg Gly Leu Glu
      245     250     255
Arg Tyr Phe Tyr Gln Leu Phe Leu Phe Phe Ala His Ile Lys Thr Thr
      260     265     270
45 Gly Leu Met Asp Ala Lys Glu Val Leu Gly Tyr Ala Pro Pro Lys Glu
      275     280     285
Ile Val Glu Asn Tyr Ala Lys Asn Ala Leu Arg Leu Lys Glu Ala Gly
      290     295     300
Tyr Lys Arg Val Phe Glu Ile Phe Arg Leu Trp His Leu Gln Ser Met
50 305     310     315     320
Gln Gly Gln Lys Glu Leu Gly Phe Leu Tyr Leu Thr Pro Ile Gln Lys
      325     330     335
Ile Ile Asn Pro
      340

```

- 241 -

(2) INFORMATION FOR SEQ ID NO:157:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 200 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
 (B) LOCATION 1...200

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:157:

```

Val Phe Met Thr Ser Ala Leu Leu Gly Leu Gln Ile Val Leu Ala Val
1           5           10           15
Leu Ile Val Val Val Val Leu Leu Gln Lys Ser Ser Ser Ile Gly Leu
25           20           25           30
Gly Ala Tyr Ser Gly Ser Asn Asp Ser Leu Phe Gly Ala Lys Gly Pro
35           40           45
Ala Ser Phe Met Ala Lys Leu Thr Met Phe Leu Gly Leu Leu Phe Val
50           55           60
Ile Asn Thr Ile Ala Leu Gly Tyr Phe Tyr Asn Lys Glu Tyr Gly Lys
30 65           70           75           80
Ser Val Leu Asp Glu Thr Lys Thr Asn Lys Glu Leu Ser Pro Leu Val
85           90           95
Pro Ala Thr Gly Thr Leu Asn Pro Thr Leu Asn Pro Thr Leu Asn Pro
35 100          105          110
Thr Leu Asn Pro Leu Glu Gln Ala Pro Thr Asn Pro Leu Met Pro Thr
115          120          125
Gln Thr Pro Lys Glu Leu Pro Lys Glu Pro Ala Lys Thr Pro Phe Val
130          135          140
Glu Ser Pro Lys Gln Asn Glu Lys Asn Glu Lys Asn Asp Ala Lys Glu
40 145          150          155          160
Asn Gly Ile Lys Gly Val Glu Lys Asn Lys Glu Asn Ala Lys Thr Pro
165          170          175
Pro Thr Thr His Gln Lys Pro Lys Thr His Ala Thr Thr Asn Ala His
45 180          185          190
Thr Asn Gln Lys Lys Asp Glu Lys
195          200

```

(2) INFORMATION FOR SEQ ID NO:158:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 159 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

- 243 -

```

Leu Asn Thr Asp Phe Ser His Ile Thr Asp Ile Glu Gly Met Arg Phe
1          5          10          15
Val Asn Glu Glu Asp Ala Leu Asn Lys Leu Ile Asn Glu Ile His Thr
5          20          25          30
Arg His Ile Asp Leu Lys Asp Ser Ile Met Leu Ala Leu Ser Phe Asn
          35          40          45
Ala Leu Tyr Leu Ala Asn Ala Leu Ala Gln Lys Phe Gly Ala Thr Tyr
          50          55          60
10 Asp Ile Leu Phe Leu Glu Pro Ile Leu Ala Pro Leu Asn Ser Lys Cys
65          70          75          80
Glu Ile Ala Leu Val Ser Glu Ser Met Asp Ile Val Met Asn Glu Ser
          85          90          95
Leu Ile Asn Ser Phe Asp Ile Ala Leu Asp Tyr Val Tyr Gly Glu Ala
15          100          105          110
Lys Arg Ala Tyr Glu Glu Asp Ile Leu Ser His Ile Tyr Gln Tyr Arg
          115          120          125
Lys Gly Asn Ala Ile Lys Ser Leu Lys Asp Lys Asn Ile Phe Ile Val
          130          135          140
20 Asp Arg Gly Ile Glu Thr Gly Phe Arg Ala Gly Leu Gly Val Gln Thr
145          150          155          160
Cys Leu Lys Lys Glu Cys Gln Asp Ile Tyr Ile Leu Thr Pro Ile Leu
          165          170          175
Ala Gln Asn Val Ala Gln Gly Leu Glu Ser Leu Cys Asp Gly Val Ile
25          180          185          190
Ser Val Tyr Arg Pro Glu Cys Phe Val Ser Val Glu His His Tyr Lys
          195          200          205
Glu Leu Lys Arg Leu Ser Asn Glu Glu Ile Glu Lys Tyr Leu Gly Ala
          210          215          220
30 Asn Asn Ala Pro Asn Leu Lys Lys Glu His
225          230

```

(2) INFORMATION FOR SEQ ID NO:160:

- 35 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 287 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
- 40 (ii) MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: YES
- (vi) ORIGINAL SOURCE:
 45 (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
 (A) NAME/KEY: misc_feature
 (B) LOCATION 1...287
- 50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:160:

```

Leu Lys Gln Ser Glu Met Ala Met Glu Phe Asn Asp Pro Arg Met Arg
1          5          10          15

```

- 244 -

Phe Phe Ile Gly Asp Val Arg Asp Leu Glu Arg Leu Asn Tyr Ala Leu
 20 25 30
 Glu Gly Val Asp Ile Cys Ile His Ala Ala Ala Leu Lys His Val Pro
 35 40 45
 5 Ile Ala Glu Tyr Asn Pro Leu Glu Cys Ile Lys Thr Asn Ile Met Gly
 50 55 60
 Ala Ser Asn Val Ile Asn Ala Cys Leu Lys Asn Glu Ile Ser Gln Val
 65 70 75 80
 10 Ile Ala Leu Ser Thr Asp Lys Ala Ala Asn Pro Ile Asn Leu Tyr Gly
 85 90 95
 Ala Thr Lys Leu Cys Ser Asp Lys Leu Phe Val Ser Ala Asn Asn Phe
 100 105 110
 Lys Gly Pro Ser Gln Thr Gln Phe Gly Val Val Arg Tyr Gly Asn Val
 115 120 125
 15 Val Gly Ser Arg Gly Ser Val Val Pro Phe Phe Lys Lys Leu Val Gln
 130 135 140
 Asn Lys Ala Ser Glu Ile Pro Ile Thr Asp Ile Arg Met Thr Arg Phe
 145 150 155 160
 20 Trp Ile Thr Leu Asp Glu Gly Val Ser Phe Val Leu Lys Ser Leu Lys
 165 170 175
 Arg Met His Gly Gly Glu Ile Phe Val Pro Lys Ile Pro Ser Met Lys
 180 185 190
 Met Ile Asp Leu Ala Lys Ala Leu Ala Pro Asn Ile Pro Thr Lys Ile
 195 200 205
 25 Ile Gly Ile Arg Pro Gly Glu Lys Leu His Glu Val Met Ile Pro Lys
 210 215 220
 Asp Glu Ser His Leu Ala Leu Glu Phe Glu Asp Phe Phe Ile Ile Gln
 225 230 235 240
 30 Pro Thr Ile Ser Phe Gln Thr Pro Lys Asp Tyr Thr Leu Thr Lys Leu
 245 250 255
 His Glu Lys Gly Gln Lys Val Ala Pro Asp Phe Glu Tyr Ser Ser His
 260 265 270
 Thr Asn Asn Gln Trp Leu Glu Pro Asp Asp Leu Leu Lys Leu Leu
 275 280 285
 35

(2) INFORMATION FOR SEQ ID NO:161:

(i) SEQUENCE CHARACTERISTICS:

40 (A) LENGTH: 201 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

45 (iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

50 (ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...201

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:161:

- 245 -

5 Met Arg Leu His Thr Ala Phe Phe Gly Ile Asn Ser Leu Leu Val Ala
 1 5 10 15
 Thr Leu Leu Ile Ser Gly Cys Ser Leu Phe Lys Lys Arg Asn Thr Asn
 20 25 30
 Ala Gln Leu Ile Pro Pro Ser Ala Asn Gly Leu Gln Ala Pro Ile Tyr
 35 40 45
 Pro Pro Thr Asn Phe Thr Pro Arg Lys Ser Ile Gln Pro Leu Pro Ser
 50 55 60
 10 Pro Arg Leu Glu Asn Asn Asp Gln Pro Ile Ile Ser Ser Asn Pro Thr
 65 70 75 80
 Asn Ala Ile Pro Asn Thr Pro Ile Leu Thr Pro Asn Asn Val Ile Glu
 85 90 95
 15 Leu Asn Ala Val Gly Met Gly Val Ala Pro Glu Ser Thr Ile Ser Pro
 100 105 110
 Ser Gln Ala Leu Ala Leu Ala Lys Arg Ala Ala Ile Val Asp Gly Tyr
 115 120 125
 Arg Gln Leu Gly Glu Lys Met Tyr Gly Ile Arg Val Asn Ala Gln Asp
 130 135 140
 20 Thr Val Lys Asp Met Val Leu Gln Asn Ser Val Ile Lys Thr Arg Val
 145 150 155 160
 Asn Ala Leu Ile Arg Asn Ala Glu Ile Thr Glu Thr Ile Tyr Lys Asp
 165 170 175
 25 Gly Leu Cys Gln Val Ser Met Glu Leu Lys Leu Asp Gly Arg Ile Trp
 180 185 190
 Tyr Arg Ile Leu Ser Gly Ser Arg Gly
 195 200

(2) INFORMATION FOR SEQ ID NO:162:

30

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 355 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

35

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

40

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

45

(A) NAME/KEY: misc_feature

(B) LOCATION 1...355

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:162:

50

Met Ser Tyr Thr Ile Asn Lys Arg Phe Ser Val Gly Val Gly Leu Arg
 1 5 10 15
 Gly Leu Tyr Ala Thr Gly Ser Phe Asn Asn Thr Val Tyr Val Pro Leu
 20 25 30
 Glu Gly Ala Ser Val Leu Ser Ala Glu Gln Ile Leu Asn Leu Pro Asn
 35 40 45

- 246 -

Asn Val Phe Ala Asp Gln Val Pro Ser Asn Met Met Thr Leu Leu Gly
 50 55 60
 Asn Ile Gly Tyr Gln Pro Ala Leu Asn Cys Gln Lys Ala Gly Gly Asp
 65 70 75 80
 5 Met Ser Asp Gln Ser Cys Gln Glu Phe Tyr Asn Gly Leu Lys Lys Ile
 85 90 95
 Met Gly Tyr Ser Gly Leu Ile Lys Ala Ser Ala Asn Leu Tyr Gly Thr
 100 105 110
 10 Thr Gln Val Val Gln Lys Ser Asn Gly Gln Gly Val Ser Gly Gly Tyr
 115 120 125
 Arg Val Gly Ser Ser Leu Arg Val Phe Asp His Gly Met Phe Ser Val
 130 135 140
 Val Tyr Asn Ser Ser Val Thr Phe Asn Met Lys Gly Gly Leu Val Ala
 145 150 155 160
 15 Ile Thr Glu Leu Gly Pro Ser Leu Gly Ser Val Leu Thr Lys Gly Ser
 165 170 175
 Leu Asn Ile Asn Val Ser Leu Pro Gln Thr Leu Ser Leu Ala Tyr Ala
 180 185 190
 20 His Gln Phe Phe Lys Asp Arg Leu Arg Val Glu Gly Val Phe Glu Arg
 195 200 205
 Thr Phe Trp Ser Gln Gly Asn Lys Phe Leu Val Thr Pro Asp Phe Ala
 210 215 220
 Asn Ala Thr Tyr Lys Gly Leu Ser Gly Thr Val Ala Ser Leu Asp Ser
 225 230 235 240
 25 Glu Thr Leu Lys Lys Met Val Gly Leu Ala Asn Phe Lys Ser Val Met
 245 250 255
 Asn Met Gly Ala Gly Trp Arg Asp Thr Asn Thr Phe Arg Leu Gly Val
 260 265 270
 30 Thr Tyr Met Gly Lys Ser Leu Arg Leu Met Gly Ala Ile Asp Tyr Asp
 275 280 285
 Gln Ala Pro Ser Pro Gln Asp Ala Ile Gly Ile Pro Asp Ser Asn Gly
 290 295 300
 Tyr Thr Val Ala Phe Gly Thr Lys Tyr Asn Phe Arg Gly Phe Asp Leu
 305 310 315 320
 35 Gly Val Ala Gly Ser Phe Thr Phe Lys Ser Asn Arg Ser Ser Leu Tyr
 325 330 335
 Gln Ser Pro Thr Ile Gly Gln Leu Arg Ile Phe Ser Ala Ser Leu Gly
 340 345 350
 40 Tyr Arg Trp
 355

(2) INFORMATION FOR SEQ ID NO:163:

45 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 587 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

50 (ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

- 247 -

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...587

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:163:

```

Met Lys Asn Phe Ser Pro Leu Tyr Cys Leu Lys Lys Leu Lys Lys Arg
1      5      10      15
10 His Leu Ile Ala Leu Ser Leu Pro Leu Leu Ser Tyr Ala Asn Gly Phe
      20      25      30
Lys Ile Gln Glu Gln Ser Leu Asn Gly Thr Ala Leu Gly Ser Ala Tyr
      35      40      45
15 Val Ala Gly Ala Arg Gly Ala Asp Ala Ser Phe Tyr Asn Pro Ala Asn
      50      55      60
Met Gly Phe Thr Asn Asp Trp Gly Glu Asn Arg Ser Glu Phe Glu Met
65      70      75      80
Thr Thr Thr Val Ile Asn Ile Pro Ala Phe Ser Phe Lys Val Pro Thr
      85      90      95
20 Thr Asn Gln Gly Leu Tyr Ser Val Thr Ser Leu Glu Ile Asp Lys Ser
      100      105      110
Gln Gln Asn Ile Leu Gly Ile Ile Asn Thr Ile Gly Leu Gly Asn Ile
      115      120      125
Leu Lys Ala Leu Gly Asn Thr Ala Ala Thr Asn Gly Leu Ser Gln Ala
25      130      135      140
Ile Asn Arg Val Gln Gly Leu Met Asn Leu Thr Asn Gln Lys Val Val
145      150      155      160
Thr Leu Ala Ser Lys Pro Asp Thr Gln Ile Val Asn Gly Trp Thr Gly
      165      170      175
30 Thr Thr Asn Phe Val Leu Pro Lys Phe Phe Tyr Lys Thr Arg Thr His
      180      185      190
Asn Gly Phe Thr Phe Gly Gly Ser Phe Thr Ala Pro Ser Gly Leu Gly
      195      200      205
Met Lys Trp Asn Gly Lys Gly Gly Glu Phe Leu His Asp Val Phe Ile
35      210      215      220
Met Met Val Glu Leu Ala Pro Ser Met Ser Tyr Thr Ile Asn Lys Arg
225      230      235      240
Phe Ser Val Gly Val Gly Leu Arg Gly Leu Tyr Ala Thr Gly Ser Phe
      245      250      255
40 Asn Asn Thr Val Tyr Val Pro Leu Glu Gly Ala Ser Val Leu Ser Ala
      260      265      270
Glu Gln Ile Leu Asn Leu Pro Asn Asn Val Phe Ala Asp Gln Val Pro
      275      280      285
Ser Asn Met Met Thr Leu Leu Gly Asn Ile Gly Tyr Gln Pro Ala Leu
45      290      295      300
Asn Cys Gln Lys Ala Gly Gly Asp Met Ser Asp Gln Ser Cys Gln Glu
305      310      315      320
Phe Tyr Asn Gly Leu Lys Lys Ile Met Gly Tyr Ser Gly Leu Ile Lys
      325      330      335
50 Ala Ser Ala Asn Leu Tyr Gly Thr Thr Gln Val Val Gln Lys Ser Asn
      340      345      350
Gly Gln Gly Val Ser Gly Gly Tyr Arg Val Gly Ser Ser Leu Arg Val
      355      360      365
Phe Asp His Gly Met Phe Ser Val Val Tyr Asn Ser Ser Val Thr Phe

```

- 248 -

```

      370              375              380
Asn Met Lys Gly Gly Leu Val Ala Ile Thr Glu Leu Gly Pro Ser Leu
385              390              395              400
Gly Ser Val Leu Thr Lys Gly Ser Leu Asn Ile Asn Val Ser Leu Pro
5              405              410              415
Gln Thr Leu Ser Leu Ala Tyr Ala His Gln Phe Phe Lys Asp Arg Leu
      420              425              430
Arg Val Glu Gly Val Phe Glu Arg Thr Phe Trp Ser Gln Gly Asn Lys
      435              440              445
10 Phe Leu Val Thr Pro Asp Phe Ala Asn Ala Thr Tyr Lys Gly Leu Ser
      450              455              460
Gly Thr Val Ala Ser Leu Asp Ser Glu Thr Leu Lys Lys Met Val Gly
465              470              475              480
Leu Ala Asn Phe Lys Ser Val Met Asn Met Gly Ala Gly Trp Arg Asp
15              485              490              495
Thr Asn Thr Phe Arg Leu Gly Val Thr Tyr Met Gly Lys Ser Leu Arg
      500              505              510
Leu Met Gly Ala Ile Asp Tyr Asp Gln Ala Pro Ser Pro Gln Asp Ala
      515              520              525
20 Ile Gly Ile Pro Asp Ser Asn Gly Tyr Thr Val Ala Phe Gly Thr Lys
      530              535              540
Tyr Asn Phe Arg Gly Phe Asp Leu Gly Val Ala Gly Ser Phe Thr Phe
545              550              555              560
Lys Ser Asn Arg Ser Ser Leu Tyr Gln Ser Pro Thr Ile Gly Gln Leu
25              565              570              575
Arg Ile Phe Ser Ala Ser Leu Gly Tyr Arg Trp
      580              585

```

(2) INFORMATION FOR SEQ ID NO:164:

30

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 205 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

35

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

40

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc_feature

45

(B) LOCATION 1...205

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:164:

50

```

Leu Ile Phe Arg Phe Phe Leu Ile Leu Ser Leu Leu Lys Gly Val Leu
1              5              10              15
Leu Ala Lys Lys Asp Trp Asn Phe Phe Lys Pro Leu Glu Pro Thr Lys
      20              25              30
Lys Tyr Phe Gly Ser Phe Lys Ile Gly Tyr Leu Tyr Gln His Ala Glu
      35              40              45

```

- 249 -

Thr Thr Lys Arg Phe Pro Ile Arg Pro Lys Asn Arg Pro Pro Ile Leu
 50 55 60
 Met Asp Lys Ile Tyr His Asp Ala Ser Leu Gly Phe Asp Ala Gly Tyr
 65 70 75 80
 5 Val Leu Lys Lys Lys Ala Leu Leu Gly Gly Tyr Leu Asp Ala Gly Met
 85 90 95
 Gly Asp Ser Tyr Phe Met Ser Ala Gly Leu Val Ala Gly Val Arg Leu
 100 105 110
 10 Phe Lys Gly Trp Val Ile Pro Lys Ile Ala Leu Gly Tyr Gln Leu Gln
 115 120 125
 Ile Leu Gly Ala Lys Ile Asp Lys Tyr Gln Phe Asn Ile Gln Ser Ala
 130 135 140
 Val Gly Ser Val Gly Leu Phe Phe Asn Ala Ala Lys Asn Phe Gly Leu
 145 150 155 160
 15 Ser Ile Glu Ala Arg Gly Gly Ile Pro Phe Tyr Phe Ile Gln Ser Arg
 165 170 175
 Phe Ser Lys Ala Phe Gly Thr Pro Arg Leu Asn Ile Tyr Ser Val Gly
 180 185 190
 20 Ile Thr Phe Thr Phe Tyr Asp Phe Thr Arg Phe Leu Gly
 195 200 205

(2) INFORMATION FOR SEQ ID NO:165:

- 25 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 253 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 30 (ii) MOLECULE TYPE: protein
 (iii) HYPOTHETICAL: YES
 (vi) ORIGINAL SOURCE:
 35 (A) ORGANISM: *Helicobacter pylori*
 (ix) FEATURE:
 (A) NAME/KEY: misc_feature
 (B) LOCATION 1...253
 40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:165:

Leu Trp His Ala Ala Phe Ser Val Gly Glu Trp Gly Trp Asn Gly Asp
 1 5 10 15
 45 Glu Ile Pro Tyr Arg Asp Cys Asp Glu Trp Gly Leu Asp Asp Phe Tyr
 20 25 30
 Gly Val Lys Pro Thr Asn Lys Ala Gly Val Leu Ser Phe Ala Arg Ser
 35 40 45
 His Arg Arg Gln Asn Gln Ala Val Leu Ser Lys Pro Lys Ser Phe Arg
 50 55 60
 50 Met Lys Lys Ile Ala Phe Ile Leu Ala Leu Trp Val Gly Leu Leu Gly
 65 70 75 80
 Ala Phe Glu Pro Lys Lys Ser His Ile Tyr Phe Gly Ala Met Val Gly
 85 90 95
 Leu Ala Pro Val Lys Ile Thr Pro Lys Pro Ala Ser Asp Ser Ser Tyr

- 250 -

100 105 110
 Thr Ala Phe Leu Trp Gly Ala Lys Gly Gly Tyr Gln Phe Ala Phe Phe
 115 120 125
 5 Lys Ala Leu Ala Leu Arg Gly Glu Phe Ser Tyr Leu Met Ala Ile Lys
 130 135 140
 Pro Thr Ala Leu His Thr Ile Asn Thr Ser Leu Ser Leu Asn Met
 145 150 155 160
 Asp Val Leu Ser Asp Phe Tyr Thr Tyr Lys Tyr Ser Phe Gly Val
 165 170 175
 10 Tyr Gly Gly Leu Gly Ile Gly Tyr Phe Tyr Gln Ser Asn His Leu Gly
 180 185 190
 Met Lys Asn Ser Ser Phe Met Gly Tyr Asn Gly Leu Phe Asn Val Gly
 195 200 205
 Leu Gly Ser Thr Ile Asp Arg His His Arg Val Glu Leu Gly Ala Lys
 15 210 215 220
 Ile Pro Phe Ser Lys Thr Arg Asn Ser Phe Lys Asn Ser Tyr Phe Leu
 225 230 235 240
 Glu Ser Val Phe Ile His Ala Ala Tyr Ser Tyr Met Phe
 245 250
 20

(2) INFORMATION FOR SEQ ID NO:166:

(i) SEQUENCE CHARACTERISTICS:

25 (A) LENGTH: 412 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

30 (iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

35 (ix) FEATURE:

(A) NAME/KEY: misc_feature
 (B) LOCATION 1...412

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:166:

40 Met Glu Ser Val Lys Thr Val Lys Thr Asn Lys Val Gly Lys Asn Thr
 1 5 10 15
 Glu Thr Ala Asn Thr Glu Ala Ser Lys Glu Thr His Phe Lys Gln Ala
 20 25 30
 45 Ser Ala Ile Thr Asn Thr Leu Arg Ser Ile Gly Gly Ile Phe Thr Lys
 35 40 45
 Ile Ala Lys Lys Val Arg Glu Leu Val Lys Lys His Pro Lys Lys Ser
 50 55 60
 Ser Val Ala Leu Val Val Leu Thr His Ile Ala Cys Lys Arg Ala Lys
 50 65 70 75 80
 Glu Leu Asp Asp Lys Val Gln Asp Lys Ser Lys Gln Ala Glu Lys Glu
 85 90 95
 Asn Gln Ile Asn Trp Trp Lys Tyr Ser Gly Leu Thr Ile Ala Ala Ser
 100 105 110

- 251 -

Leu Leu Leu Ala Ala Cys Ser Thr Gly Asp Ile Asp Lys Gln Ile Glu
 115 120 125
 Leu Glu Gln Glu Lys Lys Glu Ala Asn Lys Ser Gly Ile Lys Leu Glu
 130 135 140
 5 Gln Glu Arg Gln Lys Thr Glu Gln Glu Arg Gln Lys Thr Asn Lys Ser
 145 150 155 160
 Glu Ile Glu Leu Glu Gln Glu Arg Gln Lys Thr Asn Lys Ser Gly Ile
 165 170 175
 10 Glu Leu Ala Asn Ser Gln Ile Lys Ala Glu Gln Glu Arg Gln Lys Thr
 180 185 190
 Glu Gln Glu Lys Gln Lys Ala Asn Lys Ser Glu Ile Glu Leu Glu Gln
 195 200 205
 Gln Lys Gln Lys Thr Ile Asn Thr Gln Arg Asp Leu Ile Lys Glu Gln
 210 215 220
 15 Lys Asp Phe Ile Lys Glu Thr Glu Gln Asn Cys Gln Glu Lys His Gly
 225 230 235 240
 Gln Leu Phe Ile Lys Lys Ala Arg Ile Lys Thr Gly Ile Thr Thr Gly
 245 250 255
 20 Ile Ala Ile Glu Ile Glu Ala Glu Cys Lys Thr Pro Lys Pro Ala Lys
 260 265 270
 Thr Asn Gln Thr Pro Ile Gln Pro Lys His Leu Pro Asn Ser Lys Gln
 275 280 285
 Pro Arg Ser Gln Arg Gly Ser Lys Ala Gln Glu Leu Ile Ala Tyr Leu
 290 295 300
 25 Gln Lys Glu Leu Glu Ser Leu Pro Tyr Ser Gln Lys Ala Ile Ala Lys
 305 310 315 320
 Gln Val Asp Phe Tyr Lys Pro Ser Ser Ile Ala Tyr Leu Glu Leu Asp
 325 330 335
 30 Pro Arg Asp Phe Lys Val Thr Glu Glu Trp Gln Lys Glu Asn Leu Lys
 340 345 350
 Ile Arg Ser Lys Ala Gln Ala Lys Met Leu Glu Met Arg Asn Pro Gln
 355 360 365
 Ala His Leu Pro Thr Ser Gln Ser Leu Leu Phe Val Gln Lys Ile Phe
 370 375 380
 35 Ala Asp Ile Asn Lys Glu Ile Glu Ala Val Ala Asn Thr Glu Lys Lys
 385 390 395 400
 Thr Glu Lys Ala Gly Tyr Gly Tyr Ser Lys Arg Met
 405 410

40 (2) INFORMATION FOR SEQ ID NO:167:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 149 amino acids

(B) TYPE: amino acid

45 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

50

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

- 252 -

(A) NAME/KEY: misc_feature
(B) LOCATION 1...149

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:167:

5
Leu Asn Trp Glu His Leu Met Lys Lys Leu Ala Phe Ser Leu Leu Phe
1 5 10 15
Thr Gly Thr Phe Leu Gly Leu Phe Leu Asn Ala Ser Asp Phe Lys Ser
20 25 30
10 Met Asp Asn Lys Gln Leu Leu Glu Gln Ala Gly Lys Val Ala Pro Ser
35 40 45
Glu Val Pro Glu Phe Arg Thr Glu Val Asn Lys Arg Leu Glu Ala Met
50 55 60
Lys Glu Glu Glu Arg Gln Lys Tyr Lys Ala Asp Phe Lys Lys Ala Met
15 65 70 75 80
Asp Lys Asn Leu Ala Ser Leu Ser Gln Glu Asp Arg Asn Lys Arg Lys
85 90 95
Lys Glu Ile Leu Glu Val Ile Ala Asn Lys Lys Lys Thr Met Thr Met
100 105 110
20 Lys Glu Tyr Arg Glu Glu Gly Leu Asp Leu His Asp Cys Ala Cys Glu
115 120 125
Gly Pro Phe His Asp His Glu Lys Lys Gly Gln Lys Gly Lys Lys Pro
130 135 140
Ser His His Lys His
25 145

(2) INFORMATION FOR SEQ ID NO:168:

(i) SEQUENCE CHARACTERISTICS:
30 (A) LENGTH: 204 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
35
(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori
40
(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...204

45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:168:

Met Gln Ala Val Ile Leu Ala Asn Gly Glu Phe Pro Lys Ser Lys Lys
1 5 10 15
Cys Leu Asp Ile Leu Gln Asn Ala Pro Phe Leu Ile Ala Cys Asp Gly
50 20 25 30
Ala Val Ile Ser Leu His Ala Leu Gln Phe Lys Pro Ser Val Val Ile
35 40 45
Gly Asp Leu Asp Ser Ile Asp Ser His Leu Lys Ala Leu Tyr Asn Pro
50 55 60

- 253 -

Ile Arg Val Ser Glu Gln Asp Ser Asn Asp Leu Ser Lys Ala Phe Phe
 65 70 75 80
 Tyr Ala Leu Asn Arg Gly Cys Asp Asp Phe Ile Phe Leu Gly Leu Asn
 85 90 95
 5 Gly Lys Arg Glu Asp His Ala Leu Ala Asn Thr Phe Leu Leu Leu Glu
 100 105 110
 Tyr Phe Lys Phe Cys Lys Lys Ile Gln Ser Val Ser Asp Tyr Gly Leu
 115 120 125
 Phe Arg Val Leu Glu Thr Pro Phe Thr Leu Pro Ser Phe Lys Gly Glu
 10 130 135 140
 Gln Ile Ser Leu Phe Ser Leu Asp Leu Lys Ala Arg Phe Thr Ser Lys
 145 150 155 160
 Asn Leu Lys Tyr Pro Leu Lys Asp Leu Arg Leu Lys Thr Leu Phe Ser
 165 170 175
 15 Gly Ser Leu Asn Glu Ala Thr Asn His Cys Phe Ser Leu Ser Ser Glu
 180 185 190
 Pro Lys Ser Val Val Leu Val Tyr Gln Lys Phe Ser
 195 200

20 (2) INFORMATION FOR SEQ ID NO:169:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 280 amino acids

(B) TYPE: amino acid

25 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

30

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

35 (A) NAME/KEY: misc_feature

(B) LOCATION 1...280

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:169:

40 Val Phe Asp Ser Leu Gly Gly Phe Leu Gly Tyr Lys Thr Phe Lys Pro
 1 5 10 15
 Ile Val Asp Lys Val Lys Asn Ile Asn Ala Trp Ile Lys Asn Tyr Asp
 20 25 30
 Asn Lys Lys Ala Gln Glu Ile Met Gly Phe Ile Glu Asn Pro Thr Pro
 35 40 45
 45 Asp Phe Gln Asn Lys Phe Leu Cys Val Leu Asn Arg Gln Gly Thr
 50 55 60
 Arg His Asn Asn Tyr Leu Gly Leu Thr Ser Thr Asn Leu Leu Ile Gly
 65 70 75 80
 50 Ala Ile Tyr Phe Ser Ile Arg His Cys Ile Lys Ala Thr Trp Gln Asn
 85 90 95
 Asp Arg Asp Gln Phe Tyr Ala Pro Tyr Asp Asp Ala Phe Gln Asp Asp
 100 105 110
 Ser Glu Phe Lys Asn Asn Cys Leu Ala Phe Met Leu Phe His Thr Gln

- 254 -

```

      115      120      125
Asn Arg Ile Thr Ala Thr Gln Gly Thr Asn His Phe Ile Pro Phe Ser
  130      135      140
5  Glu Asp Glu Val Asp Ser Lys Glu Arg Tyr Leu Ser His Ala Leu Leu
  145      150      155      160
    Asp Phe Leu Lys Gly Glu Ile Lys Glu Pro Lys Lys Ser Asp Ser Leu
      165      170      175
    Phe Leu Asn Ala Lys Lys Glu Asn Lys Pro Leu Lys Phe Ser Ser Ser
      180      185      190
10  Ala Ser Lys Val Phe Asp Ala Gly Arg Glu Ile Tyr Arg Tyr Tyr His
      195      200      205
    Thr Gln Asp Phe Ile His Thr Pro Tyr Asn Ala Asn Ala Ser Leu Tyr
      210      215      220
    Asp Ile Lys Glu Phe Phe Gln Gly Arg Asn Lys Gln Gly Arg Leu Asn
15  225      230      235      240
    Ser Pro Thr Lys Ala Lys Asp Glu Tyr Tyr Lys Gln Leu Tyr Ala Asn
      245      250      255
    Leu Gln Tyr Ala Leu Lys Asp Leu Ala Lys Glu Ile Gln Pro Lys Val
      260      265      270
20  Tyr Glu Tyr Gly Phe Leu Arg Glu
      275      280

```

(2) INFORMATION FOR SEQ ID NO:170:

```

25  (i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 309 amino acids
      (B) TYPE: amino acid
      (D) TOPOLOGY: linear

```

```

30  (ii) MOLECULE TYPE: protein

```

```

      (iii) HYPOTHETICAL: YES

```

```

      (vi) ORIGINAL SOURCE:
35  (A) ORGANISM: Helicobacter pylori

```

```

      (ix) FEATURE:
      (A) NAME/KEY: misc_feature
      (B) LOCATION 1...309

```

```

40  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:170:

```

```

Cys Asp Arg Ala Ile Pro His Trp Leu Phe Ser Leu Gly Tyr Arg Tyr
1   5   10   15
45  Pro Pro Pro Leu Lys Pro Thr Asn Ala Phe Asn Leu Glu Val Phe Asp
      20   25   30
    Ser Leu Gly Gly Phe Leu Gly Tyr Lys Thr Phe Lys Pro Ile Val Asp
      35   40   45
    Lys Val Lys Asn Ile Asn Ala Trp Ile Lys Asn Tyr Asp Asn Lys Lys
50  50   55   60
    Ala Gln Glu Ile Met Gly Phe Ile Glu Asn Pro Thr Pro Asp Phe Gln
      65   70   75   80
    Asn Asn Lys Phe Leu Cys Val Leu Asn Arg Gln Gly Thr Arg His Asn
      85   90   95

```

- 255 -

```

Asn Tyr Leu Gly Leu Thr Ser Thr Asn Leu Leu Ile Gly Ala Ile Tyr
      100      105      110
Phe Ser Ile Arg His Cys Ile Lys Ala Thr Trp Gln Asn Asp Arg Asp
      115      120      125
5  Gln Phe Tyr Ala Pro Tyr Asp Asp Ala Phe Gln Asp Asp Ser Glu Phe
      130      135      140
Lys Asn Asn Cys Leu Ala Phe Met Leu Phe His Thr Gln Asn Arg Ile
      145      150      155      160
10 Thr Ala Thr Gln Gly Thr Asn His Phe Ile Pro Phe Ser Glu Asp Glu
      165      170      175
Val Asp Ser Lys Glu Arg Tyr Leu Ser His Ala Leu Leu Asp Phe Leu
      180      185      190
Lys Gly Glu Ile Lys Glu Pro Lys Lys Ser Asp Ser Leu Phe Leu Asn
      195      200      205
15 Ala Lys Lys Glu Asn Lys Pro Leu Lys Phe Ser Ser Ser Ala Ser Lys
      210      215      220
Val Phe Asp Ala Gly Arg Glu Ile Tyr Arg Tyr Tyr His Thr Gln Asp
      225      230      235      240
20 Phe Ile His Thr Pro Tyr Asn Ala Asn Ala Ser Leu Tyr Asp Ile Lys
      245      250      255
Glu Phe Phe Gln Gly Arg Asn Lys Gln Gly Arg Leu Asn Ser Pro Thr
      260      265      270
Lys Ala Lys Asp Glu Tyr Tyr Lys Gln Leu Tyr Ala Asn Leu Gln Tyr
      275      280      285
25 Ala Leu Lys Asp Leu Ala Lys Glu Ile Gln Pro Lys Val Tyr Glu Tyr
      290      295      300
Gly Phe Leu Arg Glu
305

```

30 (2) INFORMATION FOR SEQ ID NO:171:

```

      (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 187 amino acids
            (B) TYPE: amino acid
35      (D) TOPOLOGY: linear

      (ii) MOLECULE TYPE: protein

      (iii) HYPOTHETICAL: YES

40      (vi) ORIGINAL SOURCE:
            (A) ORGANISM: Helicobacter pylori

      (ix) FEATURE:
45      (A) NAME/KEY: misc_feature
            (B) LOCATION 1...187

```

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:171:

```

50 Leu Glu Thr Tyr Ile Ile Asp Ala Asp Asn Ile Asp Gly Asp Leu Phe
    1          5          10          15
Phe Tyr Asn Leu Thr Arg Asn Ser Asn Asp Phe Ser Met Leu Pro Val
    20          25          30
Phe Glu Leu Asp Arg Ile Ala Gln Lys Ile Arg Asn Ile Leu Lys Lys

```

- 256 -

```

          35          40          45
His Gly Ser Arg Lys Asp Ile Ile Leu Lys His Asn Glu Ile Lys Glu
  50          55          60
Ala Phe Phe Ser Pro Phe Lys Pro Gln Leu Lys Thr Val Gln Val Phe
5 65          70          75          80
Leu Ser His Ser His Ala Asp Lys Asn Lys Ala Leu Gly Val Lys Asp
          85          90          95
Tyr Leu Glu Ser Lys Thr Lys Arg Lys Val Phe Ile Asp Ser Leu Phe
          100          105          110
10 Trp Asp Tyr Lys Asp Asp Val Leu Asn Lys Leu Ala Lys His Asp Asp
          115          120          125
Ile Ser Lys Ile Glu Asp Ala Phe Thr Leu Ile Leu Arg Lys Ser Leu
          130          135          140
15 Gln Asp Met Ile Glu Lys Cys Pro Tyr Phe Val Phe Leu Gln Ser Lys
145          150          155          160
Asn Ser Val Ser Asn Gln Gly Leu Ser Arg Ile Thr Tyr Ser Ala Trp
          165          170          175
Ile Tyr Glu Glu Leu Lys Ile Ala Ser Phe Tyr
          180          185
20

```

(2) INFORMATION FOR SEQ ID NO:172:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 198 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...198

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:172:

```

40
Leu Glu Thr Tyr Ile Ile Asp Ala Asp Asn Ile Asp Gly Asp Leu Phe
1      5      10      15
Phe Tyr Asn Leu Thr Arg Asn Ser Asn Asp Phe Ser Met Leu Pro Val
20      25      30
45 Phe Glu Leu Asp Arg Ile Ala Gln Lys Ile Arg Asn Ile Leu Lys Lys
35      40      45
His Gly Ser Arg Lys Asp Ile Ile Leu Lys His Asn Glu Ile Lys Glu
50      55      60
Ala Phe Phe Ser Pro Phe Lys Pro Gln Leu Lys Thr Val Gln Val Phe
50 65      70      75      80
Leu Ser His Ser His Ala Asp Lys Asn Lys Ala Leu Gly Val Lys Asp
85      90      95
Tyr Leu Glu Ser Lys Thr Lys Arg Lys Val Phe Ile Asp Ser Leu Phe
100      105      110

```

- 257 -

Trp Asp Tyr Lys Asp Asp Val Leu Asn Lys Leu Ala Lys His Asp Asp
 115 120 125
 Ile Ser Lys Ile Glu Asp Ala Phe Thr Leu Ile Leu Arg Lys Ser Leu
 130 135 140
 5 Gln Asp Met Ile Glu Lys Cys Pro Tyr Phe Val Phe Leu Gln Ser Lys
 145 150 155 160
 Asn Ser Val Ser Asn Gln Gly Leu Ser Arg Ile Thr Tyr Ser Ala Trp
 165 170 175
 10 Ile Tyr Glu Glu Leu Lys Ile Ala Ser Phe Leu Leu Ala Leu Leu Thr
 180 185 190
 Arg Val Ala Gln Phe Gln
 195

(2) INFORMATION FOR SEQ ID NO:173:

15

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 189 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

20

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

25

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc_feature

30

(B) LOCATION 1...189

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:173:

Met Met Thr Lys Asn Ala Tyr Ala Phe Val Val Ile Glu Lys Ser Ile
 1 5 10 15
 Met Val Phe Lys Cys Ala Lys Asp Lys Gly Leu Ile Pro Ile Thr Glu
 20 25 30
 Gly Phe Val Pro Leu Lys Glu Gly Phe Leu Arg Ser Phe Lys Glu Arg
 35 40 45
 40 Cys Asn Leu Asp Phe Leu Glu Asn Leu Asp Leu Leu Phe Leu Tyr Asp
 50 55 60
 Tyr Gln Phe Pro Ser Glu Val Phe Ser Leu Cys Lys Asp Leu Lys Asn
 65 70 75 80
 Ser Ile Trp Asp Arg Lys Leu Val Val Val Leu Val Glu Ala Leu Glu
 85 90 95
 45 Gly Phe Lys Gly Leu Asn Leu Ser Leu Lys Ile Glu Asp Arg His Ser
 100 105 110
 Asn Ser Leu Gly Asn Gly Val Gln Lys Leu Leu Thr Asn Ala Asp Leu
 115 120 125
 50 Gly Ser Asn His Lys Pro Ile Val Ile Asp Ser Met Lys Thr Tyr His
 130 135 140
 Gln Ser Gln Gln Glu Lys Tyr Lys Arg Glu Arg Gly Glu Thr Leu Glu
 145 150 155 160
 Val Arg Pro Thr Thr Pro Pro Ser Tyr Gly Gly Gly Ser Ile Arg Ile

- 258 -

165 170 175
Ser Gly Asp Lys Lys Pro Asp Ser Asn Glu Glu Asn Phe
180 185

5 (2) INFORMATION FOR SEQ ID NO:174:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 590 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc feature.

(B) LOCATION 1...590

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:174:

25	Met	Lys	Ala	Ile	Lys	Ile	Leu	Leu	Ile	Met	Thr	Leu	Ser	Leu	Asn	Ala
	1				5					10					15	
	Ile	Ser	Val	Asn	Arg	Ala	Leu	Phe	Asp	Leu	Lys	Asp	Ser	Gln	Leu	Lys
				20					25					30		
	Gly	Glu	Leu	Thr	Pro	Lys	Ile	Val	Asp	Phe	Gly	Gly	Tyr	Lys	Ser	Asn
30			35					40					45			
	Thr	Thr	Glu	Trp	Gly	Ala	Thr	Ala	Leu	Asn	Tyr	Ile	Asn	Ala	Ala	Asn
		50					55					60				
	Gly	Asp	Ala	Lys	Lys	Phe	Ser	Ala	Leu	Val	Glu	Lys	Met	Arg	Phe	Asn
	65					70					75					80
35	Ser	Gly	Ile	Leu	Gly	Asn	Phe	Arg	Ala	His	Ala	His	Leu	Arg	Gln	Ala
					85					90					95	
	Leu	Lys	Leu	Gln	Lys	Asn	Leu	Lys	Tyr	Cys	Leu	Lys	Ile	Ile	Ala	Arg
			100						105					110		
	Asp	Ser	Phe	Tyr	Ser	Tyr	Arg	Thr	Gly	Ile	Tyr	Ile	Pro	Leu	Gly	Ile
40			115					120					125			
	Ser	Leu	Lys	Asp	Gln	Lys	Thr	Ala	Gln	Lys	Met	Leu	Ala	Asp	Leu	Ser
		130					135					140				
	Val	Val	Gly	Ala	Tyr	Leu	Lys	Lys	Gln	Gln	Glu	Asn	Glu	Lys	Ala	Gln
	145					150					155					160
45	Ser	Pro	Tyr	Tyr	Arg	Ser	Asn	Asn	Tyr	Tyr	Asn	Ser	Tyr	Tyr	Ser	Pro
					165					170					175	
	Tyr	Tyr	Gly	Met	Tyr	Gly	Met	Tyr	Gly	Met	Gly	Met	Tyr	Gly	Met	Tyr
			180						185					190		
	Gly	Met	Gly	Met	Tyr	Asp	Phe	Tyr	Asp	Phe	Tyr	Asp	Gly	Met	Tyr	Gly
50			195					200					205			
	Phe	Tyr	Pro	Asn	Met	Phe	Phe	Met	Met	Gln	Val	Gln	Asp	Tyr	Leu	Met
		210					215					220				
	Leu	Glu	Asn	Tyr	Met	Tyr	Ala	Leu	Asp	Gln	Glu	Glu	Ile	Leu	Asp	His
	225					230					235					240

-259-

Asp Ala Ser Ile Asn Gln Leu Asp Thr Pro Thr Asp Asp Asp Arg Asp
 245 250 255
 Asp Lys Asp Asp Lys Ser Ser Gln Pro Ala Asn Leu Met Ser Phe Tyr
 260 265 270
 5 Arg Asp Pro Lys Phe Ser Lys Asp Ile Gln Thr Asn Arg Leu Asn Ser
 275 280 285
 Ala Leu Val Asn Leu Asp Asn Ser His Met Leu Lys Asp Asn Ser Leu
 290 295 300
 10 Phe His Thr Lys Ala Met Pro Thr Lys Ser Val Asp Ala Ile Thr Ser
 305 310 315 320
 Gln Ala Lys Glu Leu Asn His Leu Val Gly Gln Ile Lys Glu Met Lys
 325 330 335
 Gln Asp Gly Ala Ser Pro Asn Lys Ile Asp Ser Val Val Asn Lys Ala
 340 345 350
 15 Met Glu Val Arg Asp Lys Leu Asp Asn Asn Leu Asn Gln Leu Asp Asn
 355 360 365
 Asp Leu Lys Asp Gln Lys Gly Leu Ser Ser Glu Gln Gln Ala Gln Val
 370 375 380
 20 Asp Lys Ala Leu Asp Ser Val Gln Gln Leu Ser His Ser Ser Asp Val
 385 390 395 400
 Val Gly Asn Tyr Leu Asp Gly Ser Leu Lys Ile Asp Gly Asp Asp Arg
 405 410 415
 Asp Asp Leu Asn Asp Ala Ile Asn Asn Pro Met Gln Gln Pro Ala Gln
 420 425 430
 25 Gln Thr Pro Ile Asn Asn Met Asp Asn Thr His Ala Asn Asp Ser Lys
 435 440 445
 Asp Gln Gly Gly Asn Ala Leu Ile Asn Pro Asn Asn Ala Thr Asn Asp
 450 455 460
 30 Asp His Asn Asp Asp His Met Asp Thr Asn Thr Thr Asp Thr Ser Asn
 465 470 475 480
 Ala Asn Asp Thr Pro Thr Asp Asp Lys Asp Ala Ser Gly Asn Asn Thr
 485 490 495
 Gly Asp Met Asn Asn Thr Asp Thr Gly Asn Thr Asp Thr Gly Asn Thr
 500 505 510
 35 Asp Thr Gly Asn Thr Asp Asp Met Ser Asn Met Asn Asn Gly Asn Asp
 515 520 525
 Asp Thr Gly Asn Thr Asn Asp Met Gly Asn Ser Asn Asp Met Gly
 530 535 540
 40 Asp Asp Met Asn Asn Ala Asn Asp Met Asn Asp Asp Met Gly Asn Ser
 545 550 555 560
 Asn Asp Asp Met Gly Asp Met Gly Asp Met Asn Asp Asp Met Gly Gly
 565 570 575
 Asp Met Gly Asp Met Gly Asp Met Gly Gly Asp Met Gly Asn
 580 585 590

(2) INFORMATION FOR SEQ ID NO:175:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 195 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

- 260 -

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

5

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...195

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:175:

Leu Asn Leu Arg Leu Ala Gly Ala Ser Val Leu Thr Ala Cys Val Phe
 1 5 10 15
 Ser Gly Cys Phe Phe Leu Lys Met Phe Asp Lys Lys Leu Ser Ser Asn
 15 20 25 30
 Asp Trp His Ile Gln Lys Val Glu Met Asn His Gln Val Tyr Asp Ile
 35 40 45
 Glu Thr Met Leu Ala Asp Ser Ala Phe Arg Glu His Glu Glu Glu Gln
 50 55 60
 20 Asp Ser Ser Leu Asn Thr Ala Leu Pro Glu Asp Lys Thr Ala Ile Glu
 65 70 75 80
 Ala Lys Glu Gln Glu Lys Glu Lys Arg Lys His Trp Tyr Glu Leu
 85 90 95
 Phe Lys Lys Lys Pro Lys Pro Lys Ser Ser Met Gly Glu Phe Val Phe
 100 105 110
 25 Asp Gln Lys Glu Asn Arg Ile Tyr Gly Lys Gly Tyr Cys Asn Arg Tyr
 115 120 125
 Phe Ala Ser Tyr Thr Trp Gln Gly Asp Arg His Ile Ala Ile Glu Asp
 130 135 140
 30 Ser Gly Ile Ser Arg Lys Val Cys Arg Asp Glu His Leu Met Ala Phe
 145 150 155 160
 Glu Leu Glu Phe Met Glu Asn Phe Lys Gly Asn Phe Ala Val Thr Lys
 165 170 175
 Gly Lys Asp Thr Leu Ile Leu Asp Asn Gln Lys Met Lys Ile Tyr Leu
 35 180 185 190
 Lys Thr Pro
 195

(2) INFORMATION FOR SEQ ID NO:176:

40

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 744 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

45

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

50

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc_feature

- 261 -

(B) LOCATION 1...744

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:176:

5	Met	Leu	Lys	Leu	Ala	Ser	Lys	Thr	Ile	Cys	Leu	Ser	Leu	Ile	Ser	Ser
	1				5					10				15		
	Phe	Thr	Ala	Val	Glu	Ala	Phe	Gln	Lys	His	Gln	Lys	Asp	Gly	Phe	Phe
				20				25					30			
	Ile	Glu	Ala	Gly	Phe	Glu	Thr	Gly	Leu	Leu	Gln	Gly	Thr	Gln	Thr	Gln
10			35					40					45			
	Glu	Gln	Thr	Ile	Ala	Thr	Thr	Gln	Glu	Lys	Pro	Lys	Pro	Lys	Pro	Lys
		50					55				60					
	Pro	Lys	Pro	Ile	Thr	Pro	Gln	Ser	Thr	Tyr	Gly	Lys	Tyr	Tyr	Ile	Ser
	65				70					75					80	
15	Gln	Ser	Thr	Ile	Leu	Lys	Asn	Ala	Thr	Glu	Leu	Phe	Ala	Glu	Asp	Asn
				85					90					95		
	Ile	Thr	Asn	Leu	Thr	Phe	Tyr	Ser	Gln	Asn	Pro	Val	Tyr	Val	Thr	Ala
			100						105				110			
	Tyr	Asn	Gln	Glu	Ser	Ala	Glu	Glu	Ala	Gly	Tyr	Gly	Asn	Asn	Ser	Leu
20			115					120					125			
	Ile	Met	Ile	Gln	Asn	Phe	Leu	Pro	Tyr	Asn	Leu	Asn	Asn	Ile	Glu	Leu
		130					135				140					
	Ser	Tyr	Thr	Asp	Asp	Gln	Gly	Asn	Val	Val	Ser	Leu	Gly	Val	Ile	Glu
	145			150					155						160	
25	Thr	Ile	Pro	Lys	Gln	Ser	Gln	Ile	Ile	Leu	Pro	Ala	Ser	Leu	Phe	Asn
				165					170					175		
	Asp	Pro	Gln	Leu	Asn	Ala	Asp	Gly	Phe	Gln	Gln	Leu	Gln	Thr	Asn	Thr
			180					185					190			
	Thr	Arg	Phe	Ser	Asp	Ala	Ser	Thr	Gln	Asn	Leu	Phe	Asn	Lys	Leu	Ser
30			195					200					205			
	Lys	Val	Thr	Thr	Asn	Leu	Gln	Met	Thr	Tyr	Ile	Asn	Tyr	Asn	Gln	Phe
		210					215					220				
	Ser	Ser	Gly	Asn	Gly	Ser	Gly	Ser	Lys	Pro	Pro	Cys	Pro	Pro	Tyr	Glu
	225			230					235						240	
35	Asn	Gln	Ala	Asn	Cys	Val	Ala	Lys	Val	Pro	Pro	Phe	Thr	Ser	Gln	Asp
				245					250						255	
	Ala	Lys	Asn	Leu	Thr	Asn	Leu	Met	Leu	Asn	Met	Met	Ala	Val	Phe	Asp
			260					265					270			
	Ser	Lys	Ser	Trp	Glu	Asp	Ala	Val	Leu	Asn	Ala	Pro	Phe	Gln	Phe	Ser
40			275					280					285			
	Asp	Asn	Asn	Leu	Ser	Ala	Pro	Cys	Tyr	Ser	Asp	Tyr	Leu	Thr	Cys	Val
		290					295				300					
	Asn	Pro	Tyr	Asn	Asp	Gly	Leu	Val	Asp	Pro	Lys	Leu	Ile	Ala	Lys	Asn
	305			310					315						320	
45	Lys	Gly	Asp	Glu	Tyr	Asn	Ile	Glu	Asn	Gly	Gln	Thr	Gly	Ser	Val	Ile
				325					330					335		
	Leu	Thr	Pro	Gln	Asp	Val	Ile	Tyr	Ser	Tyr	Arg	Val	Ala	Asn	Asn	Ile
			340					345					350			
	Tyr	Val	Asn	Leu	Leu	Pro	Thr	Arg	Gly	Gly	Asp	Leu	Gly	Leu	Gly	Ser
50			355				360					365				
	Gln	Tyr	Gly	Gly	Pro	Asn	Gly	Pro	Gly	Asp	Asp	Gly	Thr	Asn	Phe	Gly
		370				375					380					
	Ala	Leu	Gly	Ile	Leu	Ser	Pro	Phe	Leu	Asp	Pro	Glu	Ile	Leu	Phe	Gly
	385				390					395					400	

45

(i) SEQUENCE CHARACTERISTICS:

- 50

(ii) MOLECULE TYPE: protein

- 263 -

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

5

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...529

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:177:

```

Met Thr Tyr Ile Asn Tyr Asn Gln Phe Ser Ser Gly Asn Gly Ser Gly
1      5      10      15
Ser Lys Pro Pro Cys Pro Pro Tyr Glu Asn Gln Ala Asn Cys Val Ala
15      20      25      30
Lys Val Pro Pro Phe Thr Ser Gln Asp Ala Lys Asn Leu Thr Asn Leu
35      40      45
Met Leu Asn Met Met Ala Val Phe Asp Ser Lys Ser Trp Glu Asp Ala
50      55      60
20 Val Leu Asn Ala Pro Phe Gln Phe Ser Asp Asn Asn Leu Ser Ala Pro
65      70      75      80
Cys Tyr Ser Asp Tyr Leu Thr Cys Val Asn Pro Tyr Asn Asp Gly Leu
85      90      95
25 Val Asp Pro Lys Leu Ile Ala Lys Asn Lys Gly Asp Glu Tyr Asn Ile
100      105      110
Glu Asn Gly Gln Thr Gly Ser Val Ile Leu Thr Pro Gln Asp Val Ile
115      120      125
Tyr Ser Tyr Arg Val Ala Asn Asn Ile Tyr Val Asn Leu Leu Pro Thr
130      135      140
30 Arg Gly Gly Asp Leu Gly Leu Gly Ser Gln Tyr Gly Gly Pro Asn Gly
145      150      155      160
Pro Gly Asp Asp Gly Thr Asn Phe Gly Ala Leu Gly Ile Leu Ser Pro
165      170      175
35 Phe Leu Asp Pro Glu Ile Leu Phe Gly Lys Glu Leu Asn Lys Val Ala
180      185      190
Ile Met Gln Leu Arg Asp Ile Ile His Glu Tyr Gly His Thr Leu Gly
195      200      205
Tyr Thr His Asn Gly Asn Met Thr Tyr Gln Arg Val Arg Met Cys Glu
210      215      220
40 Glu Asn Asn Gly Pro Glu Glu Arg Cys Gln Gly Gly Arg Ile Glu Gln
225      230      235      240
Val Asp Gly Lys Glu Val Gln Val Phe Asp Asn Gly His Glu Val Arg
245      250      255
45 Asp Thr Asp Gly Ser Thr Tyr Asp Val Cys Ser Arg Phe Lys Asp Lys
260      265      270
Pro Tyr Thr Ala Gly Ser Tyr Pro Asn Ser Ile Tyr Thr Asp Cys Ser
275      280      285
Gln Val Pro Ala Gly Leu Ile Gly Val Thr Ser Ala Val Trp Gln Gln
290      295      300
50 Leu Ile Asp Gln Asn Ala Leu Pro Val Asp Phe Thr Asn Leu Ser Ser
305      310      315      320
Gln Thr Asn Tyr Leu Asn Ala Ser Leu Asn Thr Gln Asp Phe Ala Thr
325      330      335
Thr Met Leu Ser Ala Ile Ser Gln Ser Leu Ser Ser Ser Lys Ser Ser

```

- 264 -

340 345 350
 Ala Thr Thr Tyr Arg Thr Ser Lys Thr Ser Arg Pro Phe Gly Ala Pro
 355 360 365
 5 Leu Leu Gly Val Asn Leu Lys Met Gly Tyr Gln Lys Tyr Phe Asn Asp
 370 375 380
 Tyr Leu Gly Leu Ser Ser Tyr Gly Ile Ile Lys Tyr Asn Tyr Ala Gln
 385 390 395 400
 Ala Asn Asn Glu Lys Ile Gln Gln Leu Ser Tyr Gly Val Gly Met Asp
 405 410 415
 10 Val Leu Phe Asp Phe Ile Thr Asn Tyr Thr Asn Glu Lys Asn Pro Lys
 420 425 430
 Ser Asn Leu Thr Lys Lys Val Phe Thr Ser Ser Leu Gly Val Phe Gly
 435 440 445
 15 Gly Leu Arg Gly Leu Tyr Asn Ser Tyr Tyr Leu Leu Asn Gln Tyr Lys
 450 455 460
 Gly Ser Gly Asn Leu Asn Val Thr Gly Gly Leu Asn Tyr Arg Tyr Lys
 465 470 475 480
 His Ser Lys Tyr Ser Ile Gly Ile Ser Val Pro Leu Val Gln Leu Lys
 485 490 495
 20 Ser Arg Ile Val Ser Ser Asp Gly Ala Tyr Thr Asn Ser Ile Thr Leu
 500 505 510
 Asn Glu Gly Gly Ser His Phe Lys Val Phe Phe Asn Tyr Gly Trp Ile
 515 520 525
 Phe
 25

(2) INFORMATION FOR SEQ ID NO:178:

30 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 187 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

35 (ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

40 (vi) ORIGINAL SOURCE:
 (A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:
 (A) NAME/KEY: misc_feature
 (B) LOCATION 1...187

45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:178:

50 Leu Gly Cys Val Ser Met Thr Leu Gly Ile Asp Glu Ala Gly Arg Gly
 1 5 10 15
 Cys Leu Ala Gly Ser Leu Phe Val Ala Gly Val Val Cys Asn Glu Lys
 20 25 30
 Ile Ala Leu Glu Phe Leu Lys Met Gly Leu Lys Asp Ser Lys Lys Leu
 35 40 45
 Ser Pro Lys Lys Arg Phe Phe Leu Glu Asp Lys Ile Lys Thr His Gly
 50 55 60

- 265 -

Glu Val Gly Phe Phe Val Val Lys Lys Ser Ala Asn Glu Ile Asp His
 65 70 75 80
 Leu Gly Leu Gly Ala Cys Leu Lys Leu Ala Ile Glu Glu Ile Val Glu
 85 90 95
 5 Asn Gly Cys Ser Leu Ala Asn Glu Ile Lys Ile Asp Gly Asn Thr Ala
 100 105 110
 Phe Gly Leu Asn Lys Arg Tyr Pro Asn Ile Gln Thr Ile Ile Lys Gly
 115 120 125
 10 Asp Glu Thr Ile Ala Gln Ile Ala Met Ala Ser Val Leu Ala Lys Ala
 130 135 140
 Ser Lys Asp Arg Glu Met Leu Glu Leu His Ala Leu Phe Lys Glu Tyr
 145 150 155 160
 Gly Trp Asp Lys Asn Cys Gly Tyr Gly Thr Lys Gln His Ile Glu Ala
 165 170 175
 15 Ile Asn Lys Leu Gly Ala Thr Leu Ser Ser Ala
 180 185

(2) INFORMATION FOR SEQ ID NO:179:

- 20 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 204 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 25 (ii) MOLECULE TYPE: protein
 (iii) HYPOTHETICAL: YES
 (vi) ORIGINAL SOURCE:
 30 (A) ORGANISM: *Helicobacter pylori*
 (ix) FEATURE:
 (A) NAME/KEY: misc_feature
 (B) LOCATION 1...204
 35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:179:

Met Thr Leu Gly Ile Asp Glu Ala Gly Arg Gly Cys Leu Ala Gly Ser
 1 5 10 15
 40 Leu Phe Val Ala Gly Val Val Cys Asn Glu Lys Ile Ala Leu Glu Phe
 20 25 30
 Leu Lys Met Gly Leu Lys Asp Ser Lys Lys Leu Ser Pro Lys Lys Arg
 35 40 45
 Phe Phe Leu Glu Asp Lys Ile Lys Thr His Gly Glu Val Gly Phe Phe
 45 50 55 60
 Val Val Lys Lys Ser Ala Asn Glu Ile Asp His Leu Gly Leu Gly Ala
 65 70 75 80
 Cys Leu Lys Leu Ala Ile Glu Glu Ile Val Glu Asn Gly Cys Ser Leu
 85 90 95
 50 Ala Asn Glu Ile Lys Ile Asp Gly Asn Thr Ala Phe Gly Leu Asn Lys
 100 105 110
 Arg Tyr Pro Asn Ile Gln Thr Ile Ile Lys Gly Asp Glu Thr Ile Ala
 115 120 125
 Gln Ile Ala Met Ala Ser Val Leu Ala Lys Ala Ser Lys Asp Arg Glu

- 266 -

130 135 140
 Met Leu Glu Leu His Ala Leu Phe Lys Glu Tyr Gly Trp Asp Lys Asn
 145 150 155 160
 Cys Gly Tyr Gly Thr Lys Gln His Ile Glu Ala Ile Asn Lys Leu Gly
 5 165 170 175
 Ala Thr Pro Phe His Arg His Ser Phe Thr Leu Lys Asn Arg Ile Leu
 180 185 190
 Asn Pro Lys Leu Leu Glu Val Glu Gln Arg Leu Val
 195 200

10

(2) INFORMATION FOR SEQ ID NO:180:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 192 amino acids
 15 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

20 (iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
 (A) ORGANISM: Helicobacter pylori

25 (ix) FEATURE:
 (A) NAME/KEY: misc_feature
 (B) LOCATION 1...192

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:180:

Met Asn Ala Leu Lys Lys Leu Ser Phe Cys Ala Leu Leu Ser Leu Gly
 1 5 10 15
 Leu Phe Ala Gln Thr Val His Ala Gln His Leu Lys Asp Thr Ile Asn
 20 25 30
 35 Tyr Pro Asp Trp Leu Lys Ile Asn Leu Phe Asp Lys Lys Asn Pro Pro
 35 40 45
 Asn Gln Tyr Val Gly Ser Ala Ser Ile Ser Gly Lys Arg Asn Asp Phe
 50 55 60
 Tyr Ser Asn Tyr Ile Pro Tyr Asp Asp Lys Leu Pro Pro Glu Lys Asn
 40 65 70 75 80
 Ala Glu Glu Ile Ala Leu Leu Arg Ala Arg Met Asn Ala Tyr Ser Thr
 85 90 95
 Leu Glu Ser Ala Leu Leu Thr Lys Met Cys Asn Arg Ile Val Lys Ala
 100 105 110
 45 Leu Gln Val Lys Asn Asn Val Ile Ser His Leu Phe Gly Phe Val Asp
 115 120 125
 Phe Leu Thr Ser Lys Ser Ile Leu Ala Lys Arg Phe Val Asp Thr Thr
 130 135 140
 Asn His Arg Val Tyr Val Met Val Gln Phe Pro Phe Ile Gln Pro Glu
 50 145 150 155 160
 Asp Leu Ile Ala Tyr Phe Lys Ala Lys Arg Ile Asp Leu Ser Leu Ala
 165 170 175
 Ser Ala Thr Asn Leu Ser Ala Ile Leu Asn Lys Ala Leu Phe His Leu
 180 185 190

- 267 -

(2) INFORMATION FOR SEQ ID NO:181:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 86 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
 (B) LOCATION 1...86

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:181:

```

Met Asn Ala Leu Lys Lys Leu Ser Phe Cys Ala Leu Leu Ser Leu Gly
1           5           10           15
Leu Phe Ala Gln Thr Val His Ala Gln His Leu Lys Asp Thr Ile Asn
25          20          25          30
Tyr Pro Asp Trp Leu Lys Ile Asn Leu Phe Asp Lys Lys Asn Pro Pro
35          40          45
Asn Gln Tyr Val Gly Ser Ala Ser Ile Ser Gly Lys Arg Asn Asp Phe
50          55          60
Tyr Ser Asn Tyr Ile Pro Tyr Asp Asp Lys Leu Pro Pro Glu Arg Thr
30 65          70          75          80
Leu Lys Lys Ser Leu Phe
85

```

(2) INFORMATION FOR SEQ ID NO:182:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 75 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
 (B) LOCATION 1...75

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:182:

-268-

```

Leu Lys Ile Leu Thr Leu Phe Leu Ile Gly Leu Asn Ala Leu Phe Ala
1           5           10           15
Leu Asp Leu Asn Ala Leu Lys Thr Glu Ile Lys Glu Thr Tyr Leu Lys
           20           25           30
5  Glu Tyr Lys Asp Leu Lys Leu Glu Ile Glu Thr Ile Asn Leu Glu Ile
           35           40           45
Pro Glu Arg Phe Ser His Ala Ser Ile Leu Ser Tyr Glu Leu Asn Ala
           50           55           60
10 Ser Asn Lys Leu Lys Lys Asp Gly Ser Cys Phe
           65           70           75

```

(2) INFORMATION FOR SEQ ID NO:183:

```

15  (i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 211 amino acids
      (B) TYPE: amino acid
      (D) TOPOLOGY: linear

```

```

20  (ii) MOLECULE TYPE: protein

```

```

(iii) HYPOTHETICAL: YES

```

```

25  (vi) ORIGINAL SOURCE:
      (A) ORGANISM: Helicobacter pylori

```

```

      (ix) FEATURE:
            (A) NAME/KEY: misc_feature
            (B) LOCATION 1...211

```

```

30  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:183:

```

```

Met Phe Ser Ile Ile Leu Gly Gly Gly Gly Gly Asn Thr Pro Cys Gly
1           5           10           15
Leu Thr Trp Gln His Phe Lys Leu Gly Asp Leu Phe Glu Ile Glu Lys
35           20           25           30
Thr Leu Ser Phe Asn Lys Asp Ala Leu Thr Gln Gly Gln Asp Tyr Asp
           35           40           45
Tyr Ile Thr Arg Thr Ser Gln Asn Gln Gly Val Leu Gln Thr Thr Gly
           50           55           60
40 Phe Val Asn Ala Glu Asn Leu Asn Pro Pro Phe Thr Trp Ser Leu Gly
           65           70           75           80
Leu Leu Gln Met Asp Phe Phe Tyr Arg Lys Lys Ser Trp Tyr Ala Gly
           85           90           95
Gln Phe Met Arg Lys Ile Thr Pro Lys Thr Glu Ile Lys Asn Lys Ile
45           100          105          110
Asn Ser Arg Ile Ala His Tyr Phe Thr Thr Leu Leu Asn Ala Leu Lys
           115          120          125
Arg Pro Leu Leu Ser Val Leu Val Arg Asp Ile Asp Lys Thr Phe Arg
           130          135          140
50 Glu Gln Lys Ile Gln Leu Pro Leu Lys Pro Thr Ala Lys Thr Gln Ser
           145          150          155          160
Leu Asp Gly Ile Asp Phe Asp Phe Met His Thr Leu Ile Asn Ala Leu
           165          170          175
Met Lys Gln Thr Ile Gln Gly Val Val Gln Tyr Cys Asp Ala Lys Ile

```


-269-

180 185 190
 Gln Ala Thr Lys Glu Val Ile Ser Gln Glu Thr Pro Ile Gln Lys Asp
 195 200 205
 Ser Leu Phe
 5 210

(2) INFORMATION FOR SEQ ID NO:184:

10 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 406 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

15 (ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
 20 (A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
 (A) NAME/KEY: misc_feature
 (B) LOCATION 1...406

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:184:

Val Ile Gly Pro Leu Ser Ser Gln Leu Asn Ala Ile Lys Trp Gly Glu
 1 5 10 15
 Phe Lys Leu Gly Asp Leu Phe Glu Ala Ser Asn Gly Asp Phe Asp Ile
 30 20 25 30
 Gln Lys Arg His Ile Asn His Lys Gly Glu Phe Val Ile Thr Ala Gly
 35 35 40 45
 Leu Ser Asn Asn Gly Val Leu Gly Gln Ser Asp Ile Lys Ala Lys Val
 50 55 60
 35 Phe Glu Ser His Thr Ile Thr Ile Asp Met Phe Gly Cys Ala Phe Tyr
 65 70 75 80
 Arg Ser Phe Ala Tyr Lys Met Val Thr His Ala Arg Val Phe Ser Leu
 85 90 95
 40 Lys Pro Lys Phe Glu Ile Asn His Lys Ile Gly Leu Phe Leu Ser Thr
 100 105 110
 Leu Phe Phe Gly Tyr His Lys Lys Phe Gly Tyr Glu Asn Met Cys Ser
 115 120 125
 Trp Ala Lys Ile Lys Asn Asp Lys Val Ile Leu Pro Leu Lys Pro Thr
 130 135 140
 45 Ala Asn Thr Gln Thr Leu Glu Gly Ile Asp Phe Asp Phe Met Glu Lys
 145 150 155 160
 Phe Ile Ala Glu Leu Glu Gln Cys Arg Leu Ala Glu Leu Gln Ala Tyr
 165 170 175
 50 Leu Lys Ala Thr Gly Leu Glu Asn Thr Thr Leu Ser Asn Asp Glu Glu
 180 185 190
 Asn Ala Leu Asn Val Phe Asn Asn Ser Gly Gly Gly Gly Gly Asn Thr
 195 200 205
 Pro Cys Gly Leu Thr Trp Gln His Phe Lys Leu Gly Asp Leu Phe Glu
 210 215 220

-270-

Ile Glu Lys Thr Leu Ser Phe Asn Lys Asp Ala Leu Thr Gln Gly Gln
 225 230 235 240
 Asp Tyr Asp Tyr Ile Thr Arg Thr Ser Gln Asn Gln Gly Val Leu Gln
 245 250 255
 5 Thr Thr Gly Phe Val Asn Ala Glu Asn Leu Asn Pro Pro Phe Thr Trp
 260 265 270
 Ser Leu Gly Leu Leu Gln Met Asp Phe Phe Tyr Arg Lys Lys Ser Trp
 275 280 285
 Tyr Ala Gly Gln Phe Met Arg Lys Ile Thr Pro Lys Thr Glu Ile Lys
 10 290 295 300
 Asn Lys Ile Asn Ser Arg Ile Ala His Tyr Phe Thr Thr Leu Leu Asn
 305 310 315 320
 Ala Leu Lys Arg Pro Leu Leu Ser Val Leu Val Arg Asp Ile Asp Lys
 325 330 335
 15 Thr Phe Arg Glu Gln Lys Ile Gln Leu Pro Leu Lys Pro Thr Ala Lys
 340 345 350
 Thr Gln Ser Leu Asp Gly Ile Asp Phe Asp Phe Met His Thr Leu Ile
 355 360 365
 Asn Ala Leu Met Lys Gln Thr Ile Gln Gly Val Val Gln Tyr Cys Asp
 20 370 375 380
 Ala Lys Ile Gln Ala Thr Lys Glu Val Ile Ser Gln Glu Thr Pro Ile
 385 390 395 400
 Gln Lys Asp Ser Leu Phe
 405

25

(2) INFORMATION FOR SEQ ID NO:185:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 275 amino acids
 30 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

 (ii) MOLECULE TYPE: protein

 35 (iii) HYPOTHETICAL: YES

 (vi) ORIGINAL SOURCE:
 (A) ORGANISM: *Helicobacter pylori*

 40 (ix) FEATURE:
 (A) NAME/KEY: misc_feature
 (B) LOCATION 1...275

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:185:

45

Met Ser Lys Ser Leu Tyr Gln Thr Leu Asn Val Ser Glu Asn Ala Ser
 1 5 10 15
 Gln Asp Glu Ile Lys Lys Ser Tyr Arg Arg Leu Ala Arg Gln Tyr His
 20 25 30
 50 Pro Asp Leu Asn Lys Thr Lys Glu Ala Glu Glu Lys Phe Lys Glu Ile
 35 40 45
 Asn Ala Ala Tyr Glu Ile Leu Ser Asp Glu Glu Lys Arg Arg Gln Tyr
 50 55 60
 Asp Gln Phe Gly Asp Asn Met Phe Gly Gly Gln Asn Phe Ser Asp Phe

-271-

```

65              70              75              80
Ala Arg Ser Arg Gly Pro Ser Glu Asp Leu Asp Asp Ile Leu Ser Ser
85              90              95
Ile Phe Gly Lys Gly Gly Phe Ser Gln Arg Phe Ser Gln Asn Ser Gln
5      100              105              110
Gly Phe Ser Gly Phe Asn Phe Ser Asn Phe Ala Pro Glu Asn Leu Asp
115              120              125
Val Thr Ala Ile Leu Asn Val Ser Val Leu Asp Thr Leu Leu Gly Asn
130              135              140
10     Lys Lys Gln Val Ser Val Asn Asn Glu Thr Phe Ser Leu Lys Ile Pro
145              150              155              160
Ile Gly Val Glu Glu Gly Glu Lys Ile Arg Val Arg Asn Lys Gly Lys
165              170              175
15     Met Gly Arg Thr Gly Arg Gly Asp Leu Leu Leu Gln Ile His Ile Glu
180              185              190
Glu Asp Glu Met Tyr Arg Arg Glu Lys Asp Asp Ile Ile Gln Ile Phe
195              200              205
Asp Leu Pro Leu Lys Thr Ala Leu Phe Gly Gly Lys Ile Glu Ile Ala
210              215              220
20     Thr Trp His Lys Thr Leu Thr Leu Thr Ile Pro Pro Asn Thr Lys Ala
225              230              235              240
Met Gln Lys Phe Arg Ile Lys Asp Lys Gly Ile Lys Ser Arg Lys Thr
245              250              255
25     Ser His Val Gly Asp Cys Ile Ala Ser Ser Phe Asp Leu Leu Lys Leu
260              265              270
Lys Arg Phe
275

```

(2) INFORMATION FOR SEQ ID NO:186:

30

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 278 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

35

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

40

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

45

- (A) NAME/KEY: misc_feature
- (B) LOCATION 1...778

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:186:

```

50     Met Ser Lys Ser Leu Tyr Gln Thr Leu Asn Val Ser Glu Asn Ala Ser
1         5         10         15
Gln Asp Glu Ile Lys Lys Ser Tyr Arg Arg Leu Ala Arg Gln Tyr His
20         25         30
Pro Asp Leu Asn Lys Thr Lys Glu Ala Glu Glu Lys Phe Lys Glu Ile
35         40         45

```

-272-

```

    Asn Ala Ala Tyr Glu Ile Leu Ser Asp Glu Glu Lys Arg Arg Gln Tyr
      50          55          60
    Asp Gln Phe Gly Asp Asn Met Phe Gly Gly Gln Asn Phe Ser Asp Phe
      65          70          75          80
5   Ala Arg Ser Arg Gly Pro Ser Glu Asp Leu Asp Asp Ile Leu Ser Ser
      85          90          95
    Ile Phe Gly Lys Gly Gly Phe Ser Gln Arg Phe Ser Gln Asn Ser Gln
      100          105          110
    Gly Phe Ser Gly Phe Asn Phe Ser Asn Phe Ala Pro Glu Asn Leu Asp
      115          120          125
10  Val Thr Ala Ile Leu Asn Val Ser Val Leu Asp Thr Leu Leu Gly Asn
      130          135          140
    Lys Lys Gln Val Ser Val Asn Asn Glu Thr Phe Ser Leu Lys Ile Pro
      145          150          155          160
15  Ile Gly Val Glu Glu Gly Glu Lys Ile Arg Val Arg Asn Lys Gly Lys
      165          170          175
    Met Gly Arg Thr Gly Arg Gly Asp Leu Leu Leu Gln Ile His Ile Glu
      180          185          190
    Glu Asp Glu Met Tyr Arg Arg Glu Lys Asp Asp Ile Ile Gln Ile Phe
      195          200          205
20  Asp Leu Pro Leu Lys Thr Ala Leu Phe Gly Gly Lys Ile Glu Ile Ala
      210          215          220
    Thr Trp His Lys Thr Leu Thr Leu Thr Ile Pro Pro Asn Thr Lys Ala
      225          230          235          240
25  Met Gln Lys Phe Arg Ile Lys Asp Lys Gly Ile Lys Ser Arg Lys Thr
      245          250          255
    Ser His Val Gly Asp Cys Ile Ala Ser Ser Phe Asp Leu Pro Lys Ile
      260          265          270
    Glu Thr Leu Leu Met Ser
      275
30

```

(2) INFORMATION FOR SEQ ID NO:187:

- 35 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 232 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
- 40 (ii) MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: YES
- (vi) ORIGINAL SOURCE:
 (A) ORGANISM: *Helicobacter pylori*
- 45 (ix) FEATURE:
 (A) NAME/KEY: misc_feature
 (B) LOCATION 1...232
- 50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:187:

```

Val Val Gln Lys Phe Asn Phe Tyr Lys Thr Gly Gly Met Arg Leu Lys
1          5          10          15
His Phe Lys Thr Phe Leu Phe Ile Thr Met Ala Val Ile Val Ile Gly

```

-273-

```

                20                25                30
    Thr Gly Cys Ala Asn Lys Lys Lys Lys Lys Asp Glu Tyr Asn Lys Pro
      35                40                45
5   Ala Ile Phe Trp Tyr Gln Gly Ile Leu Arg Glu Ile Leu Phe Ala Asn
      50                55                60
    Leu Glu Thr Ala Asp Asn Tyr Tyr Ser Ser Leu Gln Ser Glu His Ile
    65                70                75                80
    Asn Ser Pro Leu Val Pro Glu Ala Met Leu Ala Leu Gly Gln Ala His
      85                90                95
10  Met Lys Lys Lys Glu Tyr Val Leu Ala Ser Phe Tyr Phe Asp Glu Tyr
      100               105               110
    Ile Lys Arg Phe Gly Thr Lys Asp Asn Val Asp Tyr Leu Thr Phe Leu
      115               120               125
    Lys Leu Gln Ser His Tyr Tyr Ala Phe Lys Asn His Ser Lys Asp Gln
15  130               135               140
    Glu Phe Ile Ser Asn Ser Ile Val Ser Leu Gly Glu Phe Ile Glu Lys
    145               150               155               160
    Tyr Pro Asn Ser Arg Tyr Arg Pro Tyr Val Glu Tyr Met Gln Ile Lys
      165               170               175
20  Phe Ile Leu Gly Gln Asn Glu Leu Asn Arg Ala Ile Ala Asn Val Tyr
      180               185               190
    Lys Lys Arg His Lys Pro Glu Gly Val Lys Arg Tyr Leu Glu Arg Ile
      195               200               205
    Asp Glu Thr Leu Glu Lys Glu Thr Lys Pro Lys Pro Ser His Met Pro
25  210               215               220
    Trp Tyr Val Leu Ile Phe Asp Trp
    225                230

```

(2) INFORMATION FOR SEQ ID NO:188:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 114 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...114

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:188:

```

50  Met Arg Phe Leu Asn Asn Lys His Arg Glu Lys Gly Leu Lys Ala Glu
    1                5                10                15
    Glu Glu Ala Cys Gly Phe Leu Lys Thr Leu Gly Phe Glu Met Ile Glu
      20                25                30
    Arg Asn Phe Phe Ser Gln Phe Gly Glu Ile Asp Ile Ile Ala Leu Lys
      35                40                45

```

-274-

Lys Gly Val Leu His Phe Ile Glu Val Lys Ser Gly Glu Asn Phe Asp
 50 55 60
 Pro Ile Tyr Ala Ile Thr Pro Ser Lys Leu Lys Lys Met Ile Lys Thr
 65 70 75 80
 5 Ile Arg Cys Tyr Leu Ser Gln Lys Asp Pro Asn Ser Asp Phe Cys Ile
 85 90 95
 Asp Ala Leu Ile Val Lys Asn Gly Lys Phe Glu Leu Leu Glu Asn Ile
 100 105 110
 Thr Phe

10

(2) INFORMATION FOR SEQ ID NO:189:

(i) SEQUENCE CHARACTERISTICS:
 15 (A) LENGTH: 101 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

20

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:
 (A) ORGANISM: Helicobacter pylori

25

(ix) FEATURE:
 (A) NAME/KEY: misc_feature
 (B) LOCATION 1...101

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:189:

Met Gly Ser Ile Gly Ala Met Thr Lys Gly Ser Ser Asp Arg Tyr Phe
 1 5 10 15
 Gln Glu Gly Val Ala Ser Glu Lys Leu Val Pro Glu Gly Ile Glu Gly
 35 20 25 30
 Arg Val Pro Tyr Arg Gly Lys Val Ser Asp Met Ile Phe Gln Leu Val
 35 40 45
 Gly Gly Val Arg Ser Ser Met Gly Tyr Gln Gly Ala Lys Asn Ile Leu
 50 55 60
 40 Glu Leu Tyr Gln Asn Ala Glu Phe Val Glu Ile Thr Ser Ala Gly Leu
 65 70 75 80
 Lys Lys Ser His Val His Gly Val Asp Ile Thr Lys Glu Ala Pro Asn
 85 90 95
 Ile Met Gly Glu Phe
 45 100

(2) INFORMATION FOR SEQ ID NO:190:

(i) SEQUENCE CHARACTERISTICS:
 50 (A) LENGTH: 481 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

- 275 -

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

5 (A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...481

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:190:

Met Arg Ile Leu Gln Arg Ala Leu Thr Phe Glu Asp Val Leu Met Val
 1 5 10 15
 15 Pro Arg Lys Ser Ser Val Leu Pro Lys Asp Val Ser Leu Lys Ser Arg
 20 25 30
 Leu Thr Lys Asn Ile Gly Leu Asn Ile Pro Phe Ile Ser Ala Ala Met
 35 40 45
 20 Asp Thr Val Thr Glu His Lys Thr Ala Ile Ala Met Ala Arg Leu Gly
 50 55 60
 Gly Ile Gly Ile Val His Lys Asn Met Asp Ile Gln Thr Gln Val Lys
 65 70 75 80
 Glu Ile Thr Lys Val Lys Lys Ser Glu Ser Gly Val Ile Asn Asp Pro
 85 90 95
 25 Ile Phe Ile His Ala His Arg Thr Leu Ala Asp Ala Lys Val Ile Thr
 100 105 110
 Asp Asn Tyr Lys Ile Ser Gly Val Pro Val Val Asp Asp Lys Gly Leu
 115 120 125
 30 Leu Ile Gly Ile Leu Thr Asn Arg Asp Val Arg Phe Glu Thr Asp Leu
 130 135 140
 Ser Lys Lys Val Gly Asp Val Met Thr Lys Met Pro Leu Val Thr Ala
 145 150 155 160
 His Val Gly Ile Ser Leu Asp Glu Ala Ser Asp Leu Met His Lys His
 165 170 175
 35 Lys Ile Glu Lys Leu Pro Ile Val Asp Lys Asp Asn Val Leu Lys Gly
 180 185 190
 Leu Ile Thr Ile Lys Asp Ile Gln Lys Arg Ile Glu Tyr Pro Glu Ala
 195 200 205
 40 Asn Lys Asp Asp Phe Gly Arg Leu Arg Val Gly Ala Ala Ile Gly Val
 210 215 220
 Gly Gln Leu Asp Arg Ala Glu Met Leu Val Lys Ala Gly Val Asp Ala
 225 230 235 240
 Leu Val Leu Asp Ser Ala His Gly His Ser Ala Asn Ile Leu His Thr
 245 250 255
 45 Leu Glu Glu Ile Lys Lys Ser Leu Val Val Asp Val Ile Val Gly Asn
 260 265 270
 Val Val Thr Lys Glu Ala Thr Ser Asp Leu Ile Ser Ala Gly Ala Asp
 275 280 285
 50 Ala Val Lys Val Gly Ile Gly Pro Gly Ser Ile Cys Thr Thr Arg Ile
 290 295 300
 Val Ala Gly Val Gly Met Pro Gln Val Ser Ala Ile Asp Asn Cys Val
 305 310 315 320
 Glu Val Ala Ser Lys Phe Asp Ile Pro Val Ile Ala Asp Gly Gly Ile
 325 330 335

- 276 -

Arg Tyr Ser Gly Asp Val Ala Lys Ala Leu Ala Leu Gly Ala Ser Ser
 340 345 350
 Val Met Ile Gly Ser Leu Leu Ala Gly Thr Glu Glu Ser Pro Gly Asp
 355 360 365
 5 Phe Met Ile Tyr Gln Gly Arg Gln Tyr Lys Ser Tyr Arg Gly Met Gly
 370 375 380
 Ser Ile Gly Ala Met Thr Lys Gly Ser Ser Asp Arg Tyr Phe Gln Glu
 385 390 395 400
 10 Gly Val Ala Ser Glu Lys Leu Val Pro Glu Gly Ile Glu Gly Arg Val
 405 410 415
 Pro Tyr Arg Gly Lys Val Ser Asp Met Ile Phe Gln Leu Val Gly Gly
 420 425 430
 Val Arg Ser Ser Met Gly Tyr Gln Gly Ala Lys Asn Ile Leu Glu Leu
 435 440 445
 15 Tyr Gln Asn Ala Glu Phe Val Glu Ile Thr Ser Ala Gly Leu Lys Glu
 450 455 460
 Ser His Val His Gly Val Asp Ile Thr Lys Glu Ala Pro Asn Tyr Tyr
 465 470 475 480
 20 Gly

(2) INFORMATION FOR SEQ ID NO:191:

- 25 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 204 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

 30 (ii) MOLECULE TYPE: protein

 (iii) HYPOTHETICAL: YES

 (vi) ORIGINAL SOURCE:
 (A) ORGANISM: *Helicobacter pylori*
 35
 (ix) FEATURE:
 (A) NAME/KEY: misc_feature
 (B) LOCATION 1...204

 40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:191:

Met Gln Gly Phe Leu Leu Gln Thr Gln Ser Ile Arg Asp Glu Asp Leu
 1 5 10 15
 45 Ile Val His Val Leu Thr Lys Asn Gln Leu Lys Thr Leu Tyr Arg Phe
 20 25 30
 Tyr Gly Lys Arg His Ser Val Leu Asn Val Gly Arg Lys Ile Asp Phe
 35 40 45
 Glu Glu Glu Asn Asp Asp Lys Phe Leu Pro Lys Leu Arg Asn Ile Leu
 50 55 60
 50 His Leu Gly Tyr Ile Trp Glu Arg Glu Met Glu Arg Leu Phe Phe Trp
 65 70 75 80
 Gln Arg Phe Cys Ala Leu Leu Phe Lys His Leu Glu Gly Val His Ser
 85 90 95
 Leu Asp Ser Ile Tyr Phe Asp Thr Leu Asp Asp Gly Ala Ser Lys Leu

- 276 -

Arg Tyr Ser Gly Asp Val Ala Lys Ala Leu Ala Leu Gly Ala Ser Ser
 340 345 350
 Val Met Ile Gly Ser Leu Leu Ala Gly Thr Glu Glu Ser Pro Gly Asp
 355 360 365
 5 Phe Met Ile Tyr Gln Gly Arg Gln Tyr Lys Ser Tyr Arg Gly Met Gly
 370 375 380
 Ser Ile Gly Ala Met Thr Lys Gly Ser Ser Asp Arg Tyr Phe Gln Glu
 385 390 395 400
 10 Gly Val Ala Ser Glu Lys Leu Val Pro Glu Gly Ile Glu Gly Arg Val
 405 410 415
 Pro Tyr Arg Gly Lys Val Ser Asp Met Ile Phe Gln Leu Val Gly Gly
 420 425 430
 Val Arg Ser Ser Met Gly Tyr Gln Gly Ala Lys Asn Ile Leu Glu Leu
 435 440 445
 15 Tyr Gln Asn Ala Glu Phe Val Glu Ile Thr Ser Ala Gly Leu Lys Glu
 450 455 460
 Ser His Val His Gly Val Asp Ile Thr Lys Glu Ala Pro Asn Tyr Tyr
 465 470 475 480
 20 Gly

(2) INFORMATION FOR SEQ ID NO:191:

- 25 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 204 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 30 (ii) MOLECULE TYPE: protein
 (iii) HYPOTHETICAL: YES
 (vi) ORIGINAL SOURCE:
 35 (A) ORGANISM: *Helicobacter pylori*
 (ix) FEATURE:
 (A) NAME/KEY: misc_feature
 (B) LOCATION 1...204

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:191:

Met Gln Gly Phe Leu Leu Gln Thr Gln Ser Ile Arg Asp Glu Asp Leu
 1 5 10 15
 45 Ile Val His Val Leu Thr Lys Asn Gln Leu Lys Thr Leu Tyr Arg Phe
 20 25 30
 Tyr Gly Lys Arg His Ser Val Leu Asn Val Gly Arg Lys Ile Asp Phe
 35 40 45
 Glu Glu Glu Asn Asp Asp Lys Phe Leu Pro Lys Leu Arg Asn Ile Leu
 50 55 60
 50 His Leu Gly Tyr Ile Trp Glu Arg Glu Met Glu Arg Leu Phe Phe Trp
 65 70 75 80
 Gln Arg Phe Cys Ala Leu Leu Phe Lys His Leu Glu Gly Val His Ser
 85 90 95
 Leu Asp Ser Ile Tyr Phe Asp Thr Leu Asp Asp Gly Ala Ser Lys Leu

- 277 -

```

          100          105          110
Ser Lys Gln His Pro Leu Arg Val Ile Leu Glu Met Tyr Ala Val Leu
          115          120          125
Leu Asn Phe Glu Gly Arg Leu Gln Ser Tyr Asn Ser Cys Phe Leu Cys
5   130          135          140
Asp Ala Lys Leu Glu Arg Ser Val Ala Leu Ala Gln Gly Phe Ile Leu
145          150          155          160
Ala His Pro Ser Cys Leu Lys Ala Lys Ser Leu Asp Leu Glu Lys Ile
          165          170          175
10  Gln Ala Phe Phe Arg Thr Gln Ser Thr Ile Asp Leu Glu Thr Glu Glu
          180          185          190
Val Glu Glu Leu Trp Arg Thr Leu Asn Leu Gly Phe
          195          200

```

15 (2) INFORMATION FOR SEQ ID NO:192:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 82 amino acids

(B) TYPE: amino acid

20 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

25

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

30 (A) NAME/KEY: misc_feature

(B) LOCATION 1...82

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:192:

```

35 Met Gly Val Gly Arg Val Gly Asn Met Ala Leu Leu Ala Cys Ala Gly
   1           5           10           15
Pro Met Gly Ile Gly Ala Ile Ala Ile Asn Gly Gly Arg Gln
   20           25           30
Arg Ser Arg Met Leu Val Val Asp Ile Asp Asp Lys Arg Leu Glu Gln
40   35           40           45
Val Gln Lys Met Leu Pro Gly Asn Trp Arg Pro Val Thr Ala Leu Ser
   50           55           60
Trp Cys Leu Cys Ile Pro Lys Arg Gly Ala Ile Arg Ala Arg Cys Cys
65           70           75           80
45 Glu Arg

```

(2) INFORMATION FOR SEQ ID NO:193:

50

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 67 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

- 278 -

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

5 (vi) ORIGINAL SOURCE:

(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:

(A) NAME/KEY: misc_feature

10 (B) LOCATION 1...67

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:193:

15 Leu Ser Gly Thr Ala Val Ser Cys Arg Cys Thr Cys Arg Ile Gln Leu
 1 5 10 15
 Val Leu Val Arg Thr Ser Ile Pro Val Val Ile Gly Cys Ser Cys Pro
 20 25 30
 Phe Leu Ser Ser Ile Gly Phe Thr Thr Gly Thr His Gln Ser Pro Val
 35 40 45
 20 Lys Arg Cys Gly Val Asn Ala Gly Lys Thr Pro Ser Lys Lys His Leu
 50 55 60
 His Leu Asn
 65

25 (2) INFORMATION FOR SEQ ID NO:194:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 114 amino acids

30 (B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

35

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:

40 (A) NAME/KEY: misc_feature

(B) LOCATION 1...114

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:194:

45 Val Trp Leu Ala Ala Leu Gly Phe Leu Ile Thr Ala Val Gly Leu Pro
 1 5 10 15
 Val Ile Thr Val Ile Ala Leu Ala Lys Val Gly Gly Ser Ser Thr Pro
 20 25 30
 Ser Ala Ile Arg Ser Ala Gly Met Pro Ala Ala Cys Trp Arg Arg Ser
 35 40 45
 50 Ala Thr Trp Arg Ser Ala Arg Cys Ser Pro Phe Arg Ala Pro Pro Arg
 50 55 60
 Cys Pro Ser Lys Val Ser Val Val Pro Leu Leu Gly Glu Glu Ala Ala
 65 70 75 80

- 279 -

Arg Arg Cys Ser Ser Thr Ala Trp Arg Thr Ser Ser Ser Pro Trp Pro
 85 90 95
 Ser Pro Ser Thr Pro Val Ala Cys Trp Thr Pro Ser Asp Ala Ser Ser
 100 105 110
 5 Pro Arg

(2) INFORMATION FOR SEQ ID NO:195:

10 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 20 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: circular
 15 (ii) MOLECULE TYPE: DNA (genomic)
 (iii) HYPOTHETICAL: NO
 20 (iv) ANTI-SENSE: NO
 (vi) ORIGINAL SOURCE:
 (A) ORGANISM: Helicobacter pylori
 25 (ix) FEATURE:
 (A) NAME/KEY: misc_feature
 (B) LOCATION 1...20
 30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:195:
 TATACCATGG TGGGCGCTAA 20

(2) INFORMATION FOR SEQ ID NO:196:

35 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 23 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: circular
 40 (ii) MOLECULE TYPE: DNA (genomic)
 (iii) HYPOTHETICAL: NO
 45 (iv) ANTI-SENSE: NO
 (vi) ORIGINAL SOURCE:
 (A) ORGANISM: Helicobacter pylori
 50 (ix) FEATURE:
 (A) NAME/KEY: misc_feature
 (B) LOCATION 1...23
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:196:

- 280 -

ATGAATTCGA GTAAGGATTT TTG

23

(2) INFORMATION FOR SEQ ID NO:197:

5

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 22 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: circular

10

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

15

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Helicobacter pylori

20

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...22

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:197:

TTAACCATGG TGAAAAGCGA TA

22

(2) INFORMATION FOR SEQ ID NO:198:

30

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 23 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: circular

35

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

40

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Helicobacter pylori

45

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...23

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:198:

TAGAATTCGC ATAACGATCA ATC

23

(2) INFORMATION FOR SEQ ID NO:199:

- 281 -

- 5 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- 10 (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
15 (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...22
20
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:199:

ATATCCATGG TGAGTTTGAT GA

22

- 25 (2) INFORMATION FOR SEQ ID NO:200:
- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
30 (C) STRANDEDNESS: double
(D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- 35 (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
40 (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...25
45
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:200:

ATGAATTCAA TTTTATTATT TGCCA

25

- 50 (2) INFORMATION FOR SEQ ID NO:201:
- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid

- 282 -

(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

5 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

10 (vi) ORIGINAL SOURCE:

(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:

15 (A) NAME/KEY: misc_feature

(B) LOCATION 1...21

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:201:

20 AATTCCATGG TGGGGGCTAT G

21

(2) INFORMATION FOR SEQ ID NO:202:

(i) SEQUENCE CHARACTERISTICS:

25 (A) LENGTH: 23 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: circular

30 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

35 (vi) ORIGINAL SOURCE:

(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:

40 (A) NAME/KEY: misc_feature

(B) LOCATION 1...23

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:202:

45 ATGAATTCTC GATAGCCAAA ATC

23

(2) INFORMATION FOR SEQ ID NO:203:

(i) SEQUENCE CHARACTERISTICS:

50 (A) LENGTH: 25 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

- 283 -

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

5 (vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:

10 (A) NAME/KEY: misc_feature
(B) LOCATION 1...25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:203:

15 AATTCCATGG TGCATAACTT CCATT 25

(2) INFORMATION FOR SEQ ID NO:204:

(i) SEQUENCE CHARACTERISTICS:

20 (A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

25 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

30 (vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:

35 (A) NAME/KEY: misc_feature
(B) LOCATION 1...25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:204:

40 AAGAATTCTC TAGCATCCAA ATGGA 25

(2) INFORMATION FOR SEQ ID NO:205:

(i) SEQUENCE CHARACTERISTICS:

45 (A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

50 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

- 284 -

(vi) ORIGINAL SOURCE:
 (A) ORGANISM: *Helicobacter pylori*

5 (ix) FEATURE:
 (A) NAME/KEY: misc_feature
 (B) LOCATION 1...24

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:205:
ATTTCATGG TCATGTCTCA TATT 24

(2) INFORMATION FOR SEQ ID NO:206:

15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 23 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: circular

20 (ii) MOLECULE TYPE: DNA (genomic)

 (iii) HYPOTHETICAL: NO

25 (iv) ANTI-SENSE: NO

 (vi) ORIGINAL SOURCE:
 (A) ORGANISM: *Helicobacter pylori*

30 (ix) FEATURE:
 (A) NAME/KEY: misc_feature
 (B) LOCATION 1...23

 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:206:

35 ATGAATTCCA TCTTTTATTC CAC 23

(2) INFORMATION FOR SEQ ID NO:207:

40 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 27 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: circular

45 (ii) MOLECULE TYPE: DNA (genomic)

 (iii) HYPOTHETICAL: NO

50 (iv) ANTI-SENSE: NO

 (vi) ORIGINAL SOURCE:
 (A) ORGANISM: *Helicobacter pylori*

- 285 -

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION 1...27

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:207:

AACCATGGTG ATTTTAAGCA TTGAAAG

27

(2) INFORMATION FOR SEQ ID NO:208:

10

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

15

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

20

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Helicobacter pylori

25

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION 1...28

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:208:

AAGAATTCCA CTCAAAATTT TTTAACAG

28

(2) INFORMATION FOR SEQ ID NO:209:

35

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

40

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

45

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Helicobacter pylori

50

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION 1...25

- 286 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:209:

GATCATCCAT ATGTTATCTT CTAAT

25

5 (2) INFORMATION FOR SEQ ID NO:210:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 23 base pairs

(B) TYPE: nucleic acid

10 (C) STRANDEDNESS: double

(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

15 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

20 (A) ORGANISM: Helicobacter pylori

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...23

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:210:

TGAATTCAAC CATTTTAACC CTG

23

30 (2) INFORMATION FOR SEQ ID NO:211:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 27 base pairs

(B) TYPE: nucleic acid

35 (C) STRANDEDNESS: double

(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

40 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

45 (A) ORGANISM: Helicobacter pylori

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...27

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:211:

TATACCATGG TGAAATTTTT TCTTTTA

27

- 287 -

(2) INFORMATION FOR SEQ ID NO:212:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Helicobacter pylori

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION 1...25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:212:

AGAATTCAAT TCGGTCTTGT AAAAG

25

(2) INFORMATION FOR SEQ ID NO:213:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Helicobacter pylori

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION 1...24

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:213:

TATACCATGG TGATGGACAA ACTC

24

(2) INFORMATION FOR SEQ ID NO:214:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs

- 288 -

(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

5 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

10 (iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

15 (A) NAME/KEY: misc_feature

(B) LOCATION 1...23

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:214:

20 ATGAATCCC ACTTGGGGCG ATA

23

(2) INFORMATION FOR SEQ ID NO:215:

(i) SEQUENCE CHARACTERISTICS:

25 (A) LENGTH: 25 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: circular

30 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

35 (iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

40 (A) NAME/KEY: misc_feature

(B) LOCATION 1...25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:215:

45 TTATGGATCC AAACCAATTA AAAC

25

(2) INFORMATION FOR SEQ ID NO:216:

(i) SEQUENCE CHARACTERISTICS:

50 (A) LENGTH: 23 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: circular

- 289 -

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

5 (iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

10 (ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...23

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:216:

15 TATCTCGAGT TATAGAGAAG GGC 23

(2) INFORMATION FOR SEQ ID NO:217:

20 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

25 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

30 (iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

35 (ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...22

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:217:

40 TTAACCATGG TGAAAAGCGA TA 22

(2) INFORMATION FOR SEQ ID NO:218:

45 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

50 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

- 290 -

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...24

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:218:

TAGAATTCGC CTCTAAACT TTAG

24

(2) INFORMATION FOR SEQ ID NO:219:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 22 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...22

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:219:

TTAACCATGG TGAAAAGCGA TA

22

(2) INFORMATION FOR SEQ ID NO:220:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 23 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

- 291 -

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION 1...23

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:220:

TAGAATTCGC ATAACGATCA ATC

23

10 (2) INFORMATION FOR SEQ ID NO:221:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

15

(ii) MOLECULE TYPE: DNA (genomic)

20 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

25

- (A) ORGANISM: Helicobacter pylori

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION 1...22

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:221:

ATATCCATGG TGAGTTTGAT GA

22

35 (2) INFORMATION FOR SEQ ID NO:222:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

40

(ii) MOLECULE TYPE: DNA (genomic)

45 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

50

- (A) ORGANISM: Helicobacter pylori

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION 1...25

- 292 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:222:

ATGAATTCAA TTTTATTATT TGCCA

25

5

(2) INFORMATION FOR SEQ ID NO:223:

(i) SEQUENCE CHARACTERISTICS:

10

(A) LENGTH: 23 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

15

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

20

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:

25

(A) NAME/KEY: misc_feature

(B) LOCATION 1...23

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:223:

AATTCCATGG CTATCCAAAT CCG

23

30

(2) INFORMATION FOR SEQ ID NO:224:

(i) SEQUENCE CHARACTERISTICS:

35

(A) LENGTH: 25 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

40

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

45

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:

50

(A) NAME/KEY: misc_feature

(B) LOCATION 1...25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:224:

ATGAATTCGC CAAAATCGTA GTATT

25

- 293 -

(2) INFORMATION FOR SEQ ID NO:225:

- 5 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular
- 10 (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- 15 (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori
- 20 (ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...24
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:225:

25 GATACCATGG AATTTATGAA AAAG

24

(2) INFORMATION FOR SEQ ID NO:226:

- 30 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular
- 35 (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- 40 (vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori
- 45 (ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...25
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:226:

50 TGAATTCGAA AAAGTGTAGT TATAC

25

(2) INFORMATION FOR SEQ ID NO:227:

- (i) SEQUENCE CHARACTERISTICS:

- 294 -

- (A) LENGTH: 19 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: circular
- 5 (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- 10 (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: *Helicobacter pylori*
- 15 (ix) FEATURE:
- (A) NAME/KEY: misc_feature
 - (B) LOCATION 1...19
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:227:
- 20 CCCTTCATTT TAGAAATCG 19
- (2) INFORMATION FOR SEQ ID NO:228:
- 25 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 20 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: circular
- 30 (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- 35 (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: *Helicobacter pylori*
- 40 (ix) FEATURE:
- (A) NAME/KEY: misc_feature
 - (B) LOCATION 1...20
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:228:
- 45 ATTTCAACCA ATTCAATGCG 20
- (2) INFORMATION FOR SEQ ID NO:229:
- 50 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 20 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: circular

- 295 -

- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
(A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...20
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:229:

GCCCCTTTTG ATTTGAAGCT

20

- (2) INFORMATION FOR SEQ ID NO:230:
- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
(A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...22
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:230:

TCGCTCCAAG ATACCAAGAA GT

22

- (2) INFORMATION FOR SEQ ID NO:231:
- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO

- 296 -

(iv) ANTI-SENSE: NO

5 (vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...22

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:231:

CTTGAATTAG GGGCAAAGAT CG 22

15 (2) INFORMATION FOR SEQ ID NO:232:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
20 (C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

25 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
30 (A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...22

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:232:

ATGCGTTTTT ACCCAAAGAA GT 22

40 (2) INFORMATION FOR SEQ ID NO:233:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
45 (C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

50 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- 297 -

(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...22

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:233:

ATAACGCCAC TTCCTTATTG GT

22

(2) INFORMATION FOR SEQ ID NO:234:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 19 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...19

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:234:

CTTTGGGTAA AAACGCATC

19

(2) INFORMATION FOR SEQ ID NO:235:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:

(A) NAME/KEY: misc_feature

- 298 -

(B) LOCATION 1...20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:235:

5 CGATCTTTGA TCCTAATTCA

20

(2) INFORMATION FOR SEQ ID NO:236:

(i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

15 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

20 (iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

- 25 (A) NAME/KEY: misc_feature
(B) LOCATION 1...19

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:236:

30 ATCAAGTTGC CTATGCTGA

19

(2) INFORMATION FOR SEQ ID NO:237:

(i) SEQUENCE CHARACTERISTICS:

- 35 (A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

40 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

45 (iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

- 50 (A) NAME/KEY: misc_feature
(B) LOCATION 1...22

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:237:

- 299 -

TTGAACACTT TTGATTATGC GG

22

(2) INFORMATION FOR SEQ ID NO:238:

- 5 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular
- 10 (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- 15 (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
(A) ORGANISM: *Helicobacter pylori*
- 20 (ix) FEATURE:
(A) NAME/KEY: *misc_feature*
(B) LOCATION 1...23
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:238:

25

GGATTATGCG ATTGTTTTAC AAG

23

(2) INFORMATION FOR SEQ ID NO:239:

- 30 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular
- 35 (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- 40 (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
(A) ORGANISM: *Helicobacter pylori*
- 45 (ix) FEATURE:
(A) NAME/KEY: *misc_feature*
(B) LOCATION 1...21
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:239:

50

GTCTTTAGCA AAAATGGCGT C

21

(2) INFORMATION FOR SEQ ID NO:240:

- 300 -

- 5 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- 10 (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
15 (A) ORGANISM: Helicobacter pylori
- (ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...21
- 20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:240:

AATGAGCGTA AGAGAGCCTT C

21

- 25 (2) INFORMATION FOR SEQ ID NO:241:
- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
30 (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- 35 (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
40 (A) ORGANISM: Helicobacter pylori
- (ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...18
- 45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:241:

CTTATGGGGG TATTGTCA

18

- 50 (2) INFORMATION FOR SEQ ID NO:242:
- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double

- 301 -

(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

5 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

10 (A) ORGANISM: Helicobacter pylori

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...18

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:242:

AGCATGTGGG TATCCAGC 18

20 (2) INFORMATION FOR SEQ ID NO:243:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 19 base pairs

(B) TYPE: nucleic acid

25 (C) STRANDEDNESS: double

(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

30 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

35 (A) ORGANISM: Helicobacter pylori

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...19

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:243:

AGGTTGTTGC CTAAAGACT 19

45 (2) INFORMATION FOR SEQ ID NO:244:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 base pairs

(B) TYPE: nucleic acid

50 (C) STRANDEDNESS: double

(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

- 302 -

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

5 (vi) ORIGINAL SOURCE:
(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:
10 (A) NAME/KEY: misc_feature
(B) LOCATION 1...18

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:244:

15 CTGCCTCCAC CTTTGATC 18

(2) INFORMATION FOR SEQ ID NO:245:

(i) SEQUENCE CHARACTERISTICS:
20 (A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

25 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

30 (vi) ORIGINAL SOURCE:
(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:
35 (A) NAME/KEY: misc_feature
(B) LOCATION 1...19

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:245:

40 ACCAATATCA ATTGGCACT 19

(2) INFORMATION FOR SEQ ID NO:246:

(i) SEQUENCE CHARACTERISTICS:
45 (A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

50 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

- 303 -

(vi) ORIGINAL SOURCE:
(A) ORGANISM: *Helicobacter pylori*

5 (ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...18

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:246:

10 ACTTGGA AAA GCTCTGCA

18

(2) INFORMATION FOR SEQ ID NO:247:

15 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

20 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

25 (iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: *Helicobacter pylori*

30 (ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...19

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:247:

35 CTTGCTTGTC ATATCTAGC

19

(2) INFORMATION FOR SEQ ID NO:248:

40 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

45 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

50 (iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

- 304 -

- (A) NAME/KEY: misc_feature
- (B) LOCATION 1...18

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:248:

5

GTTGAAGTGT TGGTGCTA

18

(2) INFORMATION FOR SEQ ID NO:249:

10

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 22 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: circular

15

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

20

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Helicobacter pylori

25

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...22

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:249:

30

CAAGCAAGTG GTTTGGTTTT AG

22

(2) INFORMATION FOR SEQ ID NO:250:

35

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 22 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: circular

40

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

45

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Helicobacter pylori

50

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...22

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:250:

- 305 -

TGGAAAGAGC AAATCATTGA AG

22

(2) INFORMATION FOR SEQ ID NO:251:

5

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 21 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: circular

10

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

15

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Helicobacter pylori

20

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...21

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:251:

GCCCATAATC AAAAAGCCCA T

21

(2) INFORMATION FOR SEQ ID NO:252:

30

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 24 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: circular

35

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

40

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Helicobacter pylori

45

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...24

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:252:

CTAAAACCAA ACCACTTGCT TGTC

24

(2) INFORMATION FOR SEQ ID NO:253:

- 306 -

- 5 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- 10 (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- 15 (vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori
- (ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...16
- 20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:253:

GTAAAACGAC GGCCAG

16

- 25 (2) INFORMATION FOR SEQ ID NO:254:
- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
30 (C) STRANDEDNESS: double
(D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- 35 (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- 40 (vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori
- (ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...17
- 45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:254:

CAGGAAACAG CTATGAC

17

- 50 (2) INFORMATION FOR SEQ ID NO:255:
- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid

- 307 -

(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

5 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

10 (vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
15 (A) NAME/KEY: misc_feature
(B) LOCATION 1...21

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:255:

20 ATCTTACCTA TCACCTCAA T 21

(2) INFORMATION FOR SEQ ID NO:256:

(i) SEQUENCE CHARACTERISTICS:
25 (A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

30 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

35 (vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:
40 (A) NAME/KEY: misc_feature
(B) LOCATION 1...21

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:256:

45 AGACAGCAAC ATCTTTGTGA A 21

(2) INFORMATION FOR SEQ ID NO:257:

(i) SEQUENCE CHARACTERISTICS:
50 (A) LENGTH: 50 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

- 308 -

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

5 (vi) ORIGINAL SOURCE:
(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

10 (A) NAME/KEY: misc_feature
(B) LOCATION 1...50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:257:

15 CGCGGATCCA TATGGCTGAA AAAACGCCTT TTTTAAAC TAAAACCAC 50

(2) INFORMATION FOR SEQ ID NO:258:

(i) SEQUENCE CHARACTERISTICS:

20 (A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

25 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

30 (vi) ORIGINAL SOURCE:
(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

35 (A) NAME/KEY: misc_feature
(B) LOCATION 1...34

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:258:

40 CCGGAATTCA TCAGTATTCA ATGGGAATAA AGCC 34

(2) INFORMATION FOR SEQ ID NO:259:

(i) SEQUENCE CHARACTERISTICS:

45 (A) LENGTH: 50 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

50 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

- 309 -

(vi) ORIGINAL SOURCE:
(A) ORGANISM: *Helicobacter pylori*

5 (ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...50

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:259:

CGCGGATCCA TATGAAAGAA GAAGAAAAAG AAGAAAAAAA GACAGAAAGG

50

(2) INFORMATION FOR SEQ ID NO:260:

15 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 37 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
20 (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

25 (iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: *Helicobacter pylori*

30 (ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...37

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:260:

CCGGAATTTCG CTAAAAGAA AATAGTCCCC CAAACGC

37

(2) INFORMATION FOR SEQ ID NO:261:

40 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 43 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
45 (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

50 (iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: *Helicobacter pylori*

- 310 -

- (ix) FEATURE:
 (A) NAME/KEY: misc_feature
 (B) LOCATION 1...43

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:261:

CGCGGATCCA TATGAAAGAG GTCATTCCCA CCCCTTCAAC CCC

43

(2) INFORMATION FOR SEQ ID NO:262:

10

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 36 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: circular

15

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

20

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Helicobacter pylori

25

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
(B) LOCATION 1...36

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:262:

CCGGAATTCA TATAAATATC ATATAGGCAG AAAAAC

36

(2) INFORMATION FOR SEQ ID NO:263:

35

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 37 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: circular

40

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

45

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Helicobacter pylori

50

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
(B) LOCATION 1...37

- 311 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:263:

CGCGGATCCA TATGGAGGCA GAGCTTGATG AAAAATC

37

5 (2) INFORMATION FOR SEQ ID NO:264:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 36 base pairs

(B) TYPE: nucleic acid

10 (C) STRANDEDNESS: double

(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

15 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

20 (A) ORGANISM: Helicobacter pylori

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...36

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:264:

CCGGAATTCG ATTGATTTTG TCAAATCTAA AATCCC

36

30 (2) INFORMATION FOR SEQ ID NO:265:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 25 base pairs

(B) TYPE: nucleic acid

35 (C) STRANDEDNESS: double

(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

40 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

45 (A) ORGANISM: Helicobacter pylori

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...25

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:265:

TATTATACAT ATGGAAGAAG ATGGG

25

- 312 -

(2) INFORMATION FOR SEQ ID NO:266:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Helicobacter pylori

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION 1...23

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:266:

TAATCTCGAG TTTAGAAGGC GTA

23

(2) INFORMATION FOR SEQ ID NO:267:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Helicobacter pylori

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION 1...25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:267:

TTATATTCAT ATGGAAGACG ATGGC

25

(2) INFORMATION FOR SEQ ID NO:268:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs

- 313 -

(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

5 (ii) MOLECULE TYPE: DNA (genomic)
(iii) HYPOTHETICAL: NO
(iv) ANTI-SENSE: NO
10 (vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori
(ix) FEATURE:
15 (A) NAME/KEY: misc_feature
(B) LOCATION 1...24
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:268:

20 AATTCTCGAG CCTCTTTATA AGCC

24

(2) INFORMATION FOR SEQ ID NO:269:

25 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 46 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

30 (ii) MOLECULE TYPE: DNA (genomic)
(iii) HYPOTHETICAL: NO
(iv) ANTI-SENSE: NO
35 (vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter pylori
(ix) FEATURE:
40 (A) NAME/KEY: misc_feature
(B) LOCATION 1...46

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:269:

45 CGCGGATCCA TATGGTAGAA GCCTTTCAA AACACCAAAA AGACGG

46

(2) INFORMATION FOR SEQ ID NO:270:

50 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

- 314 -

- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- 5 (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
(A) ORGANISM: *Helicobacter pylori*
- 10 (ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...32
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:270:
- 15 CCGGAATTCG GAGCCAATAG GGAGCTAAAG CC 32
- (2) INFORMATION FOR SEQ ID NO:271:
- 20 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 31 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular
- 25 (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- 30 (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
(A) ORGANISM: *Helicobacter pylori*
- 35 (ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...31
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:271:
- 40 CCGGATCCGA AGGTGATGGT GTTTATATAG G 31
- (2) INFORMATION FOR SEQ ID NO:272:
- 45 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular
- 50 (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO

- 315 -

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: *Helicobacter pylori*

5 (ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...32

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:272:
CGCATATGGA AGGTGATGGT GTTTATATAG GG 32

(2) INFORMATION FOR SEQ ID NO:273:

15 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 37 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
20 (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

25 (iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: *Helicobacter pylori*

30 (ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION 1...37

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:273:
GCGAATTCTC ACTCTTTCCA ATAGTTTGCT GCAGAGC 37

(2) INFORMATION FOR SEQ ID NO:274:

40 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 37 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
45 (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

50 (iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: *Helicobacter pylori*

- 316 -

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...37

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:274:

CCGGAATTCT TAATCCCGTT TCAAATGGTA ATAAAGG

37

10 (2) INFORMATION FOR SEQ ID NO:275:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 36 base pairs

(B) TYPE: nucleic acid

15 (C) STRANDEDNESS: double

(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

20 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

25 (A) ORGANISM: Helicobacter pylori

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...36

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:275:

GCGAATTCCC TTTTATTTAA AAAGTGTAGT TATACC

36

This Sheet Has Been Replaced

CLAIMS

- 5 1. An isolated nucleic acid comprising a nucleotide sequence encoding an
H. pylori polypeptide at least about 60% homologous to an amino acid sequence
selected from the group consisting of SEQ ID NO: 98-SEQ ID NO: 194.
2. An isolated nucleic acid comprising a nucleotide sequence encoding an
10 *H. pylori* polypeptide selected from the group consisting of SEQ ID NO: 98-SEQ ID
NO: 194.
3. An isolated nucleic acid which encodes an *H. pylori* polypeptide,
comprising a nucleotide sequence at least about 60% homologous to a nucleotide
15 sequence selected from the group consisting of SEQ ID NO: 1-SEQ ID NO: 97, or a
complement thereof.
4. The isolated nucleic acid of claim 1, comprising a nucleotide sequence
selected from the group consisting of SEQ ID NO: 1-SEQ ID NO: 97, or a complement
20 thereof.
5. An isolated nucleic acid molecule encoding an *H. pylori* polypeptide,
comprising a nucleotide sequence which hybridizes under stringent hybridization
conditions to a nucleic acid molecule comprising the nucleotide sequence selected from
25 the group consisting of SEQ ID NO: 1-SEQ ID NO: 97, or a complement thereof.
6. An isolated nucleic acid comprising a nucleotide sequence of at least 8
nucleotides in length, wherein the sequence hybridizes under stringent hybridization
conditions to a nucleic acid having a nucleotide sequence selected from the group
30 consisting of SEQ ID NO: 1-SEQ ID NO: 97, or a complement thereof.
7. An isolated nucleic acid comprising a nucleotide sequence encoding an
H. pylori cell envelope polypeptide or a fragment thereof, said nucleic acid selected
from the group consisting of SEQ ID NO: 63, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID
35 NO: 9, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID
NO: 27, SEQ ID NO: 28, SEQ ID NO: 50, SEQ ID NO: 51, SEQ ID NO: 61, SEQ ID

THIS SHEET HAS BEEN
REPLACED





NO: 79, SEQ ID NO: 80, SEQ ID NO: 84, SEQ ID NO: 85, SEQ ID NO: 91, SEQ ID NO: 94, SEQ ID NO: 5, SEQ ID NO: 11, SEQ ID NO: 26, SEQ ID NO: 36, SEQ ID NO: 42, SEQ ID NO: 52, SEQ ID NO: 22, SEQ ID NO: 29, SEQ ID NO: 30, SEQ ID NO: 65, SEQ ID NO: 66, SEQ ID NO: 48, SEQ ID NO: 49, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 43, SEQ ID NO: 44, SEQ ID NO: 38, SEQ ID NO: 39, SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 6, SEQ ID NO: 34, SEQ ID NO: 35, SEQ ID NO: 60, SEQ ID NO: 69, and SEQ ID NO: 83, or a complement thereof.

8. The isolated nucleic acid of claim 7, wherein said *H. pylori* cell envelope polypeptide or a fragment thereof is an *H. pylori* flagella-associated polypeptide or a fragment thereof comprising a nucleotide sequence of SEQ ID NO: 63, or a complement thereof.

9. The isolated nucleic acid of claim 7, wherein said *H. pylori* cell envelope polypeptide or a fragment thereof is an *H. pylori* inner membrane polypeptide or a fragment thereof encoded by a nucleic acid selected from the group consisting of SEQ ID NO: 48, SEQ ID NO: 49, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 43, SEQ ID NO: 44, SEQ ID NO: 38, and SEQ ID NO: 39, or a complement thereof.

10. The isolated nucleic acid of claim 9, wherein said *H. pylori* inner membrane polypeptide or a fragment thereof is an *H. pylori* polypeptide or a fragment thereof involved in transport encoded by a nucleic acid selected from the group consisting of SEQ ID NO: 48, SEQ ID NO: 49, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 43, and SEQ ID NO: 44, or a complement thereof.

11. The isolated nucleic acid of claim 7, wherein said *H. pylori* cell envelope polypeptide or a fragment thereof is an *H. pylori* outer membrane polypeptide or a fragment thereof encoded by a nucleic acid selected from the group consisting of SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 50, SEQ ID NO: 51, SEQ ID NO: 61, SEQ ID NO: 79, SEQ ID NO: 80, SEQ ID NO: 84, SEQ ID NO: 85, SEQ ID NO: 91, SEQ ID NO: 94, SEQ ID NO: 5, SEQ ID NO: 11, SEQ ID NO: 26, SEQ ID NO: 36, SEQ ID NO: 42, SEQ ID NO: 52, SEQ ID NO: 22, SEQ ID NO: 29, SEQ ID NO: 30, SEQ ID NO: 65, and SEQ ID NO: 66, or a complement thereof.

THIS SHEET HAS BEEN
REPLACED



This Sheet Has Been Replaced



12. The isolated nucleic acid of claim 11, wherein said *H. pylori* outer membrane polypeptide or a fragment thereof is an *H. pylori* polypeptide having a terminal phenylalanine residue or a fragment thereof encoded by a nucleic acid selected from the group consisting of SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 42, SEQ ID NO: 50, SEQ ID NO: 51, SEQ ID NO: 52, SEQ ID NO: 61, SEQ ID NO: 79, SEQ ID NO: 80, SEQ ID NO: 84, SEQ ID NO: 85, SEQ ID NO: 91, and SEQ ID NO: 94, or a complement thereof.

13. The isolated nucleic acid of claim 12, wherein said *H. pylori* outer membrane polypeptide or a fragment thereof is an *H. pylori* polypeptide having a terminal phenylalanine residue and a C-terminal tyrosine cluster or a fragment thereof encoded by a nucleic acid selected from the group consisting of SEQ ID NO: 11, SEQ ID NO: 26, SEQ ID NO: 36, SEQ ID NO: 42, and SEQ ID NO: 52, or a complement thereof.

14. An isolated nucleic acid comprising a nucleotide sequence encoding an *H. pylori* cell envelope polypeptide or a fragment thereof, wherein said polypeptide is selected from the group consisting of SEQ ID NO: 160, SEQ ID NO: 104, SEQ ID NO: 105, SEQ ID NO: 106, SEQ ID NO: 110, SEQ ID NO: 111, SEQ ID NO: 120, SEQ ID NO: 121, SEQ ID NO: 124, SEQ ID NO: 125, SEQ ID NO: 147, SEQ ID NO: 148, SEQ ID NO: 158, SEQ ID NO: 176, SEQ ID NO: 177, SEQ ID NO: 181, SEQ ID NO: 182, SEQ ID NO: 188, SEQ ID NO: 191, SEQ ID NO: 102, SEQ ID NO: 108, SEQ ID NO: 123, SEQ ID NO: 133, SEQ ID NO: 139, SEQ ID NO: 149, SEQ ID NO: 119, SEQ ID NO: 126, SEQ ID NO: 127, SEQ ID NO: 162, SEQ ID NO: 163, SEQ ID NO: 145, SEQ ID NO: 146, SEQ ID NO: 114, SEQ ID NO: 115, SEQ ID NO: 116, SEQ ID NO: 140, SEQ ID NO: 141, SEQ ID NO: 135, SEQ ID NO: 136, SEQ ID NO: 98, SEQ ID NO: 99, SEQ ID NO: 103, SEQ ID NO: 131, SEQ ID NO: 132, SEQ ID NO: 157, SEQ ID NO: 166, and SEQ ID NO: 180.

15. The isolated nucleic acid of claim 14, wherein said *H. pylori* cell envelope polypeptide or a fragment thereof is an *H. pylori* flagella-associated polypeptide or a fragment thereof comprising an amino acid sequence of SEQ ID NO: 160.

THIS SHEET HAS BEEN
REPLACED





16. The isolated nucleic acid of claim 14, wherein said *H. pylori* cell envelope polypeptide or a fragment thereof is an *H. pylori* inner membrane polypeptide or a fragment thereof selected from the group consisting of SEQ ID NO: 145, SEQ ID NO: 146, SEQ ID NO: 114, SEQ ID NO: 115, SEQ ID NO: 116, SEQ ID NO: 140, SEQ ID NO: 141, SEQ ID NO: 135, and SEQ ID NO: 136.

17. The isolated nucleic acid of claim 16, wherein said *H. pylori* inner membrane polypeptide or a fragment thereof is an *H. pylori* polypeptide or a fragment thereof involved in transport selected from the group consisting of SEQ ID NO: 145, SEQ ID NO: 146, SEQ ID NO: 114, SEQ ID NO: 115, SEQ ID NO: 116, SEQ ID NO: 140, and SEQ ID NO: 141.

18. The isolated nucleic acid of claim 14, wherein said *H. pylori* cell envelope polypeptide or a fragment thereof is an *H. pylori* outer membrane polypeptide or a fragment thereof selected from the group consisting of SEQ ID NO: 104, SEQ ID NO: 105, SEQ ID NO: 106, SEQ ID NO: 110, SEQ ID NO: 111, SEQ ID NO: 120, SEQ ID NO: 121, SEQ ID NO: 124, SEQ ID NO: 125, SEQ ID NO: 147, SEQ ID NO: 148, SEQ ID NO: 158, SEQ ID NO: 176, SEQ ID NO: 177, SEQ ID NO: 181, SEQ ID NO: 182, SEQ ID NO: 188, SEQ ID NO: 191, SEQ ID NO: 102, SEQ ID NO: 108, SEQ ID NO: 123, SEQ ID NO: 133, SEQ ID NO: 139, SEQ ID NO: 149, SEQ ID NO: 119, SEQ ID NO: 126, SEQ ID NO: 127, SEQ ID NO: 162, and SEQ ID NO: 163.

19. The isolated nucleic acid of claim 18, wherein said *H. pylori* outer membrane polypeptide or a fragment thereof is an *H. pylori* polypeptide having a terminal phenylalanine residue or a fragment thereof selected from the group consisting of SEQ ID NO: 104, SEQ ID NO: 105, SEQ ID NO: 106, SEQ ID NO: 108, SEQ ID NO: 110, SEQ ID NO: 111, SEQ ID NO: 120, SEQ ID NO: 121, SEQ ID NO: 123, SEQ ID NO: 124, SEQ ID NO: 125, SEQ ID NO: 133, SEQ ID NO: 139, SEQ ID NO: 147, SEQ ID NO: 148, SEQ ID NO: 149, SEQ ID NO: 158, SEQ ID NO: 176, SEQ ID NO: 177, SEQ ID NO: 181, SEQ ID NO: 182, SEQ ID NO: 188, and SEQ ID NO: 191.

20. The isolated nucleic acid of claim 19, wherein said *H. pylori* outer membrane polypeptide or a fragment thereof is an *H. pylori* polypeptide having a terminal phenylalanine residue and a C-terminal tyrosine cluster or a fragment thereof

THIS SHEET HAS BEEN
REPLACED





- 321 -

selected from the group consisting of SEQ ID NO: 108, SEQ ID NO: 123, SEQ ID NO: 133, SEQ ID NO: 139, and SEQ ID NO: 149.

21. An isolated nucleic acid comprising a nucleotide sequence encoding an *H. pylori* cytoplasmic polypeptide or a fragment thereof, said nucleic acid selected from the group consisting of SEQ ID NO: 57, SEQ ID NO: 58, SEQ ID NO: 86, SEQ ID NO: 87, SEQ ID NO: 88, SEQ ID NO: 89, SEQ ID NO: 92, and SEQ ID NO: 93, or a complement thereof.

22. The isolated nucleic acid of claim 21, wherein said *H. pylori* cytoplasmic polypeptide or a fragment thereof is an *H. pylori* polypeptide or a fragment thereof involved in mRNA translation, said nucleic acid selected from the group consisting of SEQ ID NO: 57 and SEQ ID NO: 58, or a complement thereof.

23. The isolated nucleic acid of claim 21, wherein said *H. pylori* cytoplasmic polypeptide or a fragment thereof is an *H. pylori* polypeptide or a fragment thereof involved in genome replication, transcription, recombination and repair, said nucleic acid selected from the group consisting of SEQ ID NO: 86 and SEQ ID NO: 87, or a complement thereof.

24. An isolated nucleic acid comprising a nucleotide sequence encoding an *H. pylori* cytoplasmic polypeptide or a fragment thereof, wherein said polypeptide is selected from the group consisting of SEQ ID NO: 154, SEQ ID NO: 155, SEQ ID NO: 183, SEQ ID NO: 184, SEQ ID NO: 185, SEQ ID NO: 186, SEQ ID NO: 189, and SEQ ID NO: 190.

25. The isolated nucleic acid of claim 24, wherein said *H. pylori* cytoplasmic polypeptide or a fragment thereof is an *H. pylori* polypeptide or a fragment thereof involved in mRNA translation selected from the group consisting of SEQ ID NO: 154 and SEQ ID NO: 155.

26. The isolated nucleic acid of claim 24, wherein said *H. pylori* cytoplasmic polypeptide or a fragment thereof is an *H. pylori* polypeptide or a fragment thereof involved in genome replication, transcription, recombination and repair selected from the group consisting of SEQ ID NO: 183 and SEQ ID NO: 184.

THIS SHEET HAS BEEN
REPLACED





27. An isolated nucleic acid comprising a nucleotide sequence encoding an *H. pylori* secreted polypeptide or a fragment thereof, said nucleic acid selected from the group consisting of SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 20, SEQ ID NO: 25, SEQ ID NO: 31, SEQ ID NO: 32, SEQ ID NO: 45, SEQ ID NO: 46, SEQ ID NO: 53, SEQ ID NO: 64, SEQ ID NO: 67, SEQ ID NO: 70, SEQ ID NO: 77, SEQ ID NO: 78, SEQ ID NO: 81, SEQ ID NO: 82, SEQ ID NO: 90, SEQ ID NO: 95, and SEQ ID NO: 97, or a complement thereof.

28. An isolated nucleic acid comprising a nucleotide sequence encoding an *H. pylori* secreted polypeptide or a fragment thereof selected from the group consisting of SEQ ID NO: 100, SEQ ID NO: 101, SEQ ID NO: 107, SEQ ID NO: 109, SEQ ID NO: 117, SEQ ID NO: 122, SEQ ID NO: 128, SEQ ID NO: 129, SEQ ID NO: 142, SEQ ID NO: 143, SEQ ID NO: 150, SEQ ID NO: 161, SEQ ID NO: 164, SEQ ID NO: 167, SEQ ID NO: 174, SEQ ID NO: 175, SEQ ID NO: 178, SEQ ID NO: 179, SEQ ID NO: 187, SEQ ID NO: 192, and SEQ ID NO: 194.

29. An isolated nucleic acid comprising a nucleotide sequence encoding an *H. pylori* cellular polypeptide or a fragment thereof, said nucleic acid selected from the group consisting of SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 21, SEQ ID NO: 33, SEQ ID NO: 37, SEQ ID NO: 40, SEQ ID NO: 41, SEQ ID NO: 47, SEQ ID NO: 54, SEQ ID NO: 55, SEQ ID NO: 56, SEQ ID NO: 59, SEQ ID NO: 62, SEQ ID NO: 68, SEQ ID NO: 71, SEQ ID NO: 72, SEQ ID NO: 73, SEQ ID NO: 74, SEQ ID NO: 75, SEQ ID NO: 76, and SEQ ID NO: 96, or a complement thereof.

30. An isolated nucleic acid comprising a nucleotide sequence encoding an *H. pylori* cellular polypeptide or a fragment thereof, wherein said polypeptide is selected from the group consisting of SEQ ID NO: 112, SEQ ID NO: 113, SEQ ID NO: 118, SEQ ID NO: 130, SEQ ID NO: 134, SEQ ID NO: 137, SEQ ID NO: 138, SEQ ID NO: 144, SEQ ID NO: 151, SEQ ID NO: 152, SEQ ID NO: 153, SEQ ID NO: 156, SEQ ID NO: 159, SEQ ID NO: 165, SEQ ID NO: 168, SEQ ID NO: 169, SEQ ID NO: 170, SEQ ID NO: 171, SEQ ID NO: 172, SEQ ID NO: 173, and SEQ ID NO: 193.

31. A probe comprising a nucleotide sequence consisting of at least 8 nucleotides of a nucleotide sequence selected from the group consisting of SEQ ID NO: 1-SEQ ID NO: 97, or a complement thereof.

THIS SHEET HAS BEEN
REPLACED





32. A recombinant expression vector comprising the nucleic acid of any of claims 1, 2, 3, 4, 5, 6, 7, 14, 21, 24, 27, 28, 29 or 30 operably linked to a transcription regulatory element.
- 5 33. A cell comprising a recombinant expression vector of claim 32.
34. A method for producing an *H. pylori* polypeptide comprising culturing a cell of claim 33 under conditions that permit expression of the polypeptide.
- 10 35. The method of claim 34, further comprising purifying the polypeptide from the cell.
36. A method for detecting the presence of a *Helicobacter* nucleic acid in a sample comprising:
- 15 (a) contacting a sample with a nucleic acid of any of claims 6 or 31 so that a hybrid can form between the probe and a *Helicobacter* nucleic acid in the sample; and
- (b) detecting the hybrid formed in step (a), wherein detection of a hybrid indicates the presence of a *Helicobacter* nucleic acid in the sample.
- 20 37. An isolated *H. pylori* polypeptide comprising an amino acid sequence at least about 60% homologous to an *H. pylori* polypeptide selected from the group consisting of SEQ ID NO: 98-SEQ ID NO: 194.
- 25 38. An isolated *H. pylori* polypeptide which is encoded by a nucleic acid comprising a nucleotide sequence at least about 60% homologous to a nucleotide sequence selected from the group consisting of SEQ ID NO: 1-SEQ ID NO: 97.
- 30 39. The isolated *H. pylori* polypeptide of claim 28, wherein said polypeptide is encoded by a nucleotide sequence selected from the group consisting of SEQ ID NO: 1-SEQ ID NO: 97.
- 35 40. An isolated *H. pylori* polypeptide which is encoded by a nucleic acid which hybridizes under stringent hybridization conditions to a nucleic acid selected from the group consisting of SEQ ID NO: 1-SEQ ID NO: 97, or a complement thereof.

THIS SHEET HAS BEEN
REPLACED





- 324 -

41. An isolated *H. pylori* polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 97-SEQ ID NO: 194.

42. An isolated *H. pylori* cell envelope polypeptide or a fragment thereof,
5 wherein said polypeptide is selected from the group consisting of SEQ ID NO: 160, SEQ ID NO: 104, SEQ ID NO: 105, SEQ ID NO: 106, SEQ ID NO: 110, SEQ ID NO: 111, SEQ ID NO: 120, SEQ ID NO: 121, SEQ ID NO: 124, SEQ ID NO: 125, SEQ ID NO: 147, SEQ ID NO: 148, SEQ ID NO: 158, SEQ ID NO: 176, SEQ ID NO: 177, SEQ ID NO: 181, SEQ ID NO: 182, SEQ ID NO: 188, SEQ ID NO: 191, SEQ ID NO: 102,
10 SEQ ID NO: 108, SEQ ID NO: 123, SEQ ID NO: 133, SEQ ID NO: 139, SEQ ID NO: 149, SEQ ID NO: 119, SEQ ID NO: 126, SEQ ID NO: 127, SEQ ID NO: 162, SEQ ID NO: 163, SEQ ID NO: 145, SEQ ID NO: 146, SEQ ID NO: 114, SEQ ID NO: 115, SEQ ID NO: 116, SEQ ID NO: 140, SEQ ID NO: 141, SEQ ID NO: 135, SEQ ID NO: 136, SEQ ID NO: 98, SEQ ID NO: 99, SEQ ID NO: 103, SEQ ID NO: 131, SEQ ID NO: 132, SEQ ID NO: 157, SEQ ID NO: 166, and SEQ ID NO: 180.
15

43. The isolated polypeptide of claim 42, wherein said *H. pylori* cell envelope polypeptide or a fragment thereof is an *H. pylori* flagella-associated polypeptide or a fragment thereof comprising an amino acid sequence of SEQ ID
20 NO:160.

44. The isolated polypeptide of claim 43, wherein said *H. pylori* cell envelope polypeptide or a fragment thereof is an *H. pylori* inner membrane polypeptide or a fragment thereof selected from the group consisting of SEQ ID NO: 145, SEQ ID
25 NO: 146, SEQ ID NO: 114, SEQ ID NO: 115, SEQ ID NO: 116, SEQ ID NO: 140, SEQ ID NO: 141, SEQ ID NO: 135, SEQ ID NO: 136.

45. The isolated polypeptide of claim 44, wherein said *H. pylori* inner membrane polypeptide or a fragment thereof is an *H. pylori* polypeptide or a fragment
30 thereof involved in transport selected from the group consisting of SEQ ID NO: 145, SEQ ID NO: 146, SEQ ID NO: 114, SEQ ID NO: 115, SEQ ID NO: 116, SEQ ID NO: 140, and SEQ ID NO: 141.

46. The isolated polypeptide of claim 43, wherein said *H. pylori* cell envelope polypeptide or a fragment thereof is an *H. pylori* outer membrane polypeptide or a fragment thereof selected from the group consisting of SEQ ID NO: 104, SEQ ID
35

THIS SHEET HAS BEEN
REPLACED





NO: 105, SEQ ID NO: 106, SEQ ID NO: 110, SEQ ID NO: 111, SEQ ID NO: 120, SEQ ID NO: 121, SEQ ID NO: 124, SEQ ID NO: 125, SEQ ID NO: 147, SEQ ID NO: 148, SEQ ID NO: 158, SEQ ID NO: 176, SEQ ID NO: 177, SEQ ID NO: 181, SEQ ID NO: 182, SEQ ID NO: 188, SEQ ID NO: 191, SEQ ID NO: 102, SEQ ID NO: 108, SEQ ID NO: 123, SEQ ID NO: 133, SEQ ID NO: 139, SEQ ID NO: 149, SEQ ID NO: 119, SEQ ID NO: 126, SEQ ID NO: 127, SEQ ID NO: 162, and SEQ ID NO: 163.

47. The isolated polypeptide of claim 46, wherein said *H. pylori* outer membrane polypeptide or a fragment thereof is an *H. pylori* polypeptide having a terminal phenylalanine residue or a fragment thereof selected from the group consisting of SEQ ID NO: 104, SEQ ID NO: 105, SEQ ID NO: 106, SEQ ID NO: 108, SEQ ID NO: 110, SEQ ID NO: 111, SEQ ID NO: 120, SEQ ID NO: 121, SEQ ID NO: 123, SEQ ID NO: 124, SEQ ID NO: 125, SEQ ID NO: 133, SEQ ID NO: 139, SEQ ID NO: 147, SEQ ID NO: 148, SEQ ID NO: 149, SEQ ID NO: 158, SEQ ID NO: 176, SEQ ID NO: 177, SEQ ID NO: 181, SEQ ID NO: 182, SEQ ID NO: 188, and SEQ ID NO: 191.

48. The isolated polypeptide of claim 47, wherein said *H. pylori* outer membrane polypeptide or a fragment thereof is an *H. pylori* polypeptide having a terminal phenylalanine residue and a C-terminal tyrosine cluster or a fragment thereof selected from the group consisting of SEQ ID NO: 108, SEQ ID NO: 123, SEQ ID NO: 133, SEQ ID NO: 139, and SEQ ID NO: 149.

49. An isolated *H. pylori* cell envelope polypeptide or a fragment thereof, wherein said polypeptide is encoded by a nucleic acid selected from the group consisting of SEQ ID NO: 63, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 50, SEQ ID NO: 51, SEQ ID NO: 61, SEQ ID NO: 79, SEQ ID NO: 80, SEQ ID NO: 84, SEQ ID NO: 85, SEQ ID NO: 91, SEQ ID NO: 94, SEQ ID NO: 5, SEQ ID NO: 11, SEQ ID NO: 26, SEQ ID NO: 36, SEQ ID NO: 42, SEQ ID NO: 52, SEQ ID NO: 22, SEQ ID NO: 29, SEQ ID NO: 30, SEQ ID NO: 65, SEQ ID NO: 66, SEQ ID NO: 48, SEQ ID NO: 49, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 43, SEQ ID NO: 44, SEQ ID NO: 38, SEQ ID NO: 39, SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 6, SEQ ID NO: 34, SEQ ID NO: 35, SEQ ID NO: 60, SEQ ID NO: 69, and SEQ ID NO: 83.

35

THIS SHEET HAS BEEN
REPLACED





50. The isolated polypeptide of claim 49, wherein said *H. pylori* cell envelope polypeptide or a fragment thereof is an *H. pylori* flagella-associated polypeptide or a fragment thereof encoded by a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 63.

5

51. The isolated polypeptide of claim 49, wherein said *H. pylori* cell envelope polypeptide or a fragment thereof is an *H. pylori* inner membrane polypeptide or a fragment thereof encoded by a nucleic acid selected from the group consisting of SEQ ID NO: 48, SEQ ID NO: 49, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19,
10 SEQ ID NO: 43, SEQ ID NO: 44, SEQ ID NO: 38, SEQ ID NO: 39.

52. The isolated polypeptide of claim 51, wherein said *H. pylori* inner membrane polypeptide or a fragment thereof is an *H. pylori* polypeptide or a fragment thereof involved in transport encoded by a nucleic acid selected from the group
15 consisting of SEQ ID NO: 48, SEQ ID NO: 49, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 43, and SEQ ID NO: 44.

53. The isolated polypeptide of claim 49, wherein said *H. pylori* cell envelope polypeptide or a fragment thereof is an *H. pylori* outer membrane polypeptide
20 or a fragment thereof encoded by a nucleic acid selected from the group consisting of SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 50, SEQ ID NO: 51, SEQ ID NO: 61, SEQ ID NO: 79, SEQ ID NO: 80, SEQ ID NO: 84, SEQ ID NO: 85, SEQ ID NO: 91, SEQ ID NO: 94, SEQ ID NO: 5, SEQ ID NO: 11, SEQ ID
25 NO: 26, SEQ ID NO: 36, SEQ ID NO: 42, SEQ ID NO: 52, SEQ ID NO: 22, SEQ ID NO: 29, SEQ ID NO: 30, SEQ ID NO: 65, and SEQ ID NO: 66.

54. The isolated polypeptide of claim 53, wherein said *H. pylori* outer membrane polypeptide or a fragment thereof is an *H. pylori* polypeptide having a
30 terminal phenylalanine residue or a fragment thereof encoded by a nucleic acid selected from the group consisting of SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 36, SEQ ID NO: 42, SEQ ID NO: 50, SEQ ID NO: 51, SEQ ID NO: 52, SEQ ID NO: 61, SEQ ID NO: 79, SEQ ID
35 NO: 80, SEQ ID NO: 84, SEQ ID NO: 85, SEQ ID NO: 91, and SEQ ID NO: 94.

THIS SHEET HAS BEEN
REPLACED





55. The isolated polypeptide of claim 54, wherein said *H. pylori* outer membrane polypeptide or a fragment thereof is an *H. pylori* polypeptide having a terminal phenylalanine residue and a C-terminal tyrosine cluster or a fragment thereof encoded by a nucleic acid selected from the group consisting of SEQ ID NO: 11, SEQ
5 ID NO: 26, SEQ ID NO: 36, SEQ ID NO: 42, and SEQ ID NO: 52.

56. An isolated *H. pylori* cytoplasmic polypeptide or a fragment thereof, wherein said polypeptide is selected from the group consisting of SEQ ID NO: 154, SEQ ID NO: 155, SEQ ID NO: 183, SEQ ID NO: 184, SEQ ID NO: 185, SEQ ID NO:
10 186, SEQ ID NO: 189, and SEQ ID NO: 190.

57. The isolated polypeptide of claim 56, wherein said *H. pylori* cytoplasmic polypeptide or a fragment thereof is an *H. pylori* polypeptide or a fragment thereof involved in mRNA translation selected from the group consisting of SEQ ID NO: 154
15 and SEQ ID NO: 155.

58. The isolated polypeptide of claim 56, wherein said *H. pylori* cytoplasmic polypeptide or a fragment thereof is an *H. pylori* polypeptide or a fragment thereof involved in genome replication, transcription, recombination and repair selected from
20 the group consisting of SEQ ID NO: 183 and SEQ ID NO: 184.

59. An isolated *H. pylori* cytoplasmic polypeptide or a fragment thereof, wherein said polypeptide is encoded by a nucleic acid selected from the group consisting of SEQ ID NO: 57, SEQ ID NO: 58, SEQ ID NO: 86, SEQ ID NO: 87, SEQ ID NO: 88,
25 SEQ ID NO: 89, SEQ ID NO: 92, and SEQ ID NO: 93.

60. The isolated polypeptide of claim 59, wherein said *H. pylori* cytoplasmic polypeptide or a fragment thereof is an *H. pylori* polypeptide or a fragment thereof involved in mRNA translation, said polypeptide encoded by a nucleic acid selected from
30 the group consisting of SEQ ID NO: 57 and SEQ ID NO: 58.

61. The isolated polypeptide of claim 59, wherein said *H. pylori* cytoplasmic polypeptide or a fragment thereof is an *H. pylori* polypeptide or a fragment thereof involved in genome replication, transcription, recombination and repair, said
35 polypeptide encoded by a nucleic acid selected from the group consisting of SEQ ID NO: 86 and SEQ ID NO: 87.

THIS SHEET HAS BEEN
REPLACED





62. An isolated *H. pylori* cellular polypeptide or a fragment thereof, wherein said polypeptide is selected from the group consisting of SEQ ID NO: 112, SEQ ID NO: 113, SEQ ID NO: 118, SEQ ID NO: 130, SEQ ID NO: 134, SEQ ID NO: 137, SEQ ID NO: 138, SEQ ID NO: 144, SEQ ID NO: 151, SEQ ID NO: 152, SEQ ID NO: 153, SEQ ID NO: 156, SEQ ID NO: 159, SEQ ID NO: 165, SEQ ID NO: 168, SEQ ID NO: 169, SEQ ID NO: 170, SEQ ID NO: 171, SEQ ID NO: 172, SEQ ID NO: 173, and SEQ ID NO: 193.

63. An isolated *H. pylori* cellular polypeptide or a fragment thereof, wherein said polypeptide is encoded by a nucleic acid selected from the group consisting of SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 21, SEQ ID NO: 33, SEQ ID NO: 37, SEQ ID NO: 40, SEQ ID NO: 41, SEQ ID NO: 47, SEQ ID NO: 54, SEQ ID NO: 55, SEQ ID NO: 56, SEQ ID NO: 59, SEQ ID NO: 62, SEQ ID NO: 68, SEQ ID NO: 71, SEQ ID NO: 72, SEQ ID NO: 73, SEQ ID NO: 74, SEQ ID NO: 75, SEQ ID NO: 76, and SEQ ID NO: 96.

64. An isolated *H. pylori* secreted polypeptide or a fragment thereof, wherein said polypeptide is selected from the group consisting of SEQ ID NO: 100, SEQ ID NO: 101, SEQ ID NO: 107, SEQ ID NO: 109, SEQ ID NO: 117, SEQ ID NO: 122, SEQ ID NO: 128, SEQ ID NO: 129, SEQ ID NO: 142, SEQ ID NO: 143, SEQ ID NO: 150, SEQ ID NO: 161, SEQ ID NO: 164, SEQ ID NO: 167, SEQ ID NO: 174, SEQ ID NO: 175, SEQ ID NO: 178, SEQ ID NO: 179, SEQ ID NO: 187, SEQ ID NO: 192, and SEQ ID NO: 194.

65. An isolated *H. pylori* secreted polypeptide or a fragment thereof, wherein said polypeptide is encoded by a nucleic acid selected from the group consisting of SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 20, SEQ ID NO: 25, SEQ ID NO: 31, SEQ ID NO: 32, SEQ ID NO: 45, SEQ ID NO: 46, SEQ ID NO: 53, SEQ ID NO: 64, SEQ ID NO: 67, SEQ ID NO: 70, SEQ ID NO: 77, SEQ ID NO: 78, SEQ ID NO: 81, SEQ ID NO: 82, SEQ ID NO: 90, SEQ ID NO: 95, and SEQ ID NO: 97.

66. A fusion protein comprising an *H. pylori* polypeptide which comprises an amino acid sequence selected from the group consisting of SEQ ID NO: 98-SEQ ID NO: 194 operatively linked to a non-*H. pylori* polypeptide.

THIS SHEET HAS BEEN
REPLACED





67. A vaccine formulation for prophylactic or therapeutic treatment of an *H. pylori* infection comprising an effective amount of at least one isolated nucleic acid of any of claims 1, 2, 3, 4, 5, 6, 7, 14, 21, 24, 27, 28, 29 or 30.

5

68. A vaccine formulation for prophylactic or therapeutic treatment of an *H. pylori* infection comprising an effective amount of at least one *H. pylori* polypeptide or a fragment thereof of any of claims 37, 38, 40, 41, 42, 49, 56, 59, 62, 63, 64 or 65.

10

69. A vaccine formulation of claim 67, further comprising a pharmaceutically acceptable carrier.

70. A vaccine formulation of claim 68, further comprising a pharmaceutically acceptable carrier.

15

71. A vaccine formulation of claim 69, wherein the pharmaceutically acceptable carrier comprises an adjuvant.

72. A vaccine formulation of claim 70, wherein the pharmaceutically acceptable carrier comprises an adjuvant.

20

73. A vaccine formulation of claim 69, wherein the pharmaceutically acceptable carrier comprises a delivery system.

74. A vaccine formulation of claim 70, wherein the pharmaceutically acceptable carrier comprises a delivery system.

25

75. A vaccine formulation of claim 73, wherein the delivery system comprises a live vector.

30

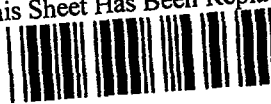
76. A vaccine formulation of claim 74, wherein the delivery system comprises a live vector.

77. A vaccine formulation of claim 75, wherein the live vector is a bacteria or a virus.

35

THIS SHEET HAS BEEN
REPLACED





78. A vaccine formulation of claim 76, wherein the live vector is a bacteria or a virus.

79. A vaccine formulation of claim 73, wherein the pharmaceutically acceptable carrier further comprises an adjuvant.

80. A vaccine formulation of claim 74, wherein the pharmaceutically acceptable carrier further comprises an adjuvant.

81. A vaccine formulation for prophylactic or therapeutic treatment of an *H. pylori* infection comprising an effective amount of at least one isolated nucleic acid encoding an *H. pylori* outer membrane polypeptide or a fragment thereof, said nucleic acid selected from the group consisting of SEQ ID NO: 28, SEQ ID NO: 50, SEQ ID NO: 24, SEQ ID NO: 11, SEQ ID NO: 52, SEQ ID NO: 42 and SEQ ID NO: 79.

82. The vaccine formulation of claim 81, wherein said nucleic acid comprises a nucleotide sequence of SEQ ID NO: 52.

83. A vaccine formulation for prophylactic or therapeutic treatment of an *H. pylori* infection comprising an effective amount of at least one *H. pylori* outer membrane polypeptide or a fragment thereof, said polypeptide selected from the group consisting of SEQ ID NO: 125, SEQ ID NO: 147, SEQ ID NO: 121, SEQ ID NO: 108, SEQ ID NO: 149, SEQ ID NO: 139 and SEQ ID NO: 176.

84. The vaccine formulation of claim 81, wherein said polypeptide comprises an amino acid sequence of SEQ ID NO: 149.

85. A vaccine formulation of claims 81 or 83, further comprising a pharmaceutically acceptable carrier.

86. A vaccine formulation of claim 85, wherein the pharmaceutically acceptable carrier comprises an adjuvant.

87. A vaccine formulation of claim 85, wherein the pharmaceutically acceptable carrier comprises a delivery system.

THIS SHEET HAS BEEN
REPLACED





88. A vaccine formulation of claim 87, wherein the delivery system comprises a live vector.
89. A vaccine formulation of claim 88, wherein the live vector is a bacteria or
5 a virus.
90. A vaccine formulation of claim 86, wherein the pharmaceutically acceptable carrier further comprises an adjuvant.
91. A method of treating or reducing a risk of *H. pylori* infection in a subject
10 comprising administering to a subject a vaccine formulation of claim 67, such that treatment or reduction of risk of *H. pylori* infection occurs.
92. A method of treating or reducing a risk of *H. pylori* infection in a subject
15 comprising administering to a subject a vaccine formulation of claim 68, such that treatment or reduction of risk of *H. pylori* infection occurs.
93. A method of treating or reducing a risk of *H. pylori* infection in a subject
20 comprising administering to a subject a vaccine formulation of claim 81, such that treatment or reduction of risk of *H. pylori* infection occurs.
94. A method of treating or reducing a risk of *H. pylori* infection in a subject
25 comprising administering to a subject a vaccine formulation of claim 83, such that treatment or reduction of risk of *H. pylori* infection occurs.
95. A method of producing a vaccine formulation comprising: combining at
least one isolated *H. pylori* polypeptide or a fragment thereof selected from the group
consisting of SEQ ID NO: 98-SEQ ID NO: 194 with a pharmaceutically acceptable
carrier to thereby form a vaccine formulation.
96. A method of producing a vaccine formulation comprising:
30 (a) providing at least one isolated *H. pylori* polypeptide or a fragment thereof selected from the group consisting of SEQ ID NO: 98-SEQ ID NO: 194; and
(b) combining at least one said isolated *H. pylori* polypeptide or a
35 fragment thereof with a pharmaceutically acceptable carrier to thereby form a vaccine formulation.

THIS SHEET HAS BEEN
REPLACED





97. A method of producing a vaccine formulation comprising:

- 5 (a) culturing a cell under condition that permit expression of an *H. pylori* polypeptide or a fragment thereof selected from the group consisting of SEQ ID NO: 98-SEQ ID NO: 194;
- (b) isolating said *H. pylori* polypeptide or a fragment thereof from said cell; and
- 10 (c) combining at least one said isolated *H. pylori* polypeptide or a fragment thereof with a pharmaceutically acceptable carrier to thereby form a vaccine formulation.

98. A chimeric *H. pylori* polypeptide comprising at least two *H. pylori* polypeptides or fragments thereof, wherein said polypeptides are encoded by nucleic acid sequences selected from the group consisting of SEQ ID NO:1-SEQ ID NO:97.

15

99. A chimeric *H. pylori* polypeptide comprising at least two *H. pylori* polypeptides or fragments thereof, wherein said polypeptides are selected from the group consisting of SEQ ID NO:98-SEQ ID NO:194.

20

THIS SHEET HAS BEEN
REPLACED

