Analysis of a *Helicobacter pylori* operon incorporating flagellar export genes

A thesis presented in partial fulfilment

of the requirements for the degree of

Doctor of Philosophy

in the Institute of Molecular BioSciences

at Massey University, New Zealand

Steffen Porwollik

May 1999

Abstract

Motility of *Helicobacter* species has been shown to be essential for successful colonization of the host. Previous studies indicated that the regulation of flagellar biosynthesis in the human gastric pathogen *Helicobacter pylori* differs from the suggested model for Gram-negative Enterobacteriaeceae.

In this study, the organization of two *H. pylori* genes involved in export of flagellar structural proteins was investigated. A 7 kb fragment of the *H. pylori* 17874 genome was cloned. Sequence determination and analysis revealed a putative operon comprising an ORF of unknown function (ORF03), and genes for the isoleucyl-tRNA synthetase (*ileS*), an *Agrobacterium tumefaciens* VirB11 homolog (*virB11*), an ATPase involved in flagellum-specific protein export (*fliI*), a presumptive flagellar export channel component (*fliQ*), and a homolog of an enzyme necessary for cell wall biosynthesis (*murB*). The genetic organization of this region was found to be conserved in a panel of clinical *H. pylori* isolates, and in *H. pylori* 915 and SS1. The locus was also identified in the genome sequences of the *H. pylori* strains J99 and 26695.

Cotranscription of ORF03, *ileS, virB11, fliI, fliQ* and *murB* was demonstrated by RT-PCR. Primer extension experiments identified the major transcription start site, which coincided with the A residue of the initiation codon of ORF03. A promoter element was inferred that resembled the *E. coli* σ^{70} consensus sequence. In addition, a minor transcription start site was detected upstream from *ileS*.

Non-polar mutation of *virB11*, *fliI* and *fliQ* was generated by an allele replacement strategy. Engineered *H. pylori fliI* and *fliQ* mutant strains were completely aflagellate and nonmotile, whereas a *virB11* mutant still produced flagella and displayed slightly greater motility. The *fliI* and *fliQ* mutant strains produced severely reduced levels of flagellin and the hook protein FlgE, although reduction was less stable in the *fliI* mutant. Production of OMP4, a member of the outer membrane protein family identified in *H. pylori* 26695, was diminished in both the *virB11* and the *fliI* mutant. This suggested related functions of the putative virulence factor transport protein (VirB11) and the flagellar export component (FliI).

Acknowledgments

I would like to express my sincere gratitude to my chief supervisor Dr Paul O'Toole for providing much scientific guidance and encouragement throughout every stage of this project. I am thankful for the many scientific discussions, and for immediate feedback concerning all the report and paper and thesis versions. Many thanks also for introducing me to real-life competetive science, and for giving me the opportunity to attend conferences both in and outside of New Zealand. And for modifying my American accent with this Irish gibberish.

I am also grateful to my second supervisor, Dr Jan Schmid, for his contribution of ideas. Thanks for always encouraging a slightly less conventional approach to problems in molecular biology, which very often helped a lot. I am glad that someone with your microbiological background took part in the effort that lead to this thesis.

Doug Hopcroft from The Horticultural and Food Research Institute of New Zealand, Palmerston North, provided the much needed expertise for the electron microscopy experiments. Thank you for your time and effort.

A special thanks goes to the people from the Helipad crew, without whom I would never have pulled through. Thanks, Tash, for being the perfect lab organizer & Kiwi slang teacher, and always a good person to talk to when experiments went down the drain. It was great having you around. To Millis, cheers for some very valuable and enjoyable dialogue I had missed for too long. To Kirsty, Basil, James & Ross, thanks for making lab work a lot funnier and more pleasant than it could be.

I thankfully acknowledge financial support from the award of a Massey University Doctoral Scholarship, and the Molecular Genetics Research Scholarship of the former Department of Microbiology and Genetics. Travel funding was gratefully received from the Massey University Graduate Research Fund, the Department of Microbiology and Genetics, and the Institute of Molecular BioSciences at Massey University.

Many people have assisted me in various ways throughout the course of this work, by providing support and words of encouragment, or just by listening when I needed to talk. Thank you to all these people, especially Martin, Andrew & Cheyenne. Thanks to my friends in Palmy. I'm glad I met you.

Danke, Katja, für all die Briefe. Und Ralf, Carola, Eike, Lutz & Betty, danke dafür, daß ich mich in Berlin noch immer zu Hause fühl'. Ich hab Euch vermißt.

Related Publications

Some of the material presented in this thesis has been published.

Porwollik, S., Noonan, B. & O'Toole, P.W. (1999). Molecular characterization of a flagellar export locus of *Helicobacter pylori*. *Infect Immun* 67:2060-2070.

Porwollik, S. & O'Toole, P.W. (1998). Molecular characterization of a flagellar export locus of *Helicobacter pylori*. *Gut* 43 (Suppl 2): A02/31

Abbreviations

A+T content of deoxyadenylate and deoxythymidylate in DNA
aa amino acid
ab antibody
ABI Applied Biosystems
AMV avian myeloblastosis virus
Ap ampicillin
APS ammonium persulphate
ATP adenosine triphosphate
BBH basal body-hook complex
BCIP 5-bromo-4-chloro3-indolyl phosphate
BLAST basic local alignment search tool
BSA bovine serum albumine
cAMP cyclic adenosine monophosphate
CAP catabolite gene activator protein
CBA columbia base agar
cDNA complementary DNA
Cm chloramphenicol
colE1 colicin E1
CRP cyclic adenosine monophosphate receptor protein
CSPD disodium 3-(4-methoxyspiro{1,2-dioxetane-3,2'-(5'-chloro)tricyclo[3.3.1.1.]decan}-4-yl)phenyl phosphate
CTP cytidine triphosphate
dATP 2' deoxyadenosine triphosphate
DEPC diethylpyrocarbonate
dGTP 2' deoxyguanosine triphosphate
DIG digoxigenin
DNA deoxyribonucleic acid
DNase deoxyribonuclease
dNTP deoxynucleoside triphosphate
dTTP 2' deoxythymidine triphosphate
dUTP 2' deoxyuridine triphosphate
EDTA ethylenediaminetetraacetic acid
EPB electroporation buffer
f1 bacteriophage f1
F_{ab} variable sequence fragment of immunoglobulin

FSB final sample buffer
FP forward primer
G+C content of deoxyguanylate and deoxycytidylate in DNA
GSP general secretory pathway
GTP guanosine triphosphate
IL interleukin
IPTG isopropyl-ß-D-galactoside
Kan kanamycin
LB Luria-Bertani broth
LBA Luria-Bertani agar
LPS lipopolysaccharide
MALT mucosa-associated lymphoid tissue
mcl monoclonal
MCS multicloning site
MOPS 3-(N-morpholino) propanesulphonic acid
n/a not applicable
NBT 4-nitro blue tetrazolium chloride
NCBI National Center for Biotechnology Information
Neo neomycin
OD optical density
ORF open reading frame
ori origin of replication
PBS phosphate-buffered saline
pcl polyclonal
PCR polymerase chain reaction
PEG polyethylene glycol
pI isoelectric point
PIR protein information resource
^{r} resistant
RNA ribonucleic acid
RNase ribonuclease
RP reverse primer
rpm revolutions per minute
RT room temperature
RT-PCR polymerase chain reaction involving an initial reverse transcriptase step
SDS sodium dodecyl sulphate
SM-TBS skim milk powder in Tris-buffered saline

SV40 simian virus 40
TB terrific broth
TBS Tris-buffered saline
TEMED NNN'N' tetramethylethylenediamine
TIGR The Institute for Genomic Research
T_m melting temperature
TNF tumor necrosis factor
Tris tris(hydroxymethyl)methylamine
TSB tryptic soy broth
U unit
UTP uridine triphosphate
Vol volume
w/v weight per volume
X-gal 5-bromo-4-chloro-3-indolyl-ß-D-galactoside

In addition, the conventional one-letter codes for amino acids, deoxyribonucleosides and ribocucleosides were applied:

amino acids: G, A, V, L, I, P, F, Y, W, S, T, C, M, N, Q, D, E, K, R, H for glycine, alanine, valine, leucine, isoleucine, proline, phenylalanine, tyrosine, tryptophan, serine, threonine, cysteine, methionine, asparagine, glutamine, aspartate, glutamate, lysine, arginine and histidine, respectively
deoxyribonucleosides: A, C, G, T for deoxyadenylate, deoxycytidylate, deoxyguanylate and deoxythymidylate, respectively
ribonucleosides: A, C, G, U for adenylate, cytidylate, guanylate and uridylate, respectively

Table of Contents

List of Figures

List of Tables

1. INTRODUCTION

1.1 A bacterium that causes gastric disease: *Helicobacter pylori*

1.1.1 History and general features

Until the mid 1980's gastritis and peptic ulcer disease were generally thought to be caused mainly by stress, bad diet and smoking. The discovery of the gastric bacterium *H. pylori* and subsequent investigations have since revolutionized the view of these and other disorders of the human gastric tract. It is now publicly acknowledged that *H. pylori* represents the main etiological factor of these diseases.

In the first half of the 20th century, small spiral organisms had been frequently observed in human stomachs (Freedburg, 1940; Palmer, 1954). Since it appeared impossible to isolate the organism, these observations never raised much interest in the scientific community. It was not until 1982 that attempts to culture the bacterium succeeded after an accidentally prolonged incubation time. B.J. Marshall and J.R. Warren were the first to isolate the observed bacterial cells from inflamed gastric tissue samples of patients at the Royal Perth Hospital, Australia (Marshall & Warren, 1984). The bacteria were abundantly present in mucosal biopsies of those patients (Fig. 1.1). Initially placed in the genus *Campylobacter,* phylogenetic analysis of the bacterium subsequently revealed that it was not in fact closely related to *Campylobacter* (Romaniuk *et al.*, 1987). Therefore, the genus *Helicobacter* was created. This genus is steadily expanding, and contains at least 17 named species to date (Jalava *et al.*, 1997), isolated from a variety of mammals.

The species *H. pylori* has been found in humans, domestic cats (Handt *et al.*, 1994) and non-human primates (Dubois *et al.*, 1994). It is a slow growing, Gram-negative, microaerophilic bacterium that is highly motile by virtue of up to six lophotrichate flagella (Dunn *et al.*, 1997). In its bacillary form, *H. pylori* is 2.5-5 μm long, 0.5-1 μm wide and spiral shaped (Stark, 1995). However, after several days of *in vitro* culture, it converts to coccoid forms, the significance of which is still controversial. It has been assumed to be either a degenerative form or a dormant stage, and some experiments suggest these forms to be viable (Benaissa *et al.*, 1996). Coccoid forms of *H. pylori* coexist with bacillary *H. pylori* cells in the gastric mucosa (Chan *et al.*, 1994). However, they display changes in DNA, RNA, membrane potential and protein synthesis consistent with cell death (Kusters *et al.*, 1997).

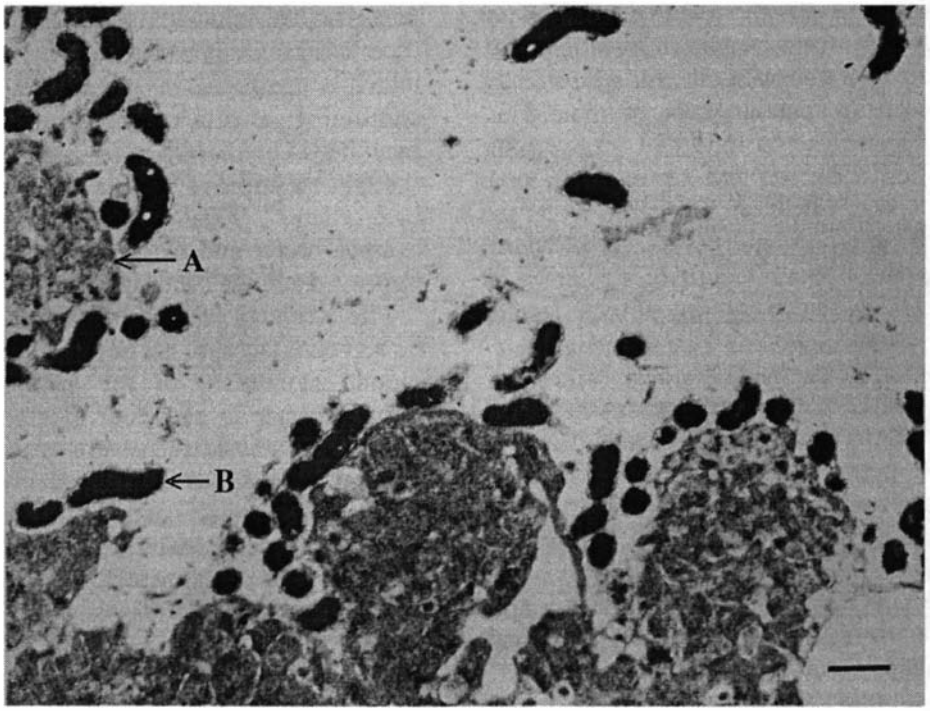


Figure 1.1. Electron micrograph of a mucosal biopsy with active chronic gastritis.

The mucus - secreting epithelial cells (A) are surrounded by *Helicobacter pylori* cells (B). Bar, 1 μm.

Figure reproduced from Marshall & Warren, 1984.

The bacterium resides in the mucus layer overlaying the gastric epithelium of the human stomach (Blaser, 1993). As with many human pathogens, the optimal growth temperature of *H. pylori* is about 37°C. Despite its ability to infect the lining of the stomach, *H. pylori* is acid-sensitive and has a pH growth range of 4.5-8.0 (Stark, 1995). Thus it must be concluded, that *in vivo* the bacterium is protected in some way from the acidic conditions present in the stomach. It obviously has to withstand a certain exposure to acid in the body of the stomach before entering the mucus layer, where a nearly neutral pH prevails (Bahari *et al.*, 1982).

Around 30% of adults in developed countries are infected with *H. pylori*, and the prevalence of infection approaches 80-90% in many developing countries (Dunn *et al.*, 1997). Most infected people do not suffer from any clinical symptoms. However, there is overwhelming evidence that *H. pylori* represents the causative agent of gastritis and ulcer disease (Blaser, 1993). Voluntary ingestion of the bacterium by two human volunteers gave rise to acute or chronic gastritis (Marshall *et al.*, 1985; Morris & Nicholson, 1987). Eradication of *H. pylori* cleared gastritis (Valle *et al.*, 1991). In an animal model, gnotobiotic piglets developed gastritis after experimental challenge with *H. pylori* (Eaton *et al.*, 1989). Eradication of *H. pylori* using antimicrobial treatment also healed duodenal and gastric ulcers (Hosking *et al.*, 1994; Lam *et al.*, 1997), and reduced ulcer recurrences (Axon *et al.*, 1997).

The natural route of transmission of the bacterium has not yet been defined. Most authors favour the gastro-oral route (Axon, 1995; Figura, 1996), but numerous studies are also conducted to investigate possible transmission of *H. pylori* by the faecal-oral or the oral-oral route (reviewed in Megraud, 1995; Feldman *et al.*, 1997).

Although *H. pylori* produces high amounts of catalase, oxidase and superoxide dismutase, it cannot tolerate higher oxygen concentrations (Stark, 1995). After laboratory passage, strains become sufficiently aerotolerant to grow in a CO_2 enriched atmosphere of 5-10% (Stark, 1995).

1.1.2 *H. pylori* pathogenesis

1.1.2.1 Pathogenic factors

H. pylori actively replicates in the gastric mucus layer and may cause disease decades after its acquisition (Blaser & Parsonnet, 1994). A number of both modified

common bacterial properties as well as unique virulence factors enable the bacterium to colonize and persist in its ecological niche and to eventually cause disease. Some of those factors have been studied extensively, while others have only recently been characterized, and many are probably still unknown.

The *H. pylori* **urease** enzyme is an indispensable colonization factor for the bacterium. The enzyme catalyzes the following chemical reaction:

$C = O(NH_2)_2 + H^+ + 2H_2O$ -----> $HCO_3^- + 2NH_4^+$.

Ammonium ions are generated by consumption of H^+ ions. The pH-elevating urease-catalyzed hydrolysis of urea is thought to protect the bacterial cells from the gastric acid. The reaction seems to be essential for survival of the bacterium in the human stomach. Urease-negative *H. pylori* mutants were unable to colonize gnotobiotic piglets (Eaton *et al.*, 1991), nude mice (Tsuda *et al.*, 1994) and *Cynomolgus* monkeys (Takahashi *et al.*, 1993). *H. pylori* produces an extraordinarily high amount of urease that has been estimated to represent 5.9% of the total cell protein (Hu & Mobley, 1990). A proportion of the urease is surface exposed (Hawtin *et al.*, 1990). The process by which it becomes associated with the bacterial surface is currently under investigation. A mechanism involving autolysis of cells with adherence of urease to remaining viable *H. pylori* cells has been suggested (Phadnis *et al.*, 1996). However, some data favour a selective release process rather than non-specific autolysis (Vanet & Labigne, 1998). Research is currently also focussing on the intracellularly located urease, the activity of which was found to be upregulated by low external pH. Intracellular *H. pylori* urease was shown to stimulate protein synthesis under acidic conditions (Scott *et al.*, 1998), and may take part in prevention of acidic death of the bacterium. However, the acid-protective role of the enzyme is apparently not its only important function during the colonization process. Urease-negative mutants failed to colonize achlorhydric piglets with a constant gastric pH of 7.0 ± 1.1 (Eaton & Krakowka, 1994), suggesting additional essential functions of the enzyme.

A second obvious virulence factor is the **motility** of *H. pylori*. It might enable the bacterium to withstand peristaltic movements in the gut. Motility is one of the few traits experimentally proven as necessary for successful colonization in an animal model, exemplified by gnotobiotic piglets (Eaton *et al.*, 1992; Eaton *et al.*, 1996). In addition, non-motile *Helicobacter mustelae* cells failed to colonize ferrets (Andrutis *et al.*, 1997).

H. pylori has been shown to adhere to the gastric epithelium, especially at intercellular junctions (Hazell *et al.*, 1986). Its high degree of species-specific tissue

tropism suggests this attachment (mediated by specific **adhesins**) to be an important colonization factor (Logan, 1996). It may allow *H. pylori* to directly target toxins to the host epithelium. A close association between the number of adherent *H. pylori* and the severity of gastritis was suggested (Rudmann *et al.*, 1992). Direct contact of *H. pylori* with gastric epithelial cells gives rise to elevated cytokine production in these cells (Rieder *et al.*, 1997), thereby causing inflammation. Only a small number of *H. pylori* cell adhesins have been identified so far. Among those are proteins binding the Lewis b blood group antigen (Ilver *et al.*, 1998), and laminin (Valkonen *et al.*, 1997). A mucin-binding neutrophil-activating protein has also been characterized (Namavar *et al.*, 1998).

H. pylori **lipopolysaccharide** (LPS) possesses low immunological activity, thereby minimizing the local inflammatory response and possibly contributing to the persistence of the infection. The reduced immunological response of the host was linked to relative underphosphorylation and underacylation of the lipid A component of *H. pylori* LPS (Moran *et al.*, 1997). Moreover, it has been shown that the O-polysaccharide chains of some *H. pylori* strains display structural similarities to Lewis x and Lewis y blood group antigens (Aspinall & Monteiro, 1996; Aspinall *et al.*, 1996). Their expression might correspond to the host Lewis phenotype (Wirth *et al.*, 1997), and is affected by phase variation (Appelmelk *et al.*, 1998). They are suspected to induce autoantibodies responsible for the pathogenesis of gastritis (Appelmelk *et al.*, 1996).

A significant finding in *H. pylori* pathogenesis research has been the characterization of a **cytotoxin VacA**. Four groups almost simultaneously reported the sequence of the cytotoxin gene (Telford *et al.*, 1994; Cover *et al.*, 1994; Phadnis *et al.*, 1994; Schmitt & Haas, 1994). This cytotoxin induced vacuolation in a variety of eukaryotic cells *in vitro*. Although not all *H. pylori* strains produce this vacuolating cytotoxin, essentially all strains possess the *vacA* gene. In the highly diverse middle region of the gene, toxigenic strains form a family, in which the sequences are highly related to one another (> 90% identity), and so do the non-toxigenic strains (Atherton *et al.*, 1995). Diversity can also be found in the N-terminal region of the gene, within the signal sequence. This diversity influences the amount of toxin produced (Atherton *et al.*, 1995). However, lack of VacA does not affect the ability of *H. pylori* cells to colonize and persist in the piglet stomach (Eaton *et al.*, 1997). VacA enhances the mucosal damage (Ghiara *et al.*, 1995), and possibly promotes progression of gastritis to peptic ulceration (Atherton *et al.*, 1995). Illustrating its putative role in ulcerogenesis, oral administration of the purified cytotoxin to mice caused ulceration (Telford *et al.*, 1994).

One of the strongest immunogenic compounds of *H. pylori* is the 132 kDa **cytotoxin associated protein CagA** (Tummuru *et al.*, 1993), which is only produced by 60-80% of all *H. pylori* strains. In contrast to *vacA*, *cagA* is not present in every *H. pylori* strain. Strains that do not produce CagA lack the gene. This gene is located within a 37 kb chromosomal region, termed *cag* pathogenicity island (see below) that is apparently absent in *cagA*$^-$ strains (Censini *et al.*, 1996; Akopyants *et al.*, 1998).

It appears that 40-60% of patients with gastritis alone are infected with *cagA*$^+$ strains, compared to 80-100% of patients with duodenal ulceration (Cover *et al.*, 1995). Apparently, *cagA*$^+$ strains are more virulent than *cagA*$^-$ strains. Infection with CagA producing *H. pylori* strains may enhance the risk of developing gastric cancer (Parsonnet *et al.*, 1997), possibly by means of an increased cell proliferation of gastric epithelial cells without elevated apoptosis (Peek *et al.*, 1997). Interestingly, almost all strains producing VacA possess *cagA*, compared with only 20% of VacA$^-$ strains (Xiang *et al.*, 1995). The *cag* pathogenicity island and *vacA* are distantly located on the *H. pylori* genome (Tomb *et al.*, 1997), and the relationship between these genes cannot be explained by genetic linkage. A division of *H. pylori* strains into type I (VacA$^+$/*cagA*$^+$) and type II (VacA$^-$/*cagA*$^-$) has been suggested, but not all *H. pylori* strains can be assigned to one of these groups (Xiang *et al.*, 1995).

Cells of the *H. pylori* type I strains trigger a more profound inflammatory response in the infected host tissue than type II strains. The marker gene *cagA* was found to be part of a 37 kb locus termed ***cag* pathogenicity island** (Censini *et al.*, 1996; Akopyants *et al.*, 1998). This DNA segment is not present in type II isolates of *H. pylori*, and has apparently been acquired by horizontal gene transfer from an unknown source into the chromosomal glutamate racemase gene (Censini *et al.*, 1996). The island is flanked by 31 bp direct repeats. In some strains, it is flanked by an insertion sequence IS605, in others a stretch of chromosomal DNA flanked by IS605 divides the island into two regions. In the genome of some *H. pylori* strains, the pathogenicity island is only partly present (Censini *et al.*, 1996). The majority of the genes located in this island displays significant similarities to elements associated with pilus or flagellum biosynthesis, translocases, permeases, conjugal gene transfer elements and proteases involved in secretion (Censini *et al.*, 1996). The observed homologies raised speculations that the pathogenicity island encodes a transporter system capable of interkingdom export of proteinaceous factors (Christie, 1997b). Six of the genes identified in the *cag* pathogenicity island of *H. pylori* 17874 are similar to components of type V protein export processes in other bacteria (section 1.2), most notably the *vir* gene cluster of *Agrobacterium tumefaciens*. This cluster is required for successful transport of T-DNA

from the bacterium into the plant cell (Christie, 1997a). Homologs of *A. tumefaciens virB4* (two genes), *virB9, virB10* and *virB11* were annotated in the *cag* pathogenicity island. In addition, a truncated *A. tumefaciens virD4* gene homolog was also present in this locus (Censini *et al.*, 1996; Akopyants *et al.*, 1998). Four of these *A. tumefaciens* elements (the two *virB4* genes, *virB11* and *virD4*) are predicted to encode nucleotide binding proteins. They are suspected to be ATPases driving the T-DNA translocation process (Christie, 1997a).

The expression of conserved **heat shock proteins** may play a role in *H. pylori* pathogenesis by inducing an autoimmune response. A monoclonal antibody raised against HspB, the GroEL-like heat shock protein of *H. pylori*, crossreacted with gastric epithelial cells (Yamaguchi *et al.*, 1997). The GroES-like HspA might contribute to virulence by providing the urease enzyme with the nickel ions necessary for its activity (Suerbaum *et al.*, 1994). However, two more candidates - Hpn (Gilbert *et al.*, 1995) and NixA (Bauerfeind *et al.*, 1996) - have also been identified which have the capacity to bind nickel ions, suggesting an interaction with urease.

A number of other putative pathogenic factors are currently under investigation. A gene of unknown function, *iceA*, expression of which is upregulated upon contact of bacterial cells with the epithelium of the host, has been implicated in ulcerogenesis (van Doorn *et al.*, 1998). Several enzymes of *H. pylori*, such as phospholipases, catalase and alcohol dehydrogenase, have also been suggested to contribute to mucosal damage (Nilius & Malfertheiner, 1996).

1.1.2.2 Clinical outcomes and treatment

Shortly after ingestion, *H. pylori* cells penetrate the gastric mucin layer in the human stomach. Some cells adhere to the epithelium and subsequently kill host cells presumably by ammonia and the vacuolating cytotoxin (for a review on those interactions, see Wadström *et al.*, 1996). During this time, an intense bacterial proliferation as well as an acute gastric inflammation occurs, termed acute superficial gastritis (Blaser & Parsonnet, 1994). Production of the cytokine IL-8 is initiated in the gastric epithelial cells after *H. pylori* adhesion (Rieder *et al.*, 1997). Other cytokines like TNF-α, IL-6 and IL-1β, reach higher output levels following *H. pylori* infection as well (Crabtree *et al.*, 1991; Yamaoka *et al.*, 1997). For IL-8 and IL-1β, this phenomenon is dependent on CagA expression by the bacterial cell (Crabtree *et al.*, 1994; Yamaoka *et al.*, 1997).

However, the initial inflammatory response is ultimately reduced to a low-level stable state (chronic diffuse superficial gastritis), and the host mounts a humoral immune response that is ineffective in eliminating *H. pylori*. The infected person becomes asymptomatic (Blaser & Parsonnet, 1994). Why *H. pylori* cannot be cleared by the immune system of the host is enigmatic. If the infection is not treated with antibiotics, *H. pylori* persists for life in the mucus layer. This is despite the acidic environment, peristaltic gut movements, and the continuous shedding of both mucus and epithelial cells.

Once successfully established, an *H. pylori* infection may broadly speaking lead to four different final outcomes: A), the infected subject has chronic superficial gastritis and remains asymptomatic throughout life; B), peptic ulcer disease occurs, including either duodenal or gastric ulcers; C), the gastritis becomes atrophic, which might lead to gastric adenocarcinoma, or D), lymphoma develops (Blaser, 1996). However, the clinically relevant outcomes B-D only develop in a minority of infected humans, and manifestations take many years to appear (Blaser, 1996).

Whereas the link between *H. pylori* and gastritis and ulcers is unequivocally proven (section 1.1.1), the association between *H. pylori* and gastric cancer is still controversial (Palli, 1997). However, in 1994, the cause and effect relationship between *H. pylori* and gastric cancer was recognized by the International Agency for Research on Cancer (World Health Organization, 1994). More evidence for this important connection was obtained in the last three years. Epidemiological studies showed a markedly increased risk for gastric cancer in humans infected with *H. pylori* compared to noninfected persons (Parsonnet *et al.*, 1997; Miehlke *et al.*, 1997). In a breakthrough study using mongolian gerbils, chronic gastritis and ulcers were observed in animals 26 weeks after an oral challenge with *H. pylori*. After 62 weeks, 37% of the infected animals had developed adenocarcinoma in the pyloric region (Watanabe *et al.*, 1998). It has been postulated that the chief role of *H. pylori* in carcinogenesis may be its ability to stimulate cell growth, thus increasing the likelihood of DNA damage in response to environmental mutagens (Tsujii *et al.*, 1993). An *H. pylori* infection also downregulated expression of E-cadherin, a protein essential for cell-cell adhesion, in gastric epithelial cells (Terres *et al.*, 1998). Germline mutations in the E-cadherin gene were found to be the molecular basis for familial gastric cancer, confirming the importance of its product in carcinogenesis (Guilford *et al.*, 1998).

The link between *H. pylori* and mucosa associated lymphoid tissue (MALT) lymphoma has been strengthened by clinical studies which have shown that early lymphoma stages regressed following eradication of *H. pylori* in up to 92% of the cases (reviewed in Wotherspoon, 1998).

Despite great efforts to reveal the factors responsible for the differences in the outcome of an *H. pylori* infection, the disease triggering events remain unclear. Differences among the *H. pylori* strains are widely accepted to have an influence on the development of ulceration. *H. pylori* type I strains containing the *cag* pathogenicity island are considered to be more virulent than type II isolates (section 1.1.2.1). Host-specific factors also contribute to the development of ulcers. It has been suggested that immunogenetic factors may influence the host's response to *H. pylori* (Azuma *et al.*, 1994). The nature of cytokines produced during the gastric immune response to *H. pylori* may represent an important factor for the outcome of the infection (D'Elios *et al.*, 1997). Environmental factors such as diet and infection duration may also play an important role in ulcerogenesis and carcinogenesis (Blaser & Parsonnet, 1994).

An *H. pylori* infection can easily be diagnosed, and the methodologies are well established (reviewed in Dunn *et al.*, 1997). Reliable tests for *H. pylori* infection include both endoscopic and non-invasive methods. The easiest and fastest test, a ^{13}C urea breath test, is 90-95% sensitive and specific.

To eradicate *H. pylori* from an infected person, triple or quadruple antimicrobial therapy is successfully used in general practice. The most common treatments consist of a combination of antimicrobial agents (amoxycillin, clarithromycin, metronidazole, tetracycline) in association with a proton-pump inhibitor (Omeprazole or Lansoprazole) and an anti-histamine 2-receptor antagonist or bismuth citrate (reviewed in Vaira *et al.*, 1998).

The development of *H. pylori* resistance to antibiotics is the major problem in eradication therapies. Metronidazole resistance rates vary regionally between 11 and 70% (Dunn *et al.*, 1997), but are even higher in developing countries (Vasquez *et al.*, 1996). Prevalence of clarithromycin resistance is usually much lower, but have been found to be as high as 50% in Peru (Vasquez *et al.*, 1996). Tetracycline resistance has also been reported, although less frequently (Midolo *et al.*, 1996).

In view of emerging resistance to antibiotics and the high prevalence of *H. pylori* in third world countries, the concept of vaccination as a strategy for population-based eradication of *H. pylori* is appealing. Practical animal models were developed to mimic

human *H. pylori* infection, such as *Helicobacter felis* infection in mice (Lee *et al.*, 1990), *H. mustelae* in ferrets (Fox *et al.*, 1990), and *H. pylori* in gnotobiotic piglets (Krakowka *et al.*, 1991) and mice (Marchetti *et al.*, 1995; Lee *et al.*, 1997a). Once the models had been established, it was possible to immunize these laboratory animals against a subsequent *Helicobacter* challenge, using whole cell lysates (Chen *et al.*, 1992), or purified recombinant *H. pylori* antigens. Among those antigens that successfully conferred immunity were both subunits of the *H. pylori* urease (Ferrero *et al.*, 1994; Cuenca *et al.*, 1996), the heat shock proteins HspA and HspB (Ferrero *et al.*, 1995), the vacuolating cytotoxin VacA (Marchetti *et al.*, 1995), and the *H. pylori* catalase (Radcliff *et al.*, 1997). The recombinant larger urease subunit of the bacterium was well tolerated by humans in phase I trials (Kreiss *et al.*, 1996).

Against the background of constantly lowering infection rates concurrent with increased living standards in most parts of the world, feasability of a vaccine is still being questioned and discussed. The cost / effect relationship of vaccine development and administration is controversial. However, research is still being conducted vigorously towards a future vaccine. Latest estimations expect this *H. pylori* vaccine to be available in approximately 10 years (Marshall, 1998).

1.1.3 The *H. pylori* genome

Just 15 years after the first successful culture of the bacterium, the fully annotated genome sequence of *H. pylori* (strain 26695) was published (Tomb *et al.*, 1997). It was the sixth prokaryotic genome to be completely decoded (Owen, 1998).

The single circular chromosome of *H. pylori* 26695 contains approximately 1.66 Mb. It comprises 1590 putative genes, two thirds of which are homologs of known genes from other organisms (Berg *et al.*, 1997). However, biological roles of genes were often assigned only on the basis of database matches of less than 50% similarity (Owen, 1998), increasing the possibility of misinterpretation and errata.

Analysis of the *H. pylori* genome provided valuable insights into *H. pylori* pathogenesis. It revealed the existence of two putative antigenic variation mechanisms which might contribute to the ability of *H. pylori* to evade the host immune response for life. Firstly, at least 20 genes encoding putative lipoproteins, and thirty two genes coding for a family of outer membrane proteins were identified. These large numbers suggested recombinational events leading to mosaic organization and effective evasion of the host

defense (Tomb *et al.*, 1997). Secondly, stretches of mono- or dinucleotide repeats were present in 17 ORFs, many of them predicted to be surface exposed. This strongly supported the theory of a slipped strand mispairing process which would result in ON/OFF switching by generating frame-shift mutations (Berg *et al.*, 1997). At least one example of this mechanism has already been characterized for the *H. pylori* lipopolysaccharide (Appelmelk *et al.*, 1998).

Different strains of *H. pylori* were originally thought to exhibit extraordinary genome diversity (Taylor *et al.*, 1992; Bukanov & Berg, 1994; Logan & Berg, 1996), with very variable gene orders (Jiang *et al.*, 1996). Extensive gene rearrangements within the *H. pylori* genome seemed to have resulted from horizontal transfer and recombination (Go *et al.*, 1996; Suerbaum *et al.*, 1998). However, unexpectedly high levels of conservation of porins between the unrelated *H. pylori* type I strains 26695 and J99 had been described (Hancock *et al.*, 1998). The genome sequence of strain J99 was subsequently analyzed. Comparison of the *H. pylori* 26695 chromosome with that of strain J99 revealed that only 6-7% of the genetic content of the *H. pylori* genome was strain-specific (Alm *et al.*, 1999). Moreover, the gene order appeared to be more conserved than expected, with approximately 85% of the genes common in both strains being flanked by the same neighbour element on each side. Complex organizational differences were restricted to 10 assigned regions of the chromosome (Alm *et al.*, 1999). The different restriction site content of genomes of different *H. pylori* strains, responsible for the previous overestimation of the genetic diversity of the bacterium, seemed to be mostly caused by single silent nucleotide changes (Alm *et al.*, 1999). A detailed analysis of the genome sequence data of the two different *H. pylori* strains will provide a fascinating insight into strain-specific features and adaptations of the bacterium in response to its host or environment. Assessment of similarities and differences to genome sequences of other microbial organisms, particularly those of gastrointestinal microaerophilic bacteria (like *Campylobacter jejuni*), will subsequently lead to a better understanding of common regulatory mechanisms in *H. pylori.*

Major consequences of the revelation of the genome sequence can be foreseen. Easy identification of new suitable targets for therapy and development of new molecular techniques for *H. pylori* are only two of them. The postgenomic era in *Helicobacter* research has begun, and the new technologies are shaping up. A high density oligonucleotide array containing tag complements (Affymetrix Biochip) has been constructed that will be useful for selection of defined *H. pylori* mutants generated by signature tagged allelic replacement (Dorrell *et al.*, 1998). A prospective *H. pylori* chip

containing signature oligonucleotides characteristic for each of the 1590 genes of the *H. pylori* 26695 genome would facilitate identification of genes only switched on *in vivo* (Lee, 1998).

Many questions remain to be answered. For example, the perplexing tissue specificity of *H. pylori*, its lifelong survival against the inflammatory response of the host, and critical bacterial and host factors determining the outcome of the infection are still not understood. The entry of the *H. pylori* genome sequence into the public scientific domain will certainly increase the speed with which answers to these questions can be found.

1.2 Protein secretion systems in bacteria

Active translocation of proteins from a bacterial cell to the extracellular environment is accomplished by several secretion pathways. The ability to transport macromolecules into the extracellular space is especially important for pathogenic bacteria since the necessary interaction between the bacterial cell and its host can only be established by secreted proteins. To date, five major pathways of macromolecular secretion have been identified in Gram-negative bacteria (Hueck, 1998), although classification of type IV secretion as an autonomous pathway is controversial and not generally acknowledged (Salmond, 1994; Christie, 1997a). The various systems of bacterial protein export have been the subject of intensive investigations in recent years. This has prompted several reclassifications that have been comprehensively reviewed in the literature. This overview of the topic is primarily based on these reviews, which are cited.

The currently assumed classification and characteristics of the translocation processes are illustrated in Fig. 1.2. Table 1.1 shows important features of the five translocation systems and represents an updated summary of a similar table originally presented five years ago by Salmond (Salmond, 1994).

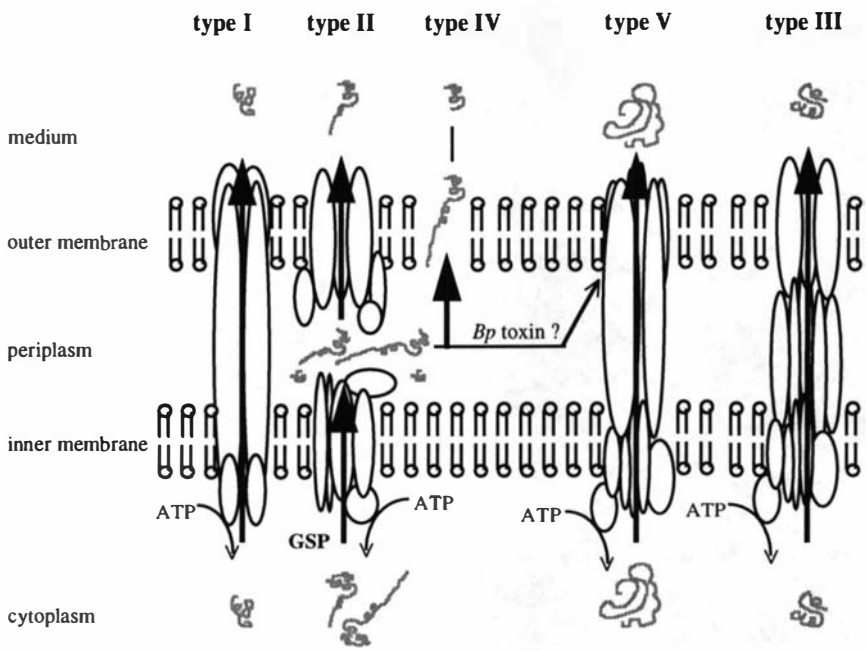


Figure 1.2. Schematic diagram of currently classified bacterial protein secretion systems.

Type I, III and V systems are one - step translocation procedures, whereas type II and type IV mechanisms proceed in two separate steps. The substrates of type IV secretion may or may not be released from the outer membrane by autocleavage. The *Bordetella pertussis* toxin is assumed to follow a type V secretion, but may also use the general secretory pathway for translocation across the cytoplasmic membrane (see text).

GSP, general secretory pathway; *Bp*, *Bordetella pertussis.*

Figure reproduced from Mecsas & Strauss, 1996, with modification.

Table 1.1. General features of protein secretion systems in Gram-negative bacteria.

Feature	Type I	Type II	Type III	Type IV	Type V
genes involved	3	~ 20		several (lacks genes for step 2 in type II)	multiple
sec dependance	-	+	-	+	- / +
number of steps during translocation	1	2	1 (?)	2	? / 2 (?)
signal sequence	C-terminal	N-terminal	5' region of mRNA (?)	N-terminal	? / N-terminal
signal cleavage	-	+	-	+	? / +
example	*E. coli* hemolysin	*K. oxycota* pullulanase	*Y. enterolitica* Yop	*S. flexneri* SepA	*A. tumefaciens* T-DNA / *B. pertussis* toxin

Type I secretion generally requires only three secretory components: an outer membrane protein, a channel protein spanning through the periplasmic space, and a transport ATPase associated with the inner membrane (Hueck, 1998). This transport ATPase contains an ATP binding cassette and is therefore generally called an ABC transporter. Type I secretion occurs in a continuous, one-step procedure without detectable periplasmic intermediates (Fath & Kolter, 1993). The substrate is probably transported through some form of channel or pore formed between the two bacterial membranes (Salmond, 1994). Toxins, proteases and lipases are known to be secreted by this mechanism (Binet *et al.*, 1997). The genes encoding the secretion apparatus and the secreted protein are usually linked, and the exporter is substrate-specific (Binet *et al.*, 1997). The signal sequence is located at the C-terminus, probably conformational rather than sequence-specific, and is not cleaved during the translocation process (Salmond, 1994). Type I secretion is exemplified by translocation of *E. coli* hemolysin (Fath & Kolter, 1993).

Type II secretion is the primary pathway for translocation of extracellular degradative enzymes, like *Pseudomonas aeruginosa* elastase, alkaline phosphatase, phospholipase C, and *Klebsiella oxytoca* pullulanase (reviewed in Pugsley, 1993). Because of the diversity of the proteins secreted via this apparatus, the type II pathway has been termed general secretory pathway (GSP). Contrary to type I secretion, type II translocation in Gram-negatives is a two-step procedure (Pugsley, 1993). It involves a separate export step across the inner membrane prior to translocation across the cell envelope, and periplasmic intermediates can be detected. These intermediates undergo conformational changes before being subjected to the second step of translocation (Pugsley *et al.*, 1997). Export of the exoprotein into the periplasmic space (step 1) is

dependent on a number of inner membrane proteins, a cytoplasmic membrane-associated ATPase, a chaperone and several accessory proteins. This first step of the type II pathway is also used during export of proteins by type IV secretion (Hueck, 1998). In addition, the *Bordetella pertussis* toxin, thought to be translocated by type V secretion, may also use the GSP to reach the periplasm (Winans *et al.*, 1996). The secreted protein contains an N-terminal signal sequence that is cleaved during transport into the periplasmic space by a signal peptidase. In *E. coli*, some components necessary for this translocation (translocases, a secretory ATPase and a secretory chaperone) are encoded by *sec* genes (Pugsley, 1993). Therefore, this secretion mechanism is generally called *sec*-dependent. The second step of type II translocation, transport across the outer membrane, requires several additional inner and outer membrane proteins. At least 14 components have already been identified that are involved in this step of pullulanase secretion in *K. oxytoca* (Pugsley, 1993). Several components of this complicated apparatus have homologs in other transport systems. Notably, proteins involved in DNA uptake and assembly of type 4 fimbriae share extensive similarities with type II secretion components (Hobbs & Mattick, 1993; Mattick & Alm, 1995). Type 4 fimbriae are polar organelles on the surface of a variety of bacteria, and mediate a form of flagella-independent surface translocation called twitching motility (Henrichsen, 1983). They are thought to be involved in adherence of the prokaryotic cell to host epithelial tissue (Ottow, 1975). Arguably, the apparatus required for fimbrial assembly may in actual fact be a variant of the GSP system (Mattick & Alm, 1995), where one subunit of this apparatus, the structural pilus protein PilA, forms an extracellular organelle, the pilus.

Type III secretion is, like the type I pathway, a *sec*-independent one-step process (reviewed in Mecsas & Strauss, 1996; Hueck, 1998). This protein secretion mechanism is specialized for translocation of virulence determinants in several animal and plant pathogens. Components known to be secreted by this pathway include a number of *Yersinia* outer proteins (Yops) of *Y. enterolitica*, *Y. pseudotuberculosis* and *Y. pestis*, invasion plasmid antigens (Ipas) of *Shigella flexneri*, *Salmonella* secreted proteins (Ssp's) of *S. typhimurium,* and so-called harpins (or Hrps) of *Erwinia* spp., *Xanthomonas campestris* pv. *vesicatoria*, and *Pseudomonas syringae* pv. *syringae* (reviewed in Salmond, 1994; Lee, 1997). These proteins are necessary for bacterial virulence, or raise the colonization rate of the pathogen. They interact directly with host cells and proteins (Mecsas & Strauss, 1996). Cleavage of a signal peptide does not occur during secretion. The signal for protein secretion via a type III pathway has long been thought to be located in the N-terminal region of the protein. However, comparison of the N-terminal regions of several Yops did not reveal any common consensus sequence or structural motif

(Wattiau *et al.*, 1996). It has been suggested that the real signal for type III secretion resides in the structure of the 5' region of the mRNA which encodes the secreted protein. Silent single nucleotide mutations in the region coding for the 15 N-terminal amino acids of *Y. enterolitica* YopN moderately affected secretion efficiencies of a fusion construct consisting of these amino acids and the cytoplasmic neomycin phosphotransferase (Anderson & Schneewind, 1997).

The genes encoding type III secretion systems are usually clustered (Hueck, 1998). Although generally regarded as a one-step procedure, it has been shown that type III transport of proteins across the inner membrane can be genetically separated from the subsequent translocation across the cell envelope (Charkowski *et al.*, 1997). Type III secretion of proteins is apparently triggered by contact of the bacterium with the host cells (Rosqvist *et al.*, 1994; Watarai *et al.*, 1995), and has therefore been called contact-dependent (Galan, 1996). This peculiarity made the system highly attractive for vaccine delivery. In a breakthrough study, delivery of viral epitopes by the *S. typhimurium* type III secretion system was shown to elicit a protective antiviral immune response in mice (Rüssmann *et al.*, 1998).

The type III secretion apparatus is composed of approximately 20 components (Hueck, 1998), and requires a cytoplasmic ATPase. Some similarities to components of other systems have been observed. Most notably, a number of the type III secretion apparatus proteins have homologs in the system responsible for flagellar assembly (reviewed in Lee, 1997) (section 1.3.2). These two systems share a number of key characteristics and are apparently closely related. Similar to proteins exported by a type III secretion pathway, most flagellar structural proteins do not contain the N-terminal *sec* signal sequence. Regulation of the *Yersinia* genes encoding Yops has been attributed to a repressor molecule YscM, that is exported out of the cell upon activation of type III secretion by cell contact (Allaoui *et al.*, 1995; Pettersson *et al.*, 1996). In *S. typhimurium*, a similar mechanism involving the repressor FlgM regulates expression of flagellar components (section 1.3.1.3).

Type IV secretion systems use the same route as type II systems for protein export to the periplasm across the inner membrane. Thus, they are *sec*-dependent, and N-terminal signal sequences are removed from the protein during export into the periplasm (Finlay & Falkow, 1997). However, the transport of the proteins across the outer membrane is different to the type II pathway. Proteins following type IV translocation are called autotransporters, and they apparently form a pore in the outer membrane through which they pass (Finlay & Falkow, 1997). Usually, autoproteolytic cleavage subsequently releases the mature protein into the exterior. This pathway is exemplified by

the secreted protein SepA from *S. flexneri* (Benjelloun-Touimi *et al.*, 1995). The *H. pylori* vacuolating cytotoxin VacA is also suspected to be transported via this pathway (Schmitt & Haas, 1994). The signal for VacA translocation across the outer membrane is located in the C-terminus of the translated proprotein. This 33 kDa fragment is cleaved, and probably remains in the outer membrane of the bacterium (Schmitt & Haas, 1994).

Type V secretion currently comprises systems for translocation of macromolecules involved in conjugal transfer of plasmids, T-DNA transfer by *A. tumefaciens,* and secretion of the *B. pertussis* toxin (Winans *et al.*, 1996). The grouping of these three processes into one pathway is based on genetic similarities rather than mechanistic resemblance. Reporting the described type IV secretion as a subgroup of type II translocation, several authors refer to the type V group as type IV (Salmond, 1994; Christie, 1997a).

Although much research has focussed on T-DNA transfer in *A. tumefaciens*, the exact mechanism of this translocation system is still unclear. The current model for the export machinery responsible for secretion of the T-DNA-nucleoprotein complex is illustrated in Fig. 1.3. It is known that single-stranded T-DNA is covalently attached to a protein VirD2 during transfer (Stachel *et al.*, 1987). A single step, *sec*-independent transfer procedure is assumed. The transport machinery consists of an aggregate of at least 11 proteins, most of which are encoded in a single operon termed *virB* (Lin & Kado, 1993; Berger & Christie, 1994; Christie, 1997a). This aggregate forms an extracellular pilus, which consists of the two structural proteins VirB2 and VirB5 (Zupan *et al.*, 1998). Three elements of the *vir* translocation complex, VirB4, VirB11 and VirD4, are nucleotide-binding proteins. ATPase activity has been shown for VirB4 and VirB11, and both are associated with the inner membrane (Christie *et al.*, 1989; Shirasu *et al.*, 1994). The *virB* gene cluster has strikingly similar counterparts in the *ptl* locus necessary for secretion of the pertussis toxin in *B. pertussis* (Weiss *et al.*, 1993), and the *tra* locus present on plasmids of the IncP and IncN incompatibility groups (Lessl *et al.*, 1992; Pohlman *et al.*, 1994). The latter is essential for conjugal distribution of these plasmids. Apparently, these processes are highly related, and use similar principles, components and structures. Two of the common features are pilus formation and intercellular nucleoprotein transfer.

The *B. pertussis* toxin is processed during translocation. An N-terminal signal sequence is removed from the mature secreted protein, and the usage of the GSP for export across the inner membrane is likely (Pugsley, 1993). However, genetic similarities of the *ptl*, *vir* and *tra* gene clusters suggest that several components necessary for pertussis toxin export are in fact associated with the cytoplasmic membrane (for example

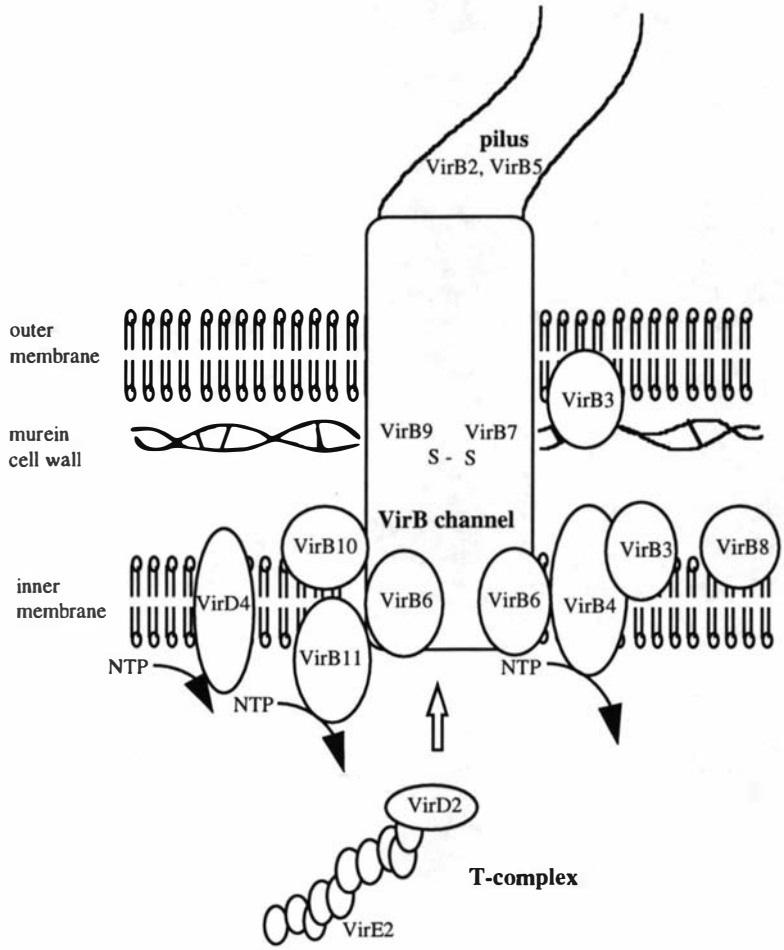


Figure 1.3. Model of the *Agrobacterium* T-complex transport apparatus.

The T-complex, consisting of T-DNA in single strand form bound to VirD2 and coated with VirE2, is exported through the transport apparatus. The exact roles of VirB3, 4, 6, 8, 10 and 11, and VirD4 are unknown. VirD4, VirB4 and VirB11 have nucleotide binding sites which are essential for their function.

VirB1 (not depicted) uses the general secretory pathway to reach the periplasmic space. The C-terminal 73 amino acids of VirB1 (VirB1*) can be found in complexes with VirB9. VirB1* is mainly secreted through the outer membrane by an unknown mechanism. The protein is suggested to have local lytic activity that allows assembly of the transporter at specific sites in the bacterial cell envelope.

Figure reproduced from Zupan *et al.*, 1998, with modification.

ptlC, ptlE and *ptlH,* the homologs of *virB4, virB8* and *virB11*, respectively) (Winans *et al.*, 1996). It is therefore currently unclear whether the export of the toxin occurs in one or two steps, or if both mechanisms apply.

A common theme in all three examples of this type V secretion mechanisms is the existence of at least two ATP binding proteins involved in energy supply for the transfer process. Recently, more genes with similarities to the transfer components mentioned above have been described, most notably in the *cag* pathogenicity island of *H. pylori* (section 1.1.2.1), but also in *Legionella pneumophila* (Vogel *et al.*, 1998).

Protein transport pathways, conjugal DNA transfer systems, type IV fimbriae and flagella all share homologous subunits and seem to be related (Hobbs & Mattick, 1993). It has therefore been speculated that all of them emerged as different systems from a common ancestor, with the conjugal transfer system most likely being the oldest of the described mechanisms (Winans *et al.*, 1996).

1.3 Flagellar biosynthesis

1.3.1 Structure and assembly of a flagellum in Gram-negative bacteria

The most intensive research on flagellar synthesis in Gram-negative bacteria has primarily focussed on two organisms, *S. typhimurium* and *Caulobacter crescentus.* Principles of flagellar gene structure, regulation and flagellar assembly were originally thought to be very similar in all Gram-negatives. However, it became apparent that some mechanisms are completely different in these two organisms, suggesting the existence of various adaptations and alterations in flagellar biology of Gram-negative bacteria.

1.3.1.1 Flagellum morphology

The bacterial flagellum is a motility apparatus that morphologically consists of the extracellular helical filament, the hook, a basal body embedded in the cell surface, and intracellular substructures such as the apparatus responsible for the export of flagellar compounds, the switch components and motor proteins (reviewed in Macnab, 1992; Macnab, 1996; Aizawa, 1996). The architecture of a flagellum of *S. typhimurium* is depicted in Fig.1.4. In this bacterium, at least 50 genes are known to be involved in flagellar synthesis and function (Kutsukake *et al.*, 1990). They are genetically clustered in 4 regions and 17 operons consisting of up to 9 genes (Macnab, 1992; Macnab, 1996).

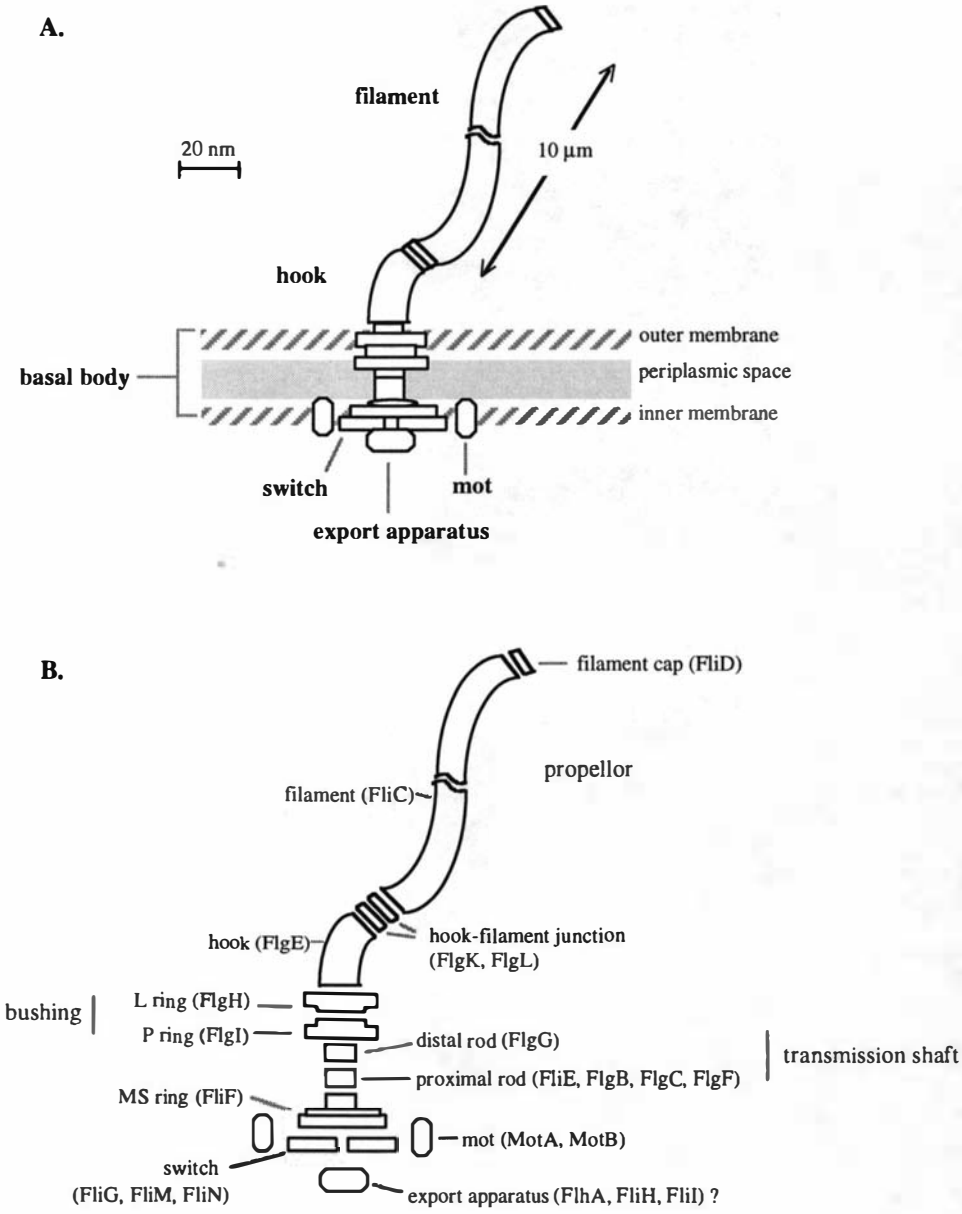


Figure 1.4. Structure of the flagellum of a Gram-negative bacterium like *S. typhimurium* or *E. coli.*

Panel A. General components of the flagellar apparatus. The location of the export apparatus has not been demonstrated.

Panel B. An exploded diagram depicting flagellar substructures and the proteins which form them.

Figure reproduced from Macnab, 1992.

However, localization on the chromosome does not always correspond to the expression hierarchy of those genes, and the significance of the genetic organization of flagellar genes remains in some parts obscure. Gene orders seem to differ from organism to organism, although some operons remain conserved.

The filament is the largest part of a flagellum. It is up to 10 μm long, and has a constant diameter of 20 nm (Namba *et al.*, 1989). It may be sheathed in some bacteria by an extension of the outer membrane, and consists of flagellin(s) and a cap protein (Macnab, 1992). Rotary filament movements result in bacterial mobility. Clockwise rotation of the filaments causes tumbling of the bacterial cell, counter-clockwise flagellar spins give rise to straight directional motion (Macnab, 1987). **The hook** connects the filament to the basal body. In *S. typhimurium*, this hook is formed by a component FlgE and two junction proteins that are responsible for proper filament ligation (Homma & Iino, 1985; Macnab, 1992). A scaffolding protein FlgD is necessary for FlgE polymerization (Ohnishi *et al.*, 1994). **The basal body** is a complex structure consisting of an L-ring in the outer membrane, an adjacent periplasmic P-ring, distal and proximal rods spanning through the periplasm, and an MS-ring structure located in the cell membrane (Macnab, 1992). In *Caulobacter*, an additional substructure termed the E-ring is located between P- and MS-rings (Stallmeyer *et al.*, 1989). Components of the switch apparatus are attached to the MS-ring and localized at the cytoplasmic side of the cell membrane. They form a bell-shaped structure termed the C-ring (Khan *et al.*, 1991). Three proteins have been identified as switch components (Kihara *et al.*, 1989). They are responsible for determination of the orientation of the flagellum rotation. At least two Mot components have been characterized that are necessary for flagellar rotation (Blair & Berg, 1988). Several proteins are suspected to be part of the so-called export apparatus that is responsible for the signal sequence independent protein translocation during flagellar assembly (section 1.3.2).

1.3.1.2 Assembly of the flagellum

The beginning of flagellation in *S. typhimurium* is characterized by the formation of the MS ring structure. The protein FliF self-assembles into this complex, which is a core structure of the basal body (Ueno *et al.*, 1994). This is the first flagellar structure detectable by electron microscopy (Kubori *et al.*, 1992). Now other components start to accumulate intracellularly on the base, starting with the switch proteins (Francis *et al.*, 1992; Kubori *et al.*, 1997). Accumulation of the export apparatus components possibly

follows (Aizawa, 1996). At least two proteins thought to be part of the export apparatus (FliP and FliR) have been proposed to be localized in the central pore of the MS ring (Fan *et al.*, 1997). However, the existence of an N-terminal signal sequence in most FliP proteins identified so far is inconsistent with its suggested localization within the inner membrane of the bacterial cell (Ohnishi *et al.*, 1997). The next step in flagellar assembly is probably the cooperative assembly of the rod proteins (Kubori *et al.*, 1992). Subsequently, the outer cylinder with the P-ring protein FlgI and the L-ring component FlgH assembles. Unlike all other flagellar proteins (except FliP), those two components are thought to be exported via the general secretory pathway. They contain a cleaved N-terminal signal peptide and presumably diffuse around their respective compartments, until they encounter a rod (Macnab, 1992).

During subsequent polymerization of the hook, nascent FlgE is inserted beneath a FlgD cap, a polymerization catalyst essential for *in vivo* hook protein subunit assembly. When the hook is mature, replacement of FlgD by FlgK occurs (Ohnishi *et al.*, 1994). At least three genes, *fliK, rflH* and *flhB* are known to be involved in hook length control in *S. typhimurium* (Kutsukake *et al.*, 1994; Hirano *et al.*, 1994; Muramoto *et al.*, 1998). FliK and FlhB are apparently necessary for both turning off hook protein export and altering the export apparatus specificity from the hook protein to flagellin and the other "late" flagellar structural proteins (Kutsukake, 1997b). RflH is thought to lock the gate for export of these proteins in cooperation with FlhB (Kutsukake, 1997b).

When hook elongation is complete, the junction proteins, the capping component and flagellin(s) are exported to adhere the filament structure to the hook. At least one of the two junction proteins, FlgL, stabilizes the filament structure (Fahrner *et al.*, 1994). The filament cap protein FliD functions like FlgD for the hook, except for the fact that it is not replaced after polymerization completion. It is essential for *in vivo* assembly of flagellin subunits, and polymerization takes place beneath this cap (Ikeda *et al.*, 1993). The filament length might be regulated at the export level (Aizawa, 1996). The molecular basis of this control mechanism is still unknown.

The motor proteins MotA and MotB are the only flagellar components that are not necessary for the assembly of the rest of the flagellar apparatus (Macnab, 1992). It is not known when they adhere to the MS-ring.

1.3.1.3 Regulation of flagellar gene expression

The genes for the flagellar proteins are expressed in a hierarchical order. A class I master operon consisting of two presumed transcription factors, FlhC and FlhD, has been proposed for a number of organisms including *S. typhimurium* (Macnab, 1992) and *Proteus mirabilis* (Dufour *et al.*, 1998), but is not established for other organisms like *C. crescentus*. In that organism, the class I gene product is the response regulator CtrA, which also governs control of initiation of chromosome replication and methylation (Quon *et al.*, 1996). Phosphorylation is crucial for its activity (Quon *et al.*, 1996). The products of the class I flagellar genes initiate transcription of the class II genes. In *S. typhimurium*, the class II subset includes genes encoding regulatory proteins as well as flagellar basal body and hook components (Kutsukake *et al.*, 1990). The expression of these genes is required for subsequent class III gene transcription (hook-filament junction proteins, flagellins). Regulation mechanisms have been extensively studied in *S. typhimurium*, but appear to be different in other organisms. In *Caulobacter*, for example, four rather than three flagellar gene classes have been characterized (Wu & Newton, 1997). The *S. typhimurium* class II flagellar gene homologs are in *C. crescentus* divided into two groups. Proteins of the MS-ring, the switch and the export complex are separated from the outer basal body components and the hook constituents (Wu & Newton, 1997). Furthermore, an additional class of flagellar regulatory genes (*flm*) has been reported in *C. crescentus*. Mutation of these genes did not affect class II or class III flagellar gene expression, but reduced production of the class IV proteins (Leclerc *et al.*, 1998).

The start signal for the process of flagellar assembly is not known. In *S. typhimurium*, flagellation is clearly cAMP/CAP regulated (Komeda *et al.*, 1976), so that glucose can exert catabolic repression by lowering cellular cAMP levels. Expression of the flagellar master operon was found to be dependent on several global regulators including cAMP-CRP (Kutsukake, 1997a). In contrast, cell cycle cues are thought to be responsible for upregulation of CtrA production in *Caulobacter* (Wu & Newton, 1997). The master operon is expressed independently from all other flagellar genes and governs the rest by activating operons in the second class (Bartlett *et al.*, 1988). In *S. typhimurium*, these genes have specific class II promoter sequences. The -10 signature sequence is identical to the flagellum-specific σ^{28} recognition site (GCCGATAA), but a consensus -35 sequence pattern is not present (Kutsukake *et al.*, 1990; Macnab, 1996). In the presence of the class I gene products FlhC and FlhD, RNA polymerase incorporating a σ^{70} factor can initiate transcription of class II genes. Usually, the σ^{70} factor facilitates transcription of housekeeping genes in bacteria like *E. coli* and *S. typhimurium*, and

recognizes the sequence motif TATAAT-N_{17}-TTGACA (Harley & Reynolds, 1987). However, when the FlhCD complex binds to a sequence motif upstream from the transcription start, the σ^{70} factor is able to recognize the class II flagellar gene recognition site (Liu & Matsumura, 1994). In *C. crescentus,* the class I gene product CtrA binds around 18-45 bp upstream of the transcription start site, and enables the housekeeping σ^{73} factor to initiate transcription of class II flagellar genes (Quon *et al.*, 1996). The promoter recognition sites of class II flagellar genes in *C. crescentus* are not strictly conserved. A specific consensus of AACNCN$_3$GTTNACCNTN$_5$GN$_3$A has been determined for three class II flagellar operons of this bacterium (Stephens & Shapiro, 1993). This sequence motif is also present upstream from two genes implicated in cell cycle control of *C. crescentus* (Quon *et al.*, 1996). In contrast, regions upstream from two other *C. crescentus* flagellar class II genes, *rpoN* (encoding the σ^{54} factor) and *flhA*, did not exhibit this consensus (Anderson *et al.*, 1995).

Further down the flagellar gene hierarchy, alternative σ factors are necessary for transcription of flagellar genes. In *S. typhimurium*, one of the class II genes, *fliA*, encodes the flagellum-specific σ^{28} factor. This factor is necessary for subsequent transcription of class III genes, and also autoregulates class II gene transcription (Kutsukake & Iino, 1994). *S. typhimurium* class III gene promoters have the same -10 signature as class II flagellar genes, but also contain a conserved -35 sequence TAAA (Macnab, 1992). However, some of the *S. typhimurium* class III flagellar genes do not absolutely require FliA for transcription. In these cases, readthrough of a class II operon upstream from a class III gene (exemplified by the class III *flgM* gene which is preceded by the class II *flgA* element), or the existence of both class II and class III promoters upstream from the respective operon (shown for the *flgKL* operon) ensures transcription in the abscence of FliA (Gillen & Hughes, 1993; Kutsukake & Ide, 1995).

In contrast, transcription of class III and class IV flagellar genes in *C. crescentus* requires the alternative σ^{54} factor. In general, this factor enables the RNA polymerase to transcribe a variety of genes upon an external stimulus (Thöny & Hennecke, 1989). In *C. crescentus*, the -13/-24 sequence (C/T)TGGC(C/G)C-N_5-TTGC was found to precede several flagellar genes (Mullin *et al.*, 1987), closely resembling the established consensus sequence for σ^{54} promoters in *E. coli* TGGYRYR-N_4-TTGCA (Y = T or C, R = G or A) (Thöny & Hennecke, 1989). The σ^{54} factor initiates transcription of these genes in association with a transcription factor FlbD. Similar to CtrA, FlbD is a response regulator, and governs transcriptional activation of various genes controlled by the σ^{54} factor (Wu *et al.*, 1995). It binds to a conserved 19 bp dyad termed *ftr* (flagellar

transcription regulator), around 100 bp upstream from the transcription start (Benson *et al.*, 1994).

In *H. pylori* (Suerbaum *et al.*, 1993) and *Vibrio cholerae* (Klose & Mekalanos, 1998), σ^{54} factors are also suspected to govern transcription of certain flagellar genes. However, in both organisms, σ^{28} recognition sequences were also present upstream from some flagellar genes (Leying *et al.*, 1992; Klose & Mekalanos, 1998). In *Rhodobacter sphaeroides*, a σ^{54} element is thought to control transcription of at least one flagellar gene (Shah & Sockett, 1995). In contrast, all five large flagellar gene operons in *Borrelia burgdorferi* are preceded by σ^{70} promoter sequences (Ge *et al.*, 1997a; Ge *et al.*, 1997b; Manson *et al.*, 1998).

The sequential transcription of flagellar genes is coupled to the status of flagellar assembly. The successful completion of the basal body-hook complex BBH is a well characterized checkpoint. When this structure has been constructed, expression of the flagellins is initiated. In *S. typhimurium*, the control is exerted by export of the anti-σ factor FlgM upon successful construction of the BBH (Hughes *et al.*, 1993; Kutsukake, 1994). This anti-σ factor binds to the flagellum-specific σ factor FliA (Ohnishi *et al.*, 1992), and prevents expression of class III flagellar genes. Upon completion of the BBH, translocation of FlgM into the exterior occurs, and FliA can then initiate transcription of the late flagellar genes. An earlier stage of flagellar assembly apparently also contributes to class III transcriptional regulation. A regulatory gene *flk* has been characterized that senses completion of the L- and P-ring attachment to the basal body structure (Karlinsey *et al.*, 1997). The model suggests that the incomplete basal body structure exerts a negative effect on FlgM translation. This signal is itself repressed by free cytoplasmic ring proteins and Flk. When the rings are attached to the basal body, the translational inhibition (and its repression) relaxes (Karlinsey *et al.*, 1998).

In *Caulobacter*, completion of the MS-ring-switch complex is required for the transition from class II to class III gene expression (Anderson & Newton, 1997). How the signal is mediated is not certain, but a gene *bfa* was found to be required for the control mechanism (Mangan *et al.*, 1995). The second checkpoint, completion of the BBH, controls expression of at least two of the three flagellins at posttranscriptional level (Anderson & Newton, 1997). The exact mechanism of this regulation is not known.

1.3.2 The flagellar protein export apparatus

The flagellar protein export apparatus is thought to be responsible for (i) the transport of the rod proteins into the periplasmic space, and (ii) translocation of extracellular flagellar compounds, possibly through a channel in hook, filament and rod (Morgan *et al.*, 1995; Mimori *et al.*, 1995). To date, our knowledge about this apparatus is very limited. Although its location is thought to be in close vicinity to the flagellar MS-ring (Dreyfus *et al.*, 1993; Fan *et al.*, 1997), and its presence is generally acknowledged, a flagellum-specific export apparatus *per se* has never been proven to exist. However, all but two flagellar compounds not located intracellularly, lack a characteristic N-terminal signal sequence (Macnab, 1992), and are thought to be transported by a process resembling type III secretion (Ohnishi *et al.*, 1997; Stephens *et al.*, 1997). Evidence has accumulated for a number of proteins to be involved in this translocation, including FliI, FliH, FliN and FlhA (Vogler *et al.*, 1991). Others, like FlhB, FliP, FliQ and FliR are suspected to be part of this apparatus as well (Minamino *et al.*, 1994; Malakooti *et al.*, 1994; Zhuang & Shapiro, 1995). However, their precise functions are not yet understood, and some components may also contribute to other structural units of the flagellum. The FliP protein, for example, is alternatively thought to be part of the basal body (Malakooti *et al.*, 1994). It contains an N-terminal signal sequence, but is nevertheless presumed to be localized in the cytoplasmic membrane (Ohnishi *et al.*, 1997). The suspected export apparatus component FlhB is known to play a role in length control of the hook as well (section 1.3.1.2).

FliI has been the subject of extensive studies in *S. typhimurium*. Its copy number per bacterial cell has been determined to be approximately 1,500 (Dreyfus *et al.*, 1993). Its amino acid sequence displayed a significant 29% identity to the catalytic β subunit of the bacterial F_0F_1 proton-translocating ATPase (Vogler *et al.*, 1991). It contained the two Walker sequence boxes A and B, motifs that are suggested to be involved in ATP binding (Walker *et al.*, 1982). Purified *S. typhimurium* FliI, which was heterologously expressed in *E. coli*, was shown to bind ATP (Dreyfus *et al.*, 1993). *S. typhimurium fliI* mutants, the flagella of which had been sheared off, regained motility only very slowly compared to the wild type cells (Vogler *et al.*, 1991). These data suggested FliI to be a flagellum-related ATPase. Its ATP-hydrolyzing activity was subsequently demonstrated *in vitro*, although the mechanism of its action seemed to be considerably different from known ATPase types (Fan & Macnab, 1996). In contrast to F_0F_1 ATPases, the His-tagged purified FliI protein was active as a monomer. Its activity was not affected by inhibitors of F-, V- or P-type ATPases (Fan & Macnab, 1996). As

expected, Mg^{2+} ions were essential for protein function. The C-terminal domain of FliI was found to be responsible for ATP hydrolysis, whereas the N-terminal part probably constitutes interaction sites engaged in an as yet undefined flagellum-specific process (Fan & Macnab, 1996). The gene for FliI has been identified in a number of organisms, including *Bacillus subtilis* (Albertini *et al.*, 1991), *R. sphaeroides* (Ballado *et al.*, 1996) and *B. burgdorferi* (Ge *et al.*, 1996). In all of these, it is located in a cluster of flagellum-related genes including *fliG* (the gene for a switch protein), *fliH* (presumably involved in flagellar export), *fliJ* (a putative chaperone implicated in flagellar export), and *fliK* (a hook length control protein gene). The presumptive FliI proteins all contained the reported Walker boxes.

FliQ, another protein suspected to be involved in flagellum-specific transport, is a small, very hydrophobic membrane protein. Its gene has been analyzed in a number of organisms, including *B. subtilis* (Bischoff *et al.*, 1992), *E. coli* (Malakooti *et al.*, 1994), *C. crescentus* (Zhuang & Shapiro, 1995), *Treponema pallidium* (Hardham *et al.*, 1995) and *S. typhimurium* (Ohnishi *et al.*, 1997). In the genomes of these organisms, *fliQ* is located upstream from *fliR* and (except in *C. crescentus*) preceded by *fliP*. In *S. typhimurium*, FliQ was predicted to span twice through the inner membrane, with both C- and N-termini facing the cytoplasm of the bacterial cell (Ohnishi *et al.*, 1997). Knockout *fliQ* mutations in *C. crescentus* disrupted not only flagellar biogenesis, but also inhibited cell division (Zhuang & Shapiro, 1995). This observation is consistent with the assumed coregulation of *Caulobacter* class II flagellar genes with other cell cycle cues (Zhuang & Shapiro, 1995). The concept of shared regulatory pathways in *C. crescentus* was illustrated by sensitivity of expression of a class II flagellar gene to interrupted DNA replication (Dingwall *et al.*, 1992).

The most interesting feature of all the assumed flagellum-specific export apparatus components is their striking similarity to known virulence factor presenting proteins in a variety of organisms. This includes, among others, proteins of the *spa* (surface presentation of invasion plasmid antigens) locus of *S. flexneri* (Venkatesan *et al.*, 1992), the *hrp* (hyposensitive response and pathogenicity) clusters of the plant pathogen *Xanthomonas campestris* (Fenselau *et al.*, 1992) and *P. syringae* (He, 1997), proteins of the *Yersinia pestis lcr* (low calcium response) system suspected to be involved in the export of the Yop antigens (Plano *et al.*, 1991), the invasion gene family *inv* of *S. typhimurium* (Galan *et al.*, 1992), and the *mop* (mobility and pathogenicity) gene cluster in *Erwinia carotovora* (Mulholland *et al.*, 1993). Some homologs of components of the flagellar export apparatus with genes involved in virulence are summarized in Table 1.2.

Table 1.2. Presumed homologs in type III secretion systems and the flagellar protein export apparatus (adapted from Lee, 1997).

Flagellar system	Virulence system			
Export apparatus candidate	*Yersinia*	*Shigella*	*Salmonella*	*Ralstonia*
FlhA	LcrD	MxiA[b]	InvA	HrpO
FliI	YscN[a]	Spa47/L	InvC/SpaL	HrpE
FliP	YscR	Spa24/P	SpaP	HrpT
FliQ	YscS	Spa9/Q	SpaQ	HrpU
FliR	YscT	Spa29/R	SpaR	HrpC
FlhB	YscU	Spa40/S	SpaS	HrpN
FliN	YscQ	Spa33/O	SpaO	HrpQ
FliJ	-	Spa13/M	InvI/SpaM	-

[a] Ysc = abbreviation for Yop secretion (Wattiau *et al.*, 1996)

[b] Mxi = abbreviation for membrane expression of invasion plasmid antigens (Andrews *et al.*, 1991)

Homologies are surprizingly high, and can approach up to 50% amino acid identity (HrpB6 of *Xanthomonas campestris* to FliI of *S. typhimurium*) (Dreyfus *et al.*, 1993). All these homologous components are known or suspected to be part of secretion mechanisms for antigens without an N-terminal signal sequence. It is tempting to suggest a superfamily of those translocation proteins which includes the flagellar-related candidates.

However, in spite of striking similarities, functional complementation between members of type III secretion and flagellar export systems has not been reported. An involvement of the presumptive flagellum-specific protein export components in secretion processes of other antigens in pathogenic bacteria is still only speculation. Is the flagellum-specific pathway exploited by other virulence factors, or *vice versa*? No publication has yet described an effect of a knockout mutation of a presumed flagellum-specific export component (like *fliI*) on the surface presentation of known antigens suspected to be secreted by a type III mechanism.

1.3.3 Unique aspects of *H. pylori* flagellum biology

The polar flagella of *H. pylori* are important factors for colonization and persistence of the microorganism in the gastric mucosa (Eaton *et al.*, 1992). Three structural flagellar genes and their products have been studied so far: the major flagellin FlaA (Leying *et al.*, 1992), the minor flagellin FlaB, which is preferentially located in the proximal part of the filament (Suerbaum *et al.*, 1993), and the hook protein FlgE

(O'Toole *et al.*, 1994). The two significantly varying flagellins are expressed under the control of different promoters. Transcription of the *flaA* gene is presumably governed by a σ^{28} factor recognizing the -35/-10 sequence TAAA-N_{13}-ACCGATA (Leying *et al.*, 1992). A σ^{54} signature sequence TGGAACR-N_4-TTGCT (R = G or A) was shown to precede the *flaB* transcription start site (Suerbaum *et al.*, 1993), closely meeting the established consensus of TGGYRYR-N_4-TTGCA (Y = T or C) (Thöny & Hennecke, 1989). The two flagellins are not genetically linked, and share only 58% identical amino acids (Suerbaum *et al.*, 1993), which is in contrast to the closely related *Campylobacter* species (Alm *et al.*, 1993). It has been hypothesized, that *H. pylori* cells can adapt the mechanical properties of the filament by varying the ratio of FlaA to FlaB in response to environmental signals (Suerbaum, 1995). Experiments using the piglet model demonstrated that expression of both *H. pylori* flagellin species is necessary for full colonization of the gastric mucosa. Knock out of either one of these genes resulted in less efficient colonization of the gastric mucosa with *H. pylori* followed by complete eradication of the bacterium after ten days (Eaton *et al.*, 1996).

The hook protein of the *H. pylori* flagellum has also been studied extensively. Isogenic mutants lacking the *flgE* gene were aflagellate and nonmotile (O'Toole *et al.*, 1994). Transcription of the gene was postulated to be controlled by a σ^{54} promoter (O'Toole *et al.*, 1994). Surprisingly, *H. pylori flgE* mutants were not compromised in flagellin production, indicating a lack of the key negative regulatory element FlgM in this bacterium (O'Toole *et al.*, 1994), and suggesting a different flagellar biogenesis process regulation. However, the gene *flbA* has been described as a regulatory element in *H. pylori* (Schmitz *et al.*, 1997). This gene is similar to *flhA*, a flagellar class II gene found in a number of bacteria including *C. crescentus* (Ramakrishnan *et al.*, 1991), *C. jejuni* (Miller *et al.*, 1993), *B. subtilis* (Carpenter & Ordal, 1993), and *P. mirabilis* (Gygi *et al.*, 1995). It has been shown to be indispensable for coordinated flagellin and FlgE expression. An isogenic *flbA* mutant had virtually no flagellin production and reduced FlgE expression in early and middle growth phase (Schmitz *et al.*, 1997).

The annotated genome sequence of *H. pylori* 26695 contains at least 40 genes likely to be needed for flagellar function and/or assembly (Tomb *et al.*, 1997). This number approximates the complement of known flagellar related genetic elements in *S. typhimurium* (Aizawa, 1996), *E. coli* (Blattner *et al.*, 1997), and *B. subtilis* (Kunst *et al.*, 1997). However, genetic organization of these elements in *H. pylori* differs dramatically from that found in those bacteria. Whereas many flagellar genes appear to be clustered in well defined genomic regions in these organisms, they are relatively evenly distributed on

the *H. pylori* genome. This obviously requires a regulation mechanism different from simple operon-based cotranscription. Furthermore, *H. pylori* exhibits a number of other atypical features in its flagellar genetics. Some well established flagellar genes could not be identified in the *H. pylori* 26695 genome. This includes the master operon components of *S. typhimurium, flhC* and *flhD*, the genes for a hook length controller, *fliK*; *flgA*, the product of which is considered to be involved in P-ring assembly; *flgF*, which encodes a basal-body rod protein; and *flgL*, encoding the hook-filament junction protein in *S. typhimurium.* As expected from experiments with isogenic *H. pylori flgE* mutants, and unsuccessful attempts to identify a homolog by low stringency hybridization (O'Toole *et al.*, 1994), a counterpart for the *S. typhimurium* anti-σ factor FlgM was not identified in *H. pylori* 26695 by genomic sequencing. A gene with strong similarity to the master regulator for flagellar gene expression in *C. crescentus*, *ctrA*, could also not be annotated in the *H. pylori* 26695 genome.

However, genes for several additional flagellar proteins exist in the *H. pylori* genome. Most surprising is the existence of second homologs for the minor flagellin and the hook protein. Furthermore, two genes with similarity to the *C. crescentus*-specific regulators of flagellin expression, *flmB* and *flmH*, are also present. In addition to pecularities possibly specific to a gastric pathogen, the altered complement of flagellar genes may be a feature of bacteria using unipolar flagella for motility. An extended subset of structural flagellar genes, in combination with a sheath, may also help the bacterium to evade the mucosal immune response of the host.

H. pylori flagella are sheathed with a relatively acid-resistant membranous structure (Suerbaum, 1995). This sheath is suspected to contain one or more structures responsible for the adhesive properties of the bacterium. An *H. pylori* lipoprotein HpaA has been reported in several strains to cause agglutination of erythrocytes (Evans *et al.*, 1993). An identical component was found to be part of the flagellar sheath, but apparently lacked adhesive properties (O'Toole *et al.*, 1995; Jones *et al.*, 1997).

Overall, our knowledge of transcriptional and translational regulation of the *H. pylori* motility apparatus is surprisingly low, considering the enormous importance of this device for the virulence of the bacterium. A lot of research still has to be done in this field to gain a deeper insight into the adaptation mechanisms performed by *H. pylori* and other gastric pathogens to create a functional motility system, which is perfectly constructed for the special environmental conditions the bacterium encounters.

1.4 Aims of this study

The process of flagellar biosynthesis in Gram-negative bacteria is still far from being understood. Extensive studies on model organisms like *S. typhimurium* and *C. crescentus* showed that the subset of proteins involved in flagellar biosynthesis is not completely conserved. Investigations in other bacteria also revealed a number of differences in their flagellar biology to these two microorganisms. Transcriptional regulation mechanisms may vary among and even within prokaryotic families.

Previous work indicated the lack of a common key element in flagellar gene regulation in the human pathogen *H. pylori* (O'Toole *et al.*, 1994), suggesting a regulatory mechanism different from the one described in the model organisms. The suspected differences in flagellar gene regulation and function are most likely an expression of the adaptation of this small-genome bacterium to its ecological niche. In parallel with the ensuing academic interest in flagellar biosynthesis of *H. pylori*, study of its flagellar apparatus was also justified by the fact that motility was an experimentally proven colonization factor of the bacterium.

At the beginning of this study, knowledge about components of the flagellar export apparatus in *H. pylori* was virtually non-existent. Considering the presumed potential of these proteins to export virulence factors as well, study of this molecular device gained even more relevance in a human pathogen.

One of the components presumed to be part of the flagellum-specific export apparatus in a number of bacteria is the FliI protein. A segment of *H. pylori* 17874 chromosomal DNA with significant sequence similarity to bacterial *fliI* genes had been cloned in Dr O'Toole's laboratory. The initial goals of this project were therefore identified as:

(i) cloning and sequencing of the complete *fliI* gene of *H. pylori* 17874,
(ii) evaluation of the genetic diversity of the *fliI* region in a panel of clinical *H. pylori* isolates,
(iii) investigation of the transcriptional regulation of *fliI*,
(iv) knockout mutagenesis of *fliI*, and finally
(v) phenotypic characterization of the established *H. pylori fliI* mutant strain.

However, it soon became apparent that genetic elements found adjacent to the *H. pylori* 17874 *fliI* gene could be included, thereby widening the scope of investigations

and insights, and bringing further aspects of *H. pylori* flagellar biology into the focus of this study. Molecular characterization of two additional genes (*H. pylori* 17874 *fliQ* and an element with significant similarities to genes involved in nucleoprotein transfer, *virB11*) complemented the original list of aims.

2. MATERIALS AND METHODS

2.1 Bacterial strains, culture and storage conditions

Bacterial strains used in this study are listed in Table 2.1.

Table 2.1. Bacterial strains.

Bacterial strain	Characteristics or genotype	Source / reference
H. pylori CCUG 17874	*H. pylori* type strain, *cagA*$^+$ /VacA$^+$ (type I), identical to NCTC 11637	culture collection University of Gothenburg, Sweden
H. pylori CCUG 915	*H. pylori* type strain, *cagA*$^-$ /VacA$^-$ (type II)	culture collection University of Gothenburg, Sweden
H. pylori SS1	mouse adapted clinical isolate, Sydney, Australia, *cagA*$^+$ /VacA$^+$ (type I)	(Lee *et al.*, 1997a)
H. pylori MU002	clinical isolate,Auckland, *cagA*$^+$ /VacA$^+$ (type I)	(Campbell *et al.*, 1997)
H. pylori MU003	clinical isolate, Auckland, *cagA*$^+$ /VacA$^+$ (type I)	(Campbell *et al.*, 1997)
H. pylori MU007	clinical isolate, Auckland, *cagA*$^+$ /VacA$^+$ (type I)	(Campbell *et al.*, 1997)
H. pylori MU015	clinical isolate, Auckland, *cagA*$^-$ /VacA$^-$ (type II)	(Campbell *et al.*, 1997)
H. pylori MU016	clinical isolate, Auckland, *cagA*$^+$ /VacA$^+$ (type I)	(Campbell *et al.*, 1997)
H. pylori MU022	clinical isolate, Auckland, *cagA*$^+$ /VacA$^+$ (type I)	(Campbell *et al.*, 1997)
H. pylori MU030	clinical isolate, Auckland, *cagA*$^+$ /VacA$^+$ (type I)	(Campbell *et al.*, 1997)
H. pylori MU038	clinical isolate, Auckland, *cagA*$^-$ /VacA$^+$ (intermediate)	(Campbell *et al.*, 1997)
H. pylori MU043	clinical isolate, Auckland, *cagA*$^+$ /VacA$^+$ (type I)	(Campbell *et al.*, 1997)
H. pylori MU044	clinical isolate, Auckland, *cagA*$^+$ /VacA$^+$ (type I)	(Campbell *et al.*, 1997)
H. pylori MU045	clinical isolate, Auckland, *cagA*$^+$ /VacA$^+$ (type I)	(Campbell *et al.*, 1997)
H. pylori MU067	clinical isolate, Auckland, *cagA*$^+$ /VacA$^-$ (intermediate)	(Campbell *et al.*, 1997)
H. pylori MU074	clinical isolate, Auckland, *cagA*$^+$ /VacA$^+$ (type I)	(Campbell *et al.*, 1997)
H. pylori MU079	clinical isolate, Auckland, *cagA*$^+$ /VacA$^+$ (type I)	(Campbell *et al.*, 1997)
H. pylori MU082	clinical isolate, Auckland, *cagA*$^+$ /VacA$^-$ (intermediate)	(Campbell *et al.*, 1997)
H. pylori MU099	clinical isolate, Auckland, *cagA*$^-$ /VacA$^-$ (type II)	(Campbell *et al.*, 1997)
H. pylori MU102	clinical isolate, Auckland, *cagA*$^+$ /VacA$^+$ (type I)	(Campbell *et al.*, 1997)
H. pylori MU103	clinical isolate, Auckland, *cagA*$^+$ /VacA$^+$ (type I)	(Campbell *et al.*, 1997)
H. mustelae 4298	laboratory-passaged strain	J.G. Fox, Division of Comparative Medicine, Massacusetts Institute of Technology, Cambridge, Ma., USA
E. coli ER2206	*endA1 thi1 supE44 mcr67 (mcrA$^-$) Δ(mcrBC-hsdRMS-mrr)114::IS10 (lac)U169/F' proAB lacIqZDM15 Tn10* (Tet^r)	New England Biolabs

Helicobacter cultures were grown at 37°C on Chocolate blood agar (CBA, section 2.2), or on *Helicobacter* motility agar (section 2.2) in an atmosphere containing 5% CO_2 maintained by a CO_2 incubator (Revco Scientific). Alternatively, liquid cultures were grown under agitation in Tryptic soy broth (TSB, section 2.2) in microaerobic conditions generated by CampyGen sachets (Oxoid). Stock cultures were stored at -70°C in TSB containing 30% glycerol.

Escherichia coli was cultured on Luria-Bertani agar (LBA, section 2.2), or in Terrific broth (TB, section 2.2) on a shaker at 200 rpm at 37°C. Stocks were maintained at -70°C in Luria-Bertani broth (LB, section 2.2) containing 35% glycerol.

2.2 Media and supplements

Media prepared for bacterial culture during this study are listed in Table 2.2. All media were made using water which had been purified by the MilliQ Reagent Water System (Millipore). They were subsequently sterilised at 121°C for 20 min. Supplements and antibiotics added to the growth media when required are illustrated in Table 2.3.

Table 2.2. Media.

Medium	Receipe	Reference / supplier
Luria-Bertani broth (LB)	1% (w/v) tryptone (Difco), 0.5% (w/v) yeast extract (Merck), 0.5% (w/v) NaCl; pH 7.0 with NaOH	(Sambrook *et al.*, 1989)
Luria-Bertani agar (LBA)	1% (w/v) tryptone (Difco), 0.5% (w/v) yeast extract (Difco), 0.5% (w/v) NaCl; pH 7.0 with NaOH, 1.5% (w/v) bacto-agar (Oxoid)	(Sambrook *et al.*, 1989)
Terrific broth (TB)	1.2% (w/v) tryptone (Difco), 2.4% (w/v) yeast extract (Merck), after sterilization add 0.4% (final concentration) glycerol and 17 mM (final concentration) KH_2PO_4 / 72 mM (final concentration) K_2HPO_4 (sterile)	(Sambrook *et al.*, 1989)
Tryptic soy broth (TSB)	3% (w/v) tryptic soy broth (Difco)	Difco
Helicobacter motility agar	3% (w/v) tryptic soy broth (Difco), 0.3%-0.4% (w/v) bacto-agar (Oxoid), after sterilization cool down to 60°C, add 5% (final concentration) defibrinated horse serum (Life Technologies)	Difco, (Josenhans *et al.*, 1995)
Chocolate blood agar (CBA)	4.4% (w/v) columbia blood agar base (Difco), after sterilization cool down to 70°C, add 5% (final concentration) defibrinated horse blood (Life Technologies), stir until hemolysis occurs (brown coloration)	Difco

Table 2.3. Media supplements and antibiotics.

Supplement	Stock concentration [mg/ml]	Final concentration [µg/ml] for:	
		H. pylori	*E. coli*
Ampicillin	100	n/a[e]	100
Tetracycline	5[c]	n/a	5
Kanamycin	50	25	50
Chloramphenicol	10[c]	10	10
IPTG[a]	240	n/a	48
X-gal[b]	20[d]	n/a	32

[a] Isopropyl-β-D-thiogalactopyranoside

[b] 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

[c] dissolved in ethanol

[d] dissolved in dimethylformamide, wrapped in foil for storage

[e] not applicable

Stock solutions of antibiotics and additives were kept at -20°C in sterile MilliQ-treated water, unless otherwise noted. When added to agar, the supplements and antibiotics were mixed with liquified media cooled to 55°C immediately before pouring.

2.3 Oligonucleotide primers

Oligonucleotides used during this study, their relative priming orientations, positions and applications are listed in Table 2.4. Excluding pUC/M13 FP and pUC/M13 RP, which were from Promega, primers were obtained from Life Technologies in deprotected form. No special purification of the oligonucleotide was requested, unless otherwise noted.

Table 2.4. Oligonucleotide primers.

Primer	Sequence 5' → 3'	Priming position	Application
pUC/M13 FP	gtt ttc cca gtc acg ac	pUC18/19; ←	*fliI* and *fliQ* mutagenesis; pSP102 sequencing
pUC/M13 RP	cag gaa aca gct atg ac	pUC18/19; →	pSP102 sequencing
HS16sF	agg cta tga cgg gta tcc ggc	*Helicobacter*-specific 16sRNA sequence; →*	operon conservation studies
HS16sR	ggt cta gca agc tag aca ctc c	*Helicobacter*-specific 16sRNA sequence; ←*	operon conservation studies
SP002	caa atc atc gcc ctc tac tag c	internal *fliI*; ←	*fliI* probe preparation; pSP102 sequencing; *fliI* transcript detection; fidelity test pHP042
SP003	atc cgc act ttc ctt att gcc	internal *fliI*; →	*fliI* probe preparation; fidelity test pSP102 and pHP042; pSP102 sequencing
SP004	aat acg cca acc act aac tgg	5' external of *fliI* fragment on pHP042	fidelity test pHP042
SP007	acg gat cca agt ttc aac cat tag gag aaa ggg	*H. pylori hspA* $aa_{2\text{-}10}$ + *Bam*HI tail	PCR amplification *hspA*
SP008	tgg gat cct tag tgt ttt ttg tga tca tga cag c	*H. pylori hspA* $aa_{STOP\text{-}111}$ + *Bam*HI tail	PCR amplification *hspA*
SP009	ccc atc aat gct ctt cac c	internal *fliI*; ←	*virB11* mutagenesis; conservation studies; pSP102 sequencing
SP010	gtg ccg att ttc aca ctg c	near *murB* 5' end; ←	fidelity test pSP102; transcript analysis; pSP102 sequencing
SP011	tga gaa cgc ttt gca gcc	near *fliI* 3' end, →	pSP102 sequencing
SP012	ggt aaa tcc acg cta atg gg	internal *fliI*; →	pSP102 sequencing
SP013	gtg ctt aac cct tta ggg c	internal *fliI*; →	operon conservation studies; characterization of *H. pylori fliI* mutants; *fliI* transcript detection; pSP102 sequencing
SP014	gca cga tca cat caa tcg c	near *virB11* 3' end; ←	characterization of *H. pylori virB11* mutants; *virB11* transcript detection; pSP102 sequencing
SP015	gtg tgc cta ggc tta agc	internal *virB11*; →	*fliQ* mutagenesis; transcript analysis; *virB11* transcript detection; pSP102 sequencing
SP016	gca cac gga tgt tga gcg	internal *virB11*; ←	transcript analysis; pSP102 sequencing
SP017	gac caa gcg tta ttt ggc	near ORF02 3' end; →	pSP102 sequencing

Table 2.4 continued

SP018	aac gcc aag tgg att gag g	internal ORF02; ←	pSP102 sequencing
SP019	gaa acg cca gca ttc acg	near *ileS* 3' end; →	*fliI* mutagenesis; genotype characterization of *H. pylori virB11* mutants; pSP102 sequencing
SP020	cca aac cag aat gat ggc c	near ORF02 5' end; →	transcript analysis; pSP102 sequencing
SP021	cca tcg tgg atc cgg gct gca tga cag gcg cta atc gc	internal *fliI*; ←; 5' tail sticky to SP022, *Bam*HI site	*fliI* mutagenesis
SP022	gca gcc cgg atc cac gat ggg aca tca tca gcg agt ctc	internal *fliI*; →; 5' tail sticky to SP021, *Bam*HI site	*fliI* mutagenesis
SP023	cca tcg tag atc tgg gct gcg cga gtt tca tga gtt gtg	near *fliQ* 5' end; ←; 5' tail sticky to SP024, *Bgl*II site	*fliQ* mutagenesis; characterization of *H. pylori fliI* mutants
SP024	gca gcc cag atc tac gat gga ccg gta tta cta gcg gg	near *fliQ* 5' end; →; 5' tail sticky to SP023, *Bgl*II site	*fliQ* mutagenesis
SP025	tgc tct tgc cat ttt gcg	near *ileS* 5' end; ←	pSP102 sequencing
SP026	ggc cat cat tct ggt ttg g	near ORF02 3' end; ←	pSP102 sequencing
SP027	gga atg agc ttg att aag g	near *fliQ* 3' end; ←	transcript analysis; characterization of *H. pylori fliQ* mutants; *fliQ* transcript detection
SP028	cac aac tca tga aac tcg c	near *fliQ* 5' end; →	transcript analysis; characterization of *H. pylori fliQ* mutants; *fliQ* transcript detection
SP032	cat cgt aga tct ggg ctg cgt tgc tcg caa aac ctc agc	internal *virB11*; ←; 5' tail sticky to SP033, *Bgl*II site	*virB11* mutagenesis
SP033	gca gcc cag atc tac gat gga tga aag ggt ggt gag cg	internal *virB11*; →; 5' tail sticky to SP032, *Bgl*II site	*virB11* mutagenesis
SP034	agg gag cga tca gca tag g	internal *ileS*; →	*virB11* mutagenesis; transcript analysis; conservation studies
SP035	gca gcc cag atc tac gat ggg gcg aat gat tta gag gg	internal *murB*; →; 5' *Bgl*II site	*murB* transcript detection
SP036	tgc tct ttg gca aag ccc	internal *murB*; ←	transcript analysis; *murB* transcript detection; conservation studies
SP039	gtg ggt tta gtg aag cgg	near ORF03 5' end; →	transcript analysis
SP040	tct tcg cat gat ctc ggc	near *ileS* 5' end; ←	transcript analysis
SP043	agc aac cct act tat ggg	near TIGR HP1417 3' end; ←	conservation studies
SP045	agg ttg ttg cag cac tag c	near *hpn* 3' end; →	conservation studies
SP048[a]	aac gcg ccg ctt cac taa acc	near ORF03 5' end; ←	primer extension
SP050[a]	act agc ctt agc gca act ccc	internal ORF03; ←	primer extension
SP051[a]	tgg ttg tgt tta agt tta ggg tgt c	near *ileS* 5' end; ←	primer extension

←, → priming orientations (relative to Fig. 3.5, except *, relative to gene orientation)

[a] purified by polyacrylamide gel electrophoresis

2.4 Vectors and recombinant plasmids

Vectors and recombinant plasmids constructed and used in this work are summarized in Table 2.5. See Appendix 1 for plasmid maps of constructs of the pHP and pSP series, and of p26K.

Table 2.5. Plasmids used in this study.

Plasmid	Characteristics	Reference / source
pUC18	Ap^r; colE1 ori; blue/white; MCS	(Yanisch-Perron *et al.*, 1985)
pUC19	Ap^r; colE1 ori; blue/white; MCS	(Yanisch-Perron *et al.*, 1985)
pBK-CMV	Neo^r; Kan^r; SV40, f1(-) and colE1 ori; blue/white; MCS	Stratagene
pHP042	Neo^r; Kan^r; 3.5 kb genomic fragment of *H. pylori* 17874 in pBK-CMV	(Beddek, 1998)
pRY109	Ap^r; Cm^r; *Campylobacter coli cat* gene in *Bam*HI site of pUC18	(Yao *et al.*, 1993)
pILL600	Kan^r; pBR322 ori; *aphA-3* of *C. coli* plasmid pIP1433 in construct containing pBR322 and *C. coli* pIP1455 sequence	(Labigne *et al.*, 1992)
pUC4ΩKm-2	Ap^r; Kan^r; colE1 ori; blue/white; *Streptococcus faecalis aphA*-3 gene in Ω element cloned in pUC4	(Perez-Casal *et al.*, 1991)
pSPT18	Ap^r; MCS; colE1 ori; SP6/T7 RNA polymerase transcription initiation sites	Boehringer Mannheim
pGEM-T	Ap^r; f1(-) and colE1 ori; blue/white; MCS; SP6/T7 RNA polymerase transcription initiation sites; 3' T overhangs at insertion site for ligation of PCR products	Promega
pHP012	Neo^r; Kan^r; 4 kb genomic fragment of *H. pylori* containing 5' end of *cdh* and complete *flgE* in pBK-CMV	(O'Toole *et al.*, 1994)
pHP014	Ap^r; 2.2 kb genomic fragment of *H. pylori* containing the 3' half of *flgE* in pUC18	(O'Toole *et al.*, 1994)
p26K	Neo^r; Kan^r; colE1 ori; 0.9 kb genomic *Hin*dIII fragment of *H. pylori* containing *tsaA* in pK18	(O'Toole *et al.*, 1991)
pSP101	Ap^r; 6.3 kb *Hin*dIII/*Bgl*II genomic fragment of *H. pylori* 17874 in *Hin*dIII/*Bam*HI cut pUC18	this work
pSP102	Ap^r; 7 kb *Hin*dIII/*Bgl*II genomic fragment of *H. pylori* 17874 in *Hin*dIII/*Bam*HI cut pUC18	this work
pSP104	Ap^r; 1.4 kb *SpHI*/*Hin*dIII fragment from pSP102 in pUC18	this work
pSP105	Ap^r; composite *fliQ* PCR fragment in *Sma*I cut pUC19	this work
pSP106	Ap^r; composite *fliI* PCR fragment in *Sma*I/*Hin*cII cut pUC19	this work
pSP107	Ap^r; Cm^r; *cat* gene from pRY109 in primer introduced *Bgl*II site of pSP105	this work
pSP108	Ap^r; Kan^r; *aphA*-3 cassette from pUC4ΩKm-2 in primer introduced *Bam*HI site of pSP106	this work
pSP109	Ap^r; Kan^r; *aphA*-3 cassette from pUC4ΩKm-2 in primer introduced *Bgl*II site of pSP105	this work
pSP110	Ap^r; Cm^r; *cat* gene from pRY109 in primer introduced *Bam*HI site of pSP106	this work
pSP112	Ap^r; PCR fragment SP002/013 (*fliI* probe) in *Sma*I digested pUC19	this work
pSP114	Ap^r; insert of pSP201 (*H. pylori* 17874 *hspA*) in *Bam*HI digested pSPT18	this work
pSP115	Ap^r; insert of pSP112 in *Eco*RI/*Hin*dIII cut pSPT18	this work
pSP116	Ap^r; composite *virB11* PCR fragment in *Sma*I cut pUC19	this work
pSP117	Ap^r; Kan^r; *aphA*-3 cassette from pILL600 in primer introduced *Bgl*II site of pSP116	this work
pSP118	Ap^r; Cm^r; *cat* gene from pRY109 in primer introduced *Bgl*II site of pSP116	this work
pSP119	Ap^r; 0.6 kb *SpH*I fragment from pHP014 (*flgE* probe) in pUC19	this work
pSP120	Ap^r; 0.45 kb *Eco*RI/*Hin*dIII fragment from pSP119 (*flgE* probe) in pSPT18	this work
pSP201	Ap^r; PCR fragment SP007/008 (*H. pylori hspA*) in pGEM-T	this work

2.5 Antisera

Table 2.6 summarizes the antisera employed in this study.

Table 2.6. Antibodies.

Name	Function	Source	Relevant characteristics
anti-DIG F_{ab}-alkaline phosphatase conjugate	detection in Southern and Northern blots	DIG Nucleic Acid Detection Kit (Boehringer Mannheim)	pcl[a] ab[b] raised against digoxygenin
anti-rabbit IgG-alkaline phosphatase conjugate	secondary ab in Western blots	Sigma	goat pcl ab raised against rabbit IgG
anti-mouse IgG-horseradish peroxidase conjugate	secondary ab in Western blots	Serotec	sheep pcl ab raised against mouse IgG
anti-CagA	primary ab in Western blots	T. Cover, Vanderbilt University, Nashville, Tn, USA	rabbit pcl ab raised against *H. pylori* CagA
anti-VacA	primary ab in Western blots	T. Cover, Vanderbilt University, Nashville, Tn, USA	rabbit pcl ab raised against *H. pylori* VacA
anti-Fla	primary ab in Western blots	P.W. O'Toole, Massey University, Palmerston North, New Zealand	rabbit pcl ab raised against *H. pylori* flagellin A (Kostrzynska *et al.*, 1991)
anti-HopB	primary ab in Western blots	P. Doig, Astra Research Center, Boston, Ma., USA	mouse mcl[c] ab raised against *H. pylori* HopB
anti-UreB	primary ab in Western blots	P. Doig, Astra Research Center, Boston, Ma., USA	mouse mcl ab raised against *H. pylori* UreB
anti-OMP4	primary ab in Western blots	P. Doig, Astra Research Center, Boston, Ma., USA	mouse mcl ab raised against *H. pylori* OMP4

[a] pcl = polyclonal
[b] ab = antibody
[c] mcl = monoclonal

2.6 DNA isolation

2.6.1 Plasmid preparation

2.6.1.1 Easy plasmid minipreparation

The Easy Plasmid Minipreparation was performed as described by Berghammer and Auer (Berghammer & Auer, 1993). Briefly, 1 ml of overnight *E. coli* TB culture was pelleted by centrifugation at 14,000 rpm (20,800 x g) in a bench-top centrifuge 5417C (Eppendorf) at RT for 30 sec, the supernatant aspirated and the pellet resuspended in 60 to 80 µl of lysis buffer {10 mM Tris-Cl pH 8.0 / 1 mM EDTA / 15% (w/v) sucrose / 2 mg/ml lysozyme / 0.2 mg/ml RNase A / 0.1 mg/ml bovine serum albumine (BSA), stored

at -20°C}. The suspension was shaken vigorously for 10 min at RT, then boiled for 60 sec, chilled on ice for 60 sec and immediately spun at 20,800 x g for 20 min at RT. The supernatant containing the plasmid was transferred into a new tube and subsequently analysed.

2.6.1.2 Wizard plasmid minipreparation

For preparation of plasmid DNA of higher purity, the Wizard™ Plus Minipreps DNA Purification System (Promega) was used according to the manufacturers instructions. This method was based on alkaline lysis of the bacterial cells followed by binding of plasmid DNA to a resin material. The captured DNA molecules were eluted with 50 µl sterile MilliQ-treated water or TE {10 mM Tris-HCl pH 7.5 / 1 mM EDTA}. Routinely, 1.7 ml overnight *E. coli* TB culture containing a high copy plasmid yielded in 15 µg pure DNA.

2.6.1.3 High Pure plasmid preparation

The High Pure™ Plasmid Isolation Kit (Boehringer Mannheim) was used as an alternative preparation method to the Wizard plasmid Minipreparation. In this method, bacterial cells were lysed, and the plasmid DNA bound in the presence of a chaotropic salt selectively to glass fiber fleece. It was subsequently eluted with sterile MilliQ-treated water or TE {10 mM Tris-HCl pH 7.5 / 1 mM EDTA}. A 1.7 ml overnight *E. coli* TB culture routinely resulted in 10 µg pure DNA of high copy plasmids when following the manufacturer's instructions.

2.6.1.4 ABI plasmid preparation

This method resulted in high yields of plasmid DNA of superior purity for subsequent sequence analysis, and is based on a recommendation of the manufacturer (in: *Taq* DyeDeoxy™ Terminator Cycle Sequencing Kit, Applied Biosystems). Briefly, 1.7 ml overnight *E. coli* TB culture were spun for 30 sec at 20,800 x g, RT. The supernatant was aspirated and the pellet resuspended in 200 µl TE {10 mM Tris-HCl pH 7.5 / 1 mM EDTA}. For alkaline lysis, 200 µl of lysis buffer {0.2 M NaOH / 1% (w/v) SDS} was added, and the tube inverted, then stored at RT for 5 min or until the suspension became clear. Subsequently, 200 µl of icecold 3 M sodium acetate were added, the tube inverted

and stored for 5 min on ice followed by 15 min centrifugation at 20,800 x g, RT. The supernatant was transferred to a new tube and 2 µl of 10 mg/ml RNase solution {in 15 mM NaCl / 10 mM Tris-HCl pH 7.5, boiled for 15 min before storage at -20°C} was added. The suspension was incubated at 37°C for 20 min. Subsequently, two chloroform extractions were performed by adding 500 µl of $CHCl_3$, mixing, spinning of the samples for 2 min at 20,800 x g, RT and transferring of the upper phase into a new tube. The DNA was precipitated by adding 1 Vol of isopropanol, mixing and spinning of the sample for 20 min at 20,800 x g, RT. The pellet was washed with 1 ml 70% ethanol, spun for 5 min at 20,800 x g, RT and airdried. It was then resuspended in 32 µl of water, and for precipitation 8 µl of icecold 4 M NaCl and 40 µl 40% PEG were added. The suspension was incubated on ice for 30 min and then centrifuged at 7,600 x g, 4°C (MicroCentaur bench top centrifuge, MSE). Finally, the resulting pellet was washed as above with 70% ethanol, airdried and resuspended in 30 µl sterile water. Usually, approximately 20 µg of extremely pure DNA of high copy plasmids were prepared following this procedure.

2.6.2 Preparation of genomic DNA from *Helicobacter*

For preparation of genomic DNA from *Helicobacter* cells the method of Pitcher *et al.* (Pitcher *et al.*, 1989) was modified following suggestions of O'Toole *et al.* (O'Toole *et al.*, 1994). A heavily streaked plate of *Helicobacter* culture was harvested using a sterile swab, and the cells resuspended in 1 ml phosphate-buffered saline (PBS) {0.1 M Na_2HPO_4 / 0.1 M NaH_2PO_4 pH 6.8}. The cells were pelleted by centrifugation for 30 sec at 20,800 x g, RT, and resuspended in 100 µl TE {10 mM Tris-HCl pH 8.0 / 1 mM EDTA}. After addition of 500 µl GES {5 M guanidium thiocyanate / 0.1 M EDTA / 0.5% (w/v) sarkosyl} and 2 µl 10 mg/ml RNase solution the suspension was mixed until clearance occured, and subsequently incubated for 10 min at RT. For DNA precipitation, 1 ml of icecold 96% ethanol was added, and the immediately apparent genomic DNA cobweb removed with a clean glass hook. The DNA was briefly submerged in 70% ethanol and then resuspended in 150 µl TE. For deproteinization, 20 µl of a 50 mg/ml proteinase K solution was added, and the suspension incubated at 50°C for 3-20 h. The clean genomic DNA was recovered and washed as described above, and finally resuspended in 300-500 µl TE.

2.7 DNA analysis methods

2.7.1 DNA agarose gel electrophoresis

For DNA visualization and size fractionation, horizontal agarose gel electrophoresis was performed following standard methodology (Sambrook *et al.*, 1989) in a Horizon® 58 mini gel apparatus (Life Technologies) or a Minicell EC370M (Savant Instruments) at RT. The gels constituted 0.8% (w/v) or 1% (w/v) agarose in 1 x TAE {40 mM Tris / 19 mM glacial acetic acid / 1 mM EDTA}. The mixture was microwaved for total solubilization and cooled down to 55°C before pouring. DNA samples to be analyzed were mixed with at least 0.2 Vol DNA loading dye {0.5% (w/v) bromophenol blue / 0.5% (w/v) xylene cyanol / 50% glycerol / 50 mM EDTA} before loading. Electrophoresis was performed at 60-100 V for 30-90 min. Subsequently, the gel was stained in 5 µg/ml ethidium bromide solution for 5-20 min, then destained in water for at least 2 min. DNA was visualized using UV irradiation on the TMW-20 Transilluminator (Alpha Innotech). Images were captured by the IS-1000 Digital Imaging System (Alpha Innotech).

The molecular mass of DNA fragments was determined by direct comparison of their migration distance to that of a set of standard fragments of known size run on the same agarose gel. Standard molecular weight markers were the 1 kb ladder (Life Technologies), the 100 bp ladder (Life Technologies), or a bacteriophage λ *Eco*RI/*Hin*dIII ladder, prepared by complete restriction of λ DNA (New England Biolabs) at 37°C for 3 h using 2 U of both restriction endonucleases per µg DNA. The custom ladders were used following the instructions of the manufacturer.

2.7.2 DNA restriction endonuclease treatment

Restriction endonucleases used in this study were either from Boehringer Mannheim, Life Technologies, or New England Biolabs. DNA was digested for 90 min according to the manufacturer's recommenations using the appropriate supplied buffer and a slight excess of restriction enzyme. To ensure low glycerol concentrations, the volume of the added enzyme never exceeded 10% of the total reaction volume. In digests where two enzymes were acting simultaneously, a convenient buffer ensuring sufficient activity (> 50%) of both endonucleases was employed. If optimal reaction temperatures differed for the two enzymes, an hour incubation at the lower of the two temperatures was followed by an hour incubation at the higher optimal temperature. Completion of the

restriction process was checked by gel electrophoresis (section 2.7.1) of a small aliquot of the digestion mix.

2.7.3 DNA quantification

DNA concentrations in solutions were estimated by comparative fluorescence evaluation on an agarose minigel after electrophoresis and staining with ethidium bromide (section 2.7.1). Standard solutions were prepared by *Hin*dIII restriction of vector DNA (pUC18, New England Biolabs) of known concentration and subsequent dilution of the mix after completed endonuclease treatment, resulting in final DNA concentrations from 5 to 30 ng/µl. These standards were run simultaneously adjacent to the DNA sample of interest during electrophoresis, and the apparent fluorescence under UV light was compared after ethidium bromide staining.

For accurate determination of DNA concentrations the samples were analyzed spectrophotometrically using a Shimadzu UV-160A spectrophotometer and quartz cuvettes (Starna) with a 1 cm light path. Absorbance of the appropriately diluted sample solution was determined at light wavelengths of 260 nm and 280 nm. DNA concentrations were calculated by multiplying the dilution factor, a constant factor 50 (double stranded DNA) or 33 (single stranded nucleic acid), and the absorbance reading at 260 nm. Purity of the sample was evaluated by the ratio of the measurements at 260 nm and 280 nm. Values ≥ 1.8 were generated by pure samples of low protein contamination.

2.7.4 Southern blotting and hybridization

For detection of specific regions or fragments on sample DNA, Southern blotting followed by hybridization and detection of digoxigenin-labelled probes was employed, based on the method described by Southern (Southern, 1975).

DNA fragments were separated on conventional agarose gel electrophoresis (section 2.7.1). Subsequently, the gel was treated according to the DIG System User's Guide for Filter Hybridization (Boehringer Mannheim) to ensure depurination (twice 5 min in 0.32% HCl) and denaturation (twice for 20 min in 1.5 M NaCl / 0.5 M NaOH) of the DNA with subsequent neutralization (twice for 20 min in 1.5 M NaCl / 1 M Tris-HCl pH 8.0). An overnight Southern capillary DNA transfer to positively charged Hybond-N^+

membrane (Amersham) was performed as described by Sambrook *et al.* (Sambrook *et al.*, 1989) with 20 x SSC {3 M NaCl / 0.3 M sodium citrate} as transfer buffer.

Alternatively, slot blotting using the BIO-DOT™ microfiltration apparatus (Bio-Rad) was performed. Wizard plasmid preparations (section 2.6.1.2) were diluted one hundred seventy-fold in TE {10 mM Tris-HCl pH 8.0 / 1 mM EDTA}. Then, 0.6 Vol 20 x SSC and 0.2 M (final concentration) NaOH were added for denaturation of the DNA. The mix was incubated for 10 min at 80°C, followed by addition of 0.25 M (final concentration) Tris-HCl pH 7.5 (neutralization). The apparatus was assembled according to the manufacturer's instructions using BIO-DOT Slot Format Filter paper (Bio-Rad) and Hybond-N$^+$ membrane (Amersham), and connected to a vacuum pump. The slots were preequilibrated with 2 x SSC {0.3 M NaCl / 30 mM sodium citrate}, and subsequently the denatured DNA was applied. The slots were then washed twice with 2 x SSC and the apparatus disassembled. Subsequently, the transferred DNA was covalently bound to the membrane by UV irradiation for 3 min on a TMW-20 transilluminator (Alpha Innotech).

For probe labelling, hybridization and detection, the DIG Nucleic Acid Detection Kit (Boehringer Mannheim) was used following instructions outlined in the DIG System User's Guide for Filter Hybridization (Boehringer Mannheim). In principle, DNA probes were prepared by DIG incorporation during polymerase chain amplification reactions (PCR, section 2.8) using DIG-dUTP which replaced one third of the dTTP in the nucleotide mix. The PCR product was precipitated by adding 1/10 Vol 3 M sodium acetate pH 5.2 and 2.5 Vol ethanol followed by storage at -20°C for 20 min and centrifugation at 20,800 x g, RT for 20 min. The pellet was washed with 70% ethanol, respun for 5 min and resuspended in sterile MilliQ-treated water.

The effectiveness of the DIG incorporation during the PCR was determined by a spot test on Hybond-N$^+$ membrane (Amersham) according the DIG System User's Guide for Filter Hybridization (Boehringer Mannheim). Briefly, the labelled probe was subjected to a serial dilution in TE containing 50 µg/ml herring sperm DNA (supplied by the manufacturer). Subsequently, 1 µl of the dilutions were spotted onto the membrane. Detection was then performed as recommended by the vendor, using anti-DIG-F_{ab} fragments conjugated with alkaline phosphatase (Table 2.6) and the chemiluminescent substrate CSPD (see below).

Immediately before usage, the labelled probe was diluted approximately fifty-fold in 5 µg/ml fragmented herring sperm DNA solution to prevent background hybridization, and boiled for 8 min. Hybridization of these probes to immobilized DNA fragments on Hybond-N$^+$ (Amersham) took place in a Bachofer hybridization oven using the standard buffer {5 x SSC {0.75 M NaCl / 75 mM sodium citrate} / 0.1% (w/v)

sarcosyl / 0.02% (w/v) SDS / 1% (w/v) blocking reagent}. First, prehybridization was performed for 1 h at 65°C without any added probe. Subsequently, hybridization took place overnight at 55°C using approximately 10 ng probe per ml hybridization buffer. Washes were performed according to the manufacturer's protocol (twice for 5 min in 2 x SSC / 0.1% (w/v) SDS at RT, followed by two washes for 15 min in 0.5 x SSC {75 mM NaCl / 7.5 mM sodium citrate} / 0.1% (w/v) SDS at 55°C).

Subsequently, the probe binding site was detected using anti-DIG-F_{ab} fragments conjugated with alkaline phosphatase and the chemiluminescent substrate CSPD. The manufacturer's recommendations and buffer instructions outlined in the DIG System User's Guide for Filter Hybridization (Boehringer Mannheim) were applied. The enzyme reaction was activated by incubation of the membrane at 37°C for 10 min before exposing to X-ray film (Fuji) for 1-20 min. The film was developed using Kodak developer and fixer. It was immersed for 5 min in the developing solution, washed briefly in water and subsequently left in fixing solution for 5 min. Finally, the film was washed in water and dried.

2.7.5 DNA sequencing

Sequence data were generated at the Waikato DNA Sequencing Facility (University of Waikato, Hamilton, New Zealand) on an ABI Prism 377 DNA Sequencer (Applied Biosystems). Templates and primers were prepared in required purity (ABI preparation of plasmid DNA, section 2.6.1.4, or Wizard plasmid preparation, section 2.6.1.2) and quantities (plasmids ≥ 12 µl at 200 ng/µl; primers ≥ 5 µl at 0.8 pmol/µl). Sequencing reactions were then performed at the Waikato facility using dye terminator chemistry based on the technique described by Sanger *et al.* (Sanger *et al.*, 1974). Generated data were visualized as chromatograms using ABI Prism software (Editview).

Sequence data were also generated using the AmpliCycle™ Sequencing Kit (Perkin Elmer) following the recommendations of the manufacturer. Approximately 10 µCi [^{35}S]dATP (Amersham) was added directly to the cycling mixes to be incorporated randomly during amplification of the target DNA. Alternatively, 20 pmol of specifically designed primers were end-labelled using 30 µCi [γ-^{32}P]ATP (Amersham) and the T4 polynucleotide kinase of the Primer Extension System-AMV Reverse Transcriptase (Promega) kit according to the manual's instructions. Subsequently, 4 pmol of the labelled priming product were used for the amplification reaction.

Sequencing reactions were separated by polyacrylamide gel electrophoresis on a Model S2 sequencing apparatus (Life Technologies). 40 cm x 30 cm separating polyacrylamide gels were prepared following the recommendations of Misra (Misra, 1987). Briefly, an acrylamide stock solution was fixed by preparing a 38% acrylamide / 2% NN'methylenebisacrylamide solution that was deionized by adding 50 g/l Amberlite MB-3 matrix and stirring for 30 min at RT. The solution was filtered through 3MM chromatography paper (Whatman) and stored in darkness at 4°C. The polyacrylamide gel was prepared by mixing 33 g urea, 7.7 ml 10 x TBE {0.89 M Tris / 0.89 M boric acid / 20 mM EDTA}, 33 ml water, 13 ml acrylamide stock solution, 370 µl 10% (w/v) ammonium persulfate, and finally 37 µl TEMED as crosslinking agent. The mix was poured onto the gel cast, stored until complete polymerization occured, and used with 1 x TBE as running buffer. Electrophoresis was performed after a prerun of ≥ 30 min at 1500-1750 V (prewarming of the separating gel) for 1h 45 min (short run) or 3h 30 min (long run) at 1500-1700 V. After completion, the gel was disassembled and fixed. When the DNA had been labelled with ^{32}P, fixation was done by immersion of the gel for 30 min in 10% acetic acid. When ^{35}S had been used, the gel was left for 30 min in 10% methanol / 10% acetic acid solution. Gels were dried in the Model 583 Gel Dryer (Bio-Rad) according to the manufacturer's instruction for 1 h at 80°C. Finally, the gel was exposed to X-ray film (Fuji) which was developed after appropriate exposure time (5 h-5 days) as outlined in section 2.7.4.

The sequencing data were assembled with the Geneworks Package (IntelliGenetics), and analyzed with MacVector (IBI) and PCGene V6.85 (IntelliGenetics) software. Alternatively, the GeneJockey analysis programme (Biosoft) was used for characterization of sequence features. Databases were scanned on the server of the National Centre for Biotechnology Information using the BLAST algorithm (Altschul *et al.*, 1990). Among the searched protein databases were the GenBank coding sequences translations, the Swissprot protein sequence database, the PIR database and the Brookhaven protein data bank. The data were also compared to the GenBank, DNA databank of Japan, Brookhaven and European Molecular Biology Laboratory nucleotide sequence databases. Multiple sequence alignments were produced with the Clustal V program (Higgins & Sharp, 1989) using the default parameters.

2.8 DNA amplification by polymerase chain reaction (PCR)

In PCR experiments, DNA segments were exponentially amplified using two oligonucleotides flanking the region to be multiplied. A template DNA was first denatured, then the temperature dropped to levels allowing specific annealing of the primers to their target sequences, and finally the optimal temperature for DNA polymerization (72°C) was adjusted. The cycle of denaturation, annealing and extension was repeated until the generated amount of specific product met the respective requirements of the experiment.

Usually, 0.5 U *Taq* polymerase (Qiagen) were used per reaction. However, if the respective product was to be cloned into plasmids, 0.5 U *Pwo* polymerase (Boehringer Mannheim) were employed resulting in blunt ended amplified fragments with slightly higher fidelity. Reaction volumes were either 10 µl or 20 µl. When *Taq* polymerase was used, the reaction mixes contained the following final concentrations: 200 µM dNTP's, 2.5 mM $MgCl_2$, 0.4 µM primer, 1 x reaction buffer. These were maintained in *Pwo* polymerase mixes except that 1.5 mM $MgSO_4$ (already present in the respective reaction buffer) replaced the $MgCl_2$. The amount of template DNA per reaction was usually 50 ng (genomic DNA), or less than 2 ng (plasmid DNA). Master mixes were employed containing all ingredients common to the reactions performed at the same time. PCR's were performed in 0.2 ml thin wall reaction tubes in an FTS-960 Microplate Thermal Sequencer (Corbett Research).

Reaction setups varied in extension time and annealing temperature. A typical reaction scheme incorporated an initial denaturation step at 94°C for 2 min, followed by 35 cycles of 15 sec at 94°C / 1 min at 54°C / 1 min at 72°C, a final incubation of 7 min at 72°C, and then storage at 4°C. The annealing temperature was always ≥ 1°C under the lower of the calculated melting temperatures T_m for the two primers used. These melting temperatures were approximated following the formula $T_m = 2 \times n(A+T) + 4 \times n(G+C)$. Extension times were designed according to the length of the expected product assuming a polymerization velocity of roughly 30-40 bp/sec (Sambrook *et al.*, 1989).

2.9 Cloning procedures

Cloning of DNA typically consisted of four steps: preparation of suitably treated fragment and vector DNA solutions, ligation of the two components, transformation of

the mixture into a host bacterium, and finally screening for desired recombinants, using the Easy plasmid minipreparation method (section 2.6.1.1)

2.9.1 DNA preparation

Vector DNA to be used in cloning experiments was prepared by one of the plasmid preparation methods described (section 2.6.1), subjected to a restriction endonuclease treatment (section 2.7.2), fractionated by agarose gel electrophoresis (section 2.7.1), and finally purified from the gel as follows. Ethidium bromide stained digested vector DNA was visualized using the UVGL-58 Mineralight lamp (Ultra-Violet Prod.), and excised from the gel. Deproteinization was then performed using the QIAEX II Gel Extraction Kit (Qiagen) according to the manufacturer's instructions. Alternatively, the vector DNA was purified using the High Pure™ PCR Product Purification Kit (Boehringer Mannheim) following recommendations outlined in the Nucleic Acid Isolation and Purification guide (Boehringer Mannheim). The purification mechanism in both kits relied on specific binding of DNA particles to silica-gel particles in the presence of chaotropic salts and subsequent elution with low salt solutions (i.e. water or TE {10 mM Tris-HCl pH 8.0 / 1 mM EDTA}). To solubilize the gel fragment containing the desired DNA prior to the High Pure deproteinization, it was resuspended in binding buffer (3 µl/µg gel) and incubated at 56°C for 10 min. Then, isopropanol was added (1.5 µl/µg gel), the suspension was mixed, and transferred into the High Pure™ Filter tube following the standard protocol in the manufacturer's instructions from then on.

Fragment DNA was isolated by digestion of plasmids (section 2.7.2), separation of the fragment of interest by agarose gel electrophoresis (section 2.7.1), and recovery from the gel as described above. Alternatively, fragments were PCR amplified using *Pwo* polymerase (section 2.8). Subsequently, these were treated as above (gel electrophoresis, gel excision, deproteinization). The termini of the fragments had to correspond to the termini of the vector DNA to ensure efficient ligation.

2.9.2 Ligation

Ligations of DNA segments with cohesive termini were performed in the supplier's 1 x ligation buffer using 0.5 U T4 DNA ligase (Boehringer Mannheim) in 10 µl. The mixes were incubated overnight at 12°C. For blunt ended fragments, 1 U T4-DNA ligase in 15 µl mix was used in an overnight reaction at 16°C in 1 x ligation buffer.

A three-fold molecular excess of fragment DNA to vector DNA was applied wherever possible to ensure efficient ligation. Approximately 20 ng plasmid DNA were usually used per ligation experiment.

2.9.3 Transformation

2.9.3.1 Preparation of competent bacterial cells

For preparation of competent *E. coli* cells, a modification of the method of Hanahan (Hanahan, 1983) was applied as follows. An overnight culture in LB was diluted fifty-fold. The fresh culture was incubated at 37°C until the optical density of the bacterial suspension corresponded to absorption values of approximately 0.3-0.4 at a wavelength of 600 nm (OD_{600}). Measurements were performed in plastic cuvettes using a Visible Spectrophotometer (Science & Technology). The cells were then harvested at 5,000 x g for 5 min at 4°C. After resuspending the pellet in 1/5 of the original culture volume of sterile ice cold CM1 {10 mM sodium acetate pH 5.6 / 5 mM NaCl / 50 mM $MnCl_2$}, the suspension was stored for 20 min on ice, and spun as above. The cells were then resuspended in 1/50 of the original culture volume of sterile icecold CM2 {10 mM sodium acetate pH 5.6 / 50 mM $MnCl_2$ / 5% glycerol / 70 mM $CaCl_2$}, aliquoted and stored at -70°C. Typically, 50 µl of this cell suspension was used per transformation. Competence of the bacterial cells was tested by transformation (section 2.9.3.2) using 50 pg pUC19 DNA (Life Technologies). Generation of approximately 10 transformant colonies per pg plasmid DNA demonstrated sufficient competence of the cells.

Electrocompetent *H. pylori* cells were prepared based on a method described by Segal and Tompkins (Segal & Tompkins, 1993). Briefly, *Helicobacter* cells grown in TSB for 24-36 h (OD_{600} of approximately 0.4) were harvested at 5,000 x g for 5 min at 4°C. The bacteria were washed three times by resuspension in 1 Vol ice cold electroporation buffer EPB {272 mM sucrose / 15% glycerol / 2.43 mM K_2HPO_4 / 0.57 mM KH_2PO_4 pH 7.4}, and subsequently spun as above. Finally, the pellet was resuspended in 1/25 Vol of the original culture volume of EPB, aliquoted and stored at -70°C. Usually, 50 µl of the suspension was used per electroporation. Competence of the cells was tested by electroporation (section 2.9.3.3) of pSP107 DNA. Resistant colonies were generated at a frequency of approximately 10^{-3} (calculated by the number of transformed colonies divided by the number of surviving *H. pylori* cells after harvesting from a non-selective CBA plate on day 1 of the procedure).

2.9.3.2 Transformation of *E. coli*

For transformation of *E. coli*, the standard methodology (Sambrook *et al.*, 1989) was modified as follows. An aliquot of competent cells (section 2.9.3.1) was thawed on ice, diluted four-fold in 0.1 M $CaCl_2$ and mixed with the respective plasmid DNA (approximately 50 ng) or half of the ligation mix (usually containing 10 ng plasmid DNA). After 30 min incubation on ice, cells were heat-shocked at 42°C for 45 sec, then put on ice for 2 min. The suspension was diluted five-fold in LB. The mixture was agitated at 37°C for 90 min, and plated in appropriate dilutions on selective media plates containing antibiotics and supplements (section 2.2). Positive and negative transformation controls (using pUC19 vector DNA or water, respectively) were regularly performed.

2.9.3.3 Transformation of *H. pylori*

For transformation of *H. pylori* cells, the natural competence of *Helicobacter* was exploited as described by O'Toole *et al.* (O'Toole *et al.*, 1994). *H. pylori* cells grown for 2 -3 days on CBA were harvested using a sterile swab, and resuspended in TSB, that had been CO_2-equilibrated by 1 h incubation in the microaerobic environment generated by the CO_2 incubator (Revco Scientific). The lid of the medium bottle or tube was left ajar to allow generation of microaerobic conditions in the tube. An OD_{600} of approximately 0.4 was adjusted by appropriate dilution of the cell suspension in CO_2-equilibrated TSB. Subsequently, 200 µl of the suspension was transferred to a screw cap tube, and 5 µg of plasmid DNA was added. The lid of the tube was left ajar during the following incubation for 4 h at 37°C in the CO_2 incubator. The tube was occasionally subjected to agitation. The cells were then plated onto non-selective medium and incubated for 24 h at 37°C, 5% CO_2. Bacterial growth on the plates was harvested into 1 ml of CO_2-equilibrated TSB. A small aliquot was used to prepare a 10^6-fold dilution, which was spread onto non-selective plates. Those colonies were used for calculation of transformation efficiencies. After removal of 100 µl cell suspension, which was plated onto selective CBA, the remaining bacteria were concentrated by centrifugation for 15 sec at RT, 20,800 x g, resuspended in residual TSB medium, and applied on selective CBA plates. Transformant colonies usually appeared after 2-3 days incubation at frequencies of $\leq 6.2 \times 10^{-6}$ (calculated by the number of transformed colonies divided by the number of surviving *H. pylori* cells).

Alternatively, electrocompetent cells (section 2.9.3.1) were used in electroporation experiments. The method described by Segal and Tompkins was applied (Segal & Tompkins, 1993). Aliquots of 50 µl electrocompetent *H. pylori* cells were thawed on ice. Plasmid suspensions containing approximately 2 µg DNA were dialyzed for ≥ 1h against 10% glycerol using Millipore filters with 25 nm pore diameters. The competent cells were then mixed with the plasmid DNA in an ice cold cuvette, and electroporation was performed on a Gene Pulser (Bio-Rad) using a voltage of 2.5 kV, 200 Ω resistance and 25 µF capacity according to the instructions of the manufacturer. Time constants for the electroporation process were recorded, and were typically at 4.5. The cells were immediately plated onto non-selective CBA plates. After 24 h of incubation at 37°C, 5% CO_2, bacterial growth on the plates was harvested into 1 ml of CO_2-equilibrated TSB. Dilutions were plated onto non-selective and selective CBA plates as described above. Transformants appeared after 2-3 days incubation at frequencies of ≤ 2.2×10^{-3}.

2.10 General precautions for RNA work

The standard precautions were followed when working with RNA to avoid ribonuclease contamination (Sambrook *et al.*, 1989). A separate workplace was established containing equipment and chemicals specifically reserved for procedures involving RNA. Clean disposable gloves were always worn when performing RNA experiments. All RNA work was done using previously untouched disposable sterile plastic labware, plasticware that had been rubbed down thoroughly with RNase*AWAY*™ reagent (Molecular Bio-Products) using fresh tissue paper, or glassware that had been baked at 180°C for at least 10 h. Water used for RNA experiments was a 0.1% solution of diethylpyrocarbonate DEPC (Sigma), that had been incubated at 37°C overnight and subsequently autoclaved at 121°C for 20 min. Solutions were prepared from previously unopened reagent stocks.

2.11 RNA isolation from *H. pylori* cells

2.11.1 Total RNA isolation using TRIzol

Total RNA of *H. pylori* cells grown in TSB or on CBA was isolated using the TRIzol reagent (Life Technologies) according to the manufacturer's instructions. The reagent (a monophasic solution of phenol and guanidine isothiocyanate) disrupted cells

and dissolved cell compounds, but maintained the integrity of RNA which subsequently could be isolated by chloroform extraction and isopropanol precipitation. During the chloroform extraction, genomic DNA was retained in the phenolic fraction or the interphase, whereas bacterial RNA resides in the aqueous fraction. Using 1 ml of TRIzol, approximately 15 µg total RNA were isolated from 1 ml of *H. pylori* cells grown in TSB at an OD_{600} of 0.4. *H. pylori* cells from heavily streaked CBA plates that had been incubated for 36 h yielded approximately 150 µg of total RNA per plate (isolated with 2 ml of the reagent).

2.11.2 RNA isolation using density gradient ultracentrifugation

In order to obtain high yields of very pure total RNA, a caesium chloride ultracentrifugation procedure described by Glisin *et al.* (Glisin *et al.*, 1974) was modified appropriately. *H. pylori* cells grown for 48 h on 4 heavily streaked CBA plates were harvested with a sterile swab into 50 ml of CO_2-equilibrated TSB. The suspension was poured into centrifugation bottles containing 15 ml frozen TSB / 10 mM sodium azide, and cells were pelleted by centrifugation at 4°C, 4,500 x g for 15 min. The pellet was resuspended in 2.7 ml 100 mM Tris-HCl pH 6.8 / 2mM EDTA, and 300 µl of 0.5 M Tris-HCl pH 6.8 / 20 mM EDTA / 10% (w/v) SDS were added for lysis. The suspension was mixed, boiled for 4 min, and finally 50 µl of β-mercaptoethanol and 1.22 g of caesium chloride were added. The mix was layered onto 2.5 ml 5.7 M CsCl in 0.1 M EDTA in a thin walled 4.4 ml polyallomer Ultra tube (DuPont), and spun in a TST 60.4 rotor using a Sorvall OTD-Combi ultracentrifuge (DuPont) for 14 h at 40,000 rpm (approximately 170,000 x g), 20°C. The pellet was then washed with 70% ethanol, airdried and resuspended in 50 µl 0.1% DEPC. Approximately 150 µg pure total *H. pylori* RNA was isolated using this procedure.

For comparison, this method was also used for total RNA preparation from *E. coli* cells. An overnight LB culture was diluted fifty-fold, and incubated at 37°C for 8 h. The cells were then harvested by centrifugation over 30 ml frozen LB / 10 mM sodium azide as described above and subsequently subjected to the same procedure. This isolation method consistently yielded approximately 200 µg RNA per 5 ml *E. coli* culture.

2.12 RNA analysis methods

2.12.1 RNA agarose gel electrophoresis

For RNA visualization, horizontal agarose gel electrophoresis was performed in a Horizon 58 mini gel apparatus (Life Technologies) at RT. The gels were 1.2% (w/v) solutions of agarose in 1 x MOPS-EA {20 mM 3-{N-morpholino} propane-sulfonic acid / 5 mM sodium acetate / 1 mM EDTA, pH 7.0 with KOH} that were microwaved for total solubilization and cooled to 55°C. Exactly 5% saturated formaldehyde (approximately 37%) was added before pouring. RNA samples were mixed with an equal volume of RNA-FSB {50% formamide / 5% glycerol / 16.7% saturated formaldehyde / 1 x MOPS-EA / 0.4 mg/ml bromophenol blue}, incubated at 65°C for 15 min and stored on ice for 10 min. Finally, 4 µl of 0.25 mg/ml ethidium bromide solution were added to the samples before loading. The gel was submerged in 1 x MOPS-EA, and electrophoresis was performed at 60-100 V for 60 min. RNA was visualized using the TMW-20 Transilluminator (Alpha Innotech). The 0.24-9.5 kb RNA ladder (Life Technologies), a set of standard RNA fragments of known molecular mass, was subjected to the same treatment before loading, and run in parallel to the samples.

2.12.2 RNA quantification

RNA concentrations were determined spectrophotometrically using a Shimadzu UV-160A spectrophotometer as described in section 2.7.3 for DNA, applying a constant factor of 33 in the calculation formula. Genomic DNA contamination was detected by DNA gel electrophoresis (section 2.7.1) of the RNA sample, where residual chromosomal DNA of high molecular weight was separated from low molecular weight RNA, and visualized by ethidium bromide staining.

2.12.3 Northern blotting and hybridization

Attempts to determine transcript sizes were made using the Northern blotting technique followed by hybridization and subsequent detection of the labelled probe on X-ray film (Fuji). The principle of the procedure (separation of nucleic acid by gel electrophoresis - transfer onto a solid medium - hybridization of a specific oligonucleotide - detection of the binding site of the probe) corresponded to the Southern blotting technique (section 2.7.4).

RNA samples were separated according to their molecular mass by RNA agarose gel electrophoresis (section 2.12.1). The gel was photographed and subsequently washed twice for 5 min in water, and then twice for 20 min in 20 x SSC {3 M NaCl / 0.3 M sodium citrate}. An overnight capillary nucleic acid transfer to positively charged nylon membrane (Boehringer Mannheim) was performed following the standard procedure (Sambrook *et al.*, 1989). The RNA was then covalently bound to the membrane by either UV radiation (3 min on transilluminator), or by baking for 2 h at 80°C between two sheets of 3MM chromatography paper (Whatman).

For probe labelling, four different techniques were employed. Probes were either DIG-labelled PCR fragments (section 2.7.4), ECL™ labelled DNA restriction fragments (section 2.12.3.2), [α-^{32}P]dCTP radiolabelled DNA restriction fragments (section 2.12.3.3), or DIG-labelled riboprobes (section 2.12.3.4).

2.12.3.1 Northern analysis using DIG-labelled DNA probes

DIG-labelling of PCR fragments was performed as described in section 2.7.4. All subsequent steps of the procedure (hybridization and detection) were also accomplished according to the descriptions in that section.

2.12.3.2 Northern analysis using ECL™-labelled DNA probes

DNA restriction fragments were ECL™ labelled with the ECL™ direct nucleic acid labelling system (Amersham) following the manufacturer's instructions. Using the recommended amount of target DNA (100 ng), a complex of horseradish peroxidase with a positively charged polymer bound covalently to the nucleic acid with the help of glutaraldehyde, and the enzyme activity was subsequently used for detection. The peroxidase produced oxygen radicals of the hydrogen peroxide supplemented in detection reagent 1. These radicals subsequently oxidized the luminol in detection reagent 2, which then emitted blue light.

The blots were processed according to the manufacturer's recommendations. The hybridization buffer was supplemented with 0.5 M (final concentration) NaCl and 5% blocking reagent, stored at -20°C and used at 0.25 ml/cm^2 membrane. The probe hybridized overnight in a Bachofer hybridization oven at 42°C. The primary wash buffer {6 M urea / 0.4% SDS / 0.5 x SSC {75 mM NaCl / 7.5 mM sodium citrate}} was used

twice for 20 min at 42°C, followed by two 5 min washes at RT in 20 x SSC {3 M NaCl / 0.3 M sodium citrate}. An excess volume of the detection mix (0.2 ml/cm^2 membrane) ensured complete submersion of the filter, and after 1 min the membrane was drained and exposed to X-ray film (Fuji). The film was developed as explained in section 2.7.4.

2.12.3.3 Northern analysis using radiolabelled DNA probes

For radiolabelling of DNA restriction fragments, the Rediprime DNA Labelling System (Amersham) was employed following the manufacturer's recommendations. The labelled specific probe was synthesized using Klenow enzyme, a mix of dATP / dGTP / dTTP, random primers (9mers), and radiolabelled Redivue [α-^{32}P]dCTP 10 mCi/ml (Amersham). In general, 25 ng of DNA was used, and denatured prior to the labelling reaction.

Hybridization was performed in 10 x Denhardts buffer {0.2% (w/v) Ficoll / 0.2% (w/v) polyvinylpyrrolidone / 0.2% (w/v) bovine serum albumin } supplemented with 6 x SSC {1 M NaCl / 0.1 M sodium citrate}, 0.5% (w/v) SDS, 1 mM EDTA and 100 µg/ml fragmented salmon sperm DNA (Sigma). After 2 h prehybridization at 65°C, the probe was added and hybridized at 65°C overnight in a Bachofer hybridization oven. The membrane was subsequently washed twice for 5 min at RT in 2 x SSC {0.3 M NaCl / 30 mM sodium citrate} / 0.1% (w/v) SDS, followed by two washes for 15 min at 65°C in 1 x SSC {0.15 M NaCl / 15 mM sodium citrate} / 0.1% (w/v) SDS, and drained before exposure to X-ray film (Fuji). The film was developed as described (section 2.7.4).

2.12.3.4 Northern analysis using DIG-labelled riboprobes

DIG labelled riboprobes were generated using the DIG RNA Labelling Kit (Boehringer Mannheim). Appropriate DNA fragments were cloned into the vector pSPT18, the construct was linearized with restriction endonucleases producing 5' overhangs, and phenol / chloroform extracted. The extraction was performed as follows: The digestion mix was diluted with 0.1% DEPC. Then, 1 Vol phenol was added, the suspension was mixed and centrifuged for 2 min at RT, 20,800 x g. The upper phase was mixed with 1 Vol of 50% phenol / 50% chloroform, mixed and respun as above. Approximately 1 Vol chloroform was added to the top fraction, and the suspension was treated as before. The upper phase after this treatment was mixed with 1/10 Vol 3 M

sodium acetate pH 5.2 and 2.5 Vol ethanol, and stored overnight at -20°C for precipitation. The DNA was subsequently pelleted by centrifugation at 20,800 x g for 20 min at RT. The pellet was washed with 70% ethanol, respun for 5 min as above, airdried, and resuspended in 0.1% DEPC.

The fragment was subsequently transcribed from approximately 500 ng template DNA using the kit's T7 RNA polymerase and nucleotide mix (ATP, CTP, GTP at 1 mM final concentration, UTP at 0.65 mM final concentration and DIG-UTP at 0.35 mM final concentration). The manufacturer's recommendations were followed, and 3-7.5 µg riboprobes were obtained after 135 min incubation of the reaction mix at 37°C and subsequent stop of the reaction with 20 mM (final concentration) EDTA. The probe was precipitated overnight at -20°C using 1/10 Vol 3 M sodium acetate pH 5.2 and 3 Vol ethanol, washed with 75% ethanol, airdried, and resuspended in 0.1% DEPC. Successful DIG incorporation into the probes was tested by a spot test following the recommendations of the vendor (for a brief description, see section 2.7.4). RNA probe dilutions were in 0.05% DEPC / approximately 7.5% (final concentration) formaldehyde / 6 x SSC {0.9 M NaCl / 90 mM sodium citrate}.

The recommendations of the DIG Easy Hyb Granules manufacturer (Boehringer Mannheim) were applied for hybridization of the riboprobes to transcripts immobilized on nylon membranes by Northern transfer (section 2.12.3). A 40 min prehybridization step in 0.2 ml/cm^2 membrane of DIG Easy Hyb solution (Boehringer Mannheim) at 68°C (Bachofer hybridization oven) was performed. Probes were denatured for 10 min at 100°C. Then, they were added to the DIG Easy Hyb solution, and hybridized with the respective target RNA during an overnight incubation at 68°C. Riboprobes were used at 100 ng/ml DIG Easy Hyb solution. Washes were performed according to the vendor's protocol (2 x 5 min in 2 x SSC {0.3 M NaCl / 30 mM sodium citrate} / 0.1% (w/v) SDS at RT, 2 x 15 min in 0.1 x SSC {15 mM NaCl / 1.5 mM sodium citrate} / 0.1% (w/v) SDS at 68°C), and detection as described in section 2.7.4.

2.12.4 Transcript analysis by RT-PCR

Total RNA preparations of *H. pylori* cells obtained with TRIzol (section 2.11.1) were used in RT-PCR assays to detect transcripts. In these assays, a cDNA was generated from mRNA by a reverse transcriptase using specifically designed primers. This cDNA fragment was subsequently amplified in a polymerase chain reaction (section 2.8).

Total RNA preparations had to be DNase treated prior to the RT-PCR experiment to remove all traces of contaminating genomic DNA. This was achieved following recommendations of Huang *et al.* (Huang *et al.*, 1996). Briefly, 2.5 mM (final concentration) $MgSO_4$ and 2 U RNase-free DNAse (Boehringer Mannheim) per µg total RNA were added to the sample, followed by incubation at 37°C for 60 min. Then, the DNase was heat inactivated at 75°C for 10 min, and the RNA was used directly in subsequent RT-PCR experiments.

RT-PCR was performed in 0.2 ml thin wall reaction tubes in a FTS-960 Microplate Thermal Sequencer (Corbett Research). The Superscript One-Step™ RT-PCR System (Life Technologies) was applied, following the manufacturer's recommendations. The kit's enzyme mix consisted of Superscript II reverse transcriptase, and *Taq* polymerase. Negative controls were performed by replacement of the enzyme mix with 1 U of regular *Taq* polymerase (Qiagen). Activity of the *Taq* polymerase was checked by simultaneous reactions containing 1 U *Taq* polymerase and genomic *H. pylori* DNA.

Alternatively, the Titan™ One Tube RT-PCR System (Boehringer Mannheim) was used. In this kit, an Expand™ High Fidelity enzyme mix containing *Taq* and *Pwo* polymerase was supplied in conjunction with an AMV reverse transcriptase. For negative controls, the reverse transcriptase was heat-killed by incubation of the whole reaction mix at 94°C for 5 min prior to cDNA synthesis.

In general, the reverse transcriptase reaction was performed for 30 min at 45°C. Then, a common PCR reaction scheme as outlined in section 2.8 followed, usually consisting of 35 cycles with appropriate annealing temperatures (52-55°C, depending on the primer pair used) and extension times at 70°C (according to the expected size of the fragment). Amplified products were analyzed by conventional gel electrophoresis (section 2.7.1)

2.12.5 Transcript analysis by primer extension

Primer extension was used to determine transcription start sites of the investigated operon. The Primer Extension System-AMV Reverse Transcriptase kit (Promega) was used following the manufacturer's protocol. In principle, the experiments consisted of three parts: (i), labelling of a primer which annealed to the coding strand downstream from the transcription start point facing it; (ii), reverse transcription of the specific mRNA (cDNA synthesis) until the reverse transcriptase reached the 5' end of the transcript; and (iii), visualization of the primer extension product on a polyacrylamide gel.

In order to determine the start point of the transcript, a sequencing reaction of the equivalent DNA region using the AmpliCycle™ Sequencing Kit (Perkin Elmer) (section 2.7.5) was performed with the same oligonucleotide that was used in the primer extension reaction. The products of the sequencing reaction were subsequently separated adjacent to the primer extension sample on a polyacrylamide gel.

Total RNA preparations from *H. pylori* cells obtained with TRIzol (section 2.11.1) were utilized at 20-30 µg per reaction. For some controls, RNA samples were DNase treated as described in section 2.12.4. However, DNase was not heat inactivated, but phenol / chloroform extracted. One Vol of 50% phenol / 50% chloroform was added to the reaction mix, the suspension was mixed for 1 min and spun for 15 min at 7,600 x g, 4°C. The upper phase was mixed with 1 Vol of chloroform and spun as above. Finally, RNA was precipitated from the upper phase by addition of 1 Vol isopropanol, incubation for 10 min at RT and centrifugation at 7,600 x g, 4°C for 10 min. The pellet was washed twice with 1 ml 75% ethanol and readhered to the bottom of the tube by spinning for 5 min at 4°C, 2,000 x g.

Approximately 20 pmol of specifically designed primers downstream from the presumptive transcript start site were end labelled using T4 polynucleotide kinase and 30 µCi [γ-^{32}P]ATP (Amersham). For cDNA synthesis, 100 fmol of the labelled primer were applied. The annealing conditions were 70 min at 58°C, extension was performed at 42°C for 30 min. Primer extension products were subsequently analyzed on a conventional sequencing gel (section 2.7.5).

2.13 Protein sample preparation

2.13.1 Whole cell lysates

H. pylori cultures grown in TSB were centrifuged at 20,800 x g, RT for 1 min. The supernatant was collected and stored at -70°C. Residual medium was removed from the cell pellet after an additional spin at 20,800 x g for 1 min at RT. Pellets were then resuspended in sterile MilliQ-treated water, boiled for 10 min and stored at -70°C.

2.13.2 Cell fractionation

Subcellular protein fractions were prepared based on a method described by O'Toole *et al.* (O'Toole *et al.*, 1994). Approximately 50 ml of *H. pylori* cultures grown for 36 h in TSB were harvested by centrifugation at 2,000 x g in a Heraeus Labofuge^{GL} bench top centrifuge at RT for 20 min. The supernatants were stored at -70°C, and aliquots were concentrated five-fold using a Nanosep^{TM} Microconcentrator (Pall Filtron) with a molecular weight cutoff at 10 kDa. Briefly, supernatant was applied onto the concentrator, and spun at 2,800 x g, RT for 15-25 min. Alternatively, concentrating was performed at 4°C in a Sorvall RC-5B centrifuge (DuPont) at 1,250 x g for 90 min. The concentrated supernatant fractions were stored at -70°C.

The cell pellets were resuspended in 1/100 Vol Tris-saline {10 mM Tris-HCl pH 7.5 / 30 mM NaCl}, and the bacterial cells lysed by sonication in a Soniprep 150 ultrasonic disintegrator (MSE) for 8 x 20 sec with 40 sec intervals. The sonicated cell suspensions were diluted four-fold with Tris-saline, and separated into crude soluble and envelope fractions by ultracentrifugation as follows. The solutions were applied to Beckman Quickseal centrifuge tubes, which were heat-sealed with the Beckman Tube Topper. The samples were then centrifuged in the TL-100 Tabletop Ultracentrifuge (Beckman Instruments) using a TLA 100.2 rotor at 40,000 rpm (45,000 x g) for 30 min at 4°C. Pellets represented the envelope fractions. They were resuspended in 1 Vol Tris-saline and stored at -70°C.

Supernatants were recentrifuged for 2 h at 60,000 rpm (100,000 x g), 4°C. Pelleted materials were partially assembled flagellar filaments and aggregated flagellin subunits. They were subsequently resuspended in 1 Vol Tris-saline, and stored at -70°C.

The clear supernatants from the second ultracentrifugation step represented the crude soluble fractions and were also stored at -70°C.

2.14 Protein analyses

2.14.1 Protein quantification

Protein quantities of the various samples were crudely standardised by wet weight determination of the cell pellets after aspiration of remaining supernatant traces, and adjusting the sample volumes in subsequent experiments.

For correct determination of protein quantities, the Bio-Rad Protein Assay was applied according to the manufacturer's microassay procedure. In principle, protein concentrations were measured based on the method of Bradford (Bradford, 1976), involving the addition of an acidic dye (Coomassie Brilliant Blue G-250) to protein solution, which bound to the protein and changed its colour depending on the protein concentration. The absorbance maximum of the mixture shifted to 595 nm, and spectrophotometric measurements at this wavelength were used for protein concentration calculations.

A standard curve was prepared using a bovine serum albumin stock solution (New England Biolabs) of 1.45 mg/ml. Duplicates of six standard solutions were prepared representing a protein concentration range between 1 and 10 µg/ml. Exactly 200 µl dye was added to 800 µl of protein solution, and the suspension was incubated for 5 min at RT. Subsequently, the absorbance at 595 nm was determined using the Visible Spectrophotometer (Science & Technology), and plastic cuvettes.

For determination of the protein concentration in cell pellet samples, the pellets were resuspended in 1/10 Vol water, and subsequently diluted appropriately, approximating the protein amount to 15% of the wet weight of the cells. Based on the measured absorbance at 595 nm, concentrations of the unknowns were calculated by comparison to the standard curve.

2.14.2 Protein electrophoresis

Protein samples were size separated by vertical electrophoresis according to the method by Laemmli (Laemmli, 1970) in denaturing polyacrylamide gels containing SDS. The Mini-Protean^{R} II Electrophoresis Cell (Bio-Rad), holding two 8 x 5 cm gels, was used. The gels consisted of two components, three quarters separating gel incorporating 12.5% acrylamide, and one quarter stacking gel incorporating 4% acrylamide. The separating gel was a suspension of 25% lower gel buffer {1.5 M Tris-HCl pH 8.8 / 0.4% (w/v) SDS}, 41.5% acrylamide-bisacrylamide stock solution {29.2% (w/v) acrylamide and 0.8% (w/v) bisacrylamide, stored at 4°C in darkness}, 0.5% of a fresh 10% (w/v) ammonium persulphate stock solution prepared in water, and 0.1% NNN'N' tetramethylethylenediamine (TEMED). After mixing, the separating gel was poured into the chamber that had been assembled according to the manufacturer's recommendations. The gel was then overlayed with water saturated isopropanol. After polymerization of the separating gel, the stacking gel was mixed, containing 25% upper gel buffer {0.5 M Tris-

HCl pH 6.8 / 0.4% (w/v) SDS}, 13.3% acrylamide-bisacrylamide stock solution, 0.25% fresh 10% (w/v) APS stock solution, and 0.1% TEMED. The isopropanol was thoroughly washed from the surface of the separating gel, the stacking gel was poured and left to polymerize.

Before loading, protein samples were mixed with at least 1 Vol final sample buffer {10% glycerol / 3% (w/v) SDS / 62.5 mM Tris-HCl pH 6.8 / 0.01% (w/v) bromophenol blue / 5% freshly added ß-mercaptoethanol}, boiled for 10 min and stored on ice. Electrophoresis was usually performed for 30 min at 55 V, followed by 60 min at 100 V, with 25 mM Tris / 0.2 M glycine / 0.1% (w/v) SDS as running buffer. Subsequently, the gel was incubated in Coomassie Brilliant Blue stain {45.4% methanol / 9.2% acetic acid / 0.25% (w/v) coomassie brilliant blue} for ≥ 1 h at RT, and finally destained in several changes of 45.4% methanol / 9.2% acetic acid until the desired staining intensity and low background coloration was achieved.

The molecular weight of proteins was determined by direct comparison of their migration distance to that of a set of standard proteins of known size run on the same gel. Standard protein markers were the High Range Prestained Protein Molecular Weight Standards (Life Technologies), reconstituted according to the manufacturer's advise.

2.14.3 Western blotting and hybridization

Western blotting and subsequent binding of specific antibodies to target proteins was performed based on methodology of Blake *et al.* (Blake *et al.*, 1984). Proteins separated by gel electrophoresis (section 2.14.2) were transferred from an unstained polyacrylamide gel to nitrocellulose paper by the methanol Tris-glycine system described by Towbin *et al.* (Towbin *et al.*, 1979). Subsequently, specific antibodies were used for detection. The primary antibody bound to the target protein of the sample, and the secondary antibody, complemented with an active enzyme, bound to the primary one. Finally, chromogenic detection was performed using a substrate of the enzyme which was linked to the respective secondary antibody.

The transfer of the separated proteins onto nitrocellulose membranes (Sartorius) was performed in a Mini Blot Module (E-C Apparatus) according to the manufacturer's instructions, for 2 h at 20 V in 25 mM Tris / 192 mM glycine / 20% methanol. After the transfer, the membranes were rinsed in Tris-buffered saline TBS {10 mM Tris-HCl pH 7.5 / 0.9% (w/v) NaCl}, then incubated under agitation in SM-TBS {1% (w/v) skim milk powder in TBS} for 30 min at RT. This treatment ensured efficient blocking of unspecific

binding of the antibody in the following steps. Subsequently, the filter was covered in a solution of the primary antibody in SM-TBS, and agitated for 90 min at RT. Dilution factors varied between 250 and 2,500, depending on the antibody used. The membrane was washed for 10 min in TBS, then for 5 min in TBS / 0.05% (w/v) Tween 80, and finally again for 10 min in TBS. Lastly, the filter was submerged in an appropriate dilution of the secondary antibody in SM-TBS and shaken for 60 min at RT, followed by the same washing regime.

Two different detection methods were used. For primary rabbit polyclonal antibodies, the secondary antibody was an anti-rabbit IgG-alkaline phosphatase conjugate (Sigma) (section 2.5), a two thousand-fold dilution of which was sufficient for subsequent detection. After the second washing procedure the filter was rinsed in detection buffer {0.1 M Tris pH 9.5 / 0.1 M NaCl / 50 mM $MgCl_2$}, and submerged in freshly prepared detection buffer containing 330 µg/ml Nitro blue tetrazolium chloride NBT (Boehringer Mannheim) and 165 µg/ml 5-bromo-4-chloro-3-indolyl-phosphate BCIP (Boehringer Mannheim). The filter was stored in this developing solution in the dark until the required intensity of detected bands was reached. Rinsing for 30 min in water stopped the developing process. The filter was then stored dry at RT.

For primary mouse monoclonal antibodies, the secondary antibody was an anti-mouse IgG-horse radish peroxidase conjugate (Serotec) (section 2.5), a one thousand-fold dilution of which was used. For detection, a solution of 0.3% (w/v) 4-chloro-1-naphthol (Sigma) in methanol (usually 10 ml) was poured into 5 Vol TBS. Then, 0.6 µl/ml of a 30% hydrogen peroxide solution was added, and the membrane submerged in this suspension. Development occured in darkness until the necessary signal intensity was reached. The filters were subsequently agitated in several changes of MilliQ-treated water to stop the reaction, and stored dry at RT.

2.15 Microscopy

2.15.1 Phase contrast microscopy

Bacterial cells were observed using a phase contrast transmitted-light microscope KM (Carl Zeiss) at 400 x magnification.

For motility evaluations, a hanging drop preparation of liquid *H. pylori* cultures was prepared by pipetting 20 µl culture onto a slide and attaching a coverslip with modelling clay on its edges thereby generating a gap between slide and coverslip. This

preparation was then examined using a Nikon Diaphot inverted phase contrast microscope at 600 x magnification.

2.15.2 Electron microscopy

For electron microscopy, *H. pylori* cells grown in TSB for 24 h were harvested by centrifugation at RT, 20,800 x g for 30 sec. The supernatant was aspirated. A second centrifugation for 10 sec followed after which the remaining supernatant was removed. The pellets were resuspended in 0.1 Vol water immediately before being subjected to negative staining. A drop of sample was applied for 1 min onto a 200-mesh grid (Agar Scientific) covered with a Formvar film (0.5% formvar in ethylene dichloride). The excess sample was withdrawn by touching the edge of the grid with a cut piece of filter paper. The grids were then floated onto a drop of 1% (w/v) phosphotungstic acid, and adjusted to pH 7.0 with potassium hydroxide. They were then examined in a Philips 201C transmission electron microscope operated at an accelerating voltage of 60 kV under conventional bright-field illumination conditions. Images were recorded at 15,300 x magnification on Agfa Copex Positive PET 10 film. Electron microscopy was kindly performed by D. Hopcroft, The Horticultural and Food Research Institute of New Zealand, Palmerston North, New Zealand.

3. RESULTS

3.1 Cloning of the *Helicobacter pylori fliI* gene

3.1.1 The λZAP excisant pHP042

Predating the release of the *H. pylori* 26695 genome sequence, a λZAP library of partially *Sau*3A digested *H. pylori* 17874 chromosomal DNA had been constructed by Dr PW O'Toole. This λ library had been converted to a plasmid-based one by superinfection of the bacterial cells with a phage that excised recombinant pBK-CMV derivatives from the λZAP phage. These excised plasmids contained *H. pylori* genomic DNA fragments. The insert ends of the constructs had been sequenced at random as part of a genome screening exercise. One of the excisants, pHP042, contained a 3.5 kb DNA fragment, one end of which exhibited significant similarity to *fliI*, the gene coding for an ATPase involved in flagellar biosynthesis in a number of bacteria (Fan & Macnab, 1996). Similarities had been detected by BLAST searches using the National Center for Biotechnology Information (NCBI) server (for databases scanned, see section 2.7.5).

The initial objective in this study was to extend the sequence information on this end of the pHP042 insert. Plasmid DNA was isolated using the ABI plasmid preparation method (section 2.6.1.4). Automated sequencing using pUC/M13 RP was performed as described in section 2.7.5. A sequence of 740 bp of the insert end was determined and analyzed. BLAST similarity searches revealed strong homologies to both *fliI* and ABC transporter genes of various bacteria. Closer analysis showed that the *fliI*-like region on the investigated pHP042 insert end was spatially distinct from the segment similar to the ABC transporters. Whereas the first 156 bp of the generated sequence exclusively exhibited *fliI* homology, only the following 584 bp resembled segments of genes encoding ABC transporter proteins. The observed change in similarity coincided with a *Sau*3A site at position 156 of the sequence.

The genetic content of the pHP042 insert was subsequently partially determined by AJ Beddek (Beddek, 1998). Remaining uncertainties were later clarified upon release of the *H. pylori* 26695 genome sequence (Tomb *et al.*, 1997). The pHP042 insert was composed of two complete genes *glnP1* and *glnP2*, and three truncated elements *glnQ*, *cstA* and *fliI* (Beddek, 1998). The latter two gene fragments were located on opposite ends of the insert, whereas the truncated *glnQ* was contiguous with the *fliI*-like gene box.

The *glnP1, glnP2* and *glnQ* genes exhibited strong similarities (55.1%, 53.9% and 72.6%, respectively) to glutamine ABC transporter genes in *Escherichia coli* (*glnP*) and *Bacillus subtilis (glnQ)*. The *cstA* gene is thought to play an important role in the carbon starvation response of the bacterium (Beddek, 1998). Fig.3.1, panel A, illustrates the genetic organization of the pHP042 insert.

The 156 bp *fliI*-like segment exhibited similarities to internal rather than 5' or 3' terminal sequences of bacterial *fliI* genes. FliI proteins usually consist of approximately 440 amino acids. The *fliI*-like fragment of pHP042 exclusively resembled regions encoding amino acids from positions 270 to 320 of the various bacterial FliI proteins. These observations, and the existence of the mentioned *Sau*3A site, suggested that an internal *fliI* gene fragment of only 156 bp had joined the rest of the pHP042 insert by scrambled cloning during library construction. Primer pair SP002 / SP003 was designed based on the *fliI* segment present on pHP042. It allowed PCR amplification of a 96 bp fragment of the *H. pylori fliI* gene on both pHP042 plasmid DNA and *H. pylori* 17874 genomic DNA. Primer SP004 hybridized to a region within the adjoining presumptive ABC transporter gene fragment *glnQ*. PCR amplification attempts with *H. pylori* genomic DNA as template using primer pair SP002 / SP004 were unsuccessful. In contrast, this primer pair readily amplified a 395 bp fragment from pHP042 plasmid DNA (Fig. 3.1, panel B). Apparently, the *glnQ* gene box was not contiguous with the *fliI*-like gene box in the chromosome of the bacterium.

3.1.2 Subcloning of an *H. pylori* DNA segment comprising *fliI*

A *fliI* specific DIG-labelled DNA probe was prepared by DIG incorporation during PCR amplification (section 2.7.4) using primer pair SP002 / SP003 on pHP042 plasmid DNA. Successful incorporation of digoxigenin during probe preparation was demonstrated by spot test (section 2.7.4). Probe amounts of ≥ 5 pg were sufficient for signal generation during chemiluminescent detection on X-ray film.

Subsequently, 5 µg *H. pylori* 17874 genomic DNA was digested with *Hin*dIII or combinations of this enzyme with either *Eco*RI or *Bgl*II. Completion of the digest was examined by DNA gel electrophoresis of 1/10 of the sample volume. The treatment resulted in fragmentation of the genomic DNA. A smear of DNA fragments ranging in size between 10 kb and 400 bp was visible after ethidium bromide staining of the gel. Southern analysis (section 2.7.4) was performed using approximately 600 ng of the

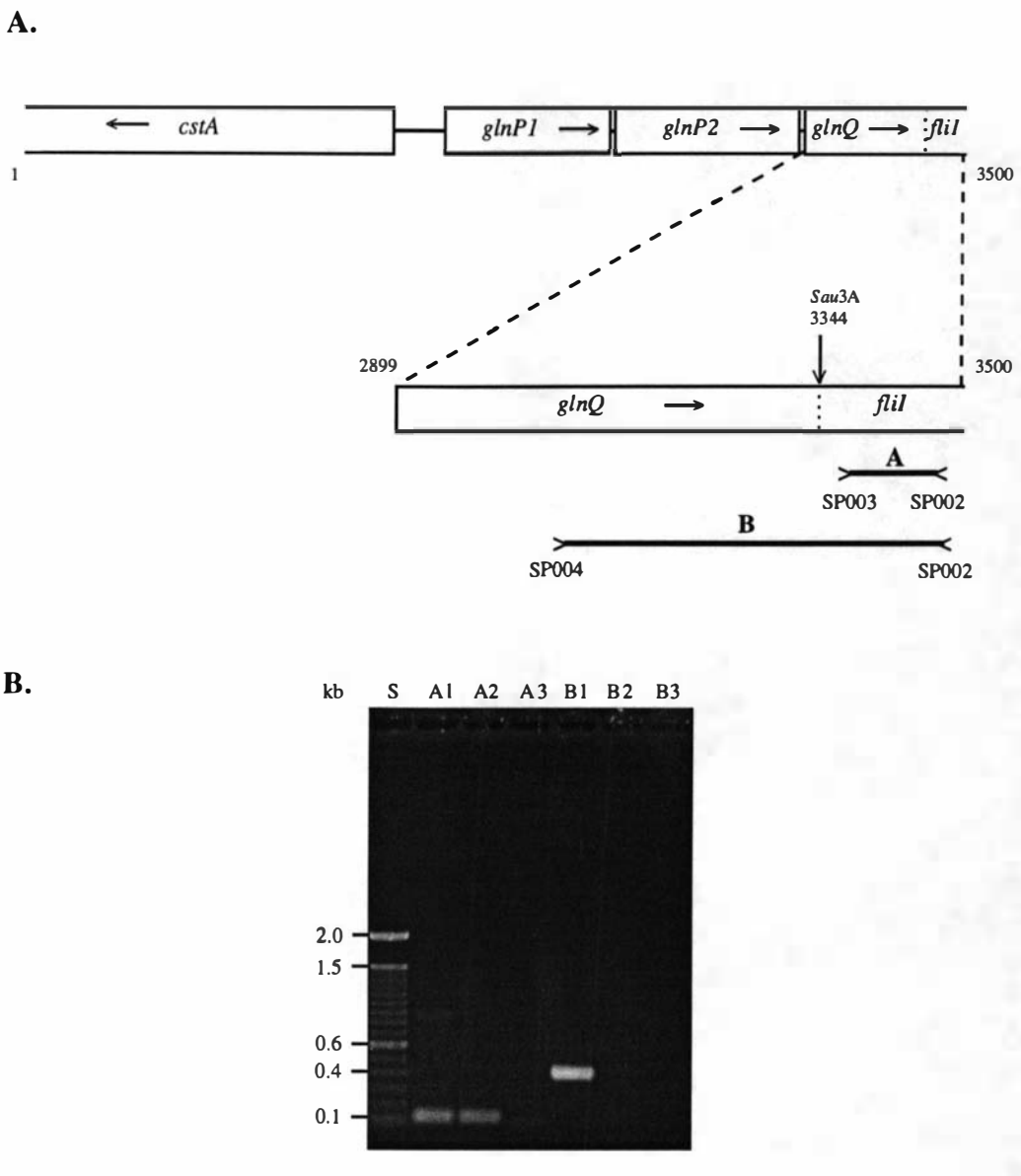


Figure 3.1. An internal segment of the *H. pylori* 17874 *fliI* gene has joined the rest of the pHP042 insert by scrambled cloning.

Panel A. Schematic representation of the genetic organization of the 3.5 kb pHP042 insert, including the *Sau*3A restriction site marking the border between the *glnQ* fragment and the *fliI*-like sequence. Primers used for linkage studies (see panel B) are indicated at their respective annealing positions.

Panel B. PCR amplifications using primer pairs SP002 / SP003 (lanes A) and SP002 / SP004 (lanes B). Templates were: lanes 1, plasmid pHP042; 2, *H. pylori* 17874 genomic DNA; 3, sterile water. Lane S, DNA molecular size standard.

digested *H. pylori* chromosomal DNA. The *fliI* specific probe hybridized to a single 7 kb *Hin*dIII fragment, and a *Hin*dIII / *Bgl*II fragment of approximately 6.5 kb (Fig. 3.2, panel A).

Approximately 2 µg of the *Hin*dIII / *Bgl*II digested *H. pylori* 17874 genomic DNA was then subjected to electrophoresis. Genomic restriction fragments of the appropriate size were gel purified and cloned into *Bam*HI / *Hin*dIII digested gel purified pUC18. Fifty two white ampicillin resistant transformants were selected for further analysis. Their plasmids were isolated by Easy plasmid minipreparation (section 2.6.1.1), and analyzed by restriction digestion using the endonucleases *Eco*RI and *Hin*dIII. Forty eight constructs contained inserts approximating the desired size. Forty six of them were slot blotted on Hybond-N^+ membrane (section 2.7.4). The *fliI* specific probe hybridized to twelve of these constructs, indicating successful cloning of *fliI* gene sequence into these plasmids (Fig. 3.2, panel B). Three more plasmids (from colonies 19, 20 and 48) reacted weakly with the *fliI* probe, but were not further investigated.

Judged from restriction analysis using *Eco*RI in combination with either *Hin*dIII or *Hin*cII, eleven of the twelve clones identified by the slot blot contained plasmids with identical 6.3 kb inserts (pSP101). However, in one of the transformants, colony 46, the insert of the plasmid appeared to be 0.7 kb bigger (pSP102). Plasmids representing these two variants were isolated by the Wizard minipreparation method (section 2.6.1.2), and analyzed further by restriction endonuclease treatment. Various enzyme combinations were used, and the resulting apparent fragment sizes of these digests are summarized in Table 3.1.

Table 3.1. Restriction analysis of the two plasmid variants identified by the *fliI* probe.

Restriction endonuclease(s)	Apparent fragment sizes (in kb)	
	pSP101	pSP102
*Sph*I	5.5; 2.3; 0.6; 0.4; 0.1	6.0; 2.3; 0.6; 0.4; 0.1
*Sph*I / *Hin*dIII	3.0; 2.3; 1.4; 0.7; 0.6; 0.4; 0.1	3.7; 2.3; 1.4; 0.7; 0.6; 0.4; 0.1
*Sph*I / *Eco*RI	4.0; 2.3; 1.1; 0.6; 0.4; 0.1	5.0; 2.3; 1.1; 0.6; 0.4; 0.1
*Sph*I / *Hin*cII	5.5; 2.3; 0.6; 0.4; 0.1	6.0; 2.3; 0.6; 0.4; 0.1
*Eco*RI / *Hin*dIII	6.0; 2.7; 0.4	6.0; 2.7; 0.7; 0.4
*Hin*dIII / *Hin*cII	5.5; 3.0; 0.55	5.5; 3.0; 0.7; 0.55
*Eco*RI / *Hin*cII	8.0; 0.9	9.0; 0.9

A restriction map was created for the two variants (Fig. 3.3, panel A), based on the restrictions outlined in Table 3.1. Plasmid pSP102 apparently contained an insert that

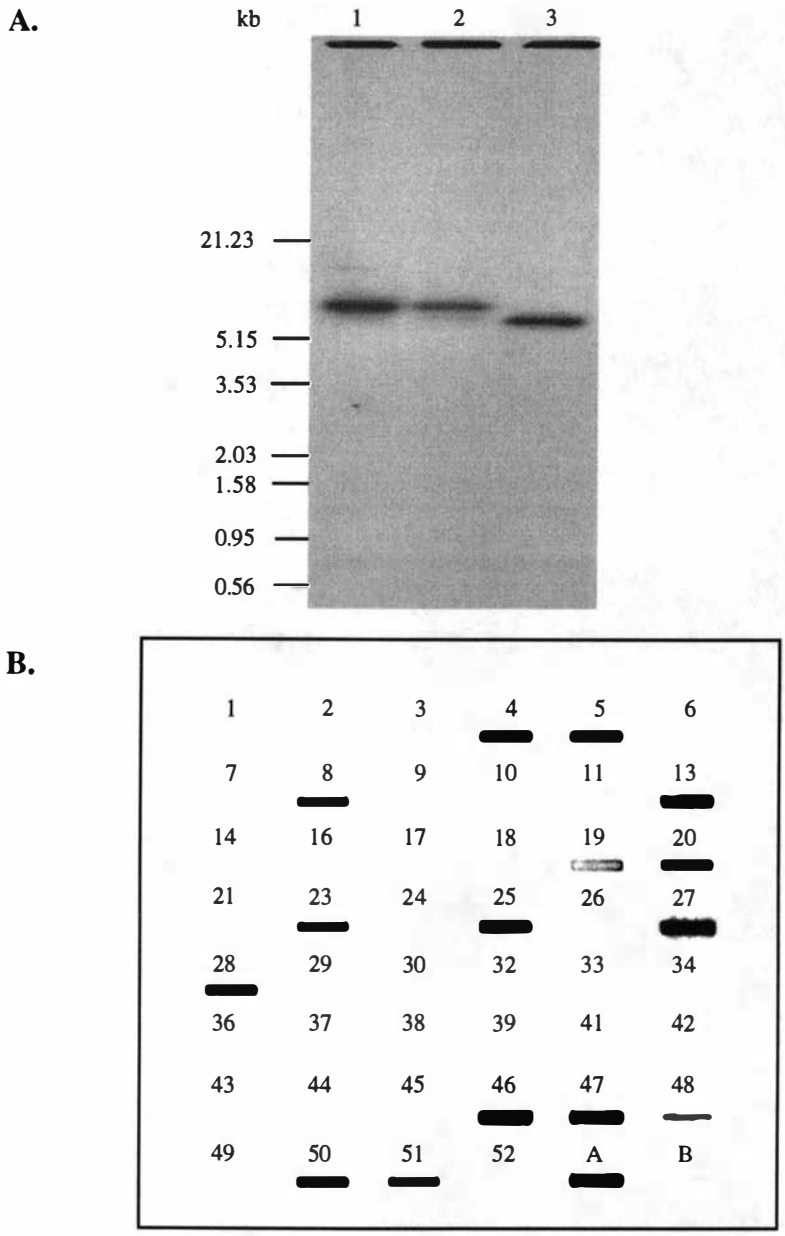


Figure 3.2. Cloning of the *H. pylori* 17874 *fliI* gene.

Panel A. Southern blotting followed by hybridization of a *fliI* specific probe on digested *H. pylori* 17874 genomic DNA. Lanes: 1, *Hin*dIII digest; 2, *Hin*dIII / *Eco*RI digest; 3, *Hin*dIII / *Bgl*II digest. The probe was prepared by DIG-PCR using primer pair SP002 / SP003 and pHP042 DNA as template.

Panel B. Slot blot of plasmids from transformant colonies harbouring a *Hin*dIII / *Bgl*II insert of the approximated desired size (6.5 kb), followed by hybridization with the *fliI* specific DIG-labelled probe. A, positive control using pHP042 plasmid DNA; B, negative control using pUC18.

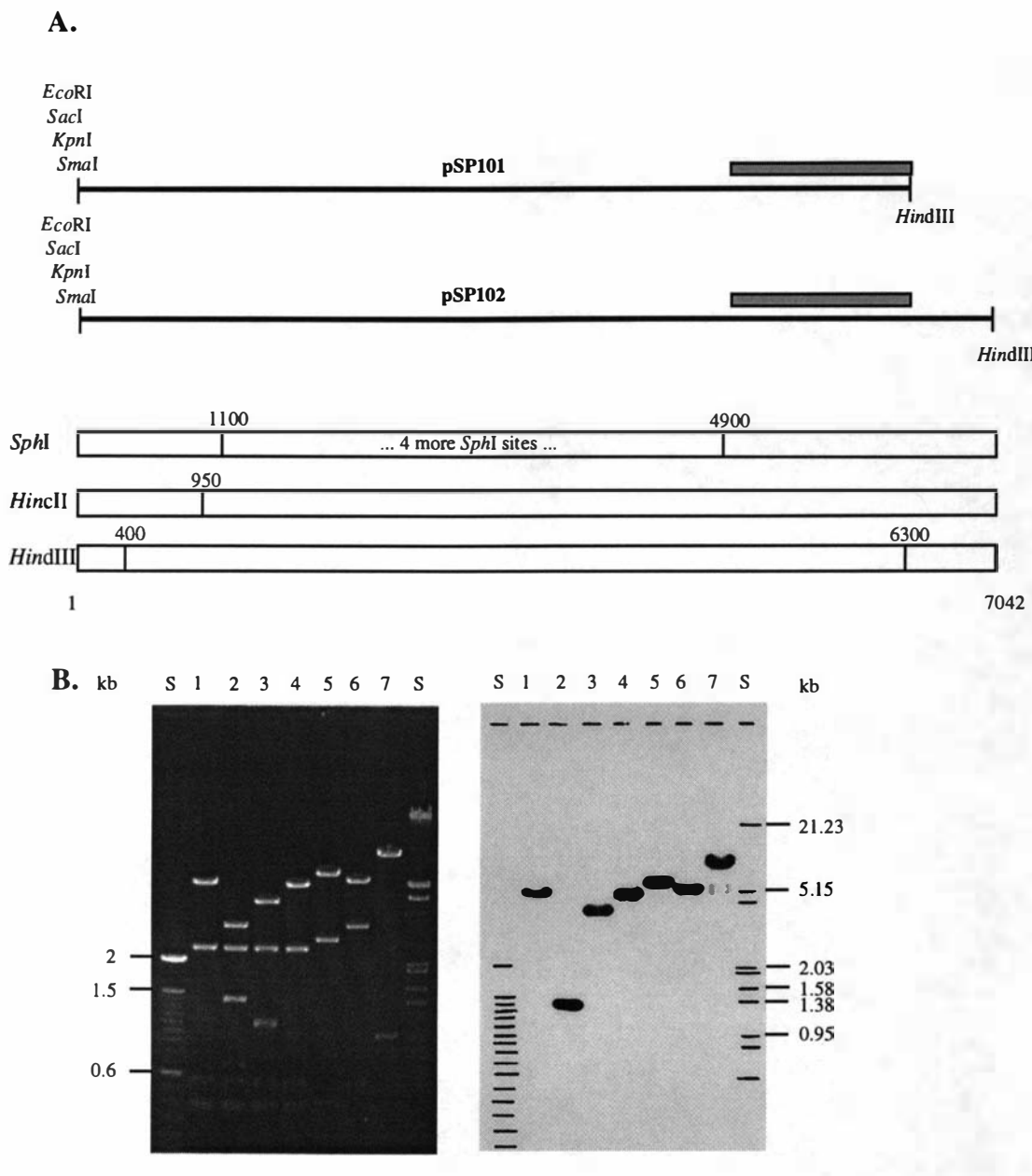


Figure 3.3. The *fliI* probe hybridized near the insert end of pSP101.

Panel A. Provisional restriction map of the pSP101 and pSP102 inserts. Restriction sites originating from the pUC18 MCS are indicated left and right of the stylized pSP101 and pSP102 inserts at the top. The grey box represents the 1.4 kb *Sph*I / *Hin*dIII restriction fragment that hybridized to the *fliI* probe (see below).

Panel B. Southern blot analysis of digested pSP101 DNA. Left: electrophoresis of the digested samples. Right: Hybridization of the transferred DNA with the *fliI* probe. Restriction enzymes were: lanes 1, *Sph*I; 2, *Sph*I / *Hin*dIII; 3, *Sph*I / *Eco*RI; 4, *Sph*I / *Hinc*II; 5, *Eco*RI / *Hin*dIII; 6, *Hin*dIII / *Hinc*II; 7, *Eco*RI / *Hinc*II. Lanes S, DNA molecular size standards.

extended that of the other eleven *fliI* positive constructs by an additional 700 bp *Hin*dIII fragment on one end.

The relative position of the *fliI* gene in the pSP101 insert was determined by Southern analysis of digested pSP101 DNA using the *fliI* specific probe (Fig. 3.3, panel B). Approximately 150 ng plasmid DNA was treated with the enzyme combinations described in Table 3.1. The fragments were separated electrophoretically, and blotted onto Hybond-N^+ membrane. The *fliI* specific probe hybridized to the 1.4 kb *Sph*I / *Hin*dIII fragment located at one end of the pSP101 insert (Fig. 3.3). Automated sequencing of this region using pUC/M13 FP on an ABI preparation (section 2.6.1.4) of plasmid pSP101 was performed. Subsequently, BLAST similarity searches revealed that the *fliI* segment present on pSP101 almost represented the complete coding sequence of the gene. However, the 3' end was missing. The additional 0.7 kb *Hin*dIII fragment present in pSP102 was expected to contain this missing 3' end of *fliI*, if it was contiguous with the insert of pSP101 in the *H. pylori* chromosome.

Primer SP010 was designed based on sequence within this extension. A PCR experiment was performed using primer pair SP003 / SP010 on *H. pylori* 17874 genomic DNA. Successful amplification of a product of the expected size (799 bp) proved generation of the extended part of the pSP102 insert by incomplete restriction rather than scrambled cloning (Fig. 3.4).

Fig. 3.4 also illustrates the presence of a *Hin*dIII site at position 395 of the pSP101 and pSP102 inserts. This site was apparently not recognized by the endonuclease during restriction of the *H. pylori* 17874 genomic DNA prior to the subcloning. It was present in the plasmids of all twelve colonies identified by the slot blot experiment to contain *H. pylori fliI* sequence. Moreover, in the initial Southern blot analysis using the *fliI* probe (Fig. 3.2), the detected *Hin*dIII fragment approximated 7 kb in size, which also suggested that the *Hin*dIII site was not recognized by the enzyme.

3.2 Sequence analysis of the pSP102 insert

3.2.1 Sequencing strategies

The genetic information present on the pSP102 insert was determined by automated sequencing on an ABI preparation of the plasmid (section 2.7.5). Both primer walking and fragment subcloning were performed. During the primer walking approach, primers were designed based on reliable chromatogram data obtained by sequencing with

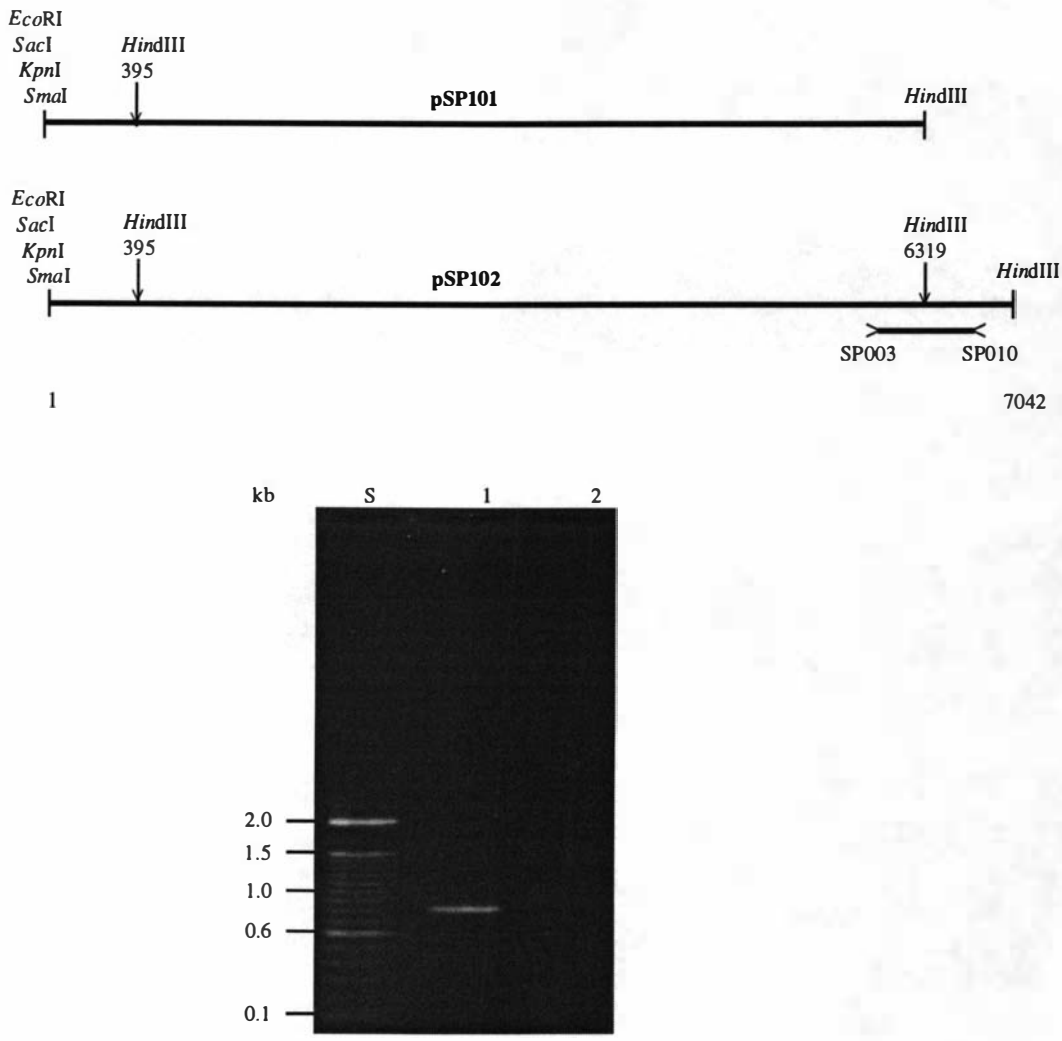


Figure 3.4. The 0.7 kb *Hin*dIII fragment of the pSP102 insert is contiguous with the pSP101 insert in the *H. pylori* 17874 genome.

The pSP101 and pSP102 inserts are represented schematically at the top. Restriction sites which originated from the pUC18 MCS are indicated left and right of the fragments. The internal *Hin*dIII sites present in the inserts are shown.

PCR was performed using primer pair SP003 / SP010 on *H. pylori* 17874 genomic DNA (lane 1), and sterile water (lane 2). Lane S, DNA molecular size standard.

the respective previous primer. The 1.4 kb *Sph*I / *Hin*dIII fragment of pSP102, which had hybridized to the *fliI* specific probe (Fig. 3.3), was subcloned. The fragment was separated from pSP102 by digestion and gel electrophoresis, gel purified, and ligated to gel purified pUC18 that had been digested appropriately. Transformants were screened by Easy plasmid preparation (section 2.6.1.1), and the desired construct pSP104 was isolated by the Wizard method (section 2.6.1.2). Primer pUC/M13 FP (Promega) was used to determine the sequence of the insert end of this construct.

Fig. 3.5 illustrates the strategy of the sequence determination of the pSP102 insert. The whole insert excluding an internal part of ORF04 was sequenced on both strands. Sequencing of the complete ORF04 was unnecessary once both ends of the gene were reached. BLAST searches revealed similarity of these ends to the same genes, namely isoleucyl-tRNA synthetase genes (*ileS*) of several organisms. The sizes of these genes were relatively conserved (approximately 2,700-3,150 bp), and corresponded to the predicted length of ORF04. This suggested that no deletion or insertion within ORF04 was present on pSP102. Primers based on the subsequently published *H. pylori* 26695 *ileS* sequence (SP034, SP040) were later used successfully in PCR experiments with *H. pylori* 17874 genomic DNA, illustrating the very high percentage of sequence identity between the *ileS* genes of those two strains.

3.2.2 Identification of the genetic elements on the pSP102 insert

Sequencing of the pSP102 insert revealed the presence of seven complete and one incomplete open reading frame. The genetic organization of the pSP102 insert is shown in Fig. 3.6. In addition, the genetic maps of the corresponding regions found in the *H. pylori* J99 chromosome (Alm *et al.*, 1999) and the *H. pylori* 26695 genome (Tomb *et al.*, 1997) are illustrated. The regions were almost identical. However, a stretch of 85 bp downstream from the first ORF *hpn* was replaced by 218 bp in *H. pylori* J99, and 1287 bp in *H. pylori* 26695. Both these replacements altered the 3' end of ORF02. In addition, two new genetic elements were present in the replacing segment in *H. pylori* 26695, the predicted products of which were not similar to any characterized proteins in the databases. Sizes of putative gene products found in pSP102, results of BLAST similarity searches and annotated putative gene names are summarized in Table 3.2.

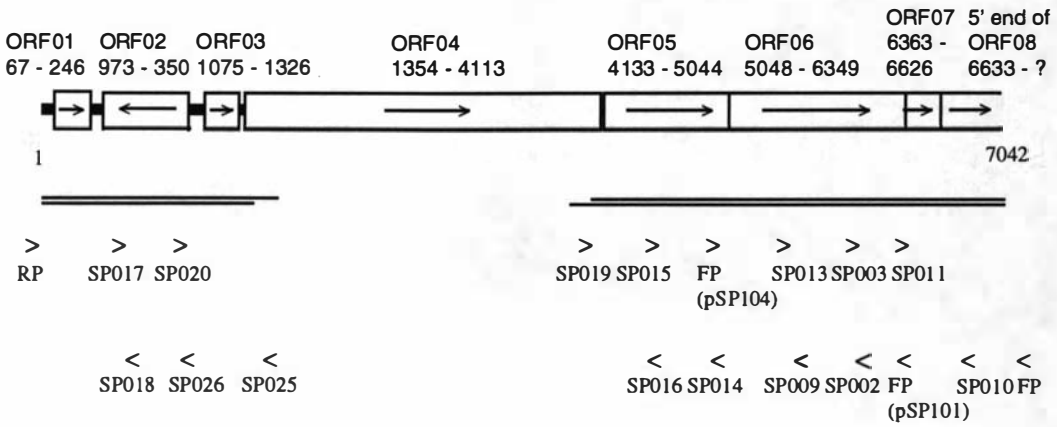


Figure 3.5. Sequence determination of the pSP102 insert.

The complete pSP102 insert is represented schematically. Coordinates and orientations of predicted ORF's on the cloned fragment are shown. Regions which were sequenced on both strands are indicated by a double line, segments which were sequenced on one strand only are marked with a single line. Primers used for sequence determination of the pSP102 insert are indicated at their respective annealing positions. FP, pUC/M13 FP; RP, pUC/M13 RP.

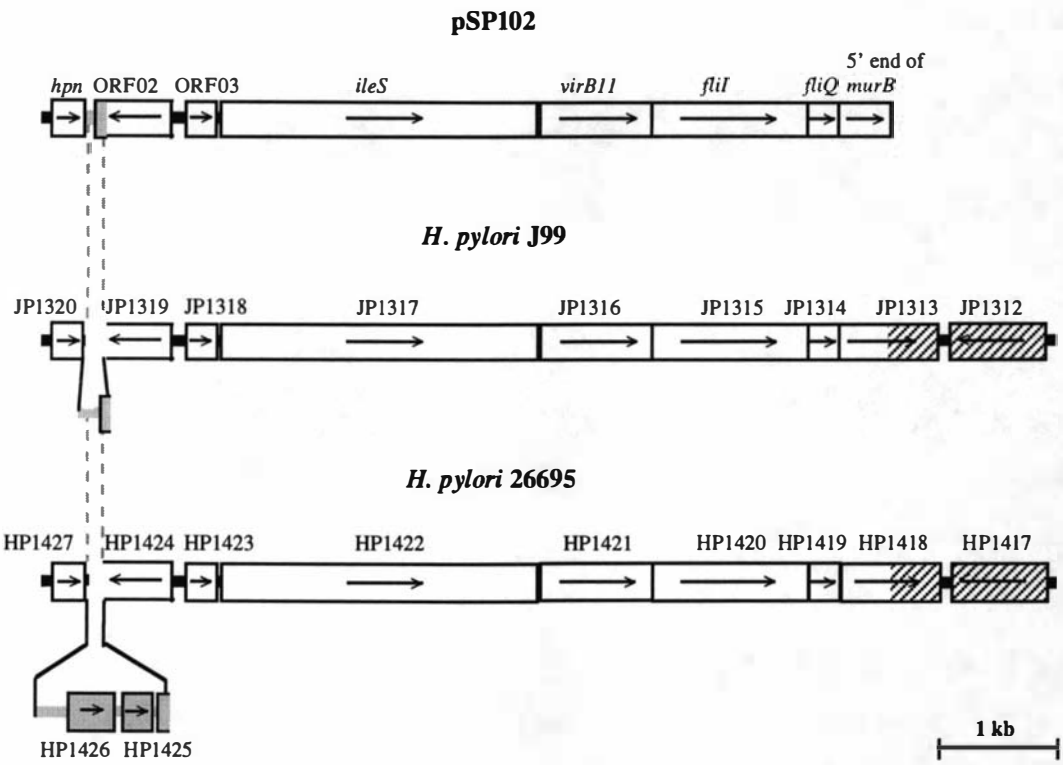


Figure 3.6. Comparison of the pSP102 insert with the corresponding regions from the *H. pylori* J99 and 26695 genomes.

The genetic organization of the pSP102 insert cloned from *H. pylori* 17874 is shown at the top. The corresponding region from the *H. pylori* J99 genome is illustrated in the middle, and the respective segment from the *H. pylori* 26695 chromosome is represented at the bottom.

Identical predicted coding sequences are white. Stripes indicate homologous regions of the chromosomes. The variable region downstream from *hpn* is shown in grey. Gene annotations are indicated on the pSP102 insert, and the respective ORF numbers of the genomes are shown.

Table 3.2. Genetic elements present in the pSP102 insert.

ORF	Size (bp)	Predicted molecular mass of gene product (kDa)	Similarity to	% identity/ % similarity	Annotated gene name
01	180	7.1	Hpn *H. pylori*	100 / 100	*hpn*
02	624	23.7	-	-	ORF02
03	251	9.4	hypothetical protein *B. subtilis*	35 / 61	ORF03
04	2760[b]	106.1[b]	IleS *C. crescentus*	49[b] / 66[b]	*ileS*
			IleS *E. coli*	39[b] / 56[b]	
05	912	34.6	TrbB *A. tumefaciens*	28 / 47	*virB11*
			TraG *E. coli* (on plasmid pMK101)	28 / 46	
			VirB11 *A. tumefaciens*	21 / 39	
			TIGR HP0525 *H. pylori* 26695	29 / 53	
06	1302	47.6	FliI *S. typhimurium*	44 / 60	*fliI*
			FliI *A. tumefaciens*	40 / 56	
			YscN *Y. enterolitica*	43 / 62	
			HrpB6 *X. campestris*	43 / 62	
			HrcN *E. amylovora*	41 / 62	
07	264	9.8	FliQ *B. subtilis*	45 / 65	*fliQ*
			FliQ *E. coli*	51 / 73	
			SpaQ *S. typhimurium*	30 / 47	
			YscS *Y. pseudotuberculosis*	34 / 58	
08[a]	780[b]	28.6[b]	MurB *B. subtilis*	31[b] / 49[b]	*murB*

[a] first 411 nt present on pSP102

[b] based on the complete gene sequences in the corresponding *H. pylori* 26695 region.

ORF01, the *H. pylori hpn* gene, has been studied by Gilbert *et al.* (Gilbert *et al.*, 1995). It encodes an extremely hydrophilic and histidine-rich Ni^{2+}- and Zn^{2+}- binding polypeptide. The protein Hpn was also found to be present in *Helicobacter mustelae*. Its function and cellular localization *in vivo* remains to be elucidated. At the amino acid level, *H. pylori* 17874 Hpn is identical to its *H. pylori* 26695 counterpart and the Hpn protein characterized by Gilbert *et al.* from a clinical *H. pylori* isolate.

The presumptive gene product of ORF02 on pSP102 displays no significant homology to any of the database entries. Its transcriptional orientation is opposite to that of all other genetic elements present on pSP102. Sequence analysis using Geneworks software showed that the putative gene product is strongly hydrophobic at its N-terminus, but relatively hydrophilic elsewhere. It exhibits a high isoelectric point pI (10.39). An N-terminal signal sequence is predicted (Tomb *et al.*, 1997).

ORF03 encodes a putative product with moderate similarities to a hypothetical 9.7 kDa protein in *B. subtilis*. The corresponding gene is located in the region between a gene coding for a protein involved in DNA repair and recombination (*mfd*) (Ayora *et al.*, 1996) and a cell division gene (*divIC*) (Levin & Losick, 1994). The C-terminus of the

predicted ORF03 gene product is hydrophilic, but hydrophobic features predominate overall. A high pI (10.1) was determined for the encoded protein by the Geneworks software programme.

ORF04 is the isoleucyl-tRNA synthetase gene *ileS* of *H. pylori*. Like the other elements of the pSP102 insert, *ileS* is a single copy element in *H. pylori* 26695 (Tomb *et al.*, 1997). It displays hydrophilic N- and C-termini, and lacks an N-terminal signal sequence. Absence of long hydrophobic regions in the predicted protein are consistent with its presumed function as an intracellular protein. The IleS protein catalyzes the transfer of isoleucine molecules onto the specific isoleucyl-tRNA. It activates the amino acid by adenylation, and subsequently transfers the adenylate onto the tRNA (Stryer, 1988).

The three elements which follow, ORF05, ORF06 and ORF07, are all similar to genes with predicted functions in protein export and are collectively referred to as a protein export locus.

ORF05, the first gene of this locus, was annotated as *virB11*. It displays significant identity to genes encoding proteins known to be involved in bacterial nucleoprotein transport (*Agrobacterium tumefaciens trbB, E. coli traG*). Both these homologs are plasmid encoded, and are thought to be essential for conjugal DNA transport (Pohlman *et al.*, 1994; Alt-Morbe *et al.*, 1996). Furthermore, ORF05 is similar to the *A. tumefaciens virB11* gene, which is necessary for interkingdom nucleoprotein transfer. The gene encodes an autophosphorylating ATPase essential for successful transfer of T-DNA from *A. tumefaciens* into the infected plant cell (Christie *et al.*, 1989; Ward *et al.*, 1990) (section 1.2; Fig. 1.3). At position 145 of the translated ORF05 sequence a putative ATP/GTP binding motif (GETSSGKT) was identified. This motif corresponds to the consensus sequence G/A-X_4-G-K-T/S of a Walker A box (Walker *et al.*, 1982; Higgins *et al.*, 1986). As illustrated in Fig. 3.7, this motif is conserved in all bacterial genes with significant homology to ORF05. Low surface probability and antigenicity of the ORF05 gene product were predicted by the GeneJockey software package.

In the *H. pylori* 26695 genome, ORF05 was annotated as *trbB* (Tomb *et al.*, 1997). However, sequence identity of 28.6% and 52.6% similarity to a *virB11* homolog present in the *cag* pathogenicity island of *H. pylori* were identified by the BLAST algorithm. This *cag-virB11* homolog has been suggested to encode a protein which is involved in the export of *H. pylori* proteinaceous factors across kingdom boundaries as part of a *vir*-like transport machinery (Christie, 1997b). The observed resemblance and

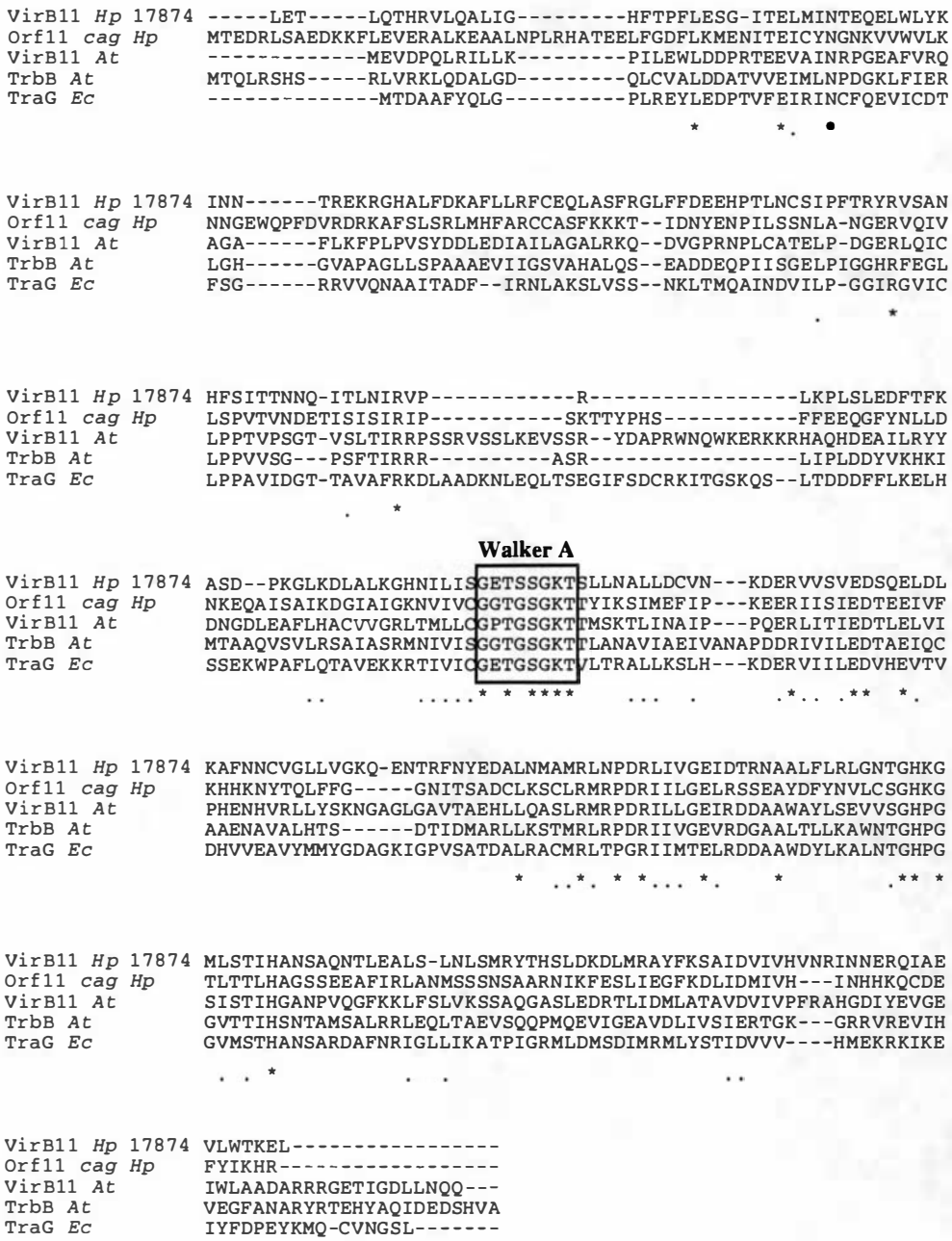


Figure 3.7. Similarity of *H. pylori* 17874 VirB11 to presumptive ATP binding nucleoprotein transport components.

Multiple sequence alignment of *H. pylori* 17874 VirB11, ORF11 of the *H. pylori* 26695 *cag* pathogenicity island (*cag*-VirB11), *A. tumefaciens* VirB11 and TrbB, and the TraG protein from the *E. coli* plasmid pKM101. The conserved nucleotide binding motif Walker A (G/A-X_4-G-K-T/S) is boxed. Identical amino acids (stars) and conservative changes (dots) are indicated. *Hp, H. pylori; At, A. tumefaciens; Ec, E. coli.*

the fact that the *ileS*-linked ORF05 is as similar to *A. tumefaciens virB11* as the *cag-virB11*, justified the suggested *virB11* annotation for ORF05.

ORF06 is the *H. pylori* 17874 *fliI* gene. A preliminary study of the *H. pylori* N6 *fliI* gene was published during the course of this thesis (Jenks *et al.*, 1997). *H. pylori* 17874 FliI displays 99% identity to its counterpart in *H. pylori* N6 at the amino acid level. BLAST searches detected high similarity of the gene to a number of *fliI* genes in other bacteria as well as to many elements whose products are known or suspected to play a role in antigen presentation by a type III secretion mechanism (*Xanthomonas campestris* HrpB6, *Yersinia enterolitica* YscN, *Erwinia amylovora* HrcN). The FliI protein also displays similarity to the β subunit of bacterial F_0F_1 ATPases. Like other ATP synthases and suggested energizers of type III protein export, the predicted protein contains the Walker box A (G/A-X_4-G-K-T/S) and Walker box B (R/K-X_{2-3}-G-X_3-L-$X_{3-5,\ \text{hydrophobic}}$-D), which have been postulated to be essential for ATP binding (Walker *et al.*, 1982) (Fig. 3.8). In the FliI amino acid sequence, the Walker A motif (AGSGVGKS) starts at position 164, the Walker B motif KNQGLDVLFIMD can be found at position 242 to 253. Further analysis using the Geneworks software programme revealed that the encoded FliI protein has low surface probability and antigenicity, and hydrophilic termini. In addition, the programme identified the presence of a conserved ATPase α and ß subunits signature sequence in FliI, starting at position 344. This motif (S/A/P-I/V-D/N-X_3-S-X-S) is also present in most of the other genes which are significantly similar to FliI (Fig. 3.8).

ORF07 is the *H. pylori fliQ* homolog. It closely resembles *fliQ* genes of other bacteria as well as elements suspected to encode inner membrane components necessary for protein export type III (exemplified by *Yersinia pseudotuberculosis yscS* and *S. typhimurium spaQ*). The Geneworks sequence analysis programme identified strong hydrophobicity of FliQ suggesting a membrane localization. Two transmembrane segments were predicted, one ranging from the amino acid residue at position 10 of the *fliQ* gene product to the residue located at position 39, the second segment starting at position 46 and extending to position 70. This hydropathicity pattern is conserved in all FliQ homologs in type III protein export systems (Fig. 3.9). As expected for a membrane protein with apparently little surface exposure, the predicted antigenicity of the gene product is very low. A relatively high pI of 9.8 was calculated for FliQ.

ORF08, the gene immediately downstream from this export locus, was annotated *murB*. This gene was truncated on the pSP102 insert, only the first 411 bp

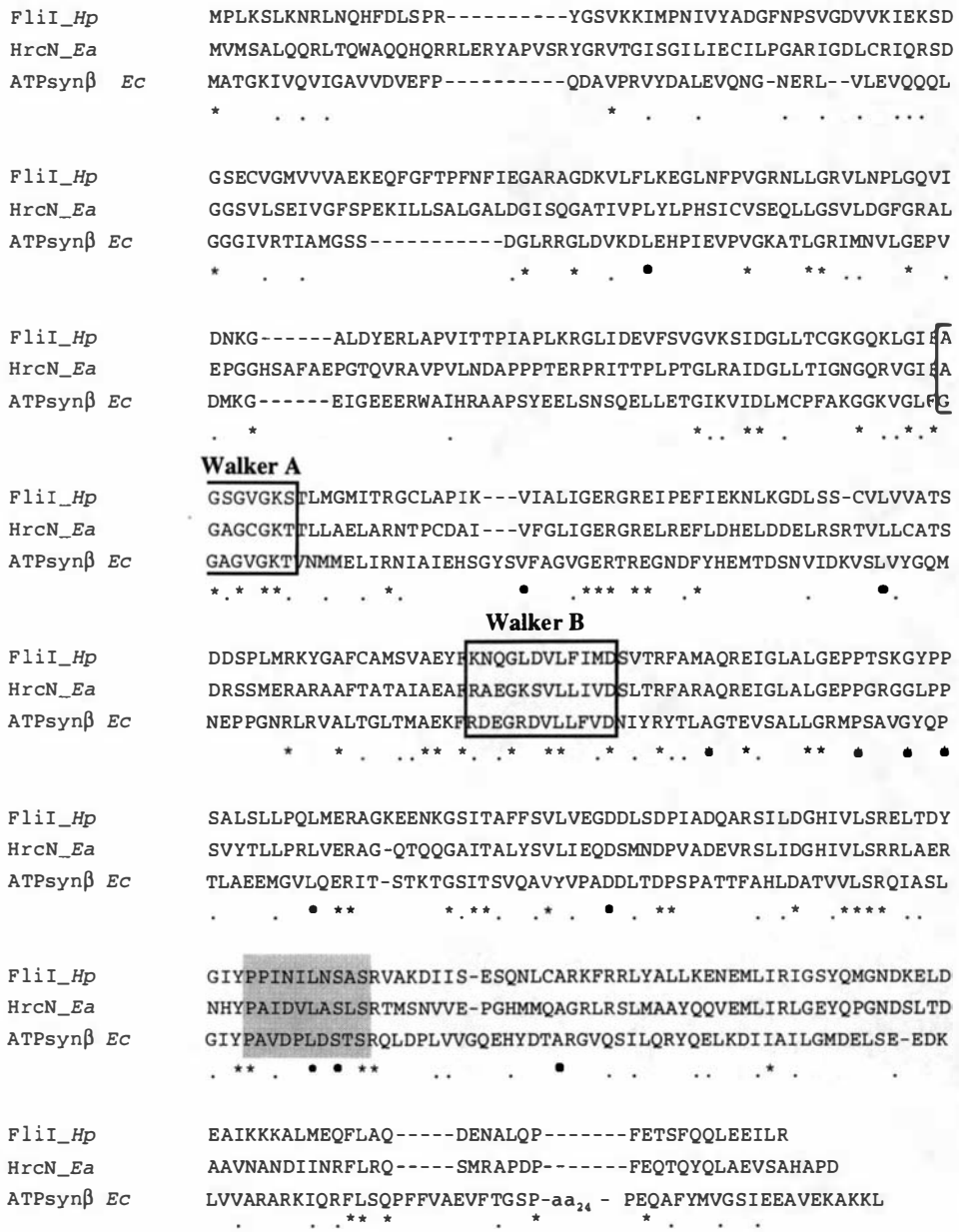


Figure 3.8. Conserved sequence motifs in *H. pylori* 17874 FliI.

Multiple sequence alignment of *H. pylori* 17874 FliI, *E. amylovora* HrcN and the ß subunit of the *E. coli* F_0F_1 ATPase. The Walker A (G/A-X_4-G-K-T/S) and Walker B (R/K-X_{2-3}-G-X_3-L-$X_{3-5,\ \text{hydrophobic}}$-D) motifs thought to be necessary for ATP binding are boxed. The conserved ATPase α and β subunits signature sequence (S/A/P-I/V-D/N-X_3-S-X-S) is shaded. Identical amino acids (stars) and conservative changes (dots) are indicated. *Hp, H. pylori; Ea, E. amylovora; Ec, E. coli.*

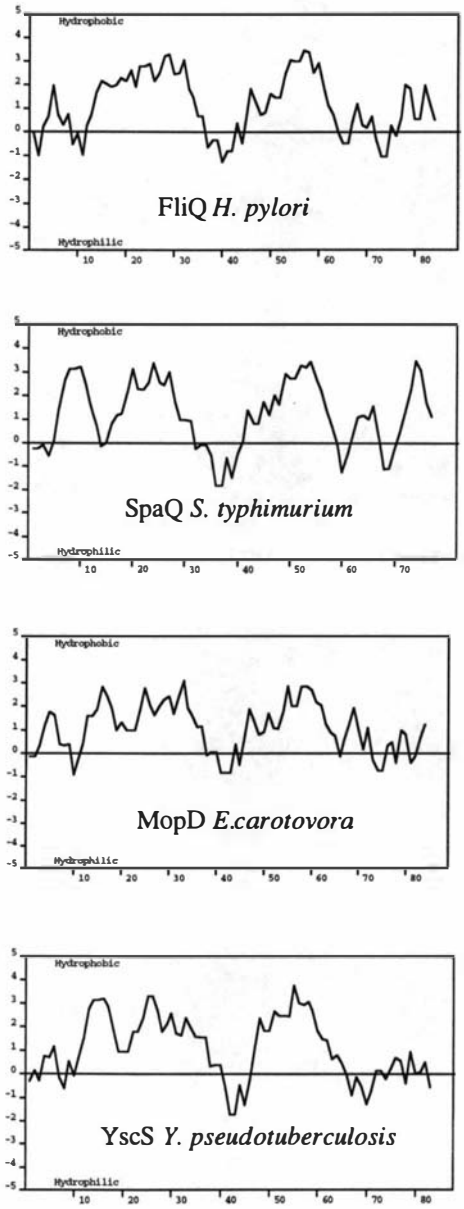


Figure 3.9. Similarity in hydropathicity plots of *H. pylori* 17874 FliQ and components of type III protein export systems.

The Kyte and Doolittle hydropathicity plots were constructed using the GeneJockey sequence analysis programme with default parameters.

being present. The complete sequence of ORF08 was subsequently obtained from the *H. pylori* 26695 genome sequence. It was found to be similar to the putative *B. subtilis* UDP-N-acetylenolpyruvoylglucosamine reductase gene, the product of which is presumed to be essential for cell wall biosynthesis (Rowland *et al.*, 1995). The MurB enzyme (E.C.1.1.1.158) catalyzes the second reaction step in the conversion of UDP-N-acetylglucose into UDP-N-acetylmuramic acid, which represents the first stage of peptidoglycan synthesis (Taku *et al.*, 1970). The translated ORF08 sequence lacks a signal sequence. Strongly hydrophobic segments are also absent.

3.2.3 Further sequence features of the pSP102 insert

The spacing of the described genes on pSP102 was noteworthy and suggested ORF03, *ileS*, *virB11*, *fliI*, *fliQ* and *murB* to be transcribed on a single mRNA. Whereas intragenic regions between *hpn*, ORF02 and ORF03 were 103 nt and 101 nt, respectively, the following open reading frames were all positioned very close to each other. The distance between ORF03 and *ileS* was 27 nt, between *ileS* and *virB11* only 19 nt, *virB11* was spaced from *fliI* by 3 nt, *fliI* from *fliQ* by 13 nt, and *fliQ* from *murB* by 6 nt, including stop codons. The next longer intragenic region appeared in the *H. pylori* 26695 genome downstream from *murB* (66 nt), and the following gene with moderate similarities to hypothetical proteins of unknown functions in *E. coli* (TIGR HP1417) is orientated opposite to the *murB* direction (Tomb *et al.*, 1997). In support of the presumed operon structure of these six elements, a search for putative transcription terminating structures between ORF03 and *murB* using PCGene V6.85 software was unsuccessful.

Ribosome binding sites for all the elements present on the pSP102 insert were detected. The respective start codons are preceded by AGGA-N_6 (*hpn* and ORF02), AGGG-N_{12} (ORF03), AGGA-N_7 (*ileS*), AGGA-N_8 (*virB11*), AGAG-N_6 (*fliI*), AGGA-N_5 (*fliQ*), and, suboptimally, AGA-N_2 (*murB*).

The G+C content of the protein export locus (*virB11* to *fliQ*) was 43%, compared to the overall 39% in *H. pylori* 26695 (Tomb *et al.*, 1997).

Predating the release of the *H. pylori* 26695 genome sequence, the determined sequence of the *H. pylori* 17874 *virB11*, *fliI* and *fliQ* genes was submitted to the GenBank database, and was allocated accession number U75584.

3.3 Conservation of the putative operon in *H. pylori*

Conservation of the genetic linkage of ORF03, *ileS*, *virB11*, *fliI*, *fliQ* and *murB* was tested in clinical *H. pylori* isolates obtained from Auckland, New Zealand (Campbell *et al.*, 1997). Investigations were performed by PCR using appropriate primer pairs and genomic DNA of the respective *H. pylori* strains. These chromosomal DNA templates were prepared as described in section 2.6.2. Negative controls without added DNA templates did not produce amplification products.

Conservation of the linkage of ORF03 to *ileS* was tested using the ORF03 internal primer SP039 in conjunction with the *ileS* internal primer SP040 which was designed based on the published *H. pylori* 26695 sequence. Eight isolates were tested, three of which were type I *H. pylori* strains (*cagA*$^+$/VacA$^+$; *H. pylori* MU003, MU030 and MU079), two were type II (*cagA*$^-$ /VacA$^-$; *H. pylori* MU015 and MU099). The three remaining isolates were intermediate strains with either *cagA*$^+$/VacA$^-$ (*H. pylori* MU082 and MU067) or *cagA*$^-$/VacA$^+$ (*H. pylori* MU038) characteristics. All eight isolates demonstrated the same expected 649 bp amplification product, consistent with the type strains *H. pylori* 17874 (type I) and 915 (type II), and the mouse adapted type I *H. pylori* SS1.

Overlapping polymerase chain reactions were performed to cover the regions from *ileS* to *fliI* (*ileS* internal primer SP034 / *fliI* internal primer SP009), and from *fliI* to *murB* (*fliI* internal primer SP013 / *murB* internal primer SP036, based on the *H. pylori* 26695 genome sequence). In addition to the previously mentioned eight clinical isolates, ten more type I strains (*cagA*$^+$/VacA$^+$) were investigated, *H. pylori* MU002, MU007, MU016, MU022, MU043, MU044, MU045, MU074, MU102 and MU103. In all isolates, amplifications resulted in exactly the same expected pattern, displaying major products of 2,438 bp and 1,932 bp, respectively. These results were consistent with the *H. pylori* strains 17874, 915 and SS1. Within the panel of strains tested, the investigated locus was highly conserved. Fig. 3.10 illustrates the conservation of the genetic linkage from ORF03 to *murB* by display of a representative selection of clinical isolates covering type I, type II and intermediate strain characteristics.

In addition to these investigations, conservation of the genetic elements flanking the presumed operon was also tested by PCR. The subset of strains described for the first set of conservation studies (linkage of ORF03 to *ileS*) was examined. When testing conservation of the genetic linkage of the 5' end of ORF02 with *ileS* using primer pair

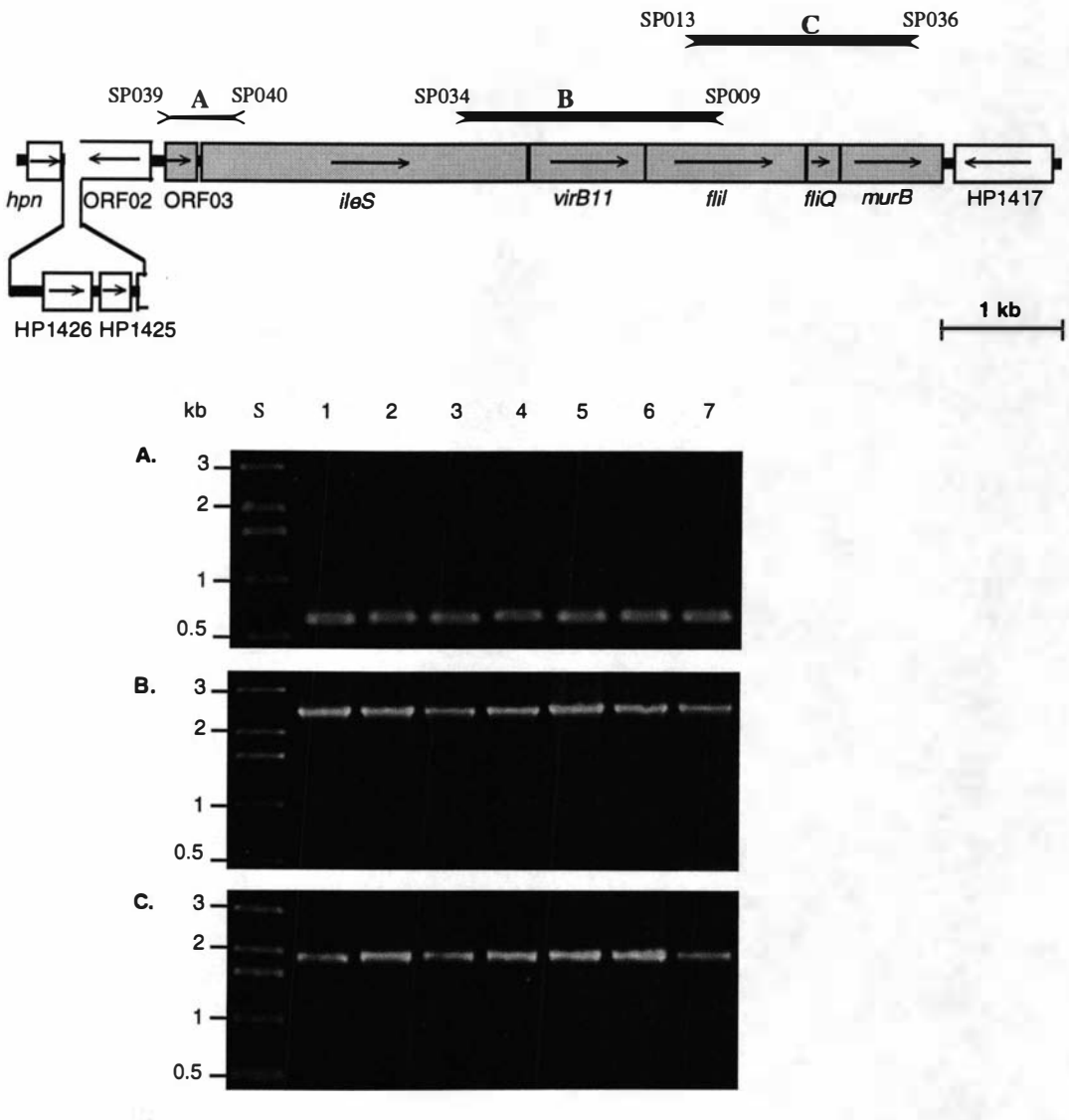


Figure 3.10. The genetic linkage of the components of the putative operon is conserved in *H. pylori.*

PCR's were performed using genomic DNA of *H. pylori* strains as template. Panels A - C correspond to reactions using primer pairs amplifying the indicated regions A - C illustrated at the top. Lanes: 1, *H. pylori* MU015 (*cagA*$^-$ / VacA$^-$); 2, *H. pylori* MU099 (*cagA*$^-$ / VacA$^-$); 3, *H. pylori* MU003 (*cagA*$^+$ / VacA$^+$); 4, *H. pylori* MU079 (*cagA*$^+$ / VacA$^+$); 5, *H. pylori* MU038 (*cagA*$^-$ / VacA$^+$); 6, *H. pylori* MU082 (*cagA*$^+$ / VacA$^-$); 7, *H. pylori* 17874. Lane S, DNA molecular size standard. No products were amplified in negative controls with sterile water as template (not shown).

The corresponding DNA region in the *H. pylori* 26695 genome is illustrated at the top. Grey boxes represent the genes of the putative operon.

SP020 (located near the 5' end of ORF02) / SP040 (*ileS* internal), all eight tested clinical isolates displayed the same expected product of 859 bp, consistent with *H. pylori* 17874, 915 and SS1, and the *H. pylori* 26695 and J99 genome sequences (Fig. 3.11, panel A).

However, investigations of the linkage of *hpn* with *ileS* revealed less consistency (Fig. 3.11, panel B). In five of the eight tested clinical isolates, amplifications using the *hpn* internal primer SP045 with the *ileS* internal primer SP040 resulted in a major product of 2,769 bp, consistent with the *H. pylori* 26695 sequence. This product was also obtained when testing the *H. pylori* strains SS1, 915 and 17874. In the latter, this result contradicted the sequencing data obtained from plasmid pSP102. Based on these data a major PCR product of 1,275 bp would have been expected. The remaining three clinical strains displayed major products of approximately 1.6 kb (Fig. 3.11, panel B). No correlation of the different amplification products with the classification of *H. pylori* strains into type I or type II was apparent. The three different strains were either type I (*H. pylori* MU079), type II (*H. pylori* MU015), or intermediate (*H. pylori* MU067). None of these strains resulted in the 1,275 bp product predicted from the pSP102 sequence. Their product sizes were also inconsistent with the *H. pylori* J99 genome sequence, where an amplification fragment of 1,408 bp would be expected.

Conservation of the linkage to the next genetic element downstream from *murB*, TIGR HP1417, was also investigated. Amplifications using the *fliQ* internal primer SP028 in conjunction with the HP1417 internal primer SP043 (based on the *H. pylori* 26695 genome sequence) on clinical isolates demonstrated a lower degree of conservation (Fig. 3.11, panel C). Four of the isolates - *H. pylori* MU030, MU067, MU079 and MU082 - exhibited a major amplification product of 1,391 bp, consistent with the *H. pylori* 26695 and J99 genome sequences, *H. pylori* 915, and *H. pylori* SS1. The four remaining clinical strains and *H. pylori* 17874 displayed a product which was approximately 700 bp bigger than expected. Two of these five "different" strains were type II, two type I and one was an intermediate (*cagA*$^-$/VacA$^+$).

Neither of these polymerase chain reactions resulted in any amplification product on *H. mustelae* 4298 genomic DNA. Moreover, attempts to amplify the individual genes *virB11*, *fliI*, *fliQ* and *murB* failed when using primer pairs designed on *H. pylori* sequence. This suggested a low degree of residue identity of those genes in *H. mustelae,* if at all present (Fig. 3.12). Integrity of the template and PCR reagents was proven by successful amplification of a 16SrRNA gene fragment using the *Helicobacter* specific

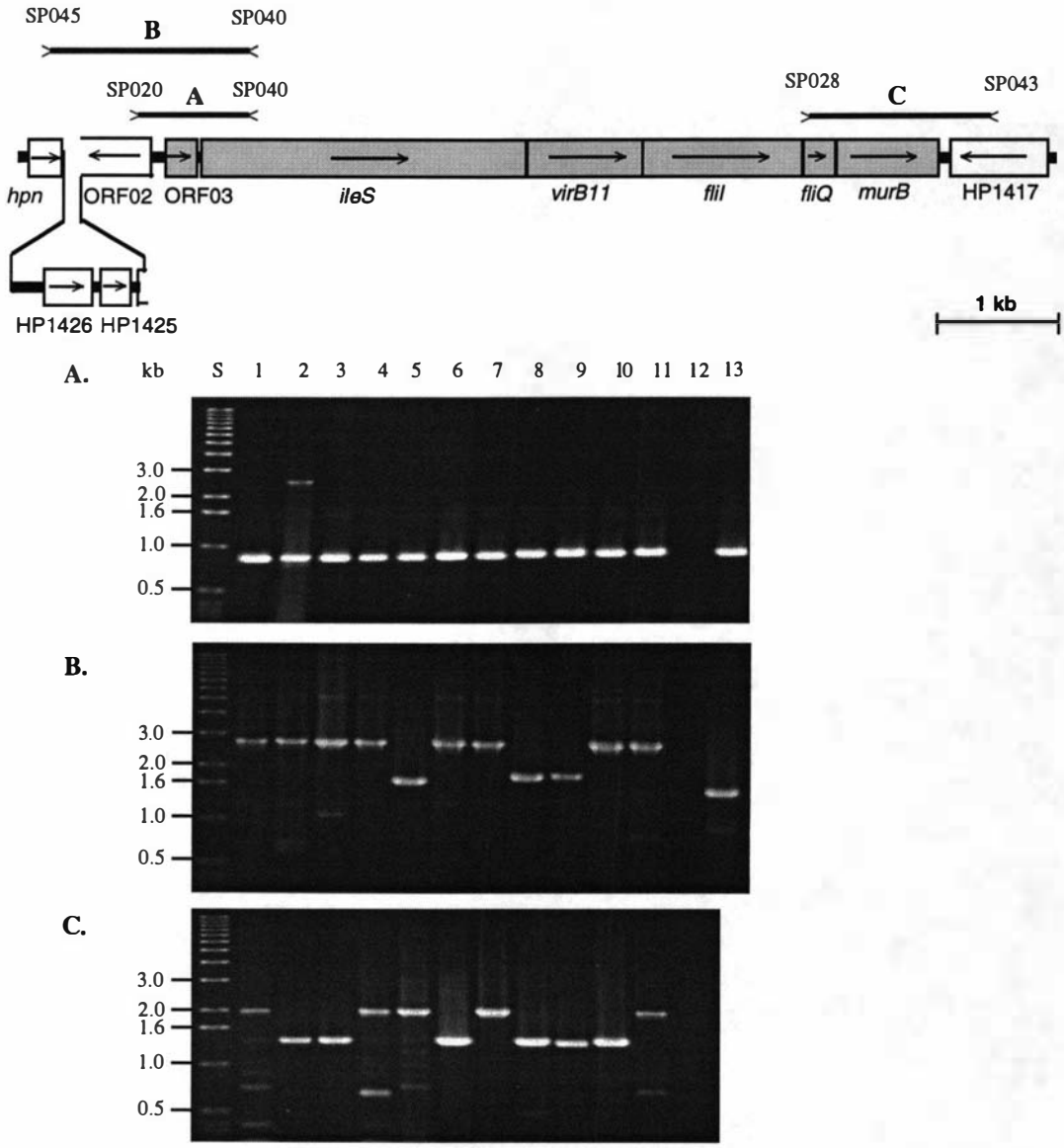


Figure 3.11. The flanking regions of the putative operon are not strictly conserved in *H. pylori*.

PCR's were performed using genomic DNA of *H. pylori* strains as template. Panels A - C correspond to reactions using primer pairs amplifying the indicated regions A - C illustrated at the top. Lanes: 1, *H. pylori* 17874; 2, *H. pylori* 915; 3, *H. pylori* SS1; 4, *H. pylori* MU003 (*cagA*$^+$ / VacA$^+$); 5, *H. pylori* MU015 (*cagA*$^-$ / VacA$^-$); 6, *H. pylori* MU030 (*cagA*$^+$ / VacA$^+$); 7, *H. pylori* MU038 (*cagA*$^-$ / VacA$^+$); 8, *H. pylori* MU067 (*cagA*$^+$ / VacA$^-$); 9, *H. pylori* MU079 (*cagA*$^-$ / VacA$^+$); 10, *H. pylori* MU082 (*cagA*$^+$ / VacA$^-$); 11, *H. pylori* MU099 (*cagA*$^-$ / VacA$^-$); 12, sterile water; 13, pSP101 plasmid DNA. Lane S, DNA molecular size standard.

The corresponding DNA region in the *H. pylori* 26695 genome is illustrated at the top. Grey boxes represent the genes of the putative operon.

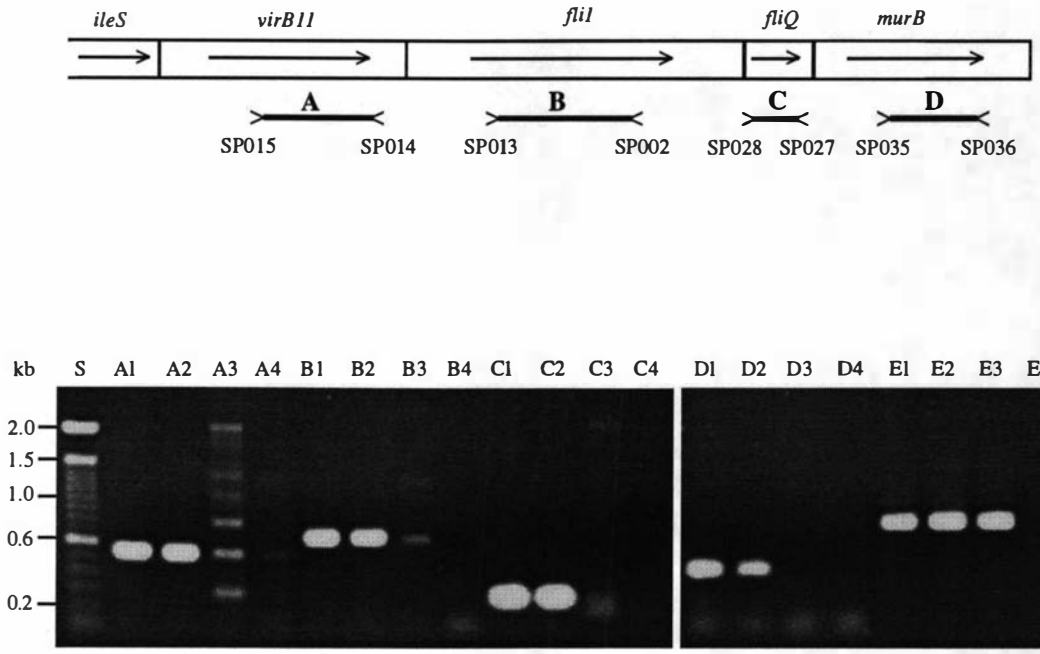


Figure 3.12 The components of the putative *H. pylori* operon are not conserved in *H. mustelae.*

PCR's were performed using genomic *Helicobacter* DNA as template. Lanes A - D correspond to reactions using primer pairs amplifying the indicated regions A - D illustrated at the top. Lane E, *Helicobacter* 16S rRNA gene specific primer pair HS16sF / HS16sR. Lanes: 1, *H. pylori* 17874; 2, *H. pylori* 915; 3, *H. mustelae* 4298; 4, sterile water. Lane S, DNA molecular size marker.

primer pair HS16sF / HS16sR on genomic DNA extracted from both *H. pylori* and *H. mustelae*.

3.4 Transcript analyses

3.4.1 Transcript detection attempts by Northern blotting

A series of Northern blots followed by hybridization with different probes was performed to detect a transcript containing *fliI* specific mRNA. In addition, detection of mRNA comprising the genetic information for three other *H. pylori* genes was attempted as controls. Transcripts of the genes encoding the *H. pylori* alkyl hydroperoxide reductase TsaA (O'Toole *et al.*, 1991), the *H. pylori* GroES homolog HspA (Suerbaum *et al.*, 1994), and the flagellar hook protein FlgE (O'Toole *et al.*, 1994) were targeted. The methodology is outlined in section 2.12.3. The recombinant plasmids and vectors used during probe preparations are described in Table 2.5. The plasmids of the pHP and pSP series and p26K are illustrated in Appendix 1.

Various gene specific probes were used. Four different *fliI* probes were constructed. Preparation of the DIG labelled DNA probe was described in section 3.1.2. In addition, *fliI* specific ECL- or [α-^{32}P]-labelled DNA probes were prepared as follows. A *Hin*dIII / *Bst*XI restriction digest of plasmid pSP104 released a 757 bp fragment containing *fliI* DNA. This fragment was gel purified, and subsequently ECL-labelled. Alternatively, it was [α-^{32}P] radiolabelled.

For construction of a *fliI* specific riboprobe, PCR was performed with pSP102 plasmid DNA as template using *Pwo* polymerase and the *fliI* internal primer pair SP002 / SP013. The amplified 641 bp fragment was gel purified and ligated to *Sma*I digested gel purified pUC19 to create plasmid pSP112. It was important to check the direction of the fragment within this construct to ensure transcription of the correct (i.e. non-coding) strand of the gene segment during probe synthesis. Orientation of the insert was determined by restriction analysis using *Bst*XI and *Eco*RI, exploiting the asymmetrically positioned *Bst*XI site present in the SP002 / SP013 fragment. This restriction site divided the fragment into subfragments of approximately 210 bp and 430 bp. An insert in the desired orientation was excised from pSP112 using the restriction endonucleases *Eco*RI and *Hin*dIII. The fragment was gel purified and ligated to accordingly digested and purified vector pSPT18 to generate plasmid pSP115. This plasmid was *Bam*HI digested to prevent run-on transcription of vector sequence during probe preparation. The digested plasmid was phenol / chloroform extracted, ethanol precipitated and subsequently used

for riboprobe generation. Successful incorporation of digoxigenin during probe preparation was shown by spot test (section 2.12.3.4). Probe amounts of ≥ 10 pg were sufficient for chemiluminescent signal generation on X-ray film.

Construction of an ECL-labelled *tsaA* specific probe was initiated by *Hin*dIII restriction of plasmid p26K. The digest released a fragment of approximately 910 bp containing the *H. pylori tsaA* gene. It was gel purified and subsequently labelled. In addition, a [α-^{32}P] radiolabelled *tsaA* probe was also prepared from this restriction fragment.

An *hspA* specific riboprobe was prepared as follows. PCR was performed with *Taq* polymerase using primer pair SP007 / SP008 on *H. pylori* 17874 genomic DNA. These primers were designed based on the 5' end and the 3' end of the *H. pylori hspA* gene. The resulting 370 bp PCR product was gel purified and cloned into vector pGEM-T to create pSP201. The primers incorporated a *Bam*HI site on their 5' ends, and the cloned fragment was subsequently excised from pSP201 by *Bam*HI digestion. It was gel purified and ligated to *Bam*HI cut, gel purified vector pSPT18. The orientation of the subcloned fragment was determined by restriction analysis exploiting the asymmetrical *Hin*dIII site in the *H. pylori hspA* gene. This site divided the gene into two fragments of approximately 30 bp and 320 bp. A construct which incorporated the *hspA* fragment in the desired orientation (pSP114) was *Xba*I digested, phenol / chloroform extracted, ethanol precipitated and used in the riboprobe generating labelling reaction. A spot test showed successful digoxigenin incorporation. Spots containing ≥ 2 pg of probe generated signals on X-ray film.

The same principles were applied when preparing the *flgE* specific riboprobe. A 0.6 kb *Sph*I restriction fragment originating from plasmid pHP014 was gel purified and ligated to appropriately treated pUC19. Conveniently positioned asymmetrical *Hin*dIII restriction sites within the fragment were used to analyze the orientation of the insert. The three sites were present very close to one end of the segment, and approximately 450 bp away from the other fragment end. The desired variant (pSP119) was digested with *Eco*RI / *Hin*dIII, which resulted in release of a 450 bp *flgE* specific fragment. It was gel purified and cloned into appropriately treated pSPT18 to form pSP120. This construct was *Hin*dIII digested, phenol / chloroform extracted, ethanol precipitated and used for riboprobe generation. DIG incorporation during probe amplification was demonstrated by the spot test. Riboprobe amounts of ≥ 20 pg were sufficient to generate a signal on X-ray film.

In addition, two *flgE* specific DNA probes were generated. A *Pst*I / *Ssp*I restriction digest of pHP012 released a 700 bp fragment containing sequence of the *H. pylori flgE* gene. The gel purified fragment was subsequently ECL-labelled. Alternatively, it was subjected to [α-^{32}P]-labelling reactions.

As well as using different probe types, a number of other conditions were varied during the Northern blot experiments. Total RNA from *H. pylori* cells grown either in TSB or on CBA was prepared by both the TRIzol method (section 2.11.1), or by density gradient ultracentrifugation (section 2.11.2). The amounts of total RNA employed ranged from 5 µg to 20 µg per lane of the separating RNA electrophoresis gel. Occasionally, RNA samples were DNase treated (section 2.12.4). Table 3.3 summarizes typical attempts to detect transcripts of the four *H. pylori* genes.

Table 3.3. Summary of transcript detection attempts by Northern analysis.

Targeted gene	Probe label	Probe characteristics	RNA preparation method	*H. pylori* culture conditions	Amount of total RNA	Signal detected?
fliI	DIG	PCR fragment	density gradient	2 d, CBA	5 µg	no
	ECL™	restriction fragment	TRIzol	2 d, CBA	6 µg	no
	ECL™	restriction fragment	density gradient	2 d, CBA	5 µg	no
	α-^{32}P	restriction fragment	density gradient	2 d, CBA	5 µg	no
	DIG	riboprobe	TRIzol, with and without DNase	24 h, TSB	5 µg	no
	DIG	riboprobe	TRIzol, with and without DNase	2 d, CBA	20 µg	no
tsaA	ECL™	restriction fragment	density gradient	2 d, CBA	5 µg	yes [a]
	ECL™	restriction fragment	TRIzol	2 d, CBA	5 µg	yes [a]
hspA	DIG	riboprobe	TRIzol, with and without DNase	24 h, TSB	5 µg	yes [b]
	DIG	riboprobe	TRIzol	2 d, CBA	20 µg	yes [b]
flgE	α-^{32}P	restriction fragment	density gradient	2 d, CBA	5 µg	no
	DIG	riboprobe	TRIzol	24 h, TSB	5 µg	no

[a] signal detected after 3 min exposure to X-ray film

[b] signal detected after 1 min exposure to X-ray film

As illustrated in Table 3.3, neither a *fliI* containing transcript nor *flgE* specific mRNA were detected in any of the trials. In contrast, detection of *hspA* and *tsaA* transcripts was successful under all conditions tested. Fig 3.13 depicts representative Northern blot analyses for the four tested *H. pylori* genes.

The approximated size of the *H. pylori hspA* transcript was 2,200 nt which corresponded to the expected length of the bicistronic mRNA incorporating *hspA* and

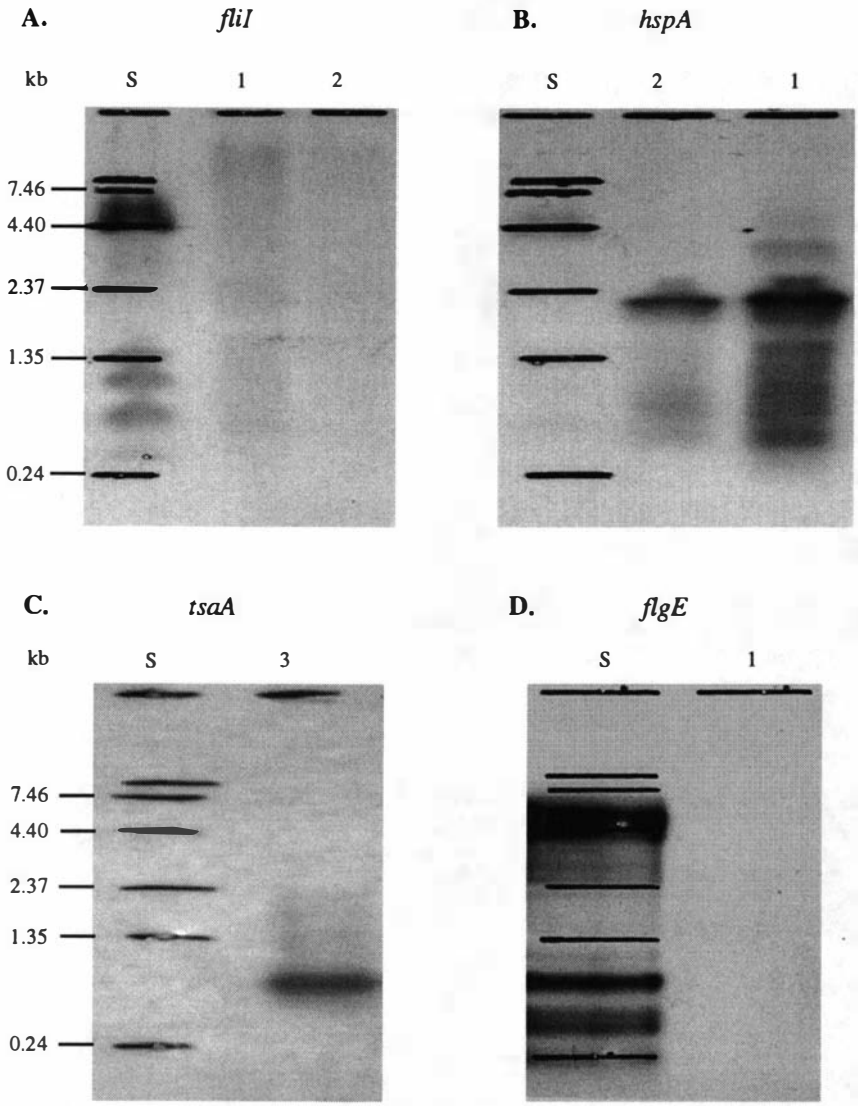


Figure 3.13. Transcript detection by Northern analysis using gene specific probes.

Panel A, *fliI* specific DIG-labelled riboprobe; Panel B, *hspA* specific DIG-labelled riboprobe; Panel C, *tsaA* specific ECL-labelled DNA probe; Panel D, *flgE* specific DIG-labelled riboprobe.

Lanes: 1, 5 µg total RNA (TRIzol) from *H. pylori* 17874 grown for 24 h in TSB; 2, 5 µg total RNA (TRIzol) from *H. pylori* 17874 grown for 24 h in TSB, DNase treated; 3, 5 µg total RNA (TRIzol) from *H. pylori* 17874 grown for 48 h on CBA; Lanes S, RNA molecular size standard (fragment pattern artificially visualized). Exposure times were: A, 10 min; B, 1 min; C, 3 min; D, 70 min.

The signals detected in lanes S of panel A and D are results of unspecific binding of the riboprobe to RNA fragments present in the molecular size standard.

hspB (Suerbaum *et al.*, 1994). The *H. pylori tsaA* gene consists of 594 bp. The size of the detected transcript was approximately 750 nt suggesting a monocistronic mRNA.

3.4.2 RT-PCR transcript analysis

Following the unsuccessful Northern analysis attempts, RT-PCR was used to confirm the existence and define the ends of the suspected large transcript comprising ORF03, *ileS*, *virB11*, *fliI*, *fliQ* and *murB*. The results of this analysis are outlined in Fig. 3.14.

Using the *ileS* internal primer SP034 in conjunction with the *murB* internal primer SP010, a product of the expected size (3,388 bp) was amplified (Fig. 3.14, lane C1). This fragment was not detected in negative controls, where the reverse transcriptase component of the enzyme mix had been inactivated prior to the reaction by incubation at 94°C for 5 min. This indicated the presence of a single continuous transcript including the five genetic elements *ileS*, *virB11*, *fliI*, *fliQ* and *murB*. However, presumptive RNA degradation during DNase treatment lead to inconsistencies during repeated attempts to amplify this large transcript. Several RT-PCR experiments amplified only small amounts of minor products, or no fragment at all.

A series of overlapping RT-PCR's of smaller fragments was therefore performed to reconfirm the existence of the transcript. These were ranging from *ileS* to *virB11* (*ileS* internal primer SP034 / *virB11* internal primer SP016), from *virB11* to *fliQ* (*virB11* internal primer SP015 / primer SP027 near the 3' end of *fliQ*), and from *fliQ* to *murB* (primer SP028 near the 5' end of *fliQ* / *murB* internal primer SP036). In all three experiments, the expected amplification products of 1,404 bp, 2,156 bp and 913 bp, respectively, were detected (reactions D-F in Fig 3.14). The consistently positive outcome of these amplifications accompanied by clean negative controls confirmed that the five genetic elements *ileS*, *virB11*, *fliI*, *fliQ* and *murB* are indeed cotranscribed.

To further confine the ends of the transcript, RT-PCR using the *ileS* internal primer SP040 in conjunction with either primer SP039, positioned near the 5' end of ORF03, or primer SP020, located at the 5' end of ORF02, was performed. The RT-PCR bridging ORF03 and *ileS* (primer pair SP039 / SP040) constantly produced the expected amplification product of 649 bp, indicating cotranscription of ORF03 with the other five members of the operon (Fig 3.14, lane B1). This product was not amplified when the

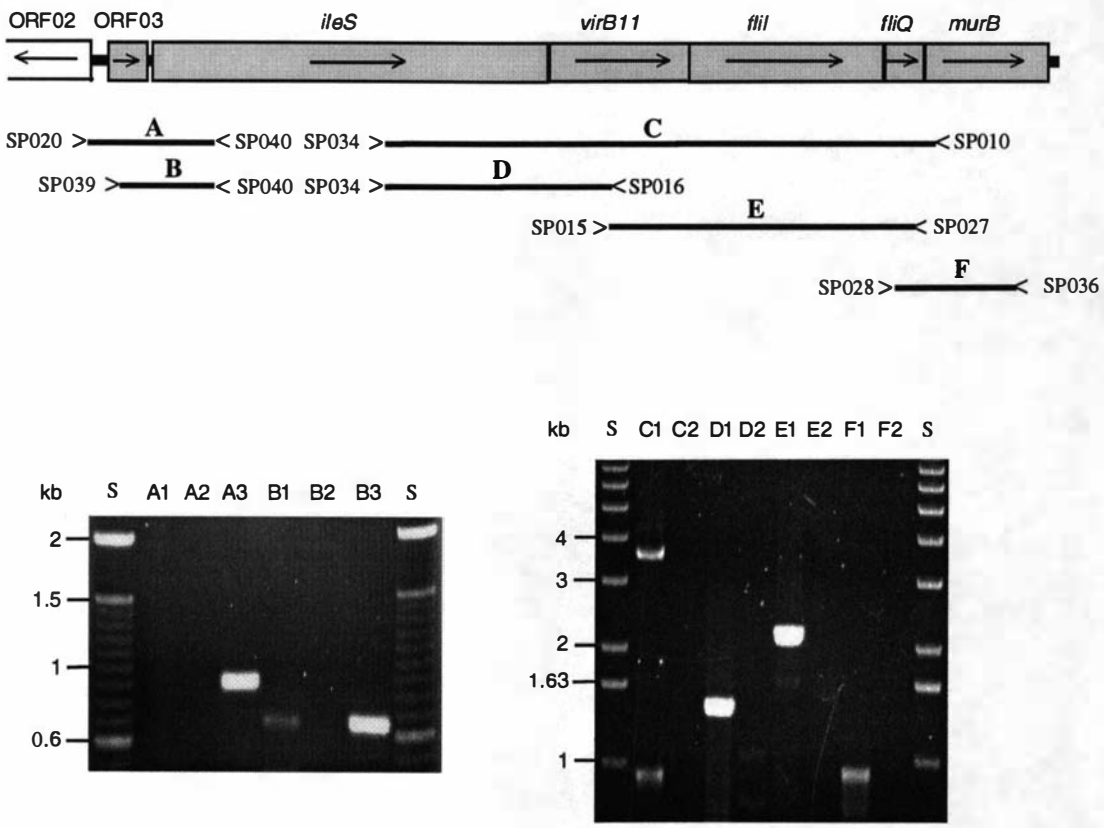


Figure 3.14. ORF03, *ileS, virB11, fliI, fliQ* and *murB* are cotranscribed.

Lanes A - F correspond to reactions using primer pairs amplifying the indicated regions A - F illustrated at the top. Lanes: 1, RT-PCR on *H. pylori* 17874 total RNA; 2, DNA contamination controls on *H. pylori* 17874 total RNA; 3, PCR on genomic *H. pylori* 17874 DNA using *Taq* polymerase. For negative controls, the supplier's enzyme mixes containing reverse transcriptase were either replaced by *Taq* polymerase (lanes A2, B2), or heat-treated prior to reverse transcription (lanes C2, D2, E2, F2). Lanes S, DNA molecular size standards.

The corresponding *H. pylori* 17874 DNA region is illustrated at the top, with grey boxes representing the genes of the operon.

enzyme mix of the Superscript One-Step™ RT-PCR System was replaced by *Taq* polymerase. This ruled out the possibility that DNA contamination was responsible for the successful transcript detection. In contrast, RT-PCR attempts using primer pair SP020 / SP040 did not amplify a cDNA product (Fig 3.14, lane A1). Simultaneously performed *Taq* polymerase controls using *H. pylori* genomic DNA as template consistently resulted in amplification of the expected fragment. These results suggested the start of the transcript comprising ORF03, *ileS, virB11, fliI, fliQ* and *murB* to be within a region of approximately 170 bp upstream from ORF03, between the annealing sites of primer SP039 and primer SP020.

3.4.3 Promoter mapping by primer extension

After the transcript start site of the operon was determined to reside within a 170 bp segment upstream from ORF03 (see above), attempts to define its exact position were initiated. Primer extension analysis was performed using *H. pylori* 17874 total RNA as described in section 2.12.5. The oligonucleotides employed had been purified by polyacrylamide gel electrophoresis to eliminate interference of primer derivatives with cDNA synthesis and detection of the transcribed product. Three primers, SP048, SP050 and SP051, were used in the experiment. The generated cDNA products are depicted in Fig. 3.15. Fig. 3.16 shows the determined transcription start sites and inferred promoter sequences for the operon.

Primers SP048 and SP050 were designed at appropriate distances from the assumed transcription start region between the annealing sites of primers SP039 and SP040. Primer SP048 bound to a region near the ORF03 5' end, primer SP050 was designed internal of ORF03, 42 nt away from the annealing site of primer SP048. Using either of these oligonucleotides, a single cDNA fragment was produced in the primer extension experiment. Primer SP048 generated a product of 50 nt, whereas a cDNA fragment of 113 nt was detected when the transcription process involved primer SP050. Negative controls, where total RNA was omitted from the reaction, did not result in any detectable product. In contrast, when using 20 µg of DNase treated total *H. pylori* 17874 RNA, the described transcription fragments were confirmed, although with lower intensity. Analysis of these unambiguous results facilitated identification of the transcript start site, which was mapped to a nucleotide corresponding to the A-residue of the presumptive ATG initiation codon of ORF03. A promoter element was identified resembling the *E. coli* σ^{70} recognition site (Fig. 3.16). The sequence of this element,

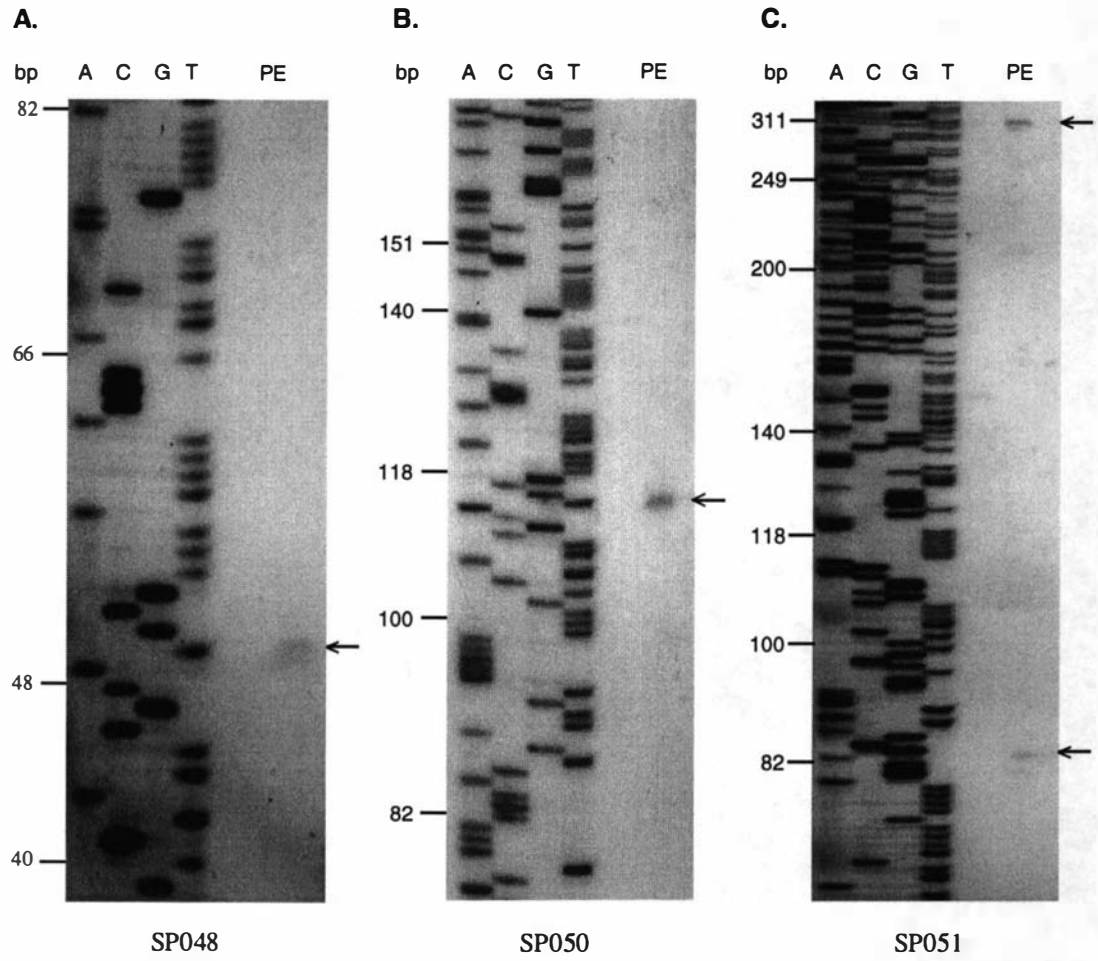


Figure 3.15. Generation of defined cDNA fragments by primer extension.

Panel A. The primer extension reaction (PE) was performed on *H. pylori* 17874 total RNA with primer SP048. The DNA sequencing products using SP048 on pSP102 are indicated. Sizes in bp are from the migration of end-labelled ØX174 restriction fragments. The primer extension product is arrowed.

Panel B. As for panel A, but using primer SP050. The arrowed primer extension product corresponds to the same transcription start site as in panel A.

Panel C. As for panel A, but using primer SP051. The upper product arrowed corresponds to the same transcription start site as in panel A and B, the lower arrowed band is a weaker product indicating a possible second promoter.

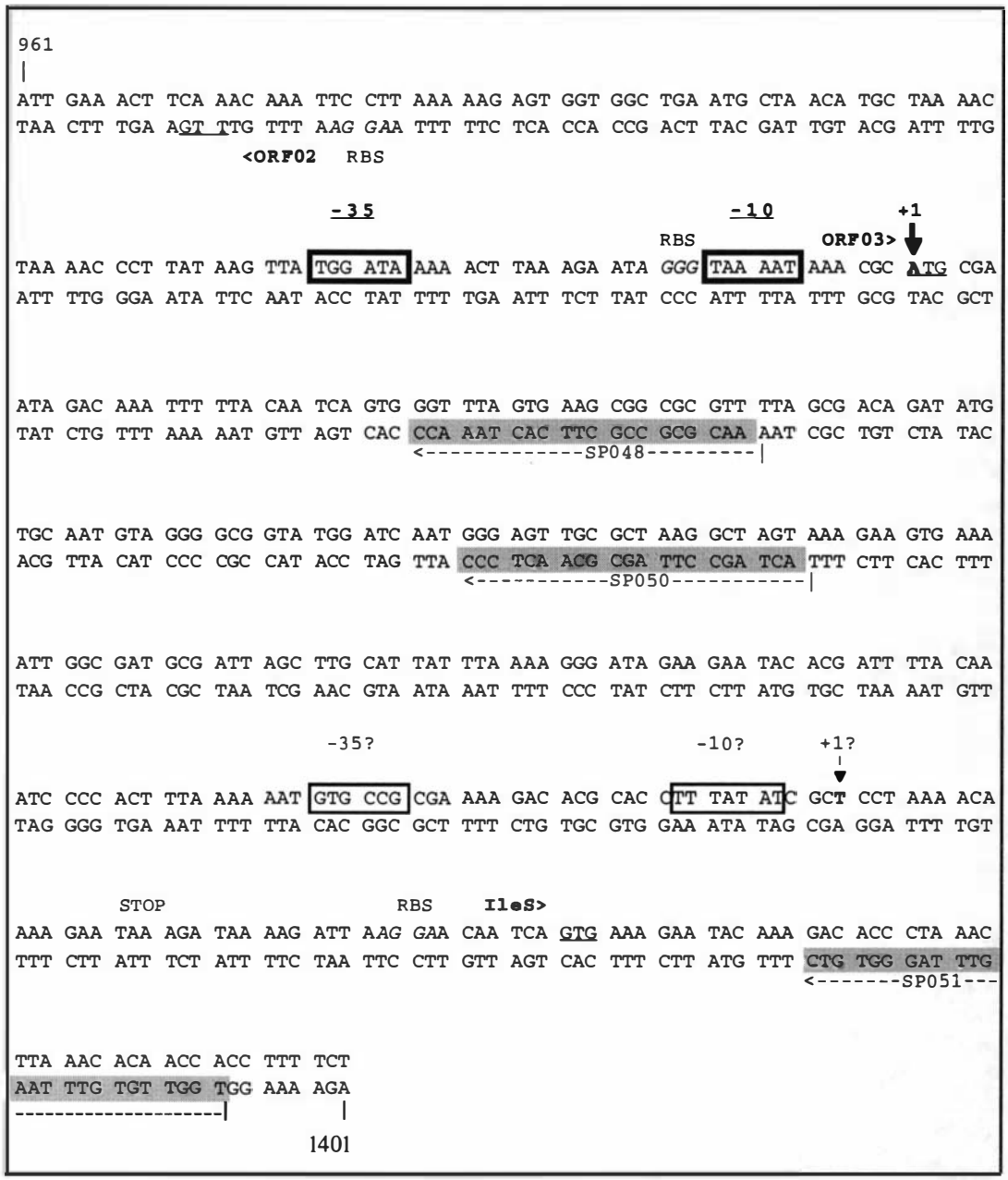


Figure 3.16. Transcription start sites and inferred promoters of the operon.

Transcription start sites identified by the primer extension reactions are arrowed, and the inferred -10 and -35 promoter sequences are boxed. The strong promoter element identified by all three primers used is labelled in bold. The weaker transcription start site, and its inferred suboptimal promoter, are indicated by question marks. Presumptive ribosome binding sites for ORF02, ORF03 and *ileS* are italicized and indicated by RBS, potential initiation codons are underlined. The ORF03 stop codon is labelled. Primer sequences are shaded. Coordinates of the represented DNA segment within the pSP102 insert are indicated at both ends.

TGGATA-N_{18}-TAAAAT, is similar to the established concensus for recognition by σ^{70} in *E. coli* (TTGACA-N_{17}-TATAAT) (Harley & Reynolds, 1987), and is almost optimally spaced from the transcription start point by 6 nt.

A third primer, SP051, was designed to test for additional promoter sites possibly initiating transcription of the operon. It hybridized to a region near the 5' end of the *ileS* gene. Using this primer, the described transcription initiation at the presumptive translation start codon of ORF03 was verified by detection of a 319 nt product (Fig. 3.15, panel C). However, this primer extension experiment also gave rise to a minor product of 83 nt. This site corresponded to position -43 with respect to the *ileS* start codon. Examination of the relevant sequence did not reveal a close match to any of the characterized consensus sequences of σ^{70} (Harley & Reynolds, 1987), σ^{28} (Helmann, 1991) or σ^{54} (Thöny & Hennecke, 1989) promoters. These three σ factors have been implicated in transcription of flagellar genes in other bacteria (section 1.3.1.3). However, a site remotely resembling a σ^{70} promoter element was found. The sequence motif GTGCCG-N_{16}-TTTATAT preceded the identified transcription start site by 3 nucleotides (Fig. 3.16). In the *H. pylori* 26695 genome sequence, the -35 region of this putative promoter gained credibility by substitution of a guanidine residue with an adenine. This substitution resulted in the motif GTGCCA (from GTGCCG), which mismatched the established σ^{70} -35 recognition motif TTGACA in two, rather than three, positions.

Both described cDNA products generated by primer SP051 were detected with very low intensity when DNase treated total RNA preparations were used as template.

3.5 Knockout mutagenesis of the genes of the export locus

3.5.1 Allele replacement strategy

For the construction of *H. pylori* mutants, a two-step PCR-based procedure followed by allelic exchange mutagenesis was employed as described by O'Toole *et al.* (O'Toole *et al.*, 1995), and shown in Fig. 3.17.

In the first PCR step, left and right adjacent regions of the respective genes to be knocked out were amplified using *Pwo* DNA polymerase. These regions were selected to result in deletion of essential intragenic segments. The two inner primers had complementary 5' ends that permitted adherence of the two amplified arms, and carried a restriction site. After gel purification, the arms were used as templates in the second PCR

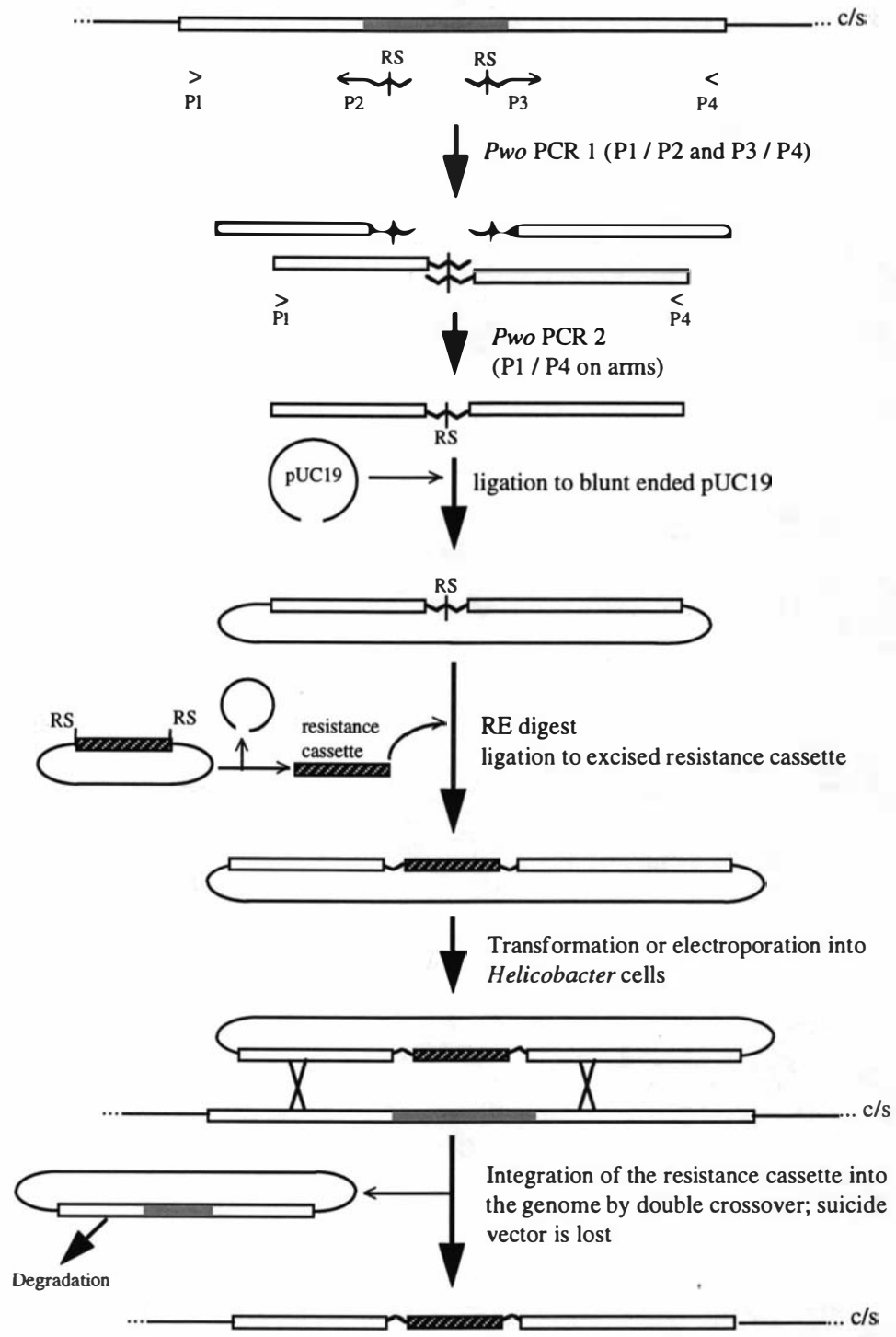


Figure 3.17. Schematic representation of the strategy for deletion-insertional knockout of *H. pylori* genes (allele replacement strategy).

The grey area represents the deleted gene region. White boxes indicate segments present in both the genome and the mutagenic plasmid. RS, restriction site; RE, restriction endonuclease; c/s, chromosomal *H. pylori* DNA.

Figure reproduced from O'Toole *et al.*, 1995, with modification.

step, performed with the two outer primers. The resulting composite fragment lacking essential gene parts was then cloned into linearized, blunt-ended pUC19. Subsequently, a resistance cassette was cloned into the restriction site introduced by the inner primers to form the mutagenic construct. The resulting plasmid was a suicide vector in this context, since it could not replicate in *Helicobacter* cells. It was used to transform *H. pylori* cells as described in section 2.9.3.3. In the bacterial cell, the resistance cassette integrated into the chromosome by a double crossover between the flanking *Helicobacter* sequences in the mutagenic construct and the corresponding homologous regions in the *Helicobacter* genome. Consequently, the functional genomic target gene copy was replaced by the constructed allele which had been disrupted with the resistance cassette. Mutants containing the introduced knockout mutation were selected on CBA incorporating the appropriate antibiotic.

3.5.2 Preparation of the mutagenic constructs

The following descriptions of the construction of mutagenic plasmids specify primers and plasmids used for directed gene knockouts by the allele replacement strategy (see above). For characteristics of these plasmids, and sequence of the oligonucleotides, refer to Table 2.5 and Table 2.4, respectively.

3.5.2.1 The *virB11* mutagenic constructs pSP117 and pSP118

For mutagenesis of the *virB11* gene, primer pairs SP034 / SP032 and SP033 / SP009 were used on pSP102 DNA in the first PCR step. Their positions are indicated in Fig. 3.18, panel B. The pairs amplified fragments of 1,285 bp and 892 bp, respectively. The PCR products were gel purified and subsequently used in approximately even molar amounts as templates in the second PCR step (20 and 25 ng, respectively). The composite fragment was amplified with primer pair SP034 / SP009. It lacked 300 bp of internal *virB11* sequence which included the predicted nucleotide binding site (Fig. 3.7), but contained a primer-derived *Bgl*II restriction site. This fragment was ligated to *Sma*I digested gel purified pUC19 to generate plasmid pSP116. The chloramphenicol resistance cassette (*cat*) of plasmid pRY109 was excised by a *Bam*HI digest. This cassette lacked a transcription terminator. It was gel purified and ligated to *Bgl*II treated gel purified pSP116. Orientation of the *cat* cassette within the new construct was determined by restriction analysis exploiting an asymmetrically positioned *Hin*dIII site. This site divided the cassette into subfragments of approximately 550 and 250 bp. A *virB11* mutagenic

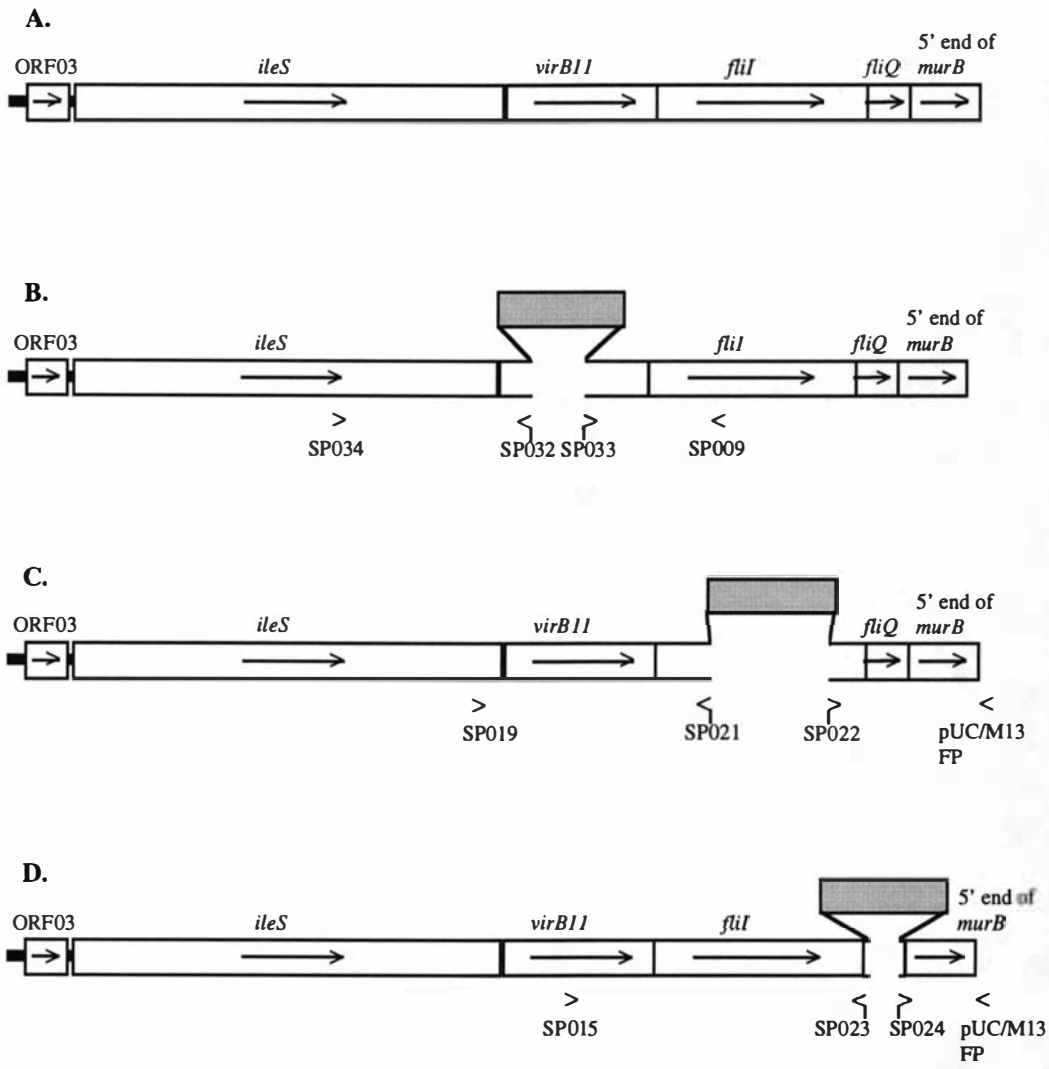


Figure 3.18. Schematic representation of the relevant genomic DNA regions in the *H. pylori* wild type and knockout mutants.

Panel A, *H. pylori* 17874 wild type; Panel B, *virB11* mutant; Panel C, *fliI* mutant; Panel D, *fliQ* mutant.

The grey boxes represent resistance cassettes. Primers used for the mutagenesis procedure are indicated at their approximate annealing sites. Tags (ı) indicate complementary 5' ends of the primers. Primer pUC/M13 FP annealed to the MCS of pUC18, the backbone of pSP102.

plasmid pSP118 was identified where the direction of the *cat* gene would correspond to the orientation of *virB11* after successful integration into the *H. pylori* chromosome.

In addition, a kanamycin resistance cassette (*aphA-3*) originating from *Campylobacter* was released from plasmid pILL600 by *Bam*HI digestion. This cassette also lacked transcription terminating structures. It was cloned into pSP116 as described above to form the mutagenic plasmid pSP117. Asymmetrically positioned *Eco*RV sites within the *aphA-3* gene were used to determine the orientation of the resistance cassette. These sites were approximately 270 and 130 bp away from the 5' and 3' ends of the cassette, respectively. They allowed determination of the cassette orientation in the new construct after digestion using *Eco*RV in conjunction with *Bst*XI. Transcriptional direction of the *aphA-3* gene present in pSP117 corresponded to the orientation of the disrupted *virB11* gene.

3.5.2.2 The *fliI* knockout plasmids pSP108 and pSP110

For mutagenesis of *fliI*, primer pairs SP019 / SP021 and SP022 / pUC/M13 FP (positions indicated in Fig 3.18, panel C) amplified arms of 1,448 bp and 962 bp, respectively, from pSP102 template DNA. Gel purified amplified arms were used as templates in the second PCR step. The composite fragment was amplified with primer pair SP019 / pUC/M13 FP. It contained the primer-derived *Bam*HI site and lacked 697 bp internal *fliI* sequence which included the ATP binding sequence motifs of the Walker boxes A and B (Fig. 3.8). The fragment was cloned into *Sma*I / *Hinc*II digested, gel purified pUC19. Restriction of pUC19 with these two enzymes was necessary to eliminate the *Bam*HI site of the vector's MCS. Consequently, the blue-white selection property of pUC19 was destroyed. The blunt end ligation of pUC19 with the composite fragment lacking *fliI* internal sequence was therefore performed in 15% polyethylene glycol (PEG) solution, following recommendations of Sambrook *et al.* (Sambrook *et al.*, 1989). The presence of PEG enhanced the probability of successful intermolecular ligation. The thus obtained construct pSP106 possessed the primer-derived *Bam*HI site in its insert. This plasmid was *Bam*HI digested, and gel purified. As before (see above), the chloramphenicol resistance cassette of pRY109 was excised by *Bam*HI digestion, and ligated to pSP106 to generate the *fliI* mutagenic plasmid pSP110. The orientation of the resistance cassette in several isolated clones was determined by *Hin*dIII restriction analysis as described above. Transcriptional direction of the *cat* cassette in *fliI* gene orientation resulted in pSP110a. When direction of the resistance cassette was opposite to *fliI* gene transcription, the corresponding mutagenic plasmid was named pSP110b.

A kanamycin resistance cassette (KanΩ) was similarly excised from plasmid pUC4ΩKm-2 by *Bam*HI digestion. This resistance cassette contains strong transcription and translation terminating hairpin structures called Ω elements upstream and downstream from the *aphA-3* gene (Perez-Casal *et al.*, 1991). Gel purified KanΩ was ligated to appropriately treated pSP106 to generate the *fliI* knockout construct pSP108.

3.5.2.3 The *fliQ* disrupting plasmids pSP107 and pSP109

FliQ mutagenesis was performed as follows: Primer pairs SP015 / SP023 and SP024 / pUC/M13 FP (positions indicated in Fig. 3.18, panel D) amplified left and right arms (1,951 bp and 656 bp, respectively). The fragments were gel purified and used as template during the second PCR step, where the composite fragment was obtained with primer pair SP015 / pUC/M13 FP. This composite lacked 36 bp internal *fliQ* sequence and contained a primer-derived *Bgl*II site instead. It was purified and subsequently cloned into *Sma*I digested, purified pUC19 to create plasmid pSP105. The pRY109-derived chloramphenicol resistance cassette was released and purified as described above, and ligated to *Bgl*II digested pSP105. The resulting *fliQ* mutagenic construct was named pSP107. The orientation of the resistance cassette corresponded to the *fliQ* gene direction, judged from *Hin*dIII restriction analysis.

The pUCΩKm-2-derived KanΩ resistance cassette was excised from its host plasmid by *Bam*HI digestion, and gel purified. Successful ligation to *Bgl*II digested pSP105 resulted in the *fliQ* mutagenic construct pSP109.

3.5.3 Generation of *H. pylori virB11*, *fliI* and *fliQ* mutants

H. pylori mutants were generated by electroporation of Wizard preparations (section 2.6.1.2) of the mutagenic plasmids pSP117 (*virB11*::*aphA-3*), pSP118 (*virB11*::*cat*), pSP110a (*fliI*::*cat*), pSP110b (*fliI*::*cat*) and pSP107 (*fliQ*::*cat*) into electrocompetent *H. pylori* 17874 cells. Alternatively, these plasmids were also taken up by bacterial cells using their natural competence (section 2.9.3.3). Natural transformation was exclusively used to generate mutants from *H. pylori* SS1 cells. Electroporation and transformation efficiencies were calculated by the number of transformed colonies divided by the number of surviving *H. pylori* 17874 cells after harvesting from a non-selective CBA plate on day 1 of the procedure. Successful mutant generation experiments were only performed once, impairing the statistical value of the calculated frequencies. However, efficiencies of pSP107 electroporation were calculated for two independent

trials, and did not exhibit great variation. After initial success, plasmid pSP107 was used as positive control in all subsequent trials. The results of the mutagenesis experiments performed are shown in Table 3.4.

Table 3.4. Mutagenesis of *H. pylori* by electroporation or natural transformation.

Disrupted gene	Mutagenic plasmid	Resistance marker gene (origin)	Method	Knockout frequency / efficiency
H. pylori 17874				
virB11	pSP117	*aphA-3* (pILL600)	electroporation	4.1×10^{-5} [c]
	pSP118	*cat* (pRY109)	electroporation	1.5×10^{-3}
flil	pSP110a	*cat* (pRY109)	electroporation	2.2×10^{-3}
	pSP110b	*cat* (pRY109)	electroporation	1.3×10^{-8} [d]
	pSP108	*aphA-3* (pUCΩKm-2)	transformation [a]	0
fliQ	pSP107	*cat* (pRY109)	electroporation	9.6×10^{-4} / 1.3×10^{-3}
	pSP109	*aphA-3* (pUCΩKm-2)	electroporation	0
	pSP107	*cat* (pRY109)	transformation [a]	6.2×10^{-6}
H. pylori SS1				
virB11	pSP117	*aphA-3* (pILL600)	transformation [b]	1.5×10^{-7}
	pSP118	*cat* (pRY109)	transformation [b]	5.6×10^{-7}
flil	pSP110a	*cat* (pRY109)	transformation [b]	3.3×10^{-8}
	pSP110b	*cat* (pRY109)	transformation [b]	0
fliQ	pSP107	*cat* (pRY109)	transformation [b]	1.6×10^{-8}

[a] *H. pylori* cells used had been grown for 2 days on CBA
[b] *H. pylori* cells used had been grown for 3 days on CBA
[c] first transformants only visible after 5 days of incubation
[d] first transformants only visible after 4 days of incubation

Where mutants had been obtained, three colonies were purified by three single colony passages on selective CBA plates. After the third purification, genomic DNA of the respective *H. pylori* culture was prepared (section 2.6.2). The desired replacement of the functional chromosomal gene with the disrupted copy of the mutagenic construct was verified by PCR using primers binding to flanking regions of the insertion site of the resistance cassette. In the event of successful replacement, the size of the generated PCR product was changed. It was expected to be bigger by the net sum of the size of the resistance cassette minus that of the deleted gene region compared to amplification products obtained with *H. pylori* 17874 wild type genomic DNA as template. The calculated changes in PCR product sizes amplified from *H. pylori* mutant genomic DNA are summarized in Table 3.5.

Table 3.5. Calculated PCR product sizes from *H. pylori* wild type and mutant genomic DNA.

Disrupted gene	Resistance cassette (origin)	Primer pair	Expected approximated PCR product size	
			wild type *H. pylori*	mutant *H. pylori*
virB11	0.8 kb *cat* (pRY109)	SP019 / SP014	1,000 bp	1,500 bp
virB11	1.4 kb *aphA-3* (pILL600)	SP019 / SP014	1,000 bp	2,100 bp
flil	0.8 kb *cat* (pRY109)	SP013 / SP023	1,050 bp	1,160 bp
fliQ	0.8 kb *cat* (pRY109)	SP027 / SP028	250 bp	1,000 bp

The PCR tests described in Table 3.5 resulted in the expected bigger product size for all purified resistent *H. pylori* colonies proving the desired disruption of the respective gene by the resistance cassette. Moreover, purity of the mutant cultures was demonstrated by the single amplification product in these tests. Fig. 3.19 shows the result of the described PCR amplification on genomic DNA of those *H. pylori* mutant colonies that were subsequently used in further studies.

3.6 Non-polarity of the mutations

Given that the genes of the described protein export locus comprized an operon, it was important to show that individual *cat* insertion mutants retained transcription of downstream genes. Theoretically, this was possible by virtue of readthrough transcription of the inserted *cat* gene originating from pRY109. Only when transcription of genes downstream from the insertion sites was proven, could phenotypic differences in the mutants be assigned to the lack of the respective mutated gene. RT-PCR was therefore performed on total RNA preparations of the *virB11::cat* (pSP118), *fliI::cat* (pSP110a) and *fliQ::cat* (pSP107) mutant strains obtained from *H. pylori* 17874. Primer pairs were used to test for transcription of the genes immediately upstream and downstream from the *cat* insertion site. In addition, primer pairs completely or partially covering the insertion sites were used to show total transcriptional knockout of the respective gene. Results of these experiments are summarized in Fig. 3.20.

In the *virB11*::*cat* mutant, transcription of *ileS* up to the insertion site was tested using the *ileS* internal primer SP019 in conjunction with the *virB11* internal primer SP032 (located just upstream from the inserted *cat* cassette). The RT-PCR generated the expected product of 351 bp. The *virB11* internal primer pair SP014 / SP015 would generate a product of 527 bp if residual *virB11* transcription would still occur. This reaction did not result in any amplification product. Transcription downstream from the insertion site was

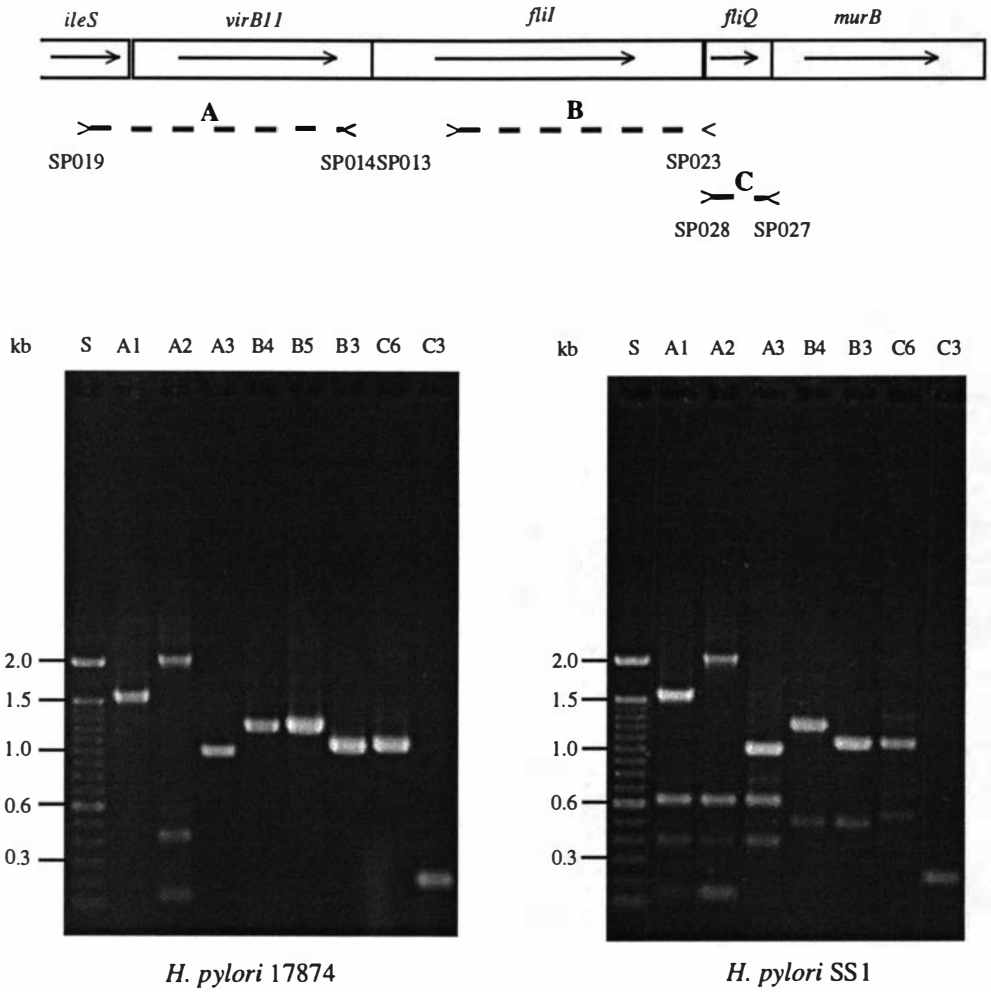


Figure 3.19. PCR verification of *H. pylori* knockout mutants.

PCR was performed using genomic DNA of *H. pylori* 17874 (left) and *H. pylori* SS1 (right) mutant and wild type colonies as template. Lanes: 1, *virB11::cat* (pSP118) mutant; 2, *virB11::aphA-3* (pSP117) mutant; 3, wild type; 4, *flil::cat* (pSP110a) mutant; 5, *flil::cat* (pSP110b) mutant; 6, *fliQ::cat* (pSP107) mutant; A, *ileS* internal primer SP019 / *virB11* internal primer SP014; B, *flil* internal primer SP013 / *fliQ* internal primer SP023; C, *fliQ* internal primer pair SP028 / SP027. Lanes S, DNA molecular size marker.

The relevant genomic region of the *H. pylori* 17874 wild type is illustrated at the top. Primers are shown at their respective annealing sites. For expected product sizes refer to Table 3.5.

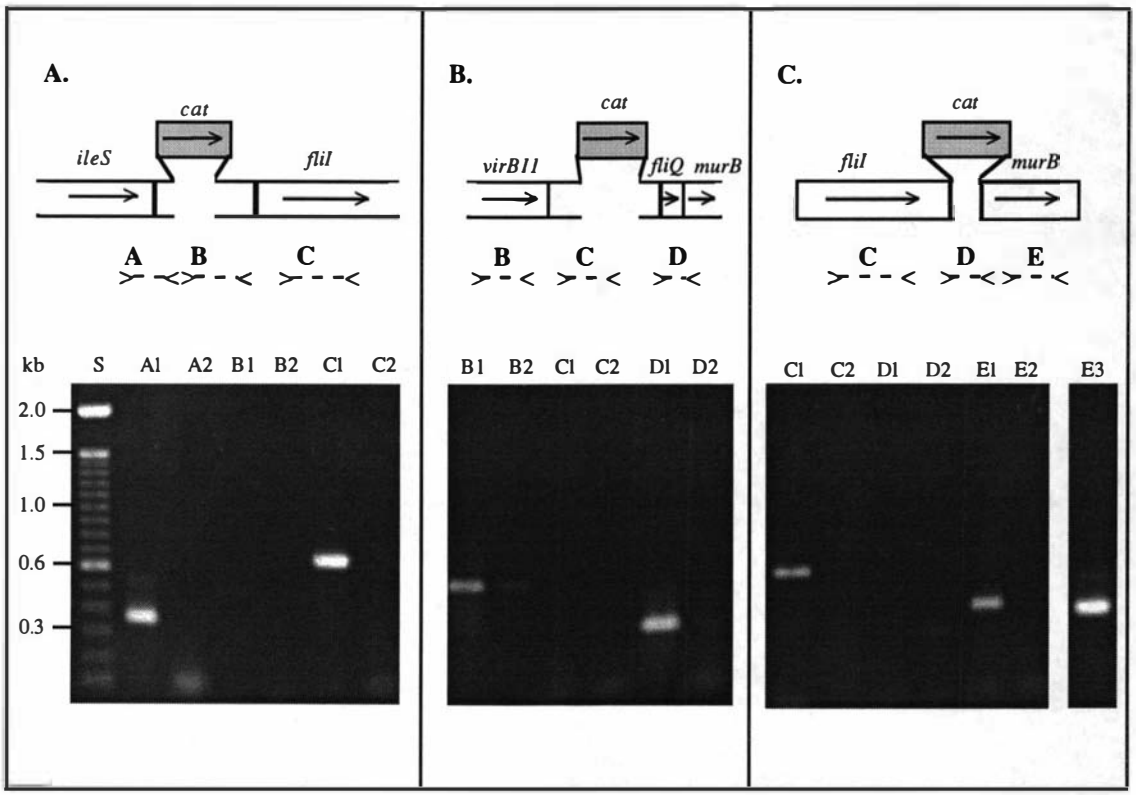


Figure 3.20. The deletion-insertion mutations are non-polar.

Panel A, *H. pylori virB11::cat* (pSP118) mutant; Panel B, *H. pylori flil::cat* (pSP110a) mutant; Panel C, *H. pylori fliQ::cat* (pSP107) mutant.

Lanes: 1, RT-PCR on *H. pylori* mutant total RNA; 2, DNA contamination controls (*Taq* polymerase only) on *H. pylori* mutant total RNA; 3, *Taq* polymerase control PCR on genomic *H. pylori* 17874 DNA; A, *ileS* internal primer SP019 / *virB11* internal primer SP032; B, *virB11* internal primer pair SP014 / SP015; C, *flil* internal primer pair SP013 / SP002; D, *fliQ* internal primer SP028 / *murB* internal primer SP010; E, *murB* internal primer pair SP035 / SP036. Lane S, DNA molecular size marker.

The relevant genomic regions from the *H. pylori* mutants are illustrated at the top. Primers for reactions A - E are indicated at their respective annealing sites.

tested using the *fliI* internal primer pair SP013 / SP002, resulting in the expected amplification fragment of 642 bp.

Primer pair SP014 / SP015 was used on total RNA from the *fliI*::*cat* mutant to test for *virB11* transcription. The RT-PCR generated the expected 527 bp product. Primer pair SP013 / SP002 failed to amplify a product indicating a lack of residual *fliI* transcription. Primer SP028 (located at the 5' end of *fliQ*) was used in conjunction with primer SP010 (hybridizing near the *murB* 5' end) to test for *fliQ* transcription, and generated the expected 316 bp product.

The *fliQ*::*cat* mutant displayed the same pattern. When testing for transcription upstream from the *cat* insertion site using the *fliI* internal primer pair SP013 / SP002, the expected 642 bp product was generated. Negative results of amplification trials using primer pair SP028 / SP010 demonstrated lack of run-on transcription of the knocked out *fliQ* gene across the *cat* insertion. Finally, the *murB* internal primer pair SP035 / SP036 indicated *murB* transcription by amplifying the anticipated 394 bp product.

In all cases, controls where the enzyme mix of the kit was replaced with *Taq* polymerase consistently failed to generate a product. Simultaneously, a *Taq* polymerase positive control using primer pair SP035 / SP036 proved activity of the enzyme with *H. pylori* 17874 chromosomal DNA as template (Fig. 3.20, lane E3). In summary, transcripts were consistently detected upstream and downstream from the insertion sites of the *cat* cassette, indicating that these genes were still transcribed.

3.7 Phenotypic characterization of the *H. pylori* mutants

The following mutants generated from *H. pylori* 17874 by electroporation were phenotypically characterized: *H. pylori virB11*::*cat* (pSP118), *H. pylori fliI*::*cat* (pSP110a), and *H. pylori fliQ*::*cat* (pSP107). Bacterial growth, motility, flagellar biosynthesis and production of selected flagellar and non-flagellar components in these mutants were investigated and compared to the respective characteristics of *H. pylori* 17874 wild type cells.

3.7.1 Growth characteristics

No difference was noticed in viability or growth characteristics of the *H. pylori* mutant strains compared to the wild type. Colony morphology and generation times were found to be unchanged. Cell densities and growth curves of *H. pylori* wild type and mutant cultures grown in TSB were similar. Microscopic observations using a phase contrast microscope (section 2.15.1) confirmed that the cell shapes and the timepoint of conversion from spiral to coccoid cell morphology were also unaffected by any of the mutations. Coccoid forms usually appeared on the third day of bacterial culture on CBA, or after approximately 48 h incubation in TSB.

3.7.2 Motility

Initially, evaluation of motility of the mutant strains was attempted using motility agar plates (Table 2.2). Agar concentrations were reduced to 0.3%, and both slopes and plates were tested. Inoculum sizes were varied. Tooth pick inoculation from *H. pylori* cells grown for 48 or 72 h on CBA plates were performed as well as inoculations with 3-10 µl of *H. pylori* cultures grown in TSB for less than 48 h. Development of halos around motile *Helicobacter* cultures on motility agar had been described in the literature (Suerbaum *et al.*, 1993; Josenhans *et al.*, 1995). However, this was not reproducibly observed for our mutant or wild type cultures. None of the aforementioned conditions resulted in reliable halo formation around colonies of any of the strains tested.

Following these unsuccessful experiments, motility of the mutants was determined by phase contrast microscopy using an inverted microscope, as described in section 2.15.1. *H. pylori* wild type cultures, that had been killed by 2 h incubation in TSB containing 1 mM potassium ferricyanide, appeared to twitch, but lacked the characteristic rotational movements frequently observed in living, motile *H. pylori* 17874 wild type cultures. This rotational movement of a fraction of the cells was therefore used as an indicator for motility.

Three independent tests were performed, using *H. pylori* cultures that had been grown in TSB for 24 h or 48 h. Assessment by three observers, two of whom had not received information about the respective mutant or wild type cultures they were evaluating, were consistent and unambiguous. Whereas the *H. pylori* 17874 wild type and the *virB11* mutant strains were clearly motile, both the *fliI* and the *fliQ* mutant strains were completely lacking any rotating bacterial cells, i.e. assessed to be non-motile.

Judged from the higher frequency of rotating cells and their higher spin velocity, the *H. pylori virB11* mutant was consistently evaluated to be more motile than the wild type strain.

3.7.3 Flagellum production

H. pylori cells from 17874 wild type and the three mutant cultures were grown for 24 h in TSB. Subsequently, they were prepared and examined by negative staining and subsequent transmission electron microscopy as described in section 2.15.2. Two different cultures of each investigated strain were examined. Fig. 3.21 displays characteristic images of the four strains.

Cell shapes of the *cat* insertion mutant strains and the wild type appeared to be identical. Sheathed unipolar flagella, usually single, occasionally a pair, were clearly visible in both the wild type and the *virB11* mutant strains. However, in some cases these flagella seemed to protrude from the side of the bacterial cell rather than the pole. The flagella observed in the *virB11* mutant cultures exhibited no detectable difference in shape, length or localization compared to those of the wild type cells.

No flagella were detected in the *fliI* and the *fliQ* mutants. In contrast, residual flagellation in less than 1% of the cells had been described in an *H. pylori* N6 *fliI* mutant (Jenks *et al.*, 1997). This observation was not confirmed in the mutant cultures generated in this study. Very rarely, though, unattached membranous structures were observed in the *fliI* mutants (not documented), which might have represented empty flagellar sheaths.

3.7.4 Expression of structural flagellar components

3.7.4.1 Flagellin and FlgE expression patterns in subcellular fractions of *H. pylori*

In order to investigate the effect of the described mutageneses on production of structural flagellar proteins, Western immunoblot analysis was performed (section 2.14.3). Subcellular protein fractions of wild type and mutant *H. pylori* cultures grown in TSB for 36 h were prepared (section 2.13.2). These cultures had been freshly inoculated from freezer stocks into 50 ml TSB. Detection was performed with a two thousand-fold dilution of the polyclonal antibody anti-Fla, raised against purified *H. pylori* flagellin (Kostrzynska *et al.*, 1991). This antibody recognized several structural components of the

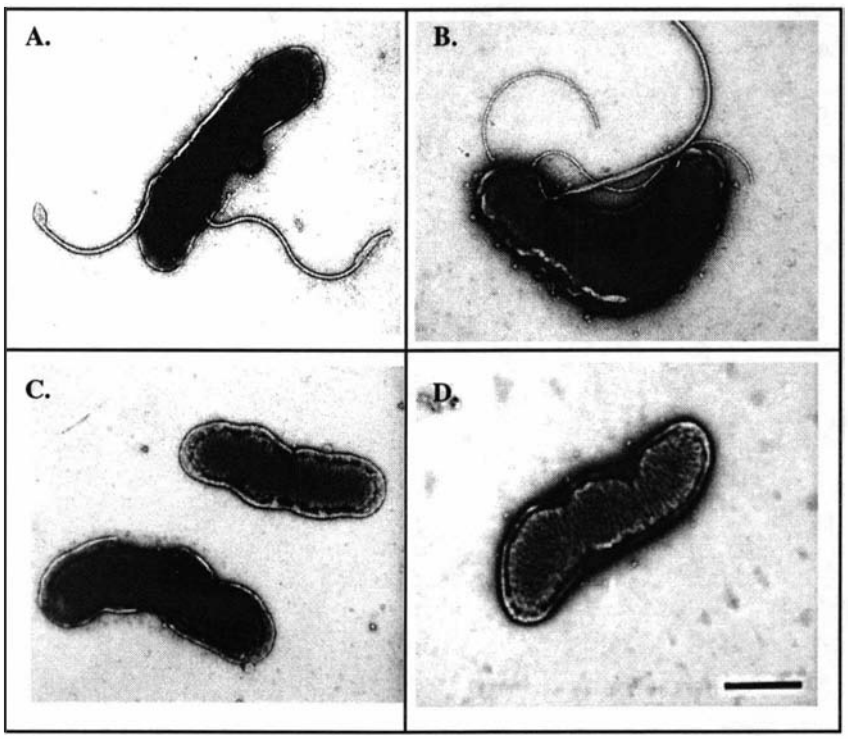


Figure 3.21. Electron micrographs of negatively stained preparations of *H. pylori* cells.

(A) Flagellated *H. pylori* 17874 wild type. (B) Flagellated *virB11::cat* (pSP118) mutant. (C) Aflagellate *fliI::cat* (pSP110a) mutant. (D) Aflagellate *fliQ::cat* (pSP107) mutant. Bar, 1 μm.

Images were recaptured from continuous-tone Agfa photographs using the IS-1000 Digital Imaging System (Alpha Innotech).

H. pylori flagellum, including the flagellins FlaA and FlaB, and the hook protein FlgE. Protein sample loadings were standardized as follows. The loading volume ratios of corresponding fractions of the different *H. pylori* cultures were calculated based on wet weight of the pelleted bacterial cells after aspiration of residual medium. Since the final total volume of the cytosolic fraction (2 ml) was four times higher than that of the envelope and the insoluble complexes (500 µl), four-fold loading volumes of this fraction were applied. The supernatants (50 ml) were five-fold concentrated as described in section 2.13.2. They were then applied in four-fold excess volumes compared to envelope and insoluble complexes. Note that, respective to the wet weight of bacterial culture, supernatant fractions were five-fold underrepresented on Western blots.

The experiments demonstrated that expression of both flagellins and FlgE was variously affected by the mutageneses. Fig. 3.22 illustrates the observed effects. In the *virB11* mutant, slightly elevated levels of the flagellins and the hook protein FlgE were detected in the supernatants, the envelopes and the cytosolic fractions compared to the wild type. In contrast, expression of these components was severely reduced in the *fliI* and the *fliQ* mutants. A dramatic reduction of flagellin and hook protein levels was demonstrated in the supernatants, the envelope protein fractions and the insoluble protein complexes of the *fliI* and the *fliQ* mutants. Reduction was not so severe in the cytosolic fractions of these mutants.

3.7.4.2 Stability of the effects of the mutageneses on flagellar protein production

To evaluate the stability of the described influence on flagellar protein production, expression of flagellins and FlgE was compared in sets of *H. pylori* wild type and mutant cultures that differed in the number of their previous passages. Whereas one set of *H. pylori* strains had been passaged three times before sample preparation, the second set had been subcultured six times. Two and three of these passages, respectively, predated the inoculation of TSB medium with the strains, and had been performed on CBA. Western immunoblot analysis was performed as described above. However, bacterial cells were not fractionated. Instead, whole cell lysates were prepared as described in section 2.13.1. Protein concentrations of the lysates were determined using the Biorad microassay (section 2.14.1). Cell lysate aliquots containing approximately 2 µg of total protein were loaded on polyacrylamide gels. Supernatants were standardized based on the calculated loading volume ratios of the cell lysate samples. Cell lysates had been resuspended in 1/10 of the original culture volume. Five-fold volumes of

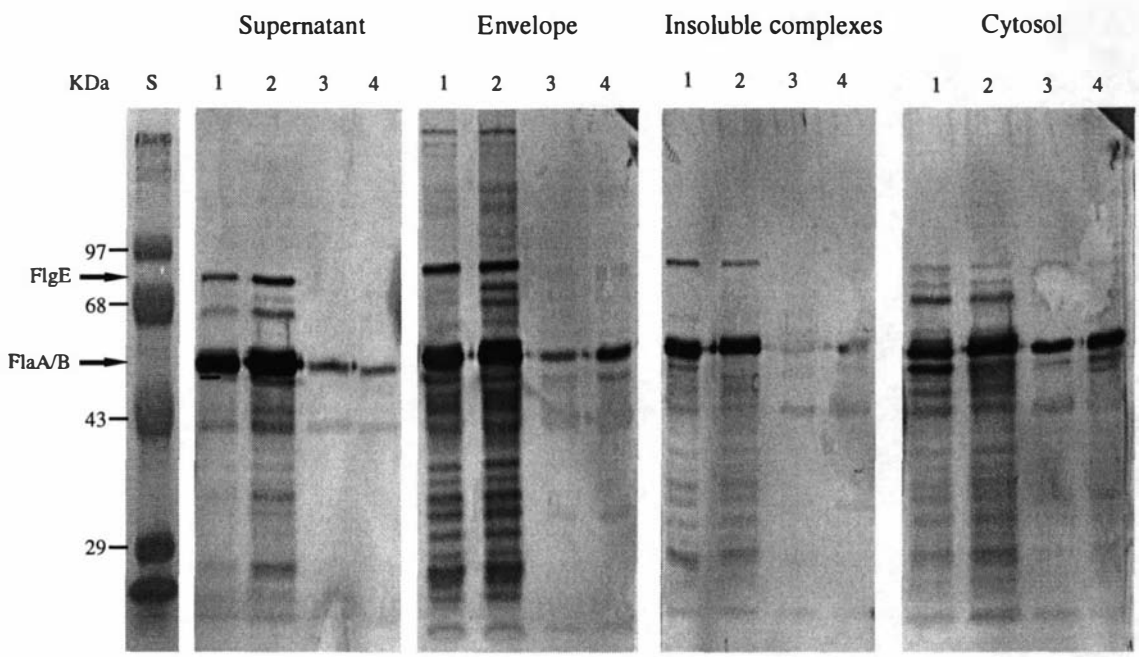


Figure 3.22. Altered flagellin and FlgE production levels in *H. pylori virB11, fliI* and *fliQ* mutants.

Western immunoblot analysis of subcellular protein fractions of *H. pylori* wild type and mutant cultures using anti-Fla. Cells were grown for 36 h in TSB. Protein samples were separated on a denaturing 12.5% polyacrylamide gel and subsequently blotted.

Lanes: 1, *H. pylori* 17874 wild type; 2, *virB11::cat* (pSP118) mutant; 3, *fliI::cat* (pSP110a) mutant; 4, *fliQ::cat* (pSP107) mutant. Lane S, protein molecular size marker.

supernatant fractions were applied on the gels compared to the cell lysates. Thus, supernatants were two-fold underrepresented on Western blots respective to the protein content per bacterial cell.

Fig. 3.23 summarizes the result of these investigations. The observed changes in flagellar protein production (section 3.7.4.1) were apparently stable during repeated subculture of the *virB11* and the *fliQ* mutants. However, the overall level of reduction was not constant in the *fliI* mutant. Repeated passage of the *fliI* mutant resulted in flagellin and FlgE levels that were still reduced, but higher than in a *fliI* mutant culture which had been subcultured less often.

3.7.5 Production of *H. pylori* virulence factors

The function of VirB11, or substrates whose export it might be involved in, was unknown. In a preliminary effort to identify proteins whose expression might have changed by mutation of *virB11*, production of three established *H. pylori* virulence factors was examined by Western immunoblotting (section 2.14.3). These were the larger urease subunit UreB, the cytotoxin associated protein CagA, and the vacuolating cytotoxin VacA. Whole cell lysates and supernatants of wild type and mutant *H. pylori* cultures grown for 24 h in TSB were used and standardized as described in section 3.7.4.2. The primary antibodies were monoclonal anti-UreB, polyclonal anti-CagA and polyclonal anti-VacA (Table 2.6). Anti-UreB and anti-CagA were diluted one thousand-fold, anti-VacA was used two thousand five hundred-fold diluted.

Western analysis demonstrated no effect of any of the mutations on expression or stability of these three established virulence factors (Fig 3.24). The investigated proteins were present in similar amounts in whole cell lysates of the three mutant cultures compared to the wild type. Concentration of the targeted components in the supernatants was lower than the sensitivity threshold of the immunoblot procedure, and no signals were obtained in those fractions.

To examine a possible effect of any of the mutations on localization patterns, distribution of the presumptive autotransporter VacA and the strongly immunoreactive CagA protein was also investigated in subcellular fractions. Samples were prepared as described in section 2.13.2. Presence of these two proteins in all four subcellular fractions was demonstrated by appearance of immunoreactive bands. However, no

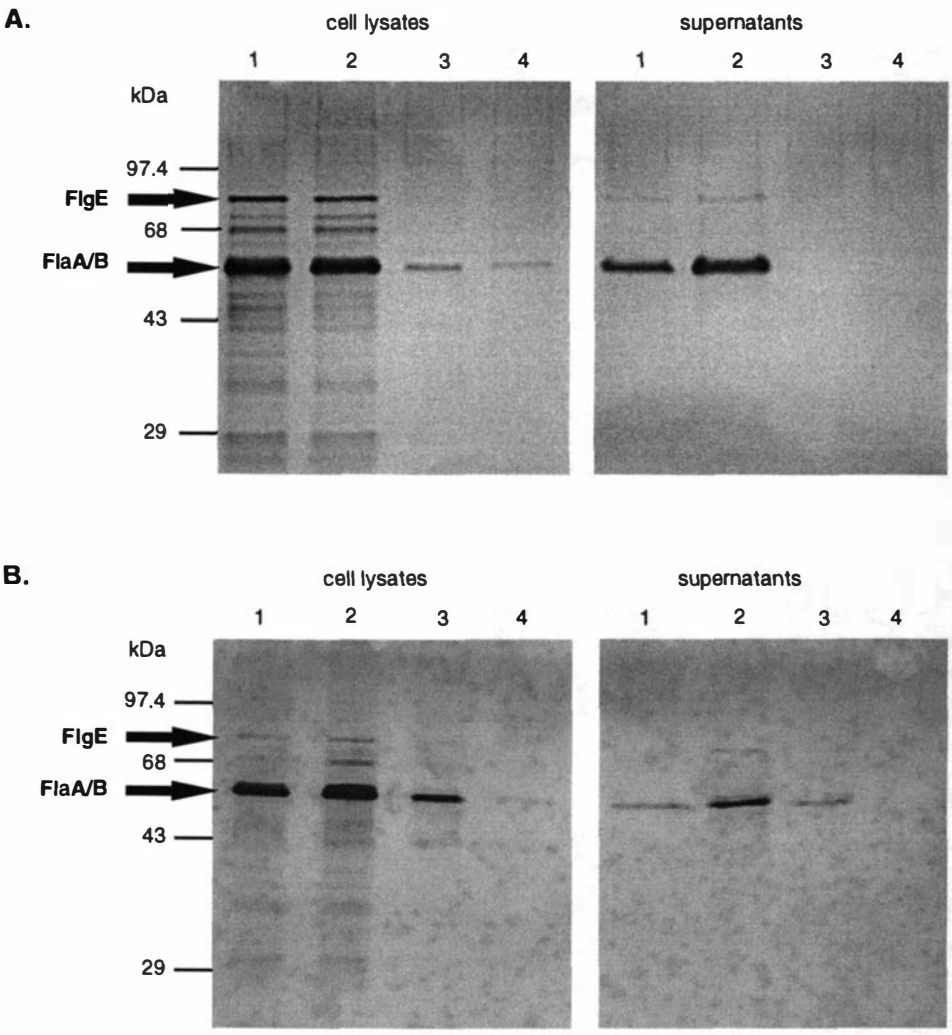


Figure 3.23. Influence of passage on flagellin and FlgE production levels in the *H. pylori* mutants.

Western immunoblot analysis of *H. pylori* cell lysates and supernatants using anti-Fla. Protein samples were separated on a denaturing 12.5% polyacrylamide gel and subsequently blotted.

Panel A, cells grown for 24 h in TSB after 3 previous passages. Panel B, cells grown for 24 h in TSB after 6 previous passages.

Lanes: 1, *H. pylori* 17874 wild type; 2, *virB11::cat* (pSP118) mutant; 3, *fliI::cat* (pSP110a) mutant; 4, *fliQ::cat* (pSP107) mutant.

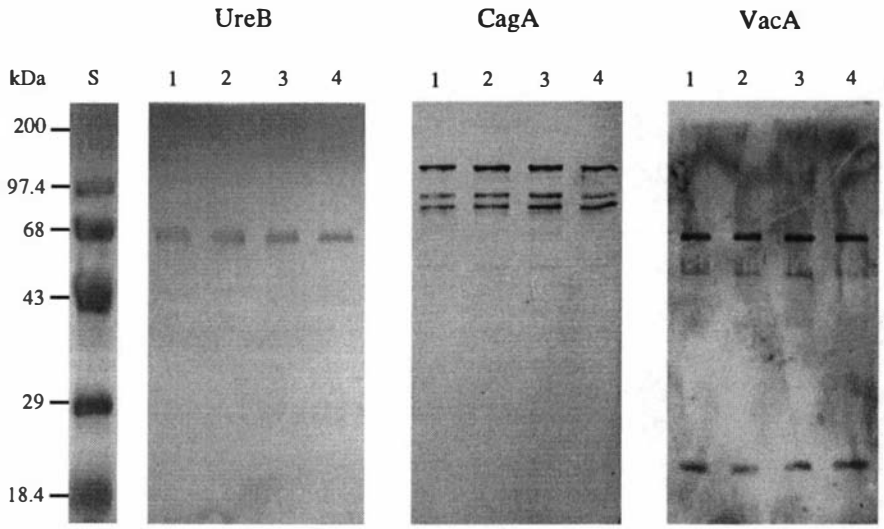


Figure 3.24. Expression of UreB, CagA and VacA is not altered in the *H. pylori virB11*, *fliI* and *fliQ* mutants compared to the 17874 wild type.

Western analysis of whole cell lysates of *H. pylori* cultures with monoclonal anti-UreB (left), polyclonal anti-CagA (middle) and polyclonal anti-VacA (right). Protein samples were separated on a denaturing 12.5% polyacrylamide gel and subsequently blotted. *H. pylori* cells had been grown for 24 h in TSB.

Lanes: 1, *H. pylori* 17874 wild type; 2, *virB11::cat* (pSP118) mutant; 3, *fliI::cat* (pSP110a) mutant; 4, *fliQ::cat* (pSP107) mutant. Lane S, protein molecular size marker.

consistent alteration of protein distribution and / or amount was detected in any of the mutants compared to the wild type culture.

3.7.6 Expression of outer membrane proteins

A knockout mutation in any of the genes involved in the export of outer membrane proteins might reasonably be expected to influence expression of those components. Anticipated intracellular accumulation of the respective protein might result in downregulation of its production. To examine a possible role of *virB11*, *fliI* or *fliQ* in the process of outer membrane protein translocalization, the effect of the knockout mutations on production of the outer membrane proteins HopB (Exner *et al.*, 1995) and OMP4 (encoded by TIGR HP0127) was investigated. Both proteins are members of the described Hop / BAB family of outer membrane components (Boren *et al.*, 1993; Exner *et al.*, 1995; Tomb *et al.*, 1997), and are thus candidates likely to play a role in virulence of the bacterium.

Whole cell lysates of wild type and mutant *H. pylori* cultures grown for 24 h in TSB were used in Western analysis experiments. The primary antibodies were polyclonal anti-HopB and monoclonal anti-OMP4, diluted one thousand-fold and two hundred and fifty-fold, respectively. Samples were prepared and loaded as described in section 3.7.4.2.

As illustrated in Fig 3.25, panel A, none of the mutations had any effect on production of HopB. Similar amounts of this protein were detected in cell lysates from all four cultures.

However, expression of OMP4 was altered. The predicted molecular mass for this antigen is 32 kDa. Reaction of *H. pylori* 17874 whole cell lysates with this antibody resulted in visualization of a protein band with an apparent molecular mass of approximately 29 kDa. Migration point and intensity of the detected immunoreactive band was unchanged in the *fliQ* mutant. In contrast, both the *virB11* and the *fliI* mutation affected the expression of protein reacting with this antibody. In these mutants, a significant reduction of immunoreactive OMP4 protein was detected by the monoclonal antibody (Fig. 3.25, panel B). The described effect was reliably reproduced in cell lysates prepared from four separate sets of *H. pylori* cultures grown for 24 h or 48 h in TSB.

Similar to UreB, VacA and CagA (section 3.7.5), no HopB or OMP4 was detected in supernatant fractions from the *H. pylori* 17874 wild type and mutant cultures.

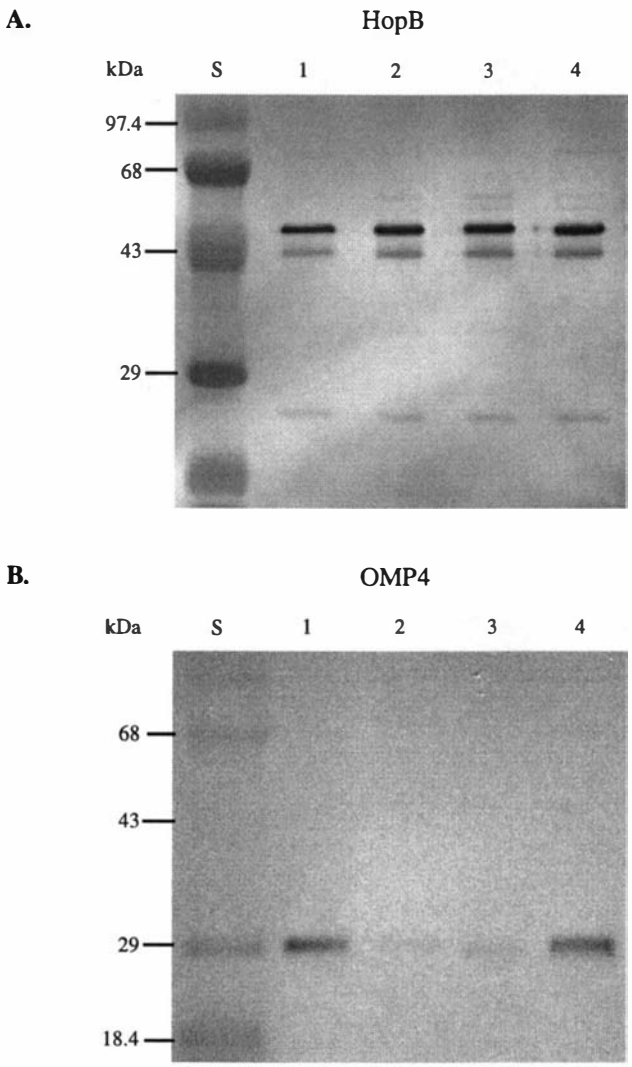


Figure 3.25. Influence of the knockout mutations on expression of two outer membrane proteins (HopB and OMP4) in *H. pylori.*

Western immunoblot analysis using anti-HopB (Panel A) or anti-OMP4 (Panel B) on whole cell lysates of *H. pylori* cells grown for 24 h in TSB. Protein samples were separated on a denaturing 12.5% polyacrylamide gel and subsequently blotted.

Lanes: 1, *H. pylori* 17874 wild type; 2, *virB11::cat* (pSP118) mutant; 3, *fliI::cat* (pSP110a) mutant; 4, *fliQ::cat* (pSP107) mutant. Lane S, protein molecular size marker.

4. DISCUSSION

4.1 Major findings

In this study, a 7 kb fragment of the *Helicobacter pylori* 17874 genome was cloned and sequenced. It revealed an operon which was comprised of six genetic elements. These were an ORF of unknown function (ORF03), the *H. pylori* isoleucyl-tRNA synthetase gene (*ileS*), an *Agrobacterium tumefaciens virB11* homolog (*virB11*), two flagellar genes (*fliI, fliQ*), and a gene probably encoding an essential cell wall metabolism enzyme (*murB*). The subsequently published genome sequences of two other *H. pylori* strains (26695 and J99) revealed the same organization of these elements which are single copy genes. Results presented in this thesis showed conservation of this operon in a panel of clinical isolates. Furthermore, RT-PCR was used to demonstrate cotranscription of the six genes. Primer extension experiments identified the major transcript start site which coincided with the A residue of the ORF03 start codon. A promoter element was thus inferred showing similarity to the *Escherichia coli* σ^{70} consensus sequence. Identification of a second weak transcription start site upstream from *ileS* suggested the presence of an additional promoter element. Non-polar knockout mutations of *virB11, fliI* and *fliQ* were engineered by an allele replacement strategy, and the created *H. pylori* mutants were analyzed. Both the *fliI* and the *fliQ* mutant strains were completely aflagellate and non-motile. In contrast, the *virB11* mutant still produced flagella and displayed slightly greater motility than the wild type. Production of the flagellar hook protein FlgE and both flagellins was severely reduced in the *fliI* and the *fliQ* mutant strains, although the reduction was less stable in the *fliI* mutant. Presence of an outer membrane protein OMP4 was diminished in the *virB11* and the *fliI* mutant, but unaffected by the *fliQ* knockout. Production of several characterized *H. pylori* virulence factors (UreB, VacA, CagA) and the outer membrane protein HopB was not altered by any of the mutations. The major results obtained in this study are summarized in Fig. 4.1.

4.2 Contribution of this study to the field

This thesis contributes to the area of flagellar biology in bacteria. Organization, function and regulation of two flagellar genes, *fliI* and *fliQ*, were investigated in the human gastric pathogen *H. pylori*. Homologs of these genes have been identified and characterized in a number of bacteria, including the Enterobacteriaeceae *Salmonella*

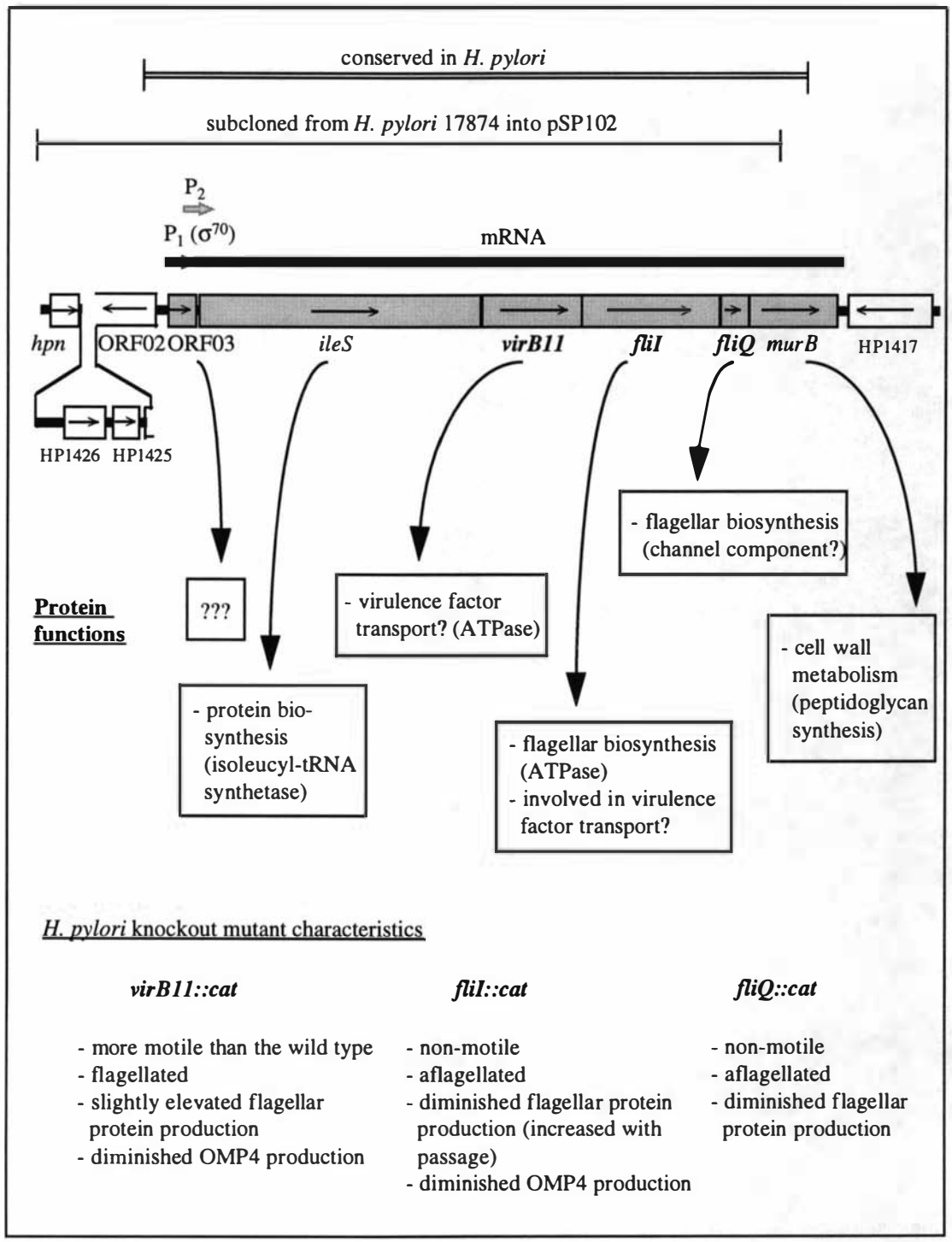


Figure 4.1. Major results of the investigations presented in this thesis.

The respective *H. pylori* 26695 genome region is illustrated. Grey boxes represent the genes of the investigated operon. The identified promoter element is arrowed in black at its approximate position. The second weak promoter identified by primer extension is represented as a grey arrow. Protein functions are suggested. The major characteristics of the constructed *H. pylori* mutants are indicated.

typhimurium and *E. coli*. While no new information regarding flagellar gene function was obtained in this study, regulation and organization of these two putative components of the flagellar export apparatus in *H. pylori* were found to be very unusual.

Upon release of the *H. pylori* 26695 genome sequence (Tomb *et al.*, 1997), it became apparent that organization of flagellar related genes in this bacterium differed profoundly from that found in other bacteria. Scattered distribution of elements involved in flagellar biosynthesis was observed on the *H. pylori* genome. This obviously ruled out the existence of operon-like specific coregulation of those flagellar genes, as described in *S. typhimurium* or *E. coli*. It was unclear what other mechanism(s) would replace this concept. The three *H. pylori* structural flagellar genes described so far, *flaA*, *flaB* and *flgE*, had been found to be regulated individually. They were either controlled by σ^{54} (*flaB* and *flgE*) or σ^{28} (*flaA*) promoter elements (Leying *et al.*, 1992; Suerbaum *et al.*, 1993; O'Toole *et al.*, 1994). In contrast, identification of the transcription start point of the regulatory element *flbA* did not reveal a conserved promoter sequence (Schmitz *et al.*, 1997).

Analysis of the *H. pylori* 26695 genome sequence suggested coregulation of at least some flagellar genes. Among others, the *fliF, fliG* and *fliH* genes might form an operon, and also *flgB, flgC* and *fliE*. However, the longest putative operon of flagellar genes on the *H. pylori* genome consists of only three elements. In comparison, the biggest motility operon identified in *Borrelia burgdorferi* comprises 26 genes (Ge *et al.*, 1997a). In *S. typhimurium*, up to 9 flagellar genes are organized in one operon (Macnab, 1996).

The major contribution of this thesis is the discovery of operon-like coregulation of flagellar genes in *H. pylori* with housekeeping elements. To my knowledge, this concept has not been demonstrated before in flagellar biology. Large motility operons have been found in most flagellated bacteria investigated so far including *Bacillus subtilis* (Albertini *et al.*, 1991), *Treponema pallidium* (Hardham *et al.*, 1995), *S. typhimurium* (Macnab, 1996) and *B. burgdorferi* (Ge *et al.*, 1997a). Occasionally, functions of some putative gene products of these operons could not be determined based on similarities to other database entries. A possible role of these unknowns in essential cell functions can therefore not be excluded. However, the described operon in *H. pylori* is possibly the first documented clear example of such coregulation. An analysis of the *H. pylori* 26695 genome sequence during preparation of this thesis suggested cotranscription of at least seven other flagellar components with housekeeping elements. Transcription of the export apparatus candidate *fliR*, for example, might be coregulated with genes which encode a

peptide chain release factor (*prf2*), and a component involved in biosynthesis of a cofactor for molybdo-enzymes (*moeA*). In addition, the annotated *H. pylori* homologs of *flhB*, *motA* and *motB* are probably also cotranscribed with elements involved in biosynthesis of prosthetic groups and cofactors. These observations, although based merely on intragenic spaces, suggest that cotranscription of flagellar genes with essential elements of unrelated functions might be quite common in *H. pylori*, and may be an adaptation of the organism to its small genome size.

The subsequent finding that transcription of the investigated operon was probably governed by a σ^{70} promoter was not surprising considering the presence of essential genes (*ileS, murB*) in the operon. Transcription of flagellar genes by RNA polymerase incorporating the housekeeping σ^{70} factor is not unknown. Conserved σ^{70} promoter sequences are located upstream from all identified *B. burgdorferi* motility operons (Manson *et al.*, 1998). In addition, class II flagellar genes are also transcribed from RNA polymerase which incorporates the housekeeping σ factor (Liu & Matsumura, 1994; Quon *et al.*, 1996). However, consensus sequences for promoters upstream from these genes are different from the established σ^{70} recognition site. Transcription of the flagellar class II genes requires certain DNA binding proteins which enable the housekeeping σ factor to recognize the promoter element (Liu & Matsumura, 1994; Quon *et al.*, 1996). These DNA binding components are encoded by class I flagellar genes. Classification of *H. pylori fliI* and *fliQ* as class II flagellar genes is therefore questionable, since their transcription is probably independent from a class I element. However, posttranscriptional regulation mechanisms (as suggested for *C. crescentus* class IV flagellar genes (Anderson & Newton, 1997)) cannot be excluded. In conclusion, the observed cotranscription of *H. pylori fliI* and *fliQ* with putative housekeeping elements represents a novel concept of flagellar gene regulation, and possibly questions the hierarchical classification of these genes as class II flagellar components. The direct involvement of σ^{70} in flagellar gene expression in *H. pylori* could indicate the importance of motility in all phases of growth of the bacterium in nature.

The second novelty revealed in these studies is the physical linkage of flagellar genes with an element sharing similarity to putative energizers of conjugal plasmid distribution in bacteria or nucleoprotein export in *A. tumefaciens*. Flagella, conjugal pili, and *vir* pili necessary for T-DNA transport are suspected to represent different manifestations of the same mechanistic principle (Winans *et al.*, 1996). These structures are apparently related, although they have different functions. They might have been evolved from the same ancestory system (Winans *et al.*, 1996). However, although

related, components of these different systems have not, to my knowledge, been shown to be cotranscribed before.

In addition, results presented in this thesis indicate for the first time a functional crosstalk between a flagellar component and a protein possibly involved in virulence factor presentation. Knockout mutagenesis of both *virB11* and *fliI* reduced the production of an *H. pylori* outer membrane protein OMP4. The mechanism whereby this reduction is achieved awaits further clarification. The effect of *fliI* mutagenesis on OMP4 production indicated non-flagellar functions of the FliI protein. In general, VirB11 and FliI are both thought to contribute to export of cellular components by ATP hydrolysis and subsequent energy transfer to the translocation process. The function of *H. pylori* VirB11 remains uncertain since the existence of *vir* pili or conjugal pili has not been confirmed in *H. pylori* to date. The protein has been found to be essential for successful colonization of *H. pylori* SS1 in mice (Noonan, 1998). This might indicate a possible role of VirB11 in virulence factor transport. One of its substrates might then be OMP4. The observed effect of the *fliI* knockout on OMP4 production could reflect involvement of FliI in OMP4 translocation, possibly in cooperation with VirB11. Further tests will have to substantiate or reject the indications for a functional crosstalk between flagellar and non-flagellar protein export machineries presented in this thesis.

4.3 Discussion of the experimental data

4.3.1 The genetic organization of the investigated operon

An operon of the *H. pylori* genome was investigated which consists of six genetic elements: ORF03, *ileS, virB11, fliI, fliQ* and *murB*. These genes are single copy elements on the *H. pylori* 26695 and J99 genomes. They encode proteins with a surprising variety of predicted functions. These include roles in protein biosynthesis, virulence factor export, flagellar biosynthesis and cell wall construction.

ORF03, the first component of this operon, displays no significant similarities to genes encoding characterized proteins. The *ileS* gene encodes isoleucyl-tRNA synthetase, an essential enzyme for protein biosynthesis transferring isoleucine molecules onto the respective tRNA (Stryer, 1988). Based on BLAST similarity analysis, the *ileS*-linked *virB11* gene is predicted to encode an ATP-hydrolyzing component of a translocation machinery transporting proteins out of the cell. *H. pylori* 17874 *fliI* and *fliQ* encode components involved in flagellar biogenesis. They are required for motility, but not survival under laboratory conditions. However, lack of motility was shown to prevent *H. pylori* infection in laboratory animals (Eaton *et al.*, 1992; Eaton *et al.*, 1996) suggesting

that these genes are essential *in vivo*. Finally, the *murB* gene shares significant similarity (49%) to *B. subtilis murB* whose product is predicted to be involved in an essential step in cell wall biosynthesis (Rowland *et al.*, 1995).

The genetic organization of the described locus is unique in three ways. Firstly, obviously essential genes with predicted housekeeping functions (*ileS, murB*) are genetically linked with elements necessary for motility (*fliI, fliQ*) and possibly virulence factor presentation (*virB11*). This arrangement appears to be sub-optimal since it implies a lack of effective regulation mechanisms matching the different requirements for the gene products encoded in this locus. To my knowledge, examples for such clustering of genes with these unrelated predicted functions have not been described in the literature.

Secondly, the genetic linkage between an element likely to be involved in virulence factor export (*virB11*) and genes necessary for transport of flagellar proteins (*fliI* and *fliQ*) is exciting. Based on sequence similarities, flagellar export has been associated with type III protein transport (Lee, 1997). However, the physical linkage of a gene similar to nucleoprotein transfer energizers contributing to T-DNA transfer or conjugal plasmid distribution and genes required for flagellar protein export has not been described previously. Obviously, coexpression and coregulation of genes involved in these similar processes would support a functional relationship between the components.

Thirdly, the two genes involved in flagellar biosynthesis in this study, *fliI* and *fliQ*, have not previously been identified as part of the same operon. The *fliI* gene has been identified in several bacteria including *S. typhimurium* (Vogler *et al.*, 1991), *Rhodobacter sphaeroides* (Ballado *et al.*, 1996) and *B. burgdorferi* (Ge *et al.*, 1996). In those genomes, it is preceded by the gene for a flagellar export apparatus candidate *fliH*, and followed by a gene encoding an acidic, hydrophilic polypeptide of 14-18 kDa, typically labelled *fliJ*. The latter was suggested to have a chaperone-like function. It might prevent premature intracellular aggregation of flagellar transport substrates (Stephens *et al.*, 1997), and may promote their interaction with the flagellar export apparatus. A *fliJ* homolog could not be detected in the *H. pylori* 26695 genome. The *fliH* gene was identified in a separate operon together with its usual preceding elements *fliF* and *fliG*, 650 kb away from *fliI*.

The *fliQ* gene is flanked by *fliP* and *fliR* in most bacterial genomes, including *B. subtilis* (Bischoff *et al.*, 1992), *E. coli* (Malakooti *et al.*, 1994), *T. pallidium* (Hardham *et al.*, 1995), *B. burgdorferi* (Ge & Charon, 1997) and *S. typhimurium* (Ohnishi *et al.*, 1997). The precise functions of FliP and FliR are not identified, but they are known to be

involved in flagellar biosynthesis (Ohnishi *et al.*, 1997). In the *H. pylori* 26695 genome, two elements with very high similarity (80.3% and 77.6%) to the *Bacillus subtilis fliP* gene are positioned in an operon-like structure more than 750 kb away from the *fliQ* gene. The putative *fliR* homolog can be found unlinked to either the *fliP* homologs or *fliQ*.

Successful subcloning of the described operon from the *H. pylori* genomic library (section 3.1.2) was achieved in surprizingly high frequency. Twelve of forty six transformants tested contained the desired *H. pylori fliI* sequence. Consequently, the locus was originally suspected to be present in several copies per bacterial cell. However, the operon was subsequently shown to contain single copy chromosomal genes. Instead, the high success rate of the subcloning trials can be rationalized from the experimental design. Both enzymes used for fragmentation of the *H. pylori* chromosome, *Hin*dIII and *Bgl*II, would statistically cut every 4^6 = 4,096 bp on a genome which had a random distribution of nucleic acid residues. These enzymes recognize A+T-rich restriction sites (AAGCTT for *Hin*dIII, AGATCT for *Bgl*II). The *H. pylori* genome has 61% A+T (Tomb *et al.*, 1997), so both enzymes would be expected to cut more frequently than the statistical value given above. This would probably hold true although some of these restriction sites on the *H. pylori* genome could be masked, for example by methylation. In fact, the *Bgl*II / *Hin*dIII digest performed on *H. pylori* 17874 genomic DNA hardly resulted in any visible fragments bigger than 3 kb after electrophoretic separation. Therefore, the excised region of 5-10 kb fragments probably did not contain many different *Bgl*II / *Hin*dIII segments which could be cloned into pUC18.

The sequence data obtained for the genes of the described operon are in agreement with predicted functions of the respective gene product. Absence of strongly hydrophobic regions within the predicted amino acid sequence of the *H. pylori* 17874 *virB11* gene product suggested VirB11 not to span through the bacterial membrane. However, the *A. tumefaciens* VirB11 protein colocalizes with the cytoplasmic membrane (Christie *et al.*, 1989; Thorstenson *et al.*, 1993). The protein is apparently intracellularly associated with the inner membrane. This assumption corresponds to the predicted function of the protein as a transport energizer. Like all of its identified homologs, the *H. pylori* 17874 VirB11 contained the Walker A motif G/A-X_4-G-K-T/S. Amino acids of this motif are known to be involved in ATP binding (Abrahams *et al.*, 1994). The presence of this motif suggested ATPase activity of the *virB11* gene product. In support of this theory, ATP hydrolysis by the *A. tumefaciens* VirB11 protein has been experimentally proven (Christie *et al.*, 1989).

In the amino acid sequence of the *H. pylori* 17874 *fliI* gene product, the two Walker boxes A and B were identified. This gene evidently also encodes an ATPase. While significance of the Walker A motif for ATP binding is undisputed, the importance of Walker B (R/K-X_{2-3}-G-X_3-L-$X_{3-5,\ hydrophobic}$-D) for enzymatic activity is controversial. However, its conserved aspartic acid residue binds the magnesium ion via a conserved water molecule (Abrahams *et al.*, 1994; Shirakihara *et al.*, 1997). A motif was also detected in the FliI sequence which is a distinguishing signature for both α and β subunits of ATPases. The motif, P-S/A/P-I/V-D/N-X_3-S-X-S, is conserved in several type III export energizers which are homologous to FliI, including *Erwinia carotovora* HrcN and *Xanthomonas campestris* HrpB6 (Fig. 3.8). In the crystal structure of the β subunit of the bovine F_1 ATPase, this motif is part of the presumed contact site between α and β subunits (Moore, 1999). It may also be involved in positioning of a conserved tyrosine residue which is located immediately upstream from the signature sequence. This residue has been shown in F_1 ß subunits to be a component of the adenine-binding pocket (Abrahams *et al.*, 1994; Shirakihara *et al.*, 1997). If subunit interaction would constitute part of the function of the signature sequence in the bovine mitochondrial F_1 ATPase , FliI would be assumed to function as a complex rather than a monomer. However, experimental data on purified *S. typhimurium* FliI suggested protein monomers to be able to hydrolyze ATP (Fan & Macnab, 1996), but these findings do not conclusively predict the state of the protein *in vivo*, nor exclude its multimerization.

The Geneworks analysis programme suggested low surface probability and antigenicity for the *H. pylori* 17874 *fliI* gene product. These characteristics are in agreement with a cytoplasmic localization of FliI. The protein might also be transiently associated to the inner membrane. In *C. crescentus*, FliI protein was localized both in the cytoplasm and in association with the membrane (Stephens *et al.*, 1997).

For the *H. pylori* 17874 *fliQ* gene product, strong hydrophobicity and two transmembrane segments were identified by the Geneworks analysis programme. Localization of FliQ in the cytoplasmic membrane is therefore assumed, which fits with its presumptive channel functions in flagellar biosynthesis. The hydropathicity plots are remarkably similar for all FliQ homologs thought to function in type III protein secretion machineries. These observations suggest a very conserved protein structure for all these channel components, and infer a similar architecture of the transport channels in flagellar and type III export systems. In addition, the high isoelectric point calculated for *H. pylori* FliQ indicates polarity and a positive charge of the protein under physiological conditions. These characteristics might facilitate interaction of FliQ with other, negatively charged flagellar components. This hypothesis is supported by the relatively low isoelectric points

of all identified hook and filament components of the *H. pylori* flagellum. Calculated pI's for FlaA, FlaB and FlgE were 5.96, 6.08 and 4.76, respectively. Similar isoelectric points were determined for the gene products of the second *flaB* and *flgE* homologs in the *H. pylori* 26695 genome (5.12 and 5.57, respectively).

A slightly elevated G+C content of 43% compared to the overall value in the *H. pylori* 26695 genome (39%) was determined for the protein export locus (*virB11, fliI* and *fliQ*). This may possibly reflect *en bloc* acquisition of motility genes by the bacterium. Several other flagellar genes including *flaA, flaB, flgE* and *fliR* share the 43% G+C content observed for the described protein export locus. However, the scattered distribution of flagellar related genes within the *H. pylori* genome argues against horizontal transfer of flagellar operons into the bacterium. It was noted, though, that a considerable number of *H. pylori* genes have sequences that are most similar to genes found in organisms that are not classified as Gram-negative bacteria (Doolittle, 1997). In conclusion of these observations, horizontal gene transfer was suggested to be a major contributor to the architecture of the *H. pylori* genome.

4.3.2 Conservation of the operon in *H. pylori*

The detected conservation of the described operon in *H. pylori* was in contrast to the originally assumed high genetic variability of the bacterium (Taylor *et al.*, 1992; Bukanov & Berg, 1994). This conservation was more consistent with a report of relatively conserved overall coding capacity (Hancock *et al.*, 1998), and corresponded to subsequent findings of relatively low strain-specific genetic content in the *H. pylori* genome (Alm *et al.*, 1999).

In all 18 strains of *H. pylori* examined, the multifunctional operon was present. In contrast, the genes immediately downstream and upstream from the described operon were less conserved.

The gene downstream from the described locus, TIGR HP1417 in *H. pylori* 26695, was separated from the flagellar export locus by additional 700 bp in several *H. pylori* strains tested. The physical linkage of *hpn* to the described operon was also not strictly conserved. Nearly half of the tested clinical *H. pylori* isolates displayed PCR amplification product sizes inconsistent with the data obtained from pSP102 or the *H. pylori* 26695 genome sequence. Surprisingly, PCR amplifications on *H. pylori* 17874 genomic DNA resulted in a product corresponding to the *H. pylori* 26695 genome

sequence, much bigger than expected from the pSP102 insert. The pSP102 insert was cloned from genomic DNA that had been prepared in 1995 from *H. pylori* 17874 cells. The conservation studies presented here were performed on chromosomal DNA of this strain that was prepared more than 18 months later. It cannot be ruled out that rearrangements had occured in the laboratory. The observed replacement of 85 bp identified in the pSP102 insert downstream from *hpn* with 1287 bp in the *H. pylori* 26695 genome, and 218 bp in the *H. pylori* J99 chromosome, was intriguing. It might be speculated that the strains which displayed differing PCR product sizes in the linkage tests of *hpn* to *ileS* may also have an altered DNA content in the same region. This locus downstream from *hpn* might thus represent a highly variable region of the *H. pylori* chromosome which may be subject to frequent rearrangements.

The two *H. pylori* 26695 genes which were lacking in the pSP102 insert (and in *H. pylori* J99), HP1425 and HP1426, did not exhibit significant similarities to previously investigated proteins.

The observed variability downstream from *hpn* almost coincided with a *Hin*dIII site on one end, but not the other, of the changed segment. This *Hin*dIII site (position 395 of the pSP102 insert) is still conserved in the *H. pylori* J99 and 26695 genomes, and their sequences continue to be identical to the insert of pSP102 for another 5 bp upstream from the restriction site. During the subcloning of the described operon, all 12 colonies that carried an insert which hybridized to the *fliI* probe were alike in this region. Protection of this restriction site from the endonuclease during the cloning procedure is a likely explanation for its presence in pSP101 and pSP102. This protection might have been achieved by either methylation, or less probably by protein interaction with this DNA region. In the *H. pylori* 26695 genome eleven restriction-modification systems and eleven specific methyltransferases were identified (Tomb *et al.*, 1997), suggesting that effective DNA modification systems exist in *H. pylori.*

The conservation of the operon, and the lower degree of sequence stability upstream and downstream from its elements, suggested that the genetic linkage of ORF03, *ileS*, the protein export locus and *murB* has been positively selected in *H. pylori.* This might be attributed to successful contribution of this arrangement to pathogenicity, or a favourable stochiometry of gene products.

4.3.3 Transcriptional regulation of the operon

The initial aim of the transcript analyses was to determine the size of the FliI encoding mRNA. Numerous attempts to detect this message by Northern analysis failed. However, the mRNA's of strongly expressed *H. pylori* genes were consistently and easily detected. The transcript size of the *tsaA* gene was relatively small (approximately 750 nt), the *hspA* message detected was 2,200 nt in size. Unfortunately, even riboprobes failed to detect the *fliI* specific message, although single strand riboprobes are thought to be one of the most sensitive means of detection (Sambrook *et al.*, 1989). Increased amounts of total mRNA in these trials did not change the negative outcome of the experiment. Apparently, the transcript copy number was too low for detection. The subsequently determined *fliI* transcript size (> 6,000 nt) was probably also a major contributing factor to the failure of the experiments. Large nucleic acid fragments were most likely transferred less efficiently from the agarose gel to the membrane. The same restricted mobility that lessened their migration distance during electrophoresis also impaired their transfer. The expected transcript size for the *H. pylori flgE* gene was approximately 2,250 nt (O'Toole *et al.*, 1994). A transcript of similar size was detected for the *hspA* gene. In contrast to *hspA*, *flgE* transcription levels were expected to be lower. Transcription of this gene was suggested to be regulated by a σ^{54} factor (O'Toole *et al.*, 1994), which might have resulted in low amounts of *flgE* specific message under the conditions tested. Moreover, the protein is assumed to be only expressed for a short time frame during active cell division. This might explain the failure to detect *flgE* transcripts.

RT-PCR was therefore performed to detect, and characterize, the FliI encoding mRNA. The superior sensitivity of this method is based on amplification of a cDNA that is generated from the transcript by the reverse transcriptase enzyme. Much lower transcript amounts can thus be detected than in Northern blotting experiments. Using this method, cotranscription of all elements of the described operon in *in vitro* grown cells was demonstrated. The RT-PCR results were consistent and unambiguous for fragment sizes lower than 3 kb. Problems were experienced during amplification of larger mRNA transcript fragments, where several attempts generated only small amounts of minor products, or no products at all. These inconsistencies could most likely be attributed to slow RNA degradation during DNase treatment of the total RNA preparations, or during heat inactivation of the DNase at 75°C for 10 min. Alternatively, residual DNase activity might have affected successful generation of cDNA fragments of this size. However, the primary aim of the RT-PCR experiments was achieved. The six genes were shown to be

an operon, and the start point of the transcript was deliminated to a 170 bp region upstream from ORF03.

Despite the unsuccessful Northern blot trials, primer extension experiments succeeded to identify the transcript start. All three primers used generated cDNA products of very defined sizes, which corresponded to the same transcription start site. This site mapped to the A-residue of the initiation codon for ORF03, making it unlikely that this gene is actually expressed. The promoter tentatively inferred for the export operon has a -10 hexamer TAAAAT, which is almost identical to the consensus sequence for $E\sigma^{70}$, the major RNA polymerase in *E. coli* (Harley & Reynolds, 1987). Its -35 sequence (TGGATA) is also reasonably similar to the consensus TTGACA. Despite the availability of the genome sequence of *H. pylori*, few promoters have been mapped and consensus sequences are not well established. On the basis of this study, promoters for essential housekeeping genes in *H. pylori* may closely resemble those recognized by the *E. coli* $E\sigma^{70}$. It is not obvious why the indicated promoter is so far away from the first gene definitely expressed (*ileS*), but it is unequivocally clear from primer extension using the primer internal to *ileS* (SP051) that most of the transcription is driven by this promoter.

Ignoring possible effects of either premature mRNA termination, processing, or different translation efficiencies, the expression of *fliI* and *fliQ* would be expected to be stochiometrically equal to that of the other gene products encoded by the polycistronic mRNA. This contrasts with the current model of transcriptional regulation of class II genes in Enterobacteriaeceae, and is consistent with the failure to annotate homologs of the established transcription regulator genes, *flhC* and *flhD* (Macnab, 1996), in the *H. pylori* 26695 genome. However, it cannot be ruled out that separate promoter elements exist upstream from some of the genes of the investigated operon. These might be turned on under specific environmental conditions, or activated by regulators yet to be identified. The constitutive basal level of gene transcription of the operon elements could thus be raised for one or more genes. However, if a strong unidentified promoter had initiated transcription for the flagellar genes present in this operon, this transcript might have been detected by Northern analysis. A second weak promoter element was experimentally detected upstream from *ileS*. This promoter only poorly resembled the σ^{70} consensus sequence, and was obviously responsible for a very low percentage of the operon message under the conditions tested. In summary, the *H. pylori fliI* and *fliQ* genes are probably transcribed from RNA polymerase incorporating a σ^{70} factor. They are part of a polycistronic mRNA that also comprises messages for housekeeping genes, and a gene probably involved in virulence factor presentation.

In addition to the previously found differences in regulation of class III flagellar components in *H. pylori* and *S. typhimurium* (O'Toole *et al.*, 1994), this study shows that transcriptional regulation mechanisms also differ substantially further up in the hierarchy of flagellar genes. The evolutionary aspect of this operon formation is controversial. An export operon including *virB11*, *fliI* and *fliQ*, might have integrated into the *H. pylori* chromosome downstream from *ileS* by chance. However, the conservation studies presented in this thesis suggest that the generated operon incorporating *ileS* and *murB* was favourable for the bacterium. Arguably, the apparent role of the identified σ^{70} promoter in transcription of a subset of *H. pylori* flagellar genes indicates the importance of bacterial motility *in vivo*, i.e. during infection. The observation that other flagellar genes like *flhB*, *fliR* and the *mot* genes might also be cotranscribed with essential genes in *H. pylori*, would be consistent with this theory. However, some *H. pylori* flagellar genes (*flaA*, *flaB*, *flgE*) are controlled by specific promoter elements (σ^{54}, σ^{28}), which is not in agreement with this suggestion. Specialization in adaptation to constant motility of the bacterium might be better manifested in *B. burgdorferi*. Here, constitutive expression of all large flagellar operons identified so far is initiated by σ^{70} promoters (Ge *et al.*, 1997a; Manson *et al.*, 1998). This illustrates another deviation from established models in flagellar gene regulation, and infers that hierarchical flagellar gene transcription might not be present at all in that bacterium.

4.3.4 Successful mutagenesis of *H. pylori* 17874 *virB11*, *fliI* and *fliQ*

A two-step PCR-based procedure was successfully applied to generate knockout mutations of *virB11*, *fliI* and *fliQ* in *H. pylori*. This technique has been used before for gene directed mutagenesis in *Helicobacter* research (Kostrzynska *et al.*, 1994; O'Toole *et al.*, 1995; Bauerfeind *et al.*, 1996; Schmitz *et al.*, 1997). However, since this method has been applied in this study to generate mutations within an operon, run-on transcription of elements downstream from the knocked out gene had to be ensured. Only then could phenotypically differing features of the mutant cells compared to the wild type colonies be attributed to the respective gene that had been mutated. In the three generated *H. pylori* mutants which were subsequently used for phenotypic analysis, the desired non-polarity of the mutations was demonstrated by RT-PCR. Transcripts were detected both upstream and downstream from the insertion site of the *cat* cassette. No residual transcription of the disrupted gene was apparent.

Mutagenesis was only successful when resistance cassettes that did not contain transcription terminating structures were used. All trials to generate mutant colonies using the KanΩ cassette, which contains effective terminators both upstream and downstream from its *aphA-3* gene, failed. The failure was not caused by inability of the excised cassette to confer resistance. The same *Bam*HI excised KanΩ fragment had been successfully applied in similar mutagenesis experiments performed in Dr O'Toole's laboratory (Beddek, 1998). However, the mutagenized region in those trials did not interfere with essential functions of the bacterium. In conclusion, transcription of at least one of the downstream elements of the investigated operon was necessary for survival of *H. pylori*. This element is probably *murB* since KanΩ was also used unsuccessfully in *fliQ* mutagenesis attempts. Consequently, the unsuccessful *fliI* mutagenesis trials using KanΩ also inferred that there is no separate promoter element active *in vitro* upstream from *fliQ* and *murB*, but downstream from the insertion site of the resistance cassette. These observations firmly linked *fliI* with these two elements.

The mutagenesis attempts demonstrated that orientation of the resistance cassette in transcriptional direction of the operon members was important for successful generation of surviving mutants. This was consistent with lethality of *murB* transcription interference. Trials to mutagenize the *H. pylori fliI* gene using constructs that differed only in the orientation of the *cat* resistance cassette (pSP110a and pSP110b) resulted in dramatically different survival rates of mutant colonies. When the *cat* orientation was opposite to that of *fliI*, mutants were generated at a rate which was lower by a factor of nearly 10^5 (Table 3.4). Severe difficulties can be predicted for transcriptional readthrough over the resistance cassette when it does not correspond to the orientation of the operon. Since promoter sites to reinitiate transcription downstream from the insertion site were missing, *murB* would not be transcribed. In addition, transcription of the resistance cassette itself would also be impeded. However, in some colonies, these obstacles were obviously overcome, possibly by rearrangement events which were not characterized.

Efficiencies of the mutagenesis trials were higher when mutagenic constructs were electroporated into electrocompetent *H. pylori* cells rather than taken up by natural competence. Using the same mutagenic construct (pSP107), knockout frequencies were approximately 1.5×10^2 higher after electroporation compared to natural DNA uptake. However, both methods resulted in sufficiently high mutant numbers. Electroporation is nevertheless favoured. This procedure is quicker, results in greater mutant numbers, and obviates the requirement for fresh healthy *H. pylori* culture at the start of the experiment.

4.3.5 Phenotypic consequences of the *virB11*, *fliI* and *fliQ* knockout

After successful mutagenesis, the generated *H. pylori* mutants were phenotypically characterized. In these analyses, both expected and unexpected differences of the mutant cells compared to the wild type were identified.

The failure of the agar assay to evaluate motility of the *H. pylori* cultures was unexpected. It might be explained by strain-specific adaptation to laboratory conditions. *H. pylori* 17874 had been passsaged in laboratories for several years, and motility has therefore not been essential for this strain for many generations. It is conceivable that flagellar production (and rotation) had been minimized. The surprisingly low amount of flagella observed per cell in the *H. pylori* 17874 wild type culture argues in favour of this theory. The *H. pylori* strain used in documented successful motility agar experiments (*H. pylori* N6) apparently produced up to six flagella per cell (Josenhans *et al.*, 1995). The highest number of flagella observed by electron microscopy on a single *H. pylori* 17874 wild type cell in this study was two. The alternatively performed microscopical evaluation of bacterial motility proved to be consistent with flagella production patterns subsequently observed by electron microscopy.

The *virB11* knockout did not inhibit flagellar biosynthesis. The slightly greater motility consistently observed, and the elevated production levels of flagellin and the hook protein, may be explained by the introduction of an additional promoter element upstream from *fliI*, namely that of the *cat* cassette. The resulting increase in FliI production may have triggered a positive feedback to production of its putative substrates. More or longer flagella were not obvious in the *virB11* mutant compared to the wild type. A statistical analysis was not performed to substantiate this observation.

HP0525, the annotated *virB11* gene in the *H. pylori* 26695 genome, is located in the *cag* pathogenicity island (*cag-virB11*). It is surrounded by genes with significant sequence identities to three other components of the *A. tumefaciens vir* gene cluster, *virB9*, *virB10*, and *virD4* (Akopyants *et al.*, 1998). The *ileS*-linked *H. pylori* 17874 *virB11* investigated in this study is almost as similar to *A. tumefaciens virB11* as the *cag* pathogenicity island *virB11* gene (21% versus 25% identity). Both *H. pylori* VirB11 candidates are similar to each other (53% amino acid similarity). The reason for the existence of these two similar elements on the *H. pylori* chromosome is currently unclear. Overlapping functions of the *cag*-VirB11 protein and the *ileS*-linked *virB11* gene product were originally assumed in the context of unidirectional virulence factor transport. However, recent data contradicted this suggestion. When attempting experimental

infection of mice with *H. pylori* SS1 wild type and mutants defective in the *ileS*-linked *virB11*, the mutants were unable to colonize the test animals (Noonan, 1998). This was surprising for the highly motile *virB11* mutant. It clearly established VirB11 as an essential colonization factor in mice. In contrast, *H. pylori* can colonize humans without *cag*-VirB11 (as almost all *H. pylori* type II strains lack this component). These observations suggest major functional or transcriptional differences which make one gene and not the other indispensable for colonization. The function of *cag*-VirB11 has not been experimentally investigated. Considering the localization of the *H. pylori* 26695 *virB11* gene as part of the *cag* pathogenicity island it has been speculated that this gene might be involved in transport of virulence factors in cooperation with the other *vir*-like elements present (Christie, 1997b). The *ileS*-linked *virB11* might in contrast be involved in protein export independent of this machinery, possibly by exploiting a different transport apparatus present in the bacterium. The two gene products, although of similar function, would work in different transport gateways. The *ileS*-linked *virB11* gene product could mediate transport processes of proteins essential for colonization of the host. These might be necessary for establishment of cell-cell contact during infection, or protect the bacterium from the host defence. One of its substrates could be OMP4. A *virB11* knockout mutation clearly reduced expression of this protein, as documented by the Western analyses performed in this study. In contrast, the *cag*-VirB11 might encode an energizer of export processes for proteins possibly enhancing the damaging effects of an *H. pylori* infection in humans. These "enhancers" would not be essential for successful colonization of the host. Alternatively, it is also possible that the *cag-virB11* is not expressed at all. The elevated virulence of *H. pylori* type I strains compared to type II strains may still be attributed to other transcribed genes present in the pathogenicity island. Currently, studies on production of *cag*-VirB11 have not been documented. It would be interesting to see whether or not *cag-virB11* is actually expressed.

Knockout of the *H. pylori flil* gene created a non-motile aflagellate mutant strain. It produced less flagellin and FlgE than the wild type, which was also observed by Jenks *et al.* (Jenks *et al.*, 1997). However, the observation of Jenks and coworkers that some *flil* mutant cells retained their ability to produce flagella, was not confirmed here. All *H. pylori flil* mutant cells examined by electron microscopy in this study were aflagellate. An early block of the flagellar assembly pathway obviously occured in this mutant. However, a second, less efficient energy transducer may function in flagellar export enabling *flil* mutants to gradually regain flagellation. Existence of this postulated second ATPase with low affinity for flagellar proteins (or low ATP hydrolyzing efficiency) would explain some of the data obtained in this study. Construction of the flagellar export channel (and

subsequently the hook and filament) would then still proceed, but more slowly, in a *fliI* mutant. This would result in gradual leakiness of this mutation with respect to the negative feedback on flagellin production, as documented by the Western experiments in section 3.7.4.2 (Fig. 3.23). It may also explain the observations of Jenks and coworkers. In support of this hypothesis, a *S. typhimurium fliI* mutant was able to rebuild sheared flagella, although more slowly than the wild type (Vogler *et al.*, 1991). Furthermore, in both *vir* pili and conjugal pili, at least two ATPases have been identified as components of the systems (Winans *et al.*, 1996). These examples suggest that a second energizer might also exist in the flagellar apparatus.

The importance, and nature, of the observed unattached membranous structures in the *H. pylori fliI* mutant culture remain uncertain. It is possible that these objects were aggregated empty flagellar sheaths, as observed in non-motile *flaA* mutants of *H. mustelae* F1 (Josenhans *et al.*, 1995). Alternatively, they might have represented preparational artefacts. At any rate, they were clearly detached from the bacterial cells, and are currently assumed not to represent functional substructures of the bacterium.

Similarly to the *virB11* knockout, the *fliI* knockout mutation suppressed OMP4 production in the bacterium. This result suggested FliI to be involved in transport of non-flagellar proteins. Similarities between components of the flagellar export apparatus and systems for virulence factor presentation in a number of organisms are well established (Lee, 1997). However, a direct crosstalk between these systems has not yet been demonstrated. The production, or stability, of OMP4 is clearly affected in both the *H. pylori virB11* and *fliI* mutants. Although OMP4 has not been proven to be a virulence factor, its sequence similarity of approximately 50% to established *H. pylori* adhesins like BabA and BabB (Ilver *et al.*, 1998) makes it a likely candidate. This study provided first experimental evidence for an overlapping function of virulence factor transport proteins and flagellar export components. The lack of influence of these mutations on HopB production suggests that not all outer membrane protein export of *H. pylori* is affected by FliI and VirB11 function, so the mechanism whereby OMP4 production or stability is reduced warrants further investigation. OMP4 has an N-terminal signal sequence making it likely to be exported in a two-step process resembling either type II or type IV secretion. HopB, in contrast, does not exhibit an obvious signal sequence. The vacuolating cytotoxin VacA is thought to be an autotransporter following type IV export (Hueck, 1998). Mutagenesis of *virB11* and *fliI* did not affect its production. Considering the standard function of FliI in export of flagellar proteins without a signal sequence, its participation in type II export seems unlikely. It is even more unlikely for the VirB11 protein, whose homolog is involved in nucleoprotein transport in *A. tumefaciens* (Winans

et al., 1996). Determination of proteins whose production is affected by mutagenesis of both *fliI* and *virB11* becomes extremely important to establish functions of these genes in protein export. The results obtained here did not allow a conclusive statement as to which translocation type was affected by the knockout mutageneses.

In the *fliQ* mutant, non-motility and loss of flagellation were observed, establishing the FliQ protein as an essential component in flagellar biosynthesis. The observed reduction of flagellin and FlgE production was more stable than in the *fliI* mutant. FliQ may be necessary for construction of the channel through which both membrane associated and exterior flagellar proteins would be transported in order to form the flagellum. If this component was missing, the channel could not be constructed. The inability of the cell to construct the flagellum-specific channel through the cytoplasmic membrane would be sensed, and a regulator would then shut off flagellin and hook protein production. Apparently, the FliQ protein cannot be replaced by a membrane component of other transport systems of the bacterium. It might be speculated that FliQ could actually be part of an interaction complex involving FliI. However, no experimental data to date support this theory.

4.4 Future studies

The results presented in this study may provide the scientific background for future research in various directions. Primarily, future investigations might be focussed on two of the sequenced elements of the operon: *fliI* and *virB11*. Both genes encode putative ATPases driving protein transport processes.

The work described in this thesis has already formed the basis for continued biochemical and structure studies on *H. pylori* FliI. The *S. typhimurium* FliI protein has been studied extensively. Its ATP hydrolyzing property has been investigated, and found to proceed by a mechanism different to other known ATPases (Fan & Macnab, 1996). Resolution of the crystal structure of *H. pylori* FliI would consequently facilitate further insights into this new reaction mechanism. Therefore, the *H. pylori* FliI protein is currently being overexpressed in *E. coli*. Subsequently, it will be solubilized and purified by appropriate affinity chromatography. Preservation of enzymatic FliI activity will be ensured before starting crystallization attempts. This project is a cooperation of Dr PW O'Toole with Dr S Moore from the crystallography unit at the Institute of Molecular BioSciences, Massey University, New Zealand.

Furthermore, site-directed mutagenesis of *fliI* gene regions suspected to be involved in substrate binding would help to characterize the interactions of the FliI protein in *H. pylori*. In addition, identification of the cellular components that interact with FliI during flagellar biosynthesis would be beneficial for flagellar-related research. An experimental approach using FLAG epitope tagging of the *fliI* gene and subsequent copurification of proteins associated with FliI may be pursued. This would require reintroduction of a FLAG-tagged *fliI* gene copy into the *H. pylori fliI* mutant strain. Alternatively, a FLAG-tagged *fliI* gene could be overexpressed in *H. pylori* using a shuttle vector (Lee *et al.*, 1997b). Purification of the labelled FliI (and possibly associated components) by affinity chromatography would follow. These studies would result in a better understanding of the export process whereby flagellar proteins are translocated into the nascent flagellum. Given its conservation within the Enterobacteriaeceae, this mechanism might be exploitable as a future therapy target.

The results presented in this thesis have also suggested an involvement of the FliI protein in the transport of other, non-flagellar components. Identification of those components would elucidate the transport system FliI might be involved in. Extended immunological analyses of the effect of a *fliI* knockout on production, or localization, of several antigenic *H. pylori* components would complement the data generated as part of this thesis. These might then substantiate the presumed multifunctionality of FliI in protein transport, or alternatively define its specificity.

The functional properties of the investigated *H. pylori virB11* gene product are not known to date. Its homologs have been shown in other bacteria to be involved in T-DNA translocation and plasmid distribution by conjugation. Thus, it might provide the energy for similar translocation processes in *H. pylori*. Considering the recent evidence that the protein is vital for successful colonization of mice by *H. pylori* SS1, the conserved *virB11* gene may be of clinical importance in the future. However, its functions have to be clearly identified. The characterization studies suggested for FliI might also be justified for VirB11. In addition, complementation experiments using an *A. tumefaciens virB11* knockout background may provide further evidence of functional similarities of *H. pylori* 17874 VirB11 with its homolog. Comprehensive investigations into the VirB11 protein structure and function would most likely result in a better understanding of the interaction of the bacterial cell with its host.

The suspected roles of *H. pylori* VirB11 and *cag*-VirB11 in virulence factor transport might be overlapping. It would be interesting to investigate an *H. pylori cag-*

virB11 knockout mutant, and analyze its possibly altered virulence and protein production. Direct comparison with the *H. pylori virB11* mutant would subsequently facilitate further statements on functional linkage or independance of these two gene products. In addition, complementation studies and double mutations might further define the relationship between the two similar genes.

The impact of *H. pylori* on gastric cells of its host is expressed by many alterations in their cellular biology. One of those changes is the enhanced production of cytokines such as IL-8. IL-8 production of the gastric cells in response to *H. pylori* exposure is a well characterized parameter for pathogenicity of the bacterium (Crabtree *et al.*, 1994). It would be interesting to see if IL-8 production (or production of other cytokines) of those gastric cells would be altered by one or more of the described knockout mutations. The information which could be obtained from these trials would be very valuable, and may provide further evidence for the importance of both FliI and VirB11 in export of virulence components of the bacterium.

An unresolved issue which arose but was left aside because of time constraints concerned those *H. pylori fliI* mutants that had been generated by electroporation of plasmid pSP110b. In this mutagenic construct, the orientation of the *cat* resistance cassette was opposite to *fliI* gene orientation. Therefore, transcription of *murB* was theoretically impossible in *H. pylori fliI::cat* (pSP110b) mutants. It is unclear what mechanism caused survival of those few mutants that were obtained. A spontaneous change of cassette orientation might have occured during the crossover event. Alternatively, a mutation allowing reinitiation of transcription downstream from the *cat* insertion site might have been introduced. The *murB* gene may also have been translocated. PCR and RT-PCR may be performed to investigate the alterations which resulted in survival of the generated mutants.

The possibly altered hierarchical nature of flagellar gene expression in *H. pylori* might also be investigated further. It would be interesting to see if FliI or FliQ expression is indeed independent from a master operon. Although apparently transcribed from a σ^{70} promoter, protein production might still be regulated posttranscriptionally, as suggested for two of the three *C. crescentus* flagellins (Anderson & Newton, 1997). In order to investigate FliI and FliQ expression, specific antibodies would have to be generated. These would subsequently facilitate expression studies in various stages of *H. pylori* culture. The expression patterns of FliI or FliQ would have to be compared with a suitable

housekeeping element, and a flagellar component which is regulated by a specific promoter (i.e. FlaA or FlaB).

With data from the entire *H. pylori* genome at hand, flagellar biology studies could now be planned more ambitiously. Genes encoding putative flagellar components could be targeted in mutagenesis experiments, and thus quickly be confirmed as part of the flagellar apparatus. However, we know little about interactions of the flagellar export apparatus components with its substrates, or possible involvement of flagellar components in virulence factor transport. The interacting molecules might be determined by identification of genes with altered expression patterns in *H. pylori* mutants which were lacking the gene for a flagellar component. The prospective chip containing signature oligonucleotides of every *H. pylori* 26695 gene (Lee, 1998) might prove immensely useful in these studies. Future investigations targeted to solving those mysteries in flagellar biology will also require structural biology, biochemical analyses and biophysical approaches. There is little doubt that the bacterial flagellum will continue to challenge, and surprise, scientists for a long time coming.

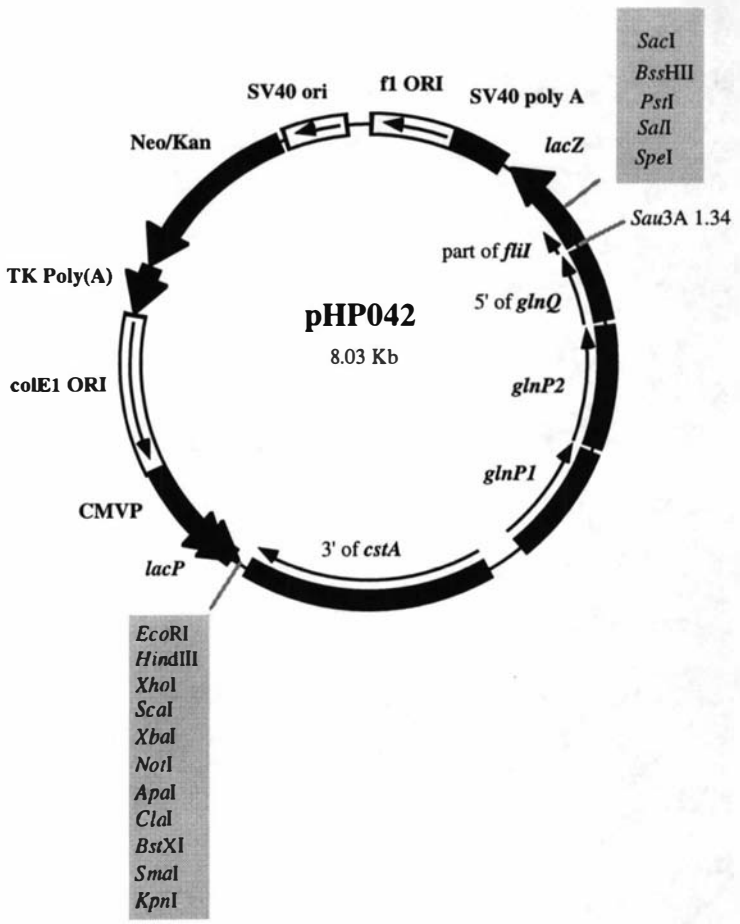
Appendix 1. Physical maps of the pHP and pSP series of plasmids.

The following plasmid maps illustrate vector sizes, genetic elements and their orientation, origins of replication and multicloning sites of the pHP and pSP plasmids used in this study. The pHP plasmids and p26K were constructed by Dr PW O'Toole (Institute of Molecular BioSciences at Massey University, Palmerston North, New Zealand). The pSP constructs were generated by the author.

The maps also depict restriction sites which were important for cloning processes or probe preparations during the presented studies. Multicloning sites are illustrated as grey boxes.

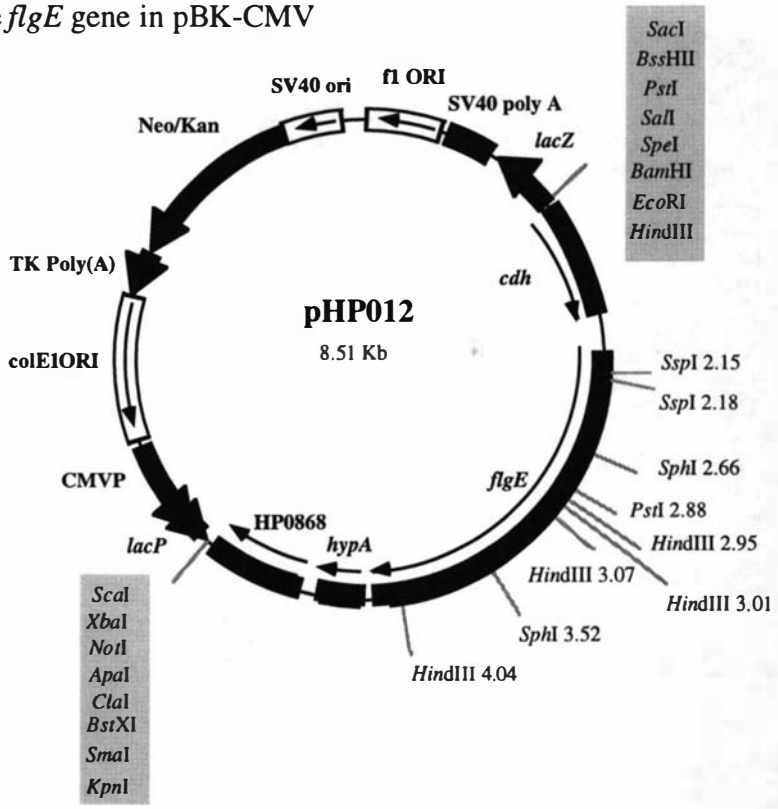
A1.1 Plasmid pHP042

a 3.5 kb genomic fragment of *H. pylori* 17874 in pBK-CMV



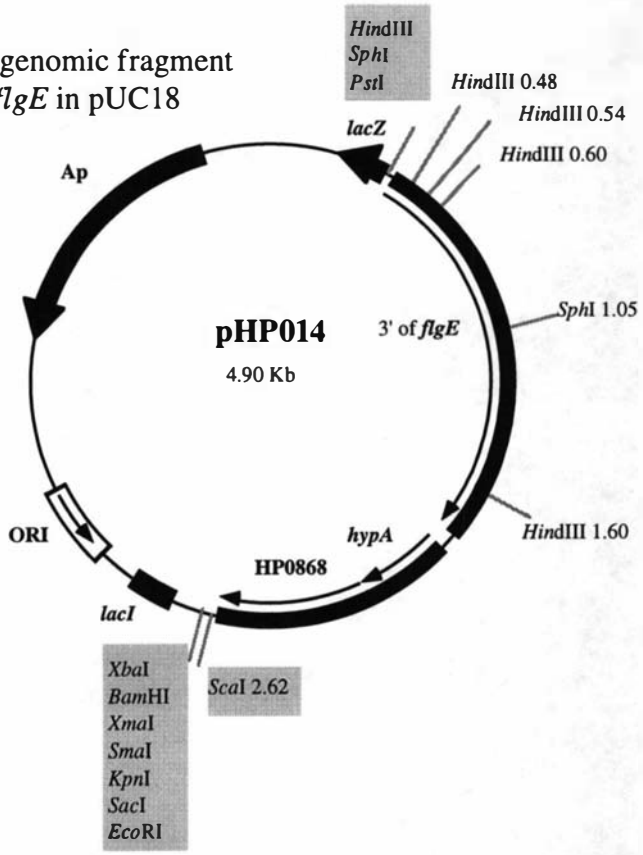
A1.2 Plasmid pHP012

a 4 kb *H. pylori* 17874 genomic fragment containing the complete *flgE* gene in pBK-CMV



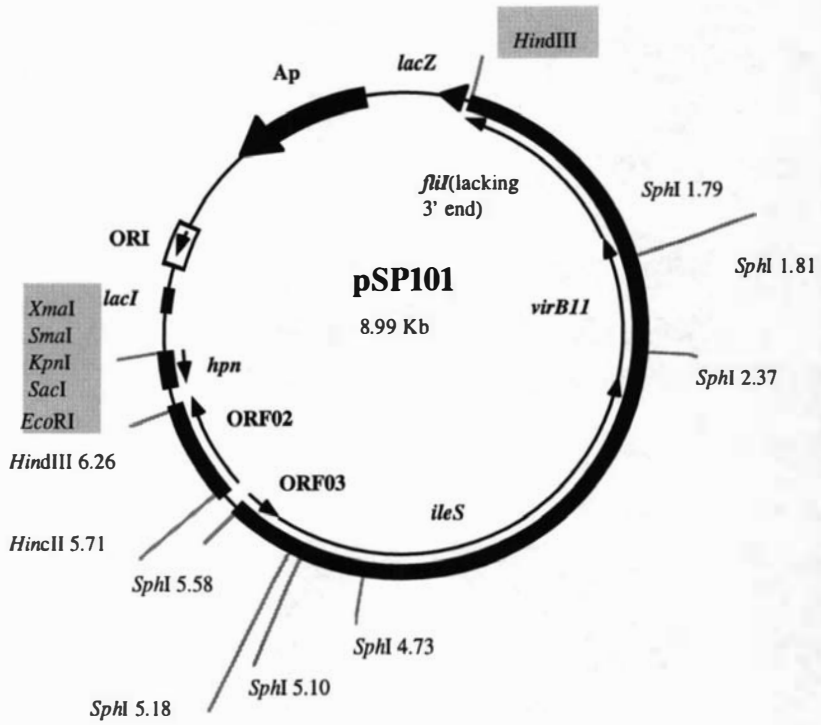
A1.3 Plasmid pHP014

a 2.2 kb *H. pylori* 17874 genomic fragment containing the 3' half of *flgE* in pUC18



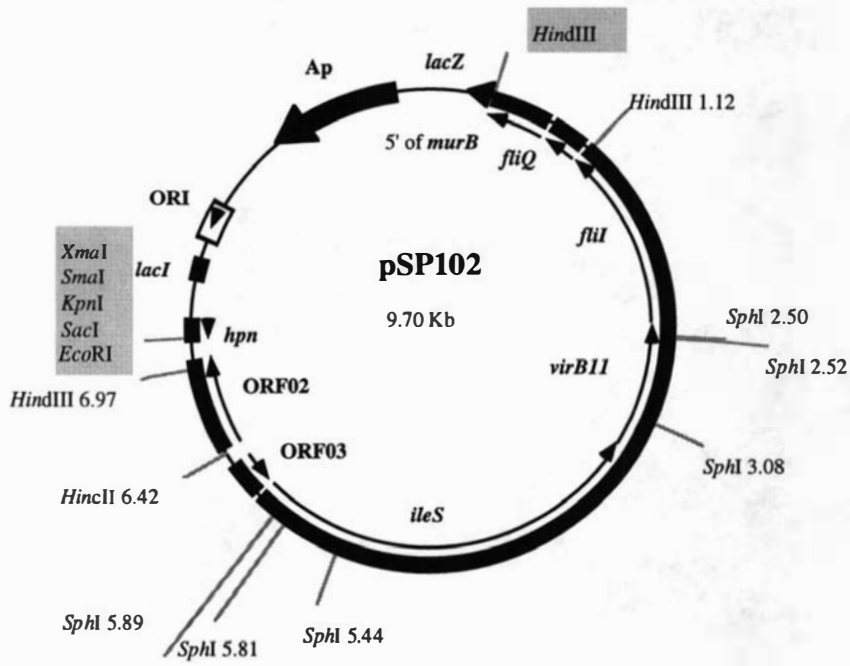
A1.4 Plasmid pSP101

a 6.3 kb *Hin*dIII/*Bgl*II genomic fragment of *H. pylori* 17874 in pUC18



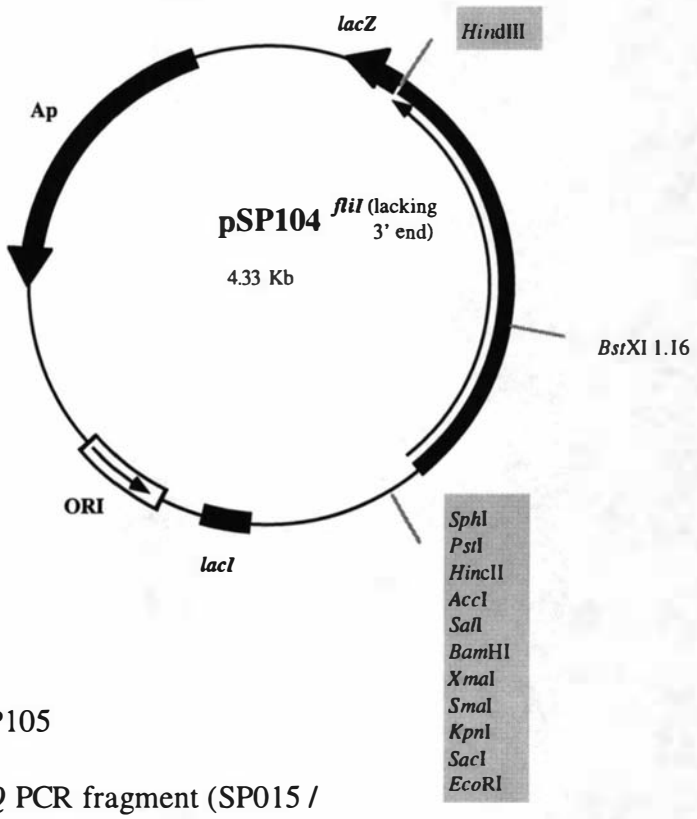
A1.5 Plasmid pSP102

a 7 kb *Hin*dIII/*Bgl*II genomic fragment of *H. pylori* 17874 in pUC18



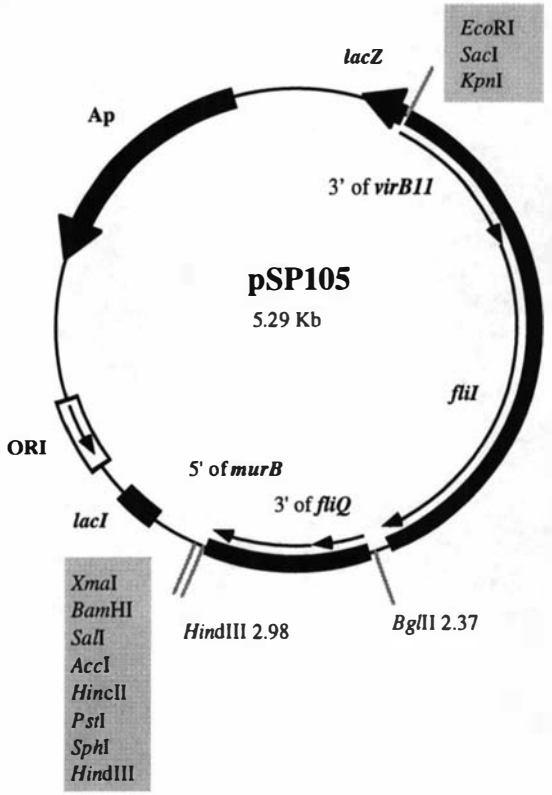
A1. 6 Plasmid pSP104

the 1.4 kb *Sph*I/*Hin*dIII fragment from pSP102 in pUC18



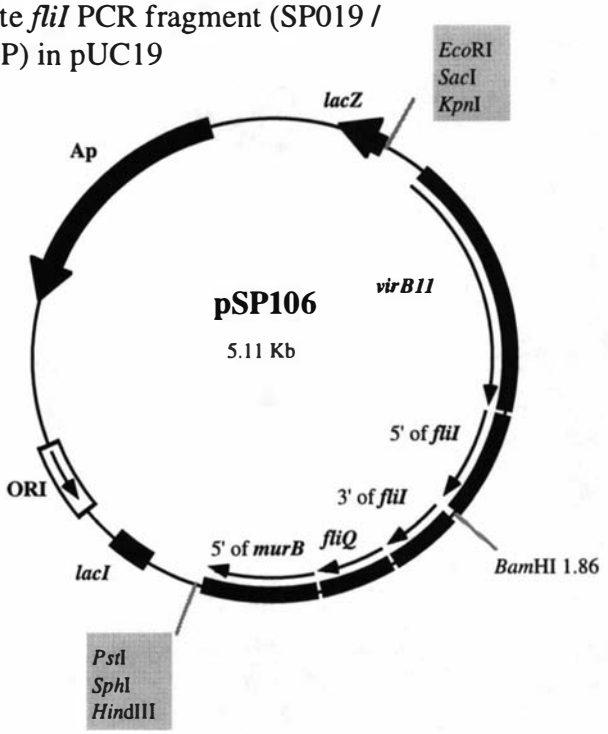
A1.7 Plasmid pSP105

the composite *fliQ* PCR fragment (SP015 / pUC/M13 FP) in pUC19



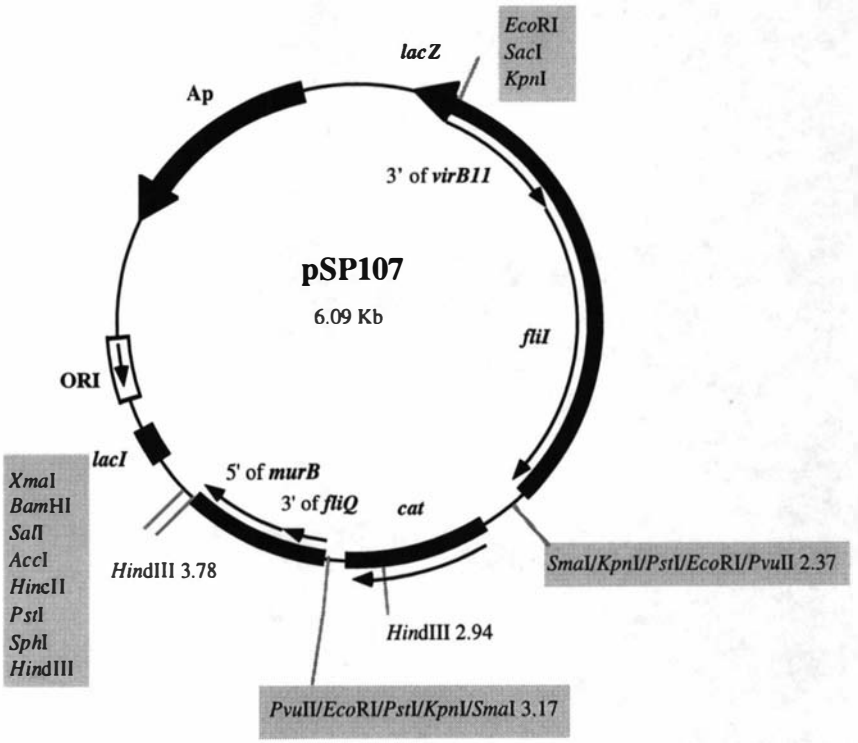
A1.8 Plasmid pSP106

the composite *fliI* PCR fragment (SP019 / pUC/M13 FP) in pUC19



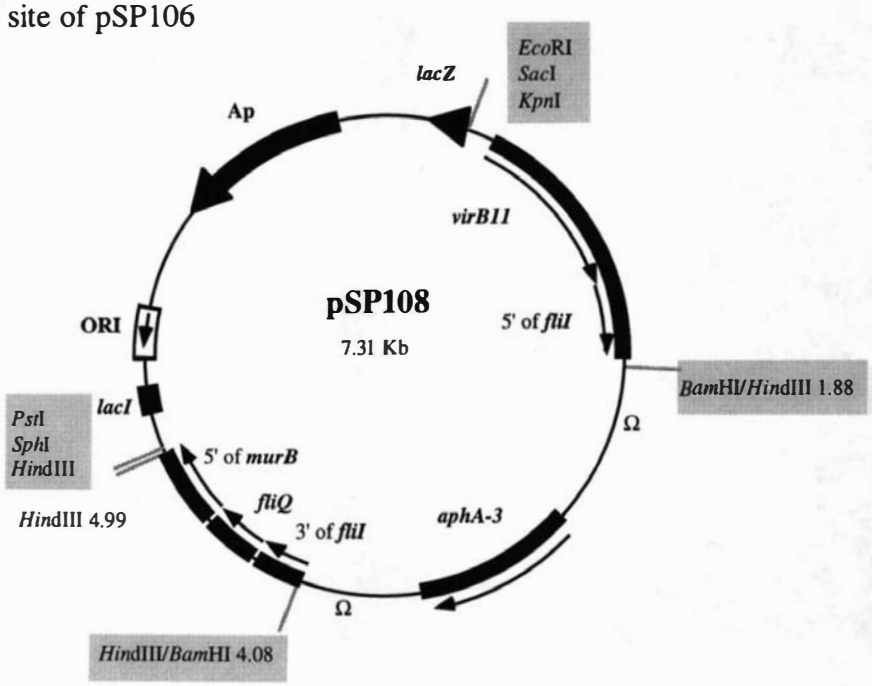
A1.9 Plasmid pSP107

the *cat* gene from pRY109cloned into the primer introduced *Bgl*II site of pSP105



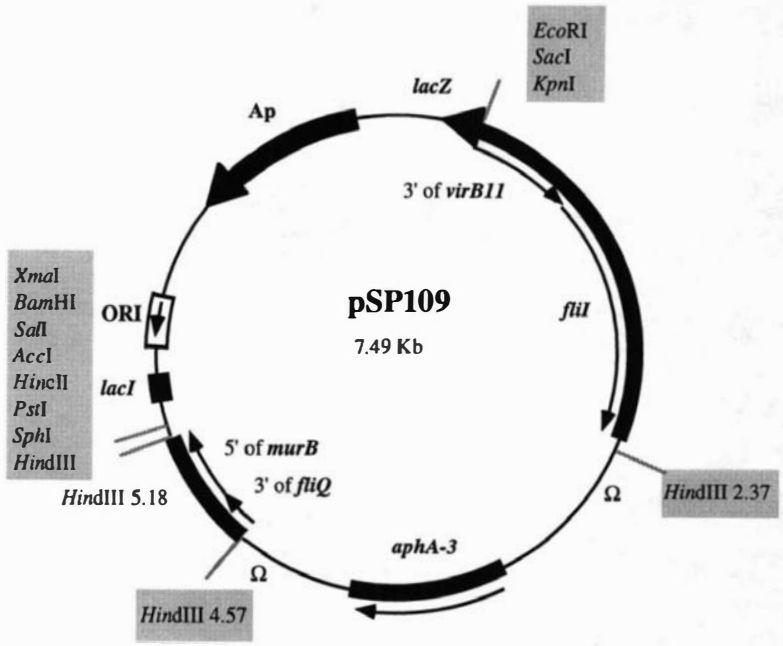
A1.10 Plasmid pSP108

the *aphA-3* cassette from pUC4ΩKm-2 cloned into the primer introduced *Bam*HI site of pSP106



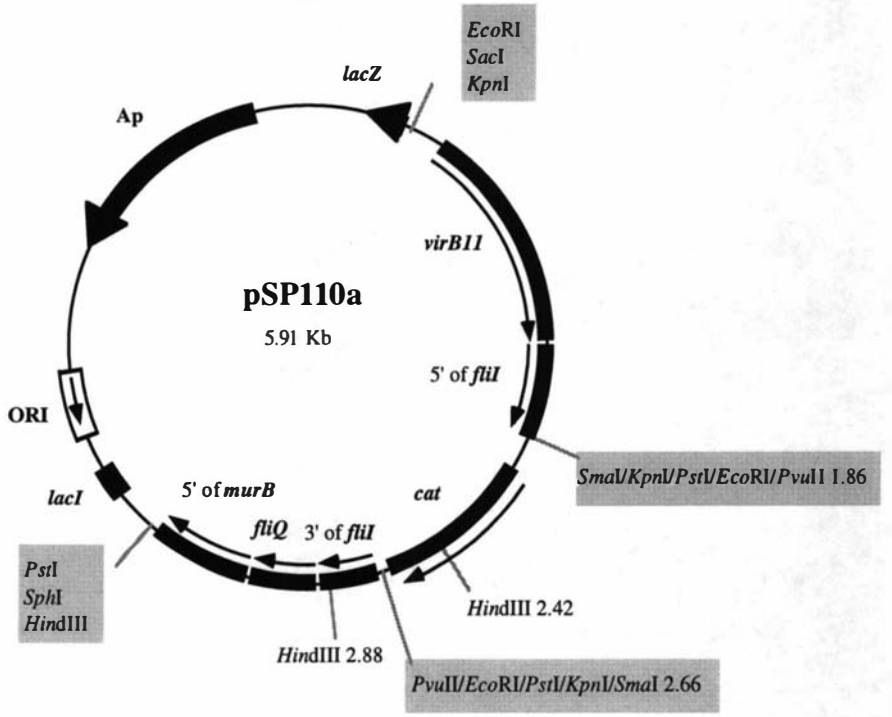
A1.11 Plasmid pSP109

the *aphA-3* cassette from pUC4ΩKm-2 cloned into the primer introduced *Bgl*II site of pSP105



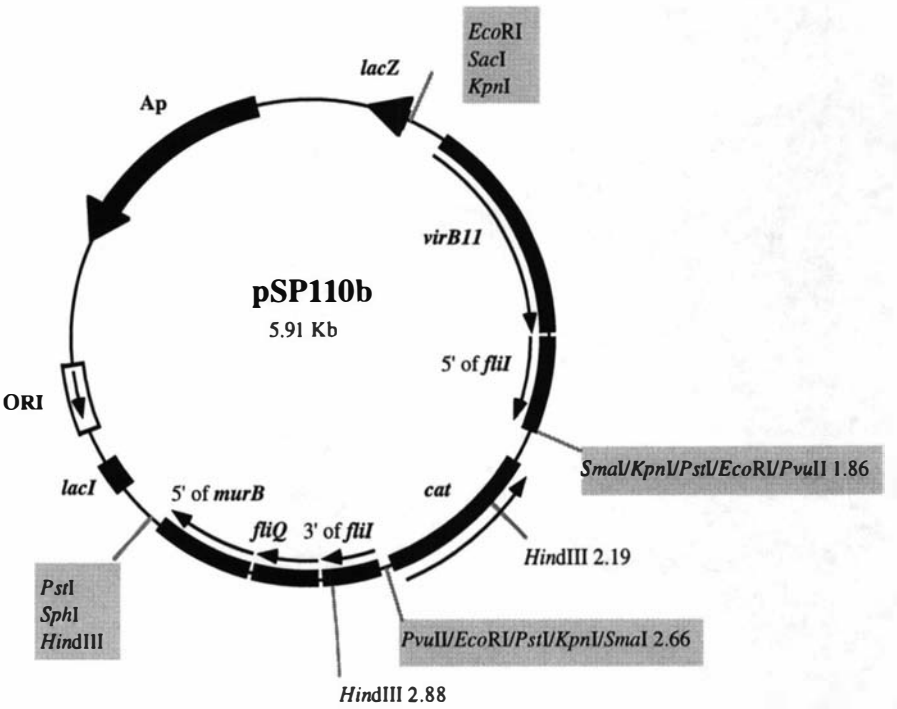
A1.12 Plasmid pSP110a

the *cat* gene from pRY109cloned into the primer introduced *Bam*HI site of pSP106 (*cat* in *fliI* orientation)



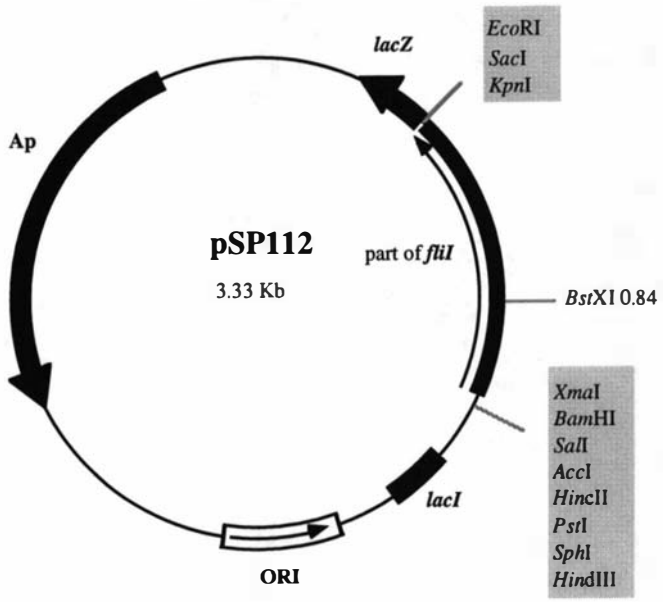
A1.13 Plasmid pSP110b

the *cat* gene from pRY109cloned into the primer introduced *Bam*HI site of pSP106 (*cat* orientation opposite to *fliI* orientation)



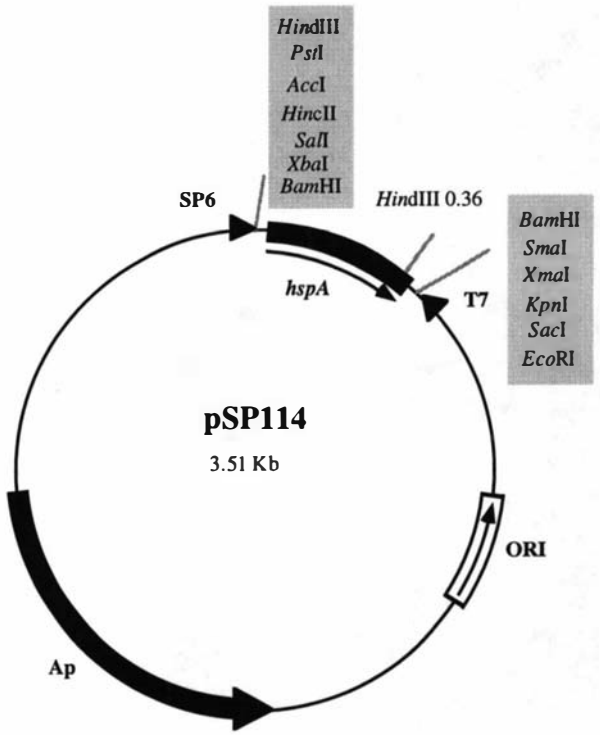
A1.14 Plasmid pSP112

PCR fragment SP002 / SP013 (*fliI* probe) in pUC19



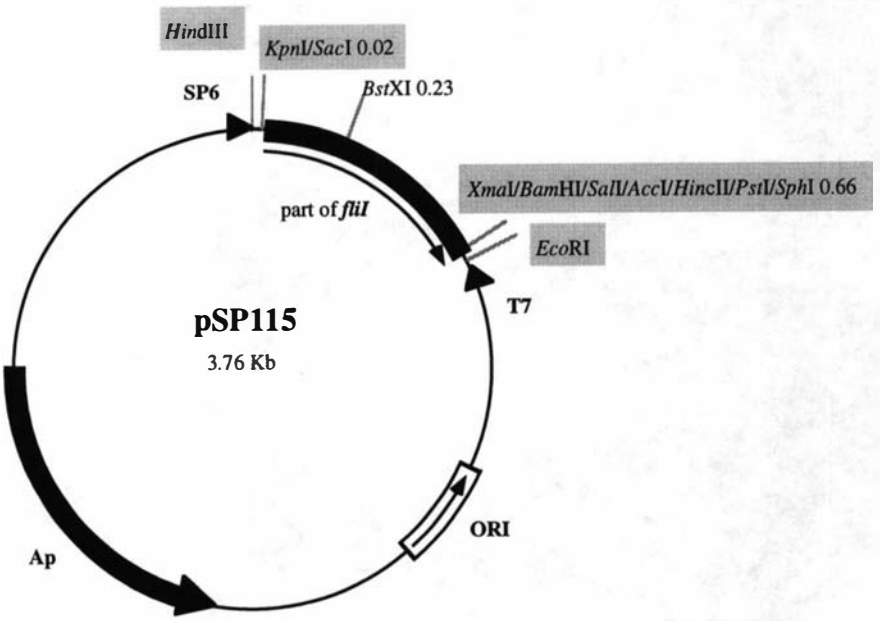
A1.15 Plasmid pSP114

the pSP201 insert (*H. pylori* 17874 *hspA*) in pSPT18



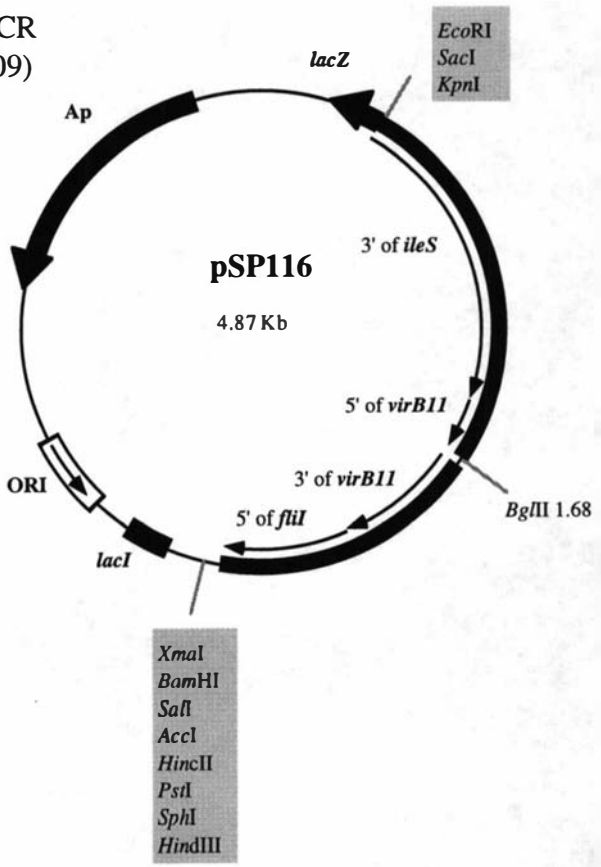
A1.16 Plasmid pSP115

the pSP112 insert in pSPT18



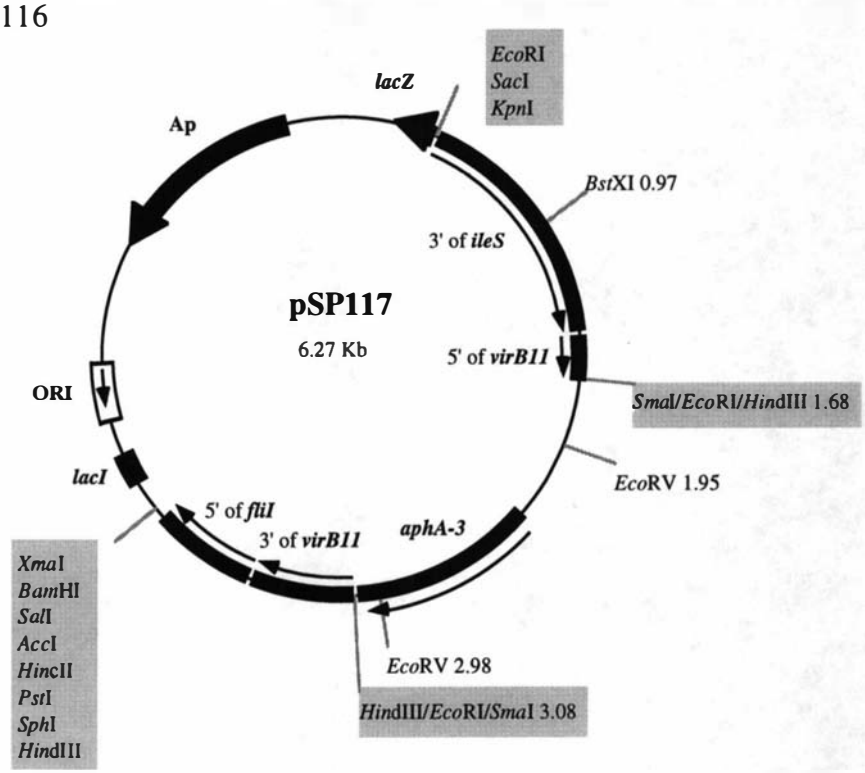
A1.17 Plasmid pSP116

the composite *virB11* PCR fragment (SP034 / SP009) in pUC19



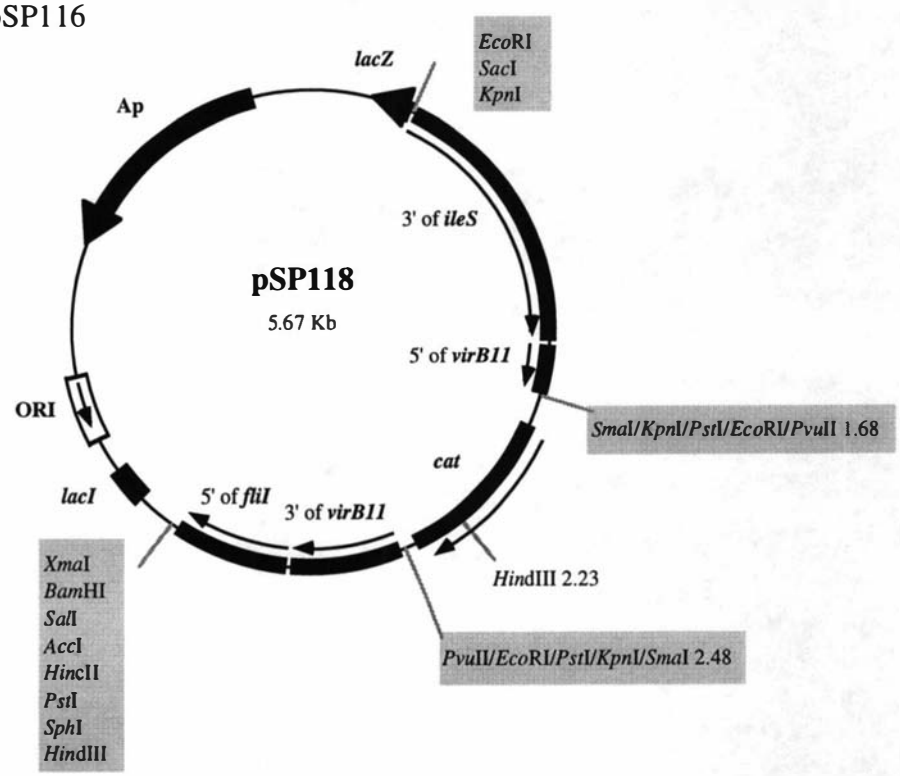
A1.18 Plasmid pSP117

the *aphA-3* cassette from pILL600 cloned into the primer introduced *Bgl*II site of pSP116



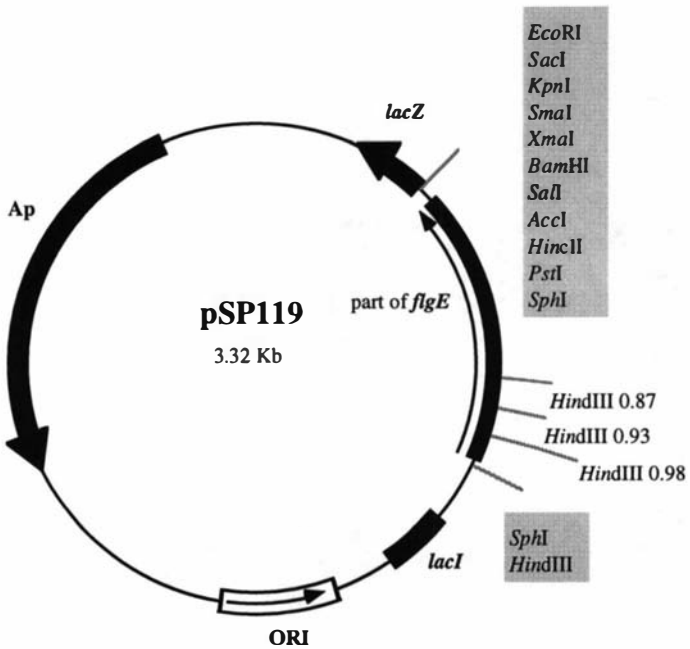
A1.19 Plasmid pSP118

the *cat* gene from pRY109cloned into the primer introduced *Bgl*II site of pSP116



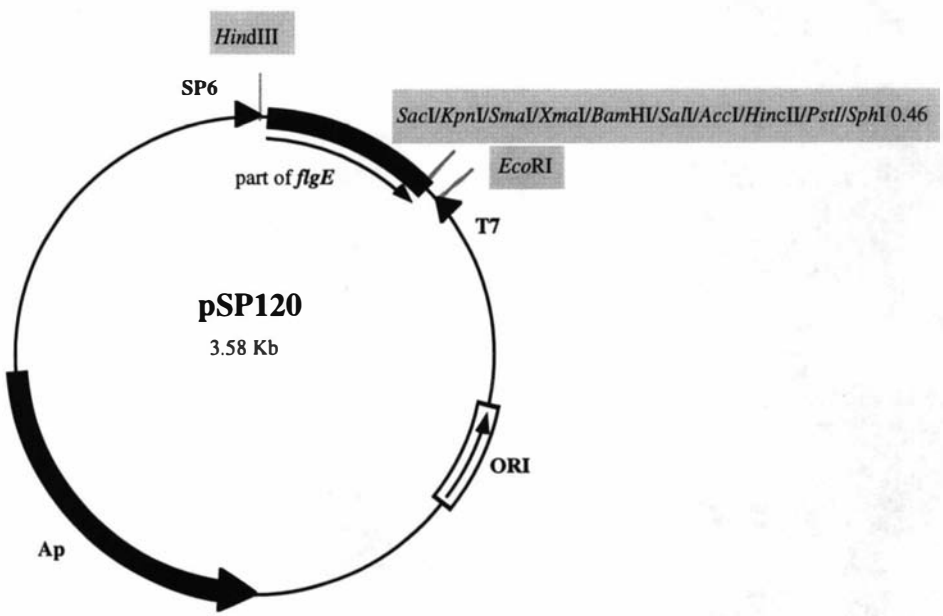
A1.20 Plasmid pSP119

the 0.6 kb *SpHI* fragment from pHP014 (*flgE* probe) in pUC19



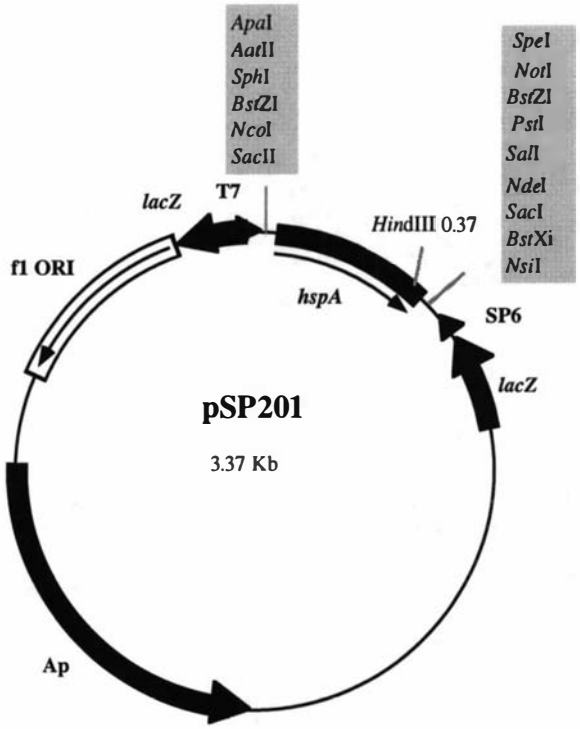
A1.21 Plasmid pSP120

the 0.45 kb *Eco*RI/*Hin*dIII fragment from pSP119 (*flgE* probe) in pSPT18



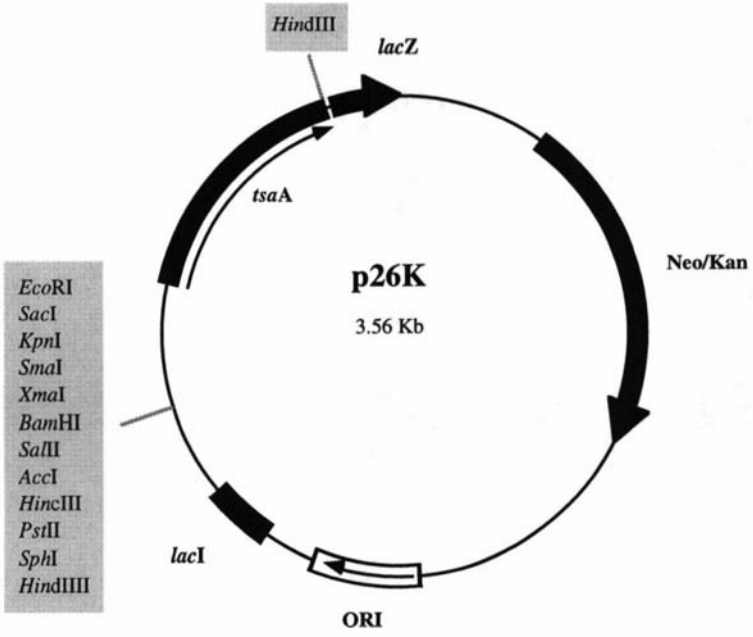
A1.22 Plasmid pSP201

PCR fragment SP007 / SP008 (*H. pylori hspA*) in pGEM-T



A1.23 Plasmid p26K

0.9 kb *H. pylori* 17874 genomic fragment containing *tsaA* in pK18



Appendix 2. The complete sequence of the pSP102 insert.

Both DNA strands are shown. Nucleotides 1585 - 3839 have been obtained from the *H. pylori* 26695 genome sequence. Numbers on the left indicate nucleotide coordinates where residue 1 corresponds to the beginning of the cloned fragment in pSP102. Amino acid sequences from putative open reading frames are illustrated. Restriction sites for *Hin*dIII, *Bst*XI, *Sph*I and *Hinc*II are underlined and indicated. Primer sequences are shaded and labelled. Primer tags: #, sticky 5' tail incorporating restriction site; $, primer purified by polyacrylamide gel electrophoresis.

```
1

CTT ATA TTG TAG AAT AAA ATC CTA GCC AGT GAG CTA GAA TTT AAT TTT TTA ATC AAA GGA
GAA TAT AAC ATC TTA TTT TAG GAT CGG TCA CTC GAT CTT AAA TTA AAA AAT TAG TTT CCT

61
   Hpn>
        M   A   H   H   E   E   Q   H   G   G   H   H   H   H   H   H   H   T
GTC ATC ATG GCA CAC CAT GAA GAA CAA CAC GGC GGA CAC CAC CAC CAT CAC CAC CAC ACA
CAG TAG TAC CGT GTG GTA CTT CTT GTT GTG CCG CCT GTG GTG GTG GTA GTG GTG GTG TGT

121
                BstXI
H   H   H   H   Y   H   G   G   E   H   H   H   H   H   H   S   S   H   H   E
CAC CAC CAC CAC TAT CAT GGC GGT GAA CAC CAC CAT CAC CAC CAC AGC TCT CAT CAT GAA
GTG GTG GTG GTG ATA GTA CCG CCA CTT GTG GTG GTA GTG GTG GTG TCG AGA GTA GTA CTT

181
          --SP045-->
E   G   C   C   S   T   S   D   S   H   H   Q   E   E   G   C   C   H   G   H
GAA GGT TGT TGC AGC ACT AGC GAT AGT CAT CAT CAA GAA GAA GGT TGT TGC CAC GGG CAT
CTT CCA ACA ACG TCG TGA TCG CTA TCA GTA GTA GTT CTT CTT CCA ACA ACG GTG CCC GTA

241

H   E   *
CAC GAG TAA TAT CGG TGT GGC TAG GGG CAA CTT GAC TAG GGT TGT CTC TGG CTT TTG ATT
GTG CTC ATT ATA GCC ACA CCG ATC CCC GTT GAA CTG ATC CCA ACA GAG ACC GAA AAC TAA

301

TTA TTT TCT CTC AAA TTT ATT CAA TCA ATT GGA TCA AGT CGC TTT TTC ACT CGC TAT GTA
AAT AAA AGA GAG TTT AAA TAA GTT AGT TAA CCT AGT TCA GCG AAA AAG TGA GCG ATA CAT
                                                        *   E   S   H   L

361
        --SP017-->                              HindIII
AAG ACC AAG CGT TAT TTG GCT TTG TTT CAT TTT TAA GCT TTA AAT TCA AGC TCC CCT CAA
TTC TGG TTC GCA ATA AAC CGA AAC AAA GTA AAA ATT CGA AAT TTA AGT TCG AGG GGA GTT
 S   W   A   N   N   P   K   T   E   N   K   L   K   L   N   L   S   G   E   I
```

```
421

TCC ACT TGG CGT TAG CCC CCT TTT CGT TAT TGT CAA ACA AGA ACA CCC CTA AAA AGC CTC
AGG TGA ACC GCA ATC GGG GGA AAA GCA ATA ACA GTT TGT TCT TGT GGG GAT TTT CGG GAG
 W   K   A   N   A   G   K   E   N   N   D   F   L   F   V   G   L   F   G   R
   <--SP018--

481

TAG GGG TTT TAA TCA GGC TCA ATT CCA CGC TAG GAA GAA CGA TCT TGC CAT CAG GTT TGA
ATC CCC AAA ATT AGT CCG AGT TAA GGT GCG ATC CTT CTT GCT AGA ACG GTA GTC CAA ACT
 P   T   K   I   L   S   L   E   V   S   P   L   V   I   K   G   D   P   K   V

541

CTT CAA AGG CCA GGA AAA AAG GGC GGT TGT TTT TGA AAT CAT AGC CGT TTT TGG TGT ATG
GAA GTT TCC GGT CCT TTT TTC CCG CCA ACA AAA ACT TTA GTA TCG GCA AAA ACC ACA TAC
 E   F   A   L   F   F   P   R   N   N   K   F   D   Y   G   N   K   T   Y   S

601

AAT AGC CAG GAG CGC TCT GCA AGA TGC CTT TTT GAG TTT CCA CGC TAA AAG AAT CCA AGT
TTA TCG GTC CTC GCG AGA CGT TCT ACG GAA AAA CTC AAA GGT GCG ATT TTC TTA GGT TCA
 Y   G   P   A   S   Q   L   I   G   K   Q   T   E   V   S   F   S   D   L   Y

661

AAT AAA ATC CAG GCT CTA GCC TAA ACG ATT GGA GTT TGG CTA AAG CGT AGC GGT TTT TCC
TTA TTT TAG GTC CGA GAT CGG ATT TGC TAA CCT CAA ACC GAT TTC GCA TCG CCA AAA AGG
 Y   F   G   P   E   L   R   F   S   Q   L   K   A   L   A   Y   R   N   K   W

721

ATA AAA GCA TGA AGC GCT CAC TCC CTA GCA TGA ACA AAG GCC CCC TAG GGT CTT TGA GCT
TAT TTT CGT ACT TCG CGA GTG AGG GAT CGT ACT TGT TTC CGG GGG ATC CCA GAA ACT CGA
 L   L   M   F   R   E   S   G   L   M   F   L   P   G   R   P   D   K   L   K

781

TTG CTT TAG GGT TTT TTT CAA GCG TTT CAT TGT GTT TGG CGA CAA CTC CTC TAT CCA CTT
AAC GAA ATC CCA AAA AAA GTT CGC AAA GTA ACA CAA ACC GCT GTT GAG GAG ATA GGT GAA
 A   K   P   N   K   E   L   T   E   N   H   K   A   V   V   G   R   D   V   K

841
                                                                  --SP020-->
TGC GCC AAT AAG TGC GCA CGC TTT GGC CTT GAT GAG CGA TAT ACA TAT TCA CCA AAC CAG
ACG CGG TTA TTC ACG CGT GCG AAA CCG GAA CTA CTC GCT ATA TGT ATA AGT GGT TTG GTC
 R   W   Y   T   R   V   S   Q   G   Q   H   A   I   Y   M   N   V   L   G   S
                                                                  <--SP026--

901
    BstXI                                                  HincII
AAT GAT GGC CTT TAG GGG GCA AAC AAG AGG CTA ATA AAA GAG CGG TTA ACA AAA TCA CAA
TTA CTA CCG GAA ATC CCC CGT TTG TTC TCC GAT TAT TTT CTC GCC AAT TGT TTT AGT GTT
 H   H   G   K   P   P   L   C   S   A   L   L   L   A   T   L   L   I   V   F

961

ATT GAA ACT TCA AAC AAA TTC CTT AAA AAG AGT GGT GGC TGA ATG CTA ACA TGC TAA AAC
TAA CTT TGA AGT TTG TTT AAG GAA TTT TTC TCA CCA CCG ACT TAC GAT TGT ACG ATT TTG
 Q   F   K   L
          <ORF02
```

1021

ORF03> *SphI*

M R

TAA AAC CCT TAT AAG TTA TGG ATA AAA ACT TAA AGA ATA GGG TAA AAT AAA CGC ATG CGA

ATT TTG GGA ATA TTC AAT ACC TAT TTT TGA ATT TCT TAT CCC ATT TTA TTT GCG TAC GCT

1081

--SP039-->

I D K F L Q S V G L V K R R V L A T D M

ATA GAC AAA TTT TTA CAA TCA GTG GGT TTA GTG AAG CGG CGC GTT TTA GCG ACA GAT ATG

TAT CTG TTT AAA AAT GTT AGT CAC CCA AAT CAC TTC GCC GCG CAA AAT CGC TGT CTA TAC

<--SP048--$

1141

C N V G A V W I N G S C A K A S K E V K

TGC AAT GTA GGG GCG GTA TGG ATC AAT GGG AGT TGC GCT AAG GCT AGT AAA GAA GTG AAA

ACG TTA CAT CCC CGC CAT ACC TAG TTA CCC TCA ACG CGA TTC CGA TCA TTT CTT CAC TTT

<--SP050--$

1201

I G D A I S L H Y L K G I E E Y T I L Q

ATT GGC GAT GCG ATT AGC TTG CAT TAT TTA AAA GGG ATA GAA GAA TAC ACG ATT TTA CAA

TAA CCG CTA CGC TAA TCG AAC GTA ATA AAT TTT CCC TAT CTT CTT ATG TGC TAA AAT GTT

1261

I P T L K N V P R K D T H L Y I A P K T

ATC CCC ACT TTA AAA AAT GTG CCG CGA AAA GAC ACG CAC CTT TAT ATC GCT CCT AAA ACA

TAG GGG TGA AAT TTT TTA CAC GGC GCT TTT CTG TGC GTG GAA ATA TAG CGA GGA TTT TGT

1321

IleS>

K E * V K E Y K D T L N

AAA GAA TAA AGA TAA AAG ATT AAG GAA CAA TCA GTG AAA GAA TAC AAA GAC ACC CTA AAC

TTT CTT ATT TCT ATT TTC TAA TTC CTT GTT AGT CAC TTT CTT ATG TTT CTG TGG GAT TTG

<--SP051--$

1381

L N T T T F S M K G N L S V N E P K T Y

TTA AAC ACA ACC ACC TTT TCT ATG AAA GGG AAT TTG AGC GTT AAT GAG CCT AAA ACT TAC

AAT TTG TGT TGG TGG AAA AGA TAC TTT CCC TTA AAC TCG CAA TTA CTC GGA TTT TGA ATG

1441

SphI

A K W Q E Q Q A F K R M Q A R K D N H G

GCA AAA TGG CAA GAG CAA CAA GCG TTC AAA CGC ATG CAA GCT AGG AAA GAT AAC CAT GGG

CGT TTT ACC GTT CTC GTT GTT CGC AAG TTT GCG TAC GTT CGA TCC TTT CTA TTG GTA CCC

<--SP025--

1501

SphI

D F T L H D G P P Y A N G H L H L G H A

GAT TTC ACC TTG CAT GAC GGG CCG CCT TAT GCG AAC GGG CAT TTG CAT TTG GGG CAT GCC

CTA AAG TGG AAC GTA CTG CCC GGC GGA ATA CGC TTG CCC GTA AAC GTA AAC CCC GTA CGG

1561

L N K I L K D I V V K R E Y F K G K K I

TTA AAT AAA ATT TTA AAA GAC ATT GTC GTT AAA AGA GAA TAT TTT AAG GGG AAG AAA ATC

AAT TTA TTT TAA AAT TTT CTG TAA CAG CAA TTT TCT CTT ATA AAA TTC CCC TTC TTT TAG

```
1621

Y   Y   T   P   G   W   D   C   H   G   L   P   I   E   Q   Q   I   L   E   R
TAT TAC ACG CCC GGT TGG GAT TGC CAT GGT TTG CCC ATT GAG CAG CAA ATT TTA GAG CGA
ATA ATG TGC GGG CCA ACC CTA ACG GTA CCA AAC GGG TAA CTC GTC GTT TAA AAT CTC GCT

1681

L   E   K   E   K   T   S   L   E   N   P   T   L   F   R   E   K   C   R   D
TTA GAA AAA GAA AAA ACA AGC CTA GAA AAC CCC ACG CTG TTT AGA GAA AAG TGC CGA GAT
AAT CTT TTT CTT TTT TGT TCG GAT CTT TTG GGG TGC GAC AAA TCT CTT TTC ACG GCT CTA
                                                          <--SP040--

1741

H   A   K   K   F   L   E   I   Q   K   N   E   F   L   Q   L   G   V   L   G
CAT GCG AAG AAA TTT TTA GAA ATC CAA AAG AAT GAA TTT TTG CAA TTG GGT GTT TTG GGG
GTA CGC TTC TTT AAA AAT CTT TAG GTT TTC TTA CTT AAA AAC GTT AAC CCA CAA AAC CCC

1801

D   F   E   D   P   Y   K   T   M   D   F   K   F   E   A   S   I   Y   R   A
GAT TTT GAA GAT CCT TAT AAA ACC ATG GAT TTT AAA TTT GAA GCG AGC ATT TAT AGA GCC
CTA AAA CTT CTA GGA ATA TTT TGG TAC CTA AAA TTT AAA CTT CGC TCG TAA ATA TCT CGG

1861

L   V   E   V   A   K   K   G   L   L   K   E   R   H   K   P   I   Y   W   S
TTA GTG GAA GTG GCT AAA AAA GGG CTT TTG AAA GAG CGC CAC AAG CCT ATT TAT TGG AGT
AAT CAC CTT CAC CGA TTT TTT CCC GAA AAC TTT CTC GCG GTG TTC GGA TAA ATA ACC TCA

1921
     SphI
Y   A   C   E   S   A   L   A   E   A   E   V   E   Y   K   M   K   K   S   P
TAT GCA TGC GAG AGC GCT TTA GCG GAA GCT GAA GTG GAA TAC AAA ATG AAA AAA TCG CCC
ATA CGT ACG CTC TCG CGA AAT CGC CTT CGA CTT CAC CTT ATG TTT TAC TTT TTT AGC GGG

1981
       BstXI
S   I   F   V   A   F   G   L   K   K   E   S   L   E   K   L   K   V   K   K
TCC ATT TTC GTG GCG TTT GGT TTG AAA AAG GAG AGT TTA GAA AAA TTA AAA GTC AAA AAA
AGG TAA AAG CAC CGC AAA CCA AAC TTT TTC CTC TCA AAT CTT TTT AAT TTT CAG TTT TTT

2041

A   S   L   V   I   W   T   T   T   P   W   T   L   Y   A   N   V   A   I   A
GCG AGC TTG GTG ATT TGG ACG ACC ACG CCT TGG ACT TTG TAT GCG AAT GTA GCG ATC GCT
CGC TCG AAC CAC TAA ACC TGC TGG TGC GGA ACC TGA AAC ATA CGC TTA CAT CGC TAG CGA

2101

L   K   K   D   A   V   Y   A   L   T   Q   K   G   Y   L   V   A   K   A   L
TTG AAA AAA GAC GCT GTT TAT GCG CTC ACC CAA AAA GGC TAT TTA GTC GCT AAA GCC TTG
AAC TTT TTT CTG CGA CAA ATA CGC GAG TGG GTT TTT CCG ATA AAT CAG CGA TTT CGG AAC

2161

H   E   K   L   A   A   L   G   V   V   D   N   E   I   T   H   E   F   N   S
CAT GAA AAA TTA GCC GCT TTA GGG GTG GTG GAT AAT GAG ATC ACA CAT GAA TTC AAT TCC
GTA CTT TTT AAT CGG CGA AAT CCC CAC CAC CTA TTA CTC TAG TGT GTA CTT AAG TTA AGG
```

2221

N D L E Y L V A T N P L N Q R D S L V A
AAT GAT TTA GAA TAT TTA GTG GCT ACA AAC CCG CTC AAT CAA AGG GAT TCG CTG GTG GCT
TTA CTA AAT CTT ATA AAT CAC CGA TGT TTG GGC GAG TTA GTT TCC CTA AGC GAC CAC CGA

2281

L G E H V G L E D G T G A V H T A P G H
TTA GGA GAG CAT GTC GGT TTA GAA GAT GGC ACA GGA GCC GTG CAT ACC GCA CCT GGG CAT
AAT CCT CTC GTA CAG CCA AAT CTT CTA CCG TGT CCT CGG CAC GTA TGG CGT GGA CCC GTA

2341

G E E D Y Y L G L R Y N L E V L M S V D
GGT GAA GAG GAC TAT TAT TTA GGC TTA AGA TAT AAT TTA GAA GTG TTA ATG TCT GTA GAT
CCA CTT CTC CTG ATA ATA AAT CCG AAT TCT ATA TTA AAT CTT CAC AAT TAC AGA CAT CTA

2401

E K G C Y D E G I I H N Q L L D E S Y L
GAG AAA GGT TGC TAT GAT GAG GGC ATT ATC CAT AAC CAA CTA TTA GAT GAA AGC TAT CTG
CTC TTT CCA ACG ATA CTA CTC CCG TAA TAG GTA TTG GTT GAT AAT CTA CTT TCG ATA GAC

2461

G E H V F K A Q K R I I E Q L G D S L L
GGC GAG CAT GTT TTT AAG GCT CAA AAA CGC ATT ATA GAG CAA TTG GGC GAT TCT TTA TTG
CCG CTC GTA CAA AAA TTC CGA GTT TTT GCG TAA TAT CTC GTT AAC CCG CTA AGA AAT AAC

2521

L E Q E I E H S Y P H C W R T H K P V I
CTA GAG CAA GAG ATT GAG CAT TCT TAT CCG CAT TGC TGG AGG ACG CAC AAG CCT GTG ATT
GAT CTC GTT CTC TAA CTC GTA AGA ATA GGC GTA ACG ACC TCC TGC GTG TTC GGA CAC TAA

2581

Y R A T T Q W F I L M D E P F I Q N D G
TAC AGA GCG ACT ACG CAA TGG TTT ATT TTA ATG GAT GAG CCT TTT ATC CAA AAT GAT GGC
ATG TCT CGC TGA TGC GTT ACC AAA TAA AAT TAC CTA CTC GGA AAA TAG GTT TTA CTA CCG

2641

S Q K T L R E V A L D A I E K V E F V P
TCT CAA AAA ACC TTA AGA GAA GTG GCT TTA GAT GCG ATT GAA AAG GTG GAA TTT GTG CCA
AGA GTT TTT TGG AAT TCT CTT CAC CGA AAT CTA CGC TAA CTT TTC CAC CTT AAA CAC GGT

2701

S S G K N R L K T M I E N R P D W C L S
AGC AGC GGG AAA AAC CGC CTA AAA ACC ATG ATA GAA AAC CGC CCT GAT TGG TGC TTG AGC
TCG TCG CCC TTT TTG GCG GAT TTT TGG TAC TAT CTT TTG GCG GGA CTA ACC ACG AAC TCG

2761

R Q R K W G V P L A F F I D K R T N K P
CGG CAA AGA AAA TGG GGC GTG CCA CTG GCC TTT TTC ATA GAC AAA CGC ACG AAT AAG CCT
GCC GTT TCT TTT ACC CCG CAC GGT GAC CGG AAA AAG TAT CTG TTT GCG TGC TTA TTC GGA

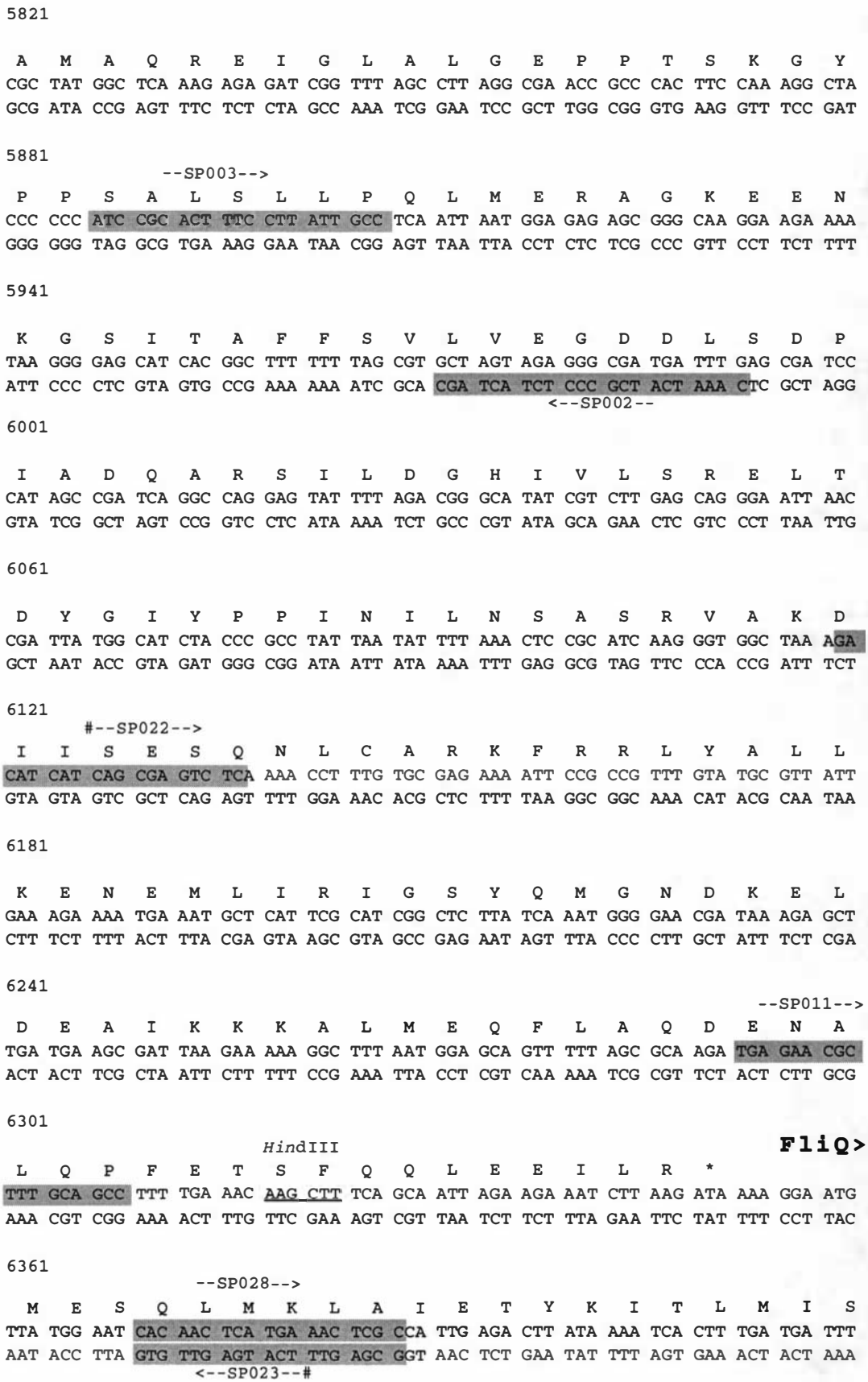
```
2821

C   F   E   S   E   V   L   E   H   V   A   N   L   F   E   K   K   G   C   D
TGT TTT GAA AGC GAA GTT TTA GAG CAT GTG GCC AAT CTT TTT GAG AAA AAA GGC TGT GAT
ACA AAA CTT TCG CTT CAA AAT CTC GTA CAC CGG TTA GAA AAA CTC TTT TTT CCG ACA CTA

2881

V   W   W   E   Y   S   V   K   D   L   L   P   P   S   Y   Q   E   D   A   K
GTG TGG TGG GAG TAT AGC GTG AAA GAT TTA TTG CCC CCT AGC TAT CAA GAG GAC GCC AAG
CAC ACC ACC CTC ATA TCG CAC TTT CTA AAT AAC GGG GGA TCG ATA GTT CTC CTG CGG TTC

2941

H   Y   E   K   I   M   H   I   L   D   V   W   F   D   S   G   S   T   F   K
CAT TAT GAG AAA ATC ATG CAC ATT TTA GAC GTG TGG TTT GAT AGT GGT AGC ACC TTT AAG
GTA ATA CTC TTT TAG TAC GTG TAA AAT CTG CAC ACC AAA CTA TCA CCA TCG TGG AAA TTC

3001

A   V   L   E   D   Y   H   G   E   K   G   Q   S   P   S   D   V   I   L   E
GCG GTT TTA GAA GAC TAT CAT GGA GAA AAA GGG CAA AGC CCT AGC GAT GTG ATC TTA GAA
CGC CAA AAT CTT CTG ATA GTA CCT CTT TTT CCC GTT TCG GGA TCG CTA CAC TAG AAT CTT

3061
       --SP034-->
G   S   D   Q   H   R   G   W   F   Q   S   S   L   L   I   G   C   V   L   N
GGG AGC GAT CAG CAT AGG GGG TGG TTT CAA AGC TCG CTT CTA ATC GGT TGT GTT TTA AAC
CCC TCG CTA GTC GTA TCC CCC ACC AAA GTT TCG AGC GAA GAT TAG CCA ACA CAA AAT TTG

3121

N   Q   A   P   F   K   K   V   I   T   H   G   F   I   V   D   E   K   G   E
AAC CAA GCC CCT TTT AAA AAG GTC ATT ACG CAT GGC TTT ATC GTA GAT GAA AAG GGC GAA
TTG GTT CGG GGA AAA TTT TTC CAG TAA TGC GTA CCG AAA TAG CAT CTA CTT TTC CCG CTT

3181

K   M   S   K   S   K   G   N   V   V   S   L   D   K   L   L   K   T   H   G
AAA ATG AGT AAA TCT AAG GGC AAT GTG GTG TCT TTG GAC AAG CTG CTC AAA ACG CAT GGG
TTT TAC TCA TTT AGA TTC CCG TTA CAC CAC AGA AAC CTG TTC GAC GAG TTT TGC GTA CCC

3241

S   D   V   V   R   L   W   V   A   F   N   D   Y   Q   N   D   L   R   V   S
AGC GAT GTG GTG CGT TTG TGG GTA GCG TTT AAT GAC TAT CAA AAC GAT TTG AGA GTC TCT
TCG CTA CAC CAC GCA AAC ACC CAT CGC AAA TTA CTG ATA GTT TTG CTA AAC TCT CAG AGA

3301

Q   T   F   F   T   Q   T   E   Q   H   Y   K   K   F   R   N   T   L   K   F
CAA ACC TTT TTC ACT CAA ACA GAA CAA CAT TAT AAA AAA TTC CGC AAC ACC CTG AAA TTC
GTT TGG AAA AAG TGA GTT TGT CTT GTT GTA ATA TTT TTT AAG GCG TTG TGG GAC TTT AAG

3361

L   L   A   N   F   S   D   M   D   L   K   N   L   E   R   P   H   N   F   S
TTA CTC GCT AAT TTT AGC GAT ATG GAT CTC AAG AAT TTA GAA CGC CCC CAT AAC TTC AGC
AAT GAG CGA TTA AAA TCG CTA TAC CTA GAG TTC TTA AAT CTT GCG GGG GTA TTG AAG TCG
```

```
3421

P   L   D   H   F   M   L   E   T   L   E   T   I   S   A   G   V   N   S   A
CCT TTA GAT CAT TTT ATG TTA GAG ACT TTA GAA ACC ATA AGC GCT GGA GTC AAT AGC GCG
GGA AAT CTA GTA AAA TAC AAT CTC TGA AAT CTT TGG TAT TCG CGA CCT CAG TTA TCG CGC

3481

F   E   E   H   D   F   V   K   G   L   N   I   L   M   A   F   V   T   N   E
TTT GAA GAG CAT GAT TTT GTG AAA GGC TTG AAT ATT TTA ATG GCG TTT GTT ACC AAT GAA
AAA CTT CTC GTA CTA AAA CAC TTT CCG AAC TTA TAA AAT TAC CGC AAA CAA TGG TTA CTT

3541

L   S   G   I   Y   L   D   A   C   K   D   S   L   Y   C   D   S   K   N   N
TTG AGC GGG ATT TAT TTA GAC GCT TGC AAG GAT AGC TTG TAT TGC GAT AGC AAA AAC AAT
AAC TCG CCC TAA ATA AAT CTG CGA ACG TTC CTA TCG AAC ATA ACG CTA TCG TTT TTG TTA

3601
                    BstXI
E   K   R   Q   A   I   Q   M   V   L   L   A   T   A   S   K   L   C   Y   F
GAA AAA CGC CAA GCC ATT CAA ATG GTT TTA CTC GCT ACA GCT AGT AAG TTG TGC TAC TTT
CTT TTT GCG GTT CGG TAA GTT TAC CAA AAT GAG CGA TGT CGA TCA TTC AAC ACG ATG AAA

3661

L   A   P   I   L   T   H   T   I   E   E   V   L   E   H   S   Q   A   L   R
TTA GCC CCG ATT TTA ACG CAC ACG ATT GAA GAA GTT TTA GAG CAT AGC CAA GCG CTT CGC
AAT CGG GGC TAA AAT TGC GTG TGC TAA CTT CTT CAA AAT CTC GTA TCG GTT CGC GAA GCG

3721

I   F   L   Q   A   K   D   V   F   D   L   K   D   I   S   V   S   E   K   L
ATT TTT TTA CAA GCC AAA GAT GTG TTT GAT TTA AAA GAC ATT AGC GTT TCA GAA AAA CTC
TAA AAA AAT GTT CGG TTT CTA CAC AAA CTA AAT TTT CTG TAA TCG CAA AGT CTT TTT GAG

3781

H   L   K   E   F   K   K   P   E   N   F   E   A   V   L   A   L   R   S   A
CAC CTC AAA GAG TTT AAA AAA CCA GAA AAT TTT GAA GCC GTT TTA GCC TTG CGT TCT GCC
GTG GAG TTT CTC AAA TTT TTT GGT CTT TTA AAA CTT CGG CAA AAT CGG AAC GCA AGA CGG

3841

F   N   E   E   L   D   R   L   K   K   E   G   V   I   K   N   S   L   E   C
TTT AAT GAA GAG TTA GAC CGA TTG AAA AAA GAA GGC GTC ATT AAA AAT TCG TTA GAG TGC
AAA TTA CTT CTC AAT CTG GCT AAC TTT TTT CTT CCG CAG TAA TTT TTA AGC AAT CTC ACG

3901

A   I   E   V   K   E   K   A   L   R   E   N   L   I   E   E   L   L   M   V
GCT ATT GAA GTA AAA GAA AAA GCG TTG CGT GAA AAT TTG ATA GAA GAG TTG CTG ATG GTA
CGA TAA CTT CAT TTT CTT TTT CGC AAC GCA CTT TTA AAC TAT CTT CTC AAC GAC TAC CAT

3961
                                                    --SP019-->
S   F   V   G   V   A   K   E   K   L   S   E   T   P   A   F   T   L   F   K
AGC TTT GTG GGG GTT GCA AAA GAA AAA TTG AGT GAA ACG CCA GCA TTC ACG CTC TTT AAA
TCG AAA CAC CCC CAA CGT TTT CTT TTT AAC TCA CTT TGC GGT CGT AAG TGC GAG AAA TTT
```

```
4021

A   P   F   Y   K   C   P   R   C   W   R   F   K   S   E   L   E   N   T   P
GCC CCC TTT TAT AAA TGC CCC AGG TGT TGG CGT TTT AAA AGC GAG CTA GAA AAC ACC CCT
CGG GGG AAA ATA TTT ACG GGG TCC ACA ACC GCA AAA TTT TCG CTC GAT CTT TTG TGG GGA

4081

                                                                VirB11>
C   K   R   C   E   E   V   L   K   E   R   *   *                   L   E   T
TGC AAG CGT TGC GAA GAG GTT TTA AAA GAG CGA TGA TAA AAG GAT AGG GCT TTT GAA AAC
ACG TTC GCA ACG CTT CTC CAA AAT TTT CTC GCT ACT ATT TTC CTA TCC CGA AAA CTT TTG

4141

 L   Q   T   H   R   V   L   Q   A   L   I   G   H   F   T   P   F   L   E   S
TTT ACA AAC CCA TAG AGT TTT ACA AGC CCT AAT CGG CCA TTT CAC CCC ATT TTT AGA AAG
AAA TGT TTG GGT ATC TCA AAA TGT TCG GGA TTA GCC GGT AAA GTG GGG TAA AAA TCT TTC

4201

 G   I   T   E   L   M   I   N   T   E   Q   E   L   W   L   Y   K   I   N   N
TGG GAT CAC CGA GTT GAT GAT CAA TAC CGA GCA GGA GCT TTG GCT TTA TAA AAT CAA TAA
ACC CTA GTG GCT CAA CTA CTA GTT ATG GCT CGT CCT CGA AAC CGA AAT ATT TTA GTT ATT

4261
                        SphI
 T   R   E   K   R   G   H   A   L   F   D   K   A   F   L   L   R   F   C   E
CAC ACG AGA AAA AAG AGG GCA TGC GCT TTT TGA TAA GGC GTT TTT GCT GAG GTT TTG CGA
GTG TGC TCT TTT TTC TCC CGT ACG CGA AAA ACT ATT CCG CAA AAA CGA CTC CAA AAC GCT
                                                            <--SP032--#
4321

 Q   L   A   S   F   R   G   L   F   F   D   E   E   H   P   T   L   N   C   S
GCA ACT AGC TAG TTT TAG GGG GTT GTT TTT TGA TGA AGA GCA CCC CAC TTT AAA TTG CTC
CGT TGA TCG ATC AAA ATC CCC CAA CAA AAA ACT ACT TCT CGT GGG GTG AAA TTT AAC GAG

4381

 I   P   F   T   R   Y   R   V   S   A   N   H   F   S   I   T   T   N   N   Q
TAT CCC TTT CAC GCG CTA TAG GGT GAG CGC GAA TCA CTT TAG CAT CAC TAC CAA TAA TCA
ATA GGG AAA GTG CGC GAT ATC CCA CTC GCG CTT AGT GAA ATC GTA GTG ATG GTT ATT AGT

4441
                            --SP015-->
 I   T   L   N   I   R   V   P   R   L   K   P   L   S   L   E   D   F   T   F
AAT CAC GCT CAA CAT CCG TGT GCC TAG GCT TAA GCC CTT AAG TTT AGA GGA TTT CAC TTT
TTA GTG CGA GTT GTA GGC ACA CGG ATC CGA ATT CGG GAA TTC AAA TCT CCT AAA GTG AAA
          <--SP016--
4501

 K   A   S   D   P   K   G   L   K   D   L   A   L   K   G   H   N   I   L   I
CAA AGC AAG CGA TCC AAA AGG TTT GAA AGA TTT AGC GCT TAA AGG GCA TAA CAT CCT CAT
GTT TCG TTC GCT AGG TTT TCC AAA CTT TCT AAA TCG CGA ATT TCC CGT ATT GTA GGA GTA

4561

 S   G   E   T   S   S   G   K   T   S   L   L   N   A   L   L   D   C   V   N
TAG CGG GGA GAC TTC AAG CGG TAA AAC AAG CCT ATT AAA CGC TCT TTT AGA TTG CGT CAA
ATC GCC CCT CTG AAG TTC GCC ATT TTG TTC GGA TAA TTT GCG AGA AAA TCT AAC GCA GTT
```

```
4621
          #--SP033-->
K   D   E   R   V   V   S   V   E   D   S   Q   E   L   D   L   K   A   F   N
TAA AGA TGA AAG GGT GGT GAG CGT TGA AGA CAG CCA AGA ATT GGA TTT AAA AGC GTT TAA
ATT TCT ACT TTC CCA CCA CTC GCA ACT TCT GTC GGT TCT TAA CCT AAA TTT TCG CAA ATT

4681

 N   C   V   G   L   L   V   G   K   Q   E   N   T   R   F   N   Y   E   D   A
TAA TTG CGT GGG GCT TTT AGT GGG CAA GCA AGA AAA CAC GCG CTT TAA TTA TGA AGA CGC
ATT AAC GCA CCC CGA AAA TCA CCC GTT CGT TCT TTT GTG CGC GAA ATT AAT ACT TCT GCG

4741

 L   N   M   A   M   R   L   N   P   D   R   L   I   V   G   E   I   D   T   R
TCT CAA TAT GGC CAT GCG CTT AAA CCC GGA CAG GCT CAT TGT GGG CGA GAT TGA CAC CAG
AGA GTT ATA CCG GTA CGC GAA TTT GGG CCT GTC CGA GTA ACA CCC GCT CTA ACT GTG GTC

4801
                                                        SphI
 N   A   A   L   F   L   R   L   G   N   T   G   H   K   G   M   L   S   T   I
GAA TGC GGC ACT CTT TTT GCG TTT AGG AAA CAC CGG GCA TAA GGG CAT GCT CTC AAC CAT
CTT ACG CCG TGA GAA AAA CGC AAA TCC TTT GTG GCC CGT ATT CCC GTA CGA GAG TTG GTA

4861
                                                            SphI
 H   A   N   S   A   Q   N   T   L   E   A   L   S   L   N   L   S   M   R   Y
TCA CGC TAA TAG CGC TCA AAA CAC TTT AGA AGC CCT TTC GCT GAA TTT GAG CAT GCG TTA
AGT GCG ATT ATC GCG AGT TTT GTG AAA TCT TCG GGA AAG CGA CTT AAA CTC GTA CGC AAT

4921

 T   H   S   L   D   K   D   L   M   R   A   Y   F   K   S   A   I   D   V   I
TAC GCA TTC TTT AGA TAA GGA TTT GAT GCG AGC GTA TTT TAA AAG CGC GAT TGA TGT GAT
ATG CGT AAG AAA TCT ATT CCT AAA CTA CGC TCG CAT AAA ATT TTC GCG CTA ACT ACA CTA
                                                          <--SP014--
4981

 V   H   V   N   R   I   N   N   E   R   Q   I   A   E   V   L   W   T   K   E
CGT GCA TGT GAA TAG AAT CAA TAA TGA GCG CCA AAT CGC TGA AGT CTT ATG GAC TAA AGA
GCA CGT ACA CTT ATC TTA GTT ATT ACT CGC GGT TTA GCG ACT TCA GAA TAC CTG ATT TCT

5041
  FliI>
 L   *   M   P   L   K   S   L   K   N   R   L   N   Q   H   F   D   L   S   P
GCT TTA AAT GCC CCT AAA ATC CTT AAA AAA CCG CTT GAA TCA GCA TTT TGA TCT ATC GCC
CGA AAT TTA CGG GGA TTT TAG GAA TTT TTT GGC GAA CTT AGT CGT AAA ACT AGA TAG CGG

5101

 R   Y   G   S   V   K   K   I   M   P   N   I   V   Y   A   D   G   F   N   P
TCG CTA CGG GAG CGT GAA AAA AAT CAT GCC CAA TAT CGT TTA TGC GGA TGG TTT TAA CCC
AGC GAT GCC CTC GCA CTT TTT TTA GTA CGG GTT ATA GCA AAT ACG CCT ACC AAA ATT GGG

5161

 S   V   G   D   V   V   K   I   E   K   S   D   G   S   E   C   V   G   M   V
CTC TGT GGG CGA TGT GGT GAA GAT TGA AAA AAG CGA TGG CAG CGA ATG CGT GGG AAT GGT
GAG ACA CCC GCT ACA CCA CTT CTA ACT TTT TTC GCT ACC GTC GCT TAC GCA CCC TTA CCA
```

```
5221

 V   V   A   E   K   E   Q   F   G   F   T   P   F   N   F   I   E   G   A   R
GGT GGT GGC AGA AAA AGA GCA GTT TGG TTT CAC GCC CTT TAA CTT TAT AGA GGG GGC TAG
CCA CCA CCG TCT TTT TCT CGT CAA ACC AAA GTG CGG GAA ATT GAA ATA TCT CCC CCG ATC

5281

 A   G   D   K   V   L   F   L   K   E   G   L   N   F   P   V   G   R   N   L
GGC TGG CGA TAA GGT GCT GTT TTT AAA AGA GGG GTT GAA TTT CCC TGT GGG CCG TAA TCT
CCG ACC GCT ATT CCA CGA CAA AAA TTT TCT CCC CAA CTT AAA GGG ACA CCC GGC ATT AGA

5341
                  --SP013-->
 L   G   R   V   L   N   P   L   G   Q   V   I   D   N   K   G   A   L   D   Y
TTT AGG GAG GGT GCT TAA CCC TTT AGG GCA AGT CAT TGA CAA TAA GGG GGC ATT GGA TTA
AAA TCC CTC CCA CGA ATT GGG AAA TCC CGT TCA GTA ACT GTT ATT CCC CCG TAA CCT AAT

5401

 E   R   L   A   P   V   I   T   T   P   I   A   P   L   K   R   G   L   I   D
TGA GCG ATT AGC GCC TGT CAT TAC AAC GCC TAT AGC CCC TTT AAA AAG AGG CTT GAT TGA
ACT CGC TAA TCG CGG ACA GTA ATG TTG CGG ATA TCG GGG AAA TTT TTC TCC GAA CTA ACT
        <--SP021--#
5461

 E   V   F   S   V   G   V   K   S   I   D   G   L   L   T   C   G   K   G   Q
TGA GGT TTT TAG CGT GGG GGT GAA GAG CAT TGA TGG GCT TTT GAC TTG CGG TAA GGG GCA
ACT CCA AAA ATC GCA CCC CCA CTT CTC GTA ACT ACC CGA AAA CTG AAC GCC ATT CCC CGT
                              <--SP009--
5521
                                                    BstXI
 K   L   G   I   F   A   G   S   G   V   G   K   S   T   L   M   G   M   I   T
AAA ACT GGG CAT TTT TGC CGG CTC TGG GGT GGG TAA ATC CAC GCT AAT GGG CAT GAT CAC
TTT TGA CCC GTA AAA ACG GCC GAG ACC CCA CCC ATT TAG GTG CGA TTA CCC GTA CTA GTG

5581

 R   G   C   L   A   P   I   K   V   I   A   L   I   G   E   R   G   R   E   I
TAG GGG TTG CTT AGC GCC CAT TAA AGT GAT CGC TTT GAT TGG GGA AAG GGG CAG AGA AAT
ATC CCC AAC GAA TCG CGG GTA ATT TCA CTA GCG AAA CTA ACC CCT TTC CCC GTC TCT TTA

5641

 P   E   F   I   E   K   N   L   K   G   D   L   S   S   C   V   L   V   V   A
CCC TGA ATT TAT AGA GAA AAA CCT GAA AGG GGA TTT AAG CTC TTG CGT GTT GGT GGT CGC
GGG ACT TAA ATA TCT CTT TTT GGA CTT TCC CCT AAA TTC GAG AAC GCA CAA CCA CCA GCG

5701

 T   S   D   D   S   P   L   M   R   K   Y   G   A   F   C   A   M   S   V   A
TAC GAG CGA TGA TAG CCC TTT GAT GCG CAA ATA CGG GGC CTT TTG CGC GAT GAG CGT GGC
ATG CTC GCT ACT ATC GGG AAA CTA CGC GTT TAT GCC CCG GAA AAC GCG CTA CTC GCA CCG

5761

 E   Y   F   K   N   Q   G   L   D   V   L   F   I   M   D   S   V   T   R   F
GGA GTA TTT TAA AAA CCA AGG GCT AGA TGT GTT ATT CAT CAT GGA TTC AGT GAC TCG TTT
CCT CAT AAA ATT TTT GGT TCC CGA TCT ACA CAA TAA GTA GTA CCT AAG TCA CTG AGC AAA
```


5821
A M A Q R E I G L A L G E P P T S K G Y
CGC TAT GGC TCA AAG AGA GAT CGG TTT AGC CTT AGG CGA ACC GCC CAC TTC CAA AGG CTA
GCG ATA CCG AGT TTC TCT CTA GCC AAA TCG GAA TCC GCT TGG CGG GTG AAG GTT TCC GAT
5881
--SP003-->
P P S A L S L L P Q L M E R A G K E E N
CCC CCC ATC CGC ACT TTC CTT ATT GCC TCA ATT AAT GGA GAG AGC GGG CAA GGA AGA AAA
GGG GGG TAG GCG TGA AAG GAA TAA CGG AGT TAA TTA CCT CTC TCG CCC GTT CCT TCT TTT
5941
K G S I T A F F S V L V E G D D L S D P
TAA GGG GAG CAT CAC GGC TTT TTT TAG CGT GCT AGT AGA GGG CGA TGA TTT GAG CGA TCC
ATT CCC CTC GTA GTG CCG AAA AAA ATC GCA CGA TCA TCT CCC GCT ACT AAA CTC GCT AGG
<--SP002--
6001
I A D Q A R S I L D G H I V L S R E L T
CAT AGC CGA TCA GGC CAG GAG TAT TTT AGA CGG GCA TAT CGT CTT GAG CAG GGA ATT AAC
GTA TCG GCT AGT CCG GTC CTC ATA AAA TCT GCC CGT ATA GCA GAA CTC GTC CCT TAA TTG
6061
D Y G I Y P P I N I L N S A S R V A K D
CGA TTA TGG CAT CTA CCC GCC TAT TAA TAT TTT AAA CTC CGC ATC AAG GGT GGC TAA AGA
GCT AAT ACC GTA GAT GGG CGG ATA ATT ATA AAA TTT GAG GCG TAG TTC CCA CCG ATT TCT
6121
#--SP022-->
I I S E S Q N L C A R K F R R L Y A L L
CAT CAT CAG CGA GTC TCA AAA CCT TTG TGC GAG AAA ATT CCG CCG TTT GTA TGC GTT ATT
GTA GTA GTC GCT CAG AGT TTT GGA AAC ACG CTC TTT TAA GGC GGC AAA CAT ACG CAA TAA
6181
K E N E M L I R I G S Y Q M G N D K E L
GAA AGA AAA TGA AAT GCT CAT TCG CAT CGG CTC TTA TCA AAT GGG GAA CGA TAA AGA GCT
CTT TCT TTT ACT TTA CGA GTA AGC GTA GCC GAG AAT AGT TTA CCC CTT GCT ATT TCT CGA
6241
--SP011-->
D E A I K K K A L M E Q F L A Q D E N A
TGA TGA AGC GAT TAA GAA AAA GGC TTT AAT GGA GCA GTT TTT AGC GCA AGA TGA GAA CGC
ACT ACT TCG CTA ATT CTT TTT CCG AAA TTA CCT CGT CAA AAA TCG CGT TCT ACT CTT GCG
6301
HindIII
FliQ>
L Q P F E T S F Q Q L E E I L R *
TTT GCA GCC TTT TGA AAC AAG CTT TCA GCA ATT AGA AGA AAT CTT AAG ATA AAA GGA ATG
AAA CGT CGG AAA ACT TTG TTC GAA AGT CGT TAA TCT TCT TTA GAA TTC TAT TTT CCT TAC
6361
--SP028-->
M E S Q L M K L A I E T Y K I T L M I S
TTA TGG AAT CAC AAC TCA TGA AAC TCG CCA TTG AGA CTT ATA AAA TCA CTT TGA TGA TTT
AAT ACC TTA GTG TTG AGT ACT TTG AGC GGT AAC TCT GAA TAT TTT AGT GAA ACT ACT AAA
<--SP023--#

```
6421
          #--SP024-->
  L   P   V   L   L   A   G   L   V   V   G   L   L   V   S   I   F   Q   A   T
CTT TAC CGG TAT TAC TAG CGG GCT TAG TGG TGG GGC TAT TAG TCA GTA TTT TTC AAG CGA
GAA ATG GCC ATA ATG ATC GCC CGA ATC ACC ACC CCG ATA ATC AGT CAT AAA AAG TTC GCT

6481

  T   Q   I   N   E   M   T   L   S   F   V   P   K   I   L   A   V   I   G   V
CCA CTC AAA TCA ATG AAA TGA CCT TGT CTT TTG TGC CTA AGA TTT TAG CCG TGA TTG GGG
GGT GAG TTT AGT TAC TTT ACT GGA ACA GAA AAC ACG GAT TCT AAA ATC GGC ACT AAC CCC

6541

  L   I   L   T   M   P   W   M   T   N   M   L   L   D   Y   T   K   T   L   I
TGC TGA TTT TAA CCA TGC CGT GGA TGA CGA ACA TGC TTT TAG ATT ACA CCA AAA CCT TAA
ACG ACT AAA ATT GGT ACG GCA CCT ACT GCT TGT ACG AAA ATC TAA TGT GGT TTT GGA AT

6601
                                    MurB>
  K   L   I   P   K   I   I   G   *       M   L   E   T   I   I   D   F   S   R
TCA AGC TCA TTC CCA AAA TCA TCG GCT AGA AAA TGC TAG AAA CCA TTA TTG ATT TTT CTC
AGT TCG AGT AAG GGT TTT AGT AGC CGA TCT TTT ACG ATC TTT GGT AAT AAC TAA AAA GAG
  <--SP027--
6661

  Y   S   S   V   K   I   G   T   P   L   E   V   S   V   L   E   N   D   D   E
GTT ACA GCA GCG TGA AAA TCG GCA CGC CTT TAG AAG TGA GCG TTT TAG AAA ACG ATG ATG
CAA TGT CGT CGC ACT TTT AGC CGT GCG GAA ATC TTC ACT CGC AAA ATC TTT TGC TAC TAC
            <--SP010--
6721

  I   S   Q   E   H   Q   I   I   G   L   A   N   N   L   L   I   A   P   S   V
AAA TCT CTC AAG AAC ACC AGA TCA TAG GCT TAG CGA ACA ACC TTT TAA TCG CTC CTA GCG
TTT AGA GAG TTC TTG TGG TCT AGT ATC CGA ATC GCT TGT TGG AAA ATT AGC GAG GAT CGC

6781

  K   N   L   A   L   L   G   K   N   Y   D   Y   I   C   D   Q   G   E   W   V
TGA AAA ATC TCG CTT TAT TAG GAA AAA ACT ACG ATT ATA TTT GCG ATC AGG GTG AGT GCG
ACT TTT TAG AGC GAA ATA ATC CTT TTT TGA TGC TAA TAT AAA CGC TAG TCC CAC TCA CGC

6841
                                                                #--SP035-->
  E   V   G   G   A   A   N   A   S   K   I   F   N   Y   F   R   A   N   D   L
TGG AGG TAG GGG GAG CGG CCA ATG CGT CTA AAA TTT TTA ATT ATT TTA GGG CGA ATG ATT
ACC TCC ATC CCC CTC GCC GGT TAC GCA GAT TTT AAA AAT TAA TAA AAT CCC GCT TAC TAA

6901

  E   G   L   E   F   L   G   Q   L   P   G   T   L   G   A   L   V   K   M   N
TAG AGG GTT TGG AAT TTT TAG GGC AAT TGC CTG GCA CTT TAG GAG CGT TAG TTA AAA TGA
ATC TCC CAA ACC TTA AAA ATC CCG TTA ACG GAC CGT GAA ATC CTC GCA ATC AAT TTT ACT

6961

  A   G   M   K   E   F   E   I   K   N   V   L   E   S   A   C   I   N   N   E
ATG CCG GCA TGA AAG AAT TTG AAA TCA AAA ATG TTT TAG AAA GCG CTT GCA TTA ATA ATG
TAC GGC CGT ACT TTC TTA AAC TTT AGT TTT TAC AAA ATC TTT CGC GAA CGT AAT TAT TAC
```

7021

*Hin*dIII

W L E K E A >

AAT GGC TAG AAA AAG AAG CTT

TTA CCG ATC TTT TTC TTC GAA

References

Abrahams, J. P., Leslie, A. G., Lutter, R. & Walker, J. E. (1994). Structure at 2.8 A resolution of F_1-ATPase from bovine heart mitochondria. *Nature* 370(6491): 621-628.

Aizawa, S. I. (1996). Flagellar assembly in *Salmonella typhimurium*. *Mol Microbiol* 19(1): 1-5.

Akopyants, N. S., Clifton, S. W., Kersulyte, D., Crabtree, J. E., Youree, B. E., Reece, C. A., Bukanov, N. O., Drazek, E. S., Roe, B. A. & Berg, D. E. (1998). Analyses of the *cag* pathogenicity island of *Helicobacter pylori*. *Mol Microbiol* 28: 37-53.

Albertini, A. M., Caramori, T., Crabb, W. D., Scoffone, F. & Galizzi, A. (1991). The *flaA* locus of *Bacillus subtilis* is part of a large operon coding for flagellar structures, motility functions, and an ATPase-like polypeptide. *J Bacteriol* 173(11): 3573-3579.

Allaoui, A., Schulte, R. & Cornelis, G. R. (1995). Mutational analysis of the *Yersinia enterocolitica virC* operon: characterization of *yscE, F, G, I, J, K* required for Yop secretion and *yscH* encoding YopR. *Mol Microbiol* 18(2): 343-355.

Alm, R. A., Guerry, P. & Trust, T. J. (1993). Significance of duplicated flagellin genes in *Campylobacter*. *J Mol Biol* 230: 359-363.

Alm, R. A., Ling, L. S., Moir, D. T., King, B. L., Brown, E. D., Doig, P. C., Smith, D. R., Noonan, B., Guild, B. C., deJonge, B. L., Carmel, G., Tummino, P. J., Caruso, A., Uria-Nickelsen, M., Mills, D. M., Ives, C., Gibson, R., Merberg, D., Mills, S. D., Jiang, Q., Taylor, D. E., Vovis, G. F. & Trust, T. J. (1999). Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori*. *Nature* 397(6715): 176-180.

Alt-Morbe, J., Stryker, J. L., Fuqua, C., Li, P. L., Farrand, S. K. & Winans, S. C. (1996). The conjugal transfer system of *Agrobacterium tumefaciens* octopine-type Ti plasmids is closely related to the transfer system of an IncP plasmid and distantly related to Ti plasmid *vir* genes. *J Bacteriol* 178(14): 4248-57.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). Basic local alignment search tool. *J Mol Biol* 215(3): 403-10.

Anderson, D. K. & Newton, A. (1997). Posttranscriptional regulation of *Caulobacter* flagellin genes by a late flagellum assembly checkpoint. *J Bacteriol* 179(7): 2281-2288.

Anderson, D. K., Ohta, N., Wu, J. & Newton, A. (1995). Regulation of the *Caulobacter crescentus rpoN* gene and function of the purified sigma 54 in flagellar gene transcription. *Mol Gen Genet* 246(6): 697-706.

Anderson, D. M. & Schneewind, O. (1997). A mRNA signal for the type III secretion of Yop proteins by *Yersinia enterocolitica*. *Science* 278(5340): 1140-1143.

Andrews, G. P., Hromockyj, A. E., Coker, C. & Maurelli, A. T. (1991). Two novel virulence loci, *mxiA* and *mxiB*, in *Shigella flexneri* 2a facilitate excretion of invasion plasmid antigens. *Infect Immun* 59(6): 1997-2005.

Andrutis, K. A., Fox, J. G., Schauer, D. B., Marini, R. P., Li, X. T., Yan, L. L., Josenhans, C. & Suerbaum, S. (1997). Infection of the ferret stomach by isogenic flagellar mutant strains of *Helicobacter mustelae*. *Infect Immun* 65(5): 1962-1966.

Appelmelk, B. J., Shiberu, B., Trinks, C., Tapsi, N., Zheng, P. Y., Verboom, T., Maaskant, J., Hokke, C. H., Schiphorst, W. E., Blanchard, D., Simoons-Smit, I. M., van den Eijnden, D. H. & Vandenbroucke-Grauls, C. M. (1998). Phase variation in *Helicobacter pylori* lipopolysaccharide. *Infect Immun* 66(1): 70-76.

Appelmelk, B. J., Simoons-Smit, I., Negrini, R., Moran, A. P., Aspinall, G. O., Forte, J. G., De Vries, T., Quan, H., Verboom, T., Maaskant, J. J., Ghiara, P., Kuipers, E. J., Bloemena, E., Tadema, T. M., Townsend, R. R., Tyagarajan, K., Crothers, J. M., Jr., Monteiro, M. A., Savio, A. & De Graaff, J. (1996). Potential role of molecular mimicry between *Helicobacter pylori* lipopolysaccharide and host Lewis blood group antigens in autoimmunity. *Infect Immun* 64(6): 2031-2040.

Aspinall, G. O. & Monteiro, M. A. (1996). Lipopolysaccharides of *Helicobacter pylori* strains P466 and MO19: structures of the O antigen and core oligosaccharide regions. *Biochemistry* 35(7): 2498-2504.

Aspinall, G. O., Monteiro, M. A., Pang, H., Walsh, E. J. & Moran, A. P. (1996). Lipopolysaccharide of the *Helicobacter pylori* type strain NCTC 11637 (ATCC 43504): structure of the O antigen chain and core oligosaccharide regions. *Biochemistry* 35(7): 2489-2497.

Atherton, J. C., Cao, P., Peek, R. M., Jr., Tummuru, M. K., Blaser, M. J. & Cover, T. L. (1995). Mosaicism in vacuolating cytotoxin alleles of *Helicobacter pylori*. Association of specific *vacA* types with cytotoxin production and peptic ulceration. *J Biol Chem* 270(30): 17771-17777.

Axon, A. T., O'Morain, C. A., Bardhan, K. D., Crowe, J. P., Beattie, A. D., Thompson, R. P., Smith, P. M., Hollanders, F. D., Baron, J. H., Lynch, D. A., Dixon, M. F., Tompkins, D. S., Birrell, H. & Gillon, K. R. (1997). Randomised double blind controlled study of recurrence of gastric ulcer after treatment for eradication of *Helicobacter pylori* infection. *Br Med J* 314(7080): 565-568.

Axon, A. T. R. (1995). Is *Helicobacter pylori* transmitted by the gastro-oral route. *Aliment Pharmacol Ther* 9(6): 585-588.

Ayora, S., Rojo, F., Ogasawara, N., Nakai, S. & Alonso, J. C. (1996). The Mfd protein of *Bacillus subtilis* 168 is involved in both transcription-coupled DNA repair and DNA recombination. *J Mol Biol* 256(2): 301-318.

Azuma, T., Konishi, J., Tanaka, Y., Hirai, M., Ito, S., Kato, T. & Kohli, Y. (1994). Contribution of HLA-DQA gene to host's response against *Helicobacter pylori*. *Lancet* 343(8896): 542-543.

Bahari, H. M., Ross, I. N. & Turnberg, L. A. (1982). Demonstration of a pH gradient across the mucus layer on the surface of human gastric mucosa *in vitro*. *Gut* 23(6): 513-516.

Ballado, T., Campos, A., Camarena, L. & Dreyfus, G. (1996). Flagellar genes from *Rhodobacter sphaeroides* are homologous to genes of the *fliF* operon of *Salmonella typhimurium* and to the type-III secretion system. *Gene* 170(1): 69-72.

Bartlett, D. H., Frantz, B. B. & Matsumura, P. (1988). Flagellar transcriptional activators FlbB and FlaI: gene sequences and 5' consensus sequences of operons under FlbB and FlaI control. *J Bacteriol* 170(4): 1575-1581.

Bauerfeind, P., Garner, R. M. & Mobley, L. T. (1996). Allelic exchange mutagenesis of *nixA* in *Helicobacter pylori* results in reduced nickel transport and urease activity. *Infect Immun* 64(7): 2877-2880.

Beddek, A. J. (1998). The genetic characterization of a carbon starvation gene of *Helicobacter pylori*. MSc Thesis, Massey University.

Benaissa, M., Babin, P., Quellard, N., Pezennec, L., Cenatiempo, Y. & Fauchere, J. L. (1996). Changes in *Helicobacter pylori* ultrastructure and antigens during conversion from the bacillary to the coccoid form. *Infect Immun* 64(6): 2331-2335.

Benjelloun-Touimi, Z., Sansonetti, P. J. & Parsot, C. (1995). SepA, the major extracellular protein of *Shigella flexneri*: autonomous secretion and involvement in tissue invasion. *Mol Microbiol* 17(1): 123-135.

Benson, A. K., Ramakrishnan, G., Ohta, N., Feng, J., Ninfa, A. J. & Newton, A. (1994). The *Caulobacter crescentus* FlbD protein acts at *ftr* sequence elements both to activate and to repress transcription of cell cycle-regulated flagellar genes. *Proc Natl Acad Sci U S A* 91(11): 4989-4993.

Berg, D. E., Hoffman, P. S., Appelmelk, B. J. & Kusters, J. G. (1997). The *Helicobacter pylori* genome sequence: genetic factors for long life in the gastric mucosa. *Trends Microbiol* 5(12): 468-474.

Berger, B. R. & Christie, P. J. (1994). Genetic complementation analysis of the *Agrobacterium tumefaciens virB* operon: *virB2* through *virB11* are essential virulence genes. *J Bacteriol* 176(12): 3646-3660.

Berghammer, H. & Auer, B. (1993). "Easypreps": fast and easy plasmid minipreparation for analysis of recombinant clones in *E. coli*. *Biotechniques* 14(4): 524, 528.

Binet, R., Letoffe, S., Ghigo, J. M., Delepelaire, P. & Wandersman, C. (1997). Protein secretion by Gram-negative bacterial ABC exporters - a review. *Gene* 192(1): 7-11.

Bischoff, D. S., Weinreich, M. D. & Ordal, G. W. (1992). Nucleotide sequences of *Bacillus subtilis* flagellar biosynthetic genes *fliP* and *fliQ* and identification of a novel flagellar gene, *fliZ*. *J Bacteriol* 174(12): 4017-4025.

Blair, D. F. & Berg, H. C. (1988). Restoration of torque in defective flagellar motors. *Science* 242(4886): 1678-1681.

Blake, M. S., Johnston, K. H., Russell-Jones, G. J. & Gotschlich, E. C. (1984). A rapid, sensitive method for detection of alkaline phosphatase-conjugated anti-antibody on western blots. *Anal Biochem* 136: 175-179.

Blaser, M. J. (1993). *Helicobacter pylori*: microbiology of a 'slow' bacterial infection. *Trends Microbiol* 1(7): 255-260.

Blaser, M. J. (1996). The bacteria behind ulcers. *Scientific American* 2: 92-97.

Blaser, M. J. & Parsonnet, J. (1994). Parasitism by the "slow" bacterium *Helicobacter pylori* leads to altered gastric homeostasis and neoplasia. *J Clin Invest* 94(1): 4-8.

Blattner, F. R., Plunkett, G. R., Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K., Mayhew, G. F., Gregor, J., Davis, N. W., Kirkpatrick, H. A., Goeden, M. A., Rose, D. J., Mau, B. & Shao, Y. (1997). The complete genome sequence of *Escherichia coli* K-12. *Science* 277(5331): 1453-1474.

Boren, T., Falk, P., Roth, K. A., Larson, G. & Normark, S. (1993). Attachment of *Helicobacter pylori* to human gastric epithelium mediated by blood group antigens. *Science* 262: 1892-1895.

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 72: 248-254.

Bukanov, N. O. & Berg, D. E. (1994). Ordered cosmid library and high resolution physical-genetic map of *Helicobacter pylori* strain NCTC 11638. *Mol Microbiol* 11: 509-523.

Campbell, S., Fraser, A., Holliss, B., Schmid, J. & O'Toole, P. W. (1997). Evidence for ethnic tropism of *Helicobacter pylori*. *Infect Immun* 65(9): 3708-3712.

Carpenter, P. B. & Ordal, G. W. (1993). *Bacillus subtilis* FlhA: a flagellar protein related to a new family of signal-transducing receptors. *Mol Microbiol* 7(5): 735-743.

Censini, S., Lange, C., Xiang, Z. Y., Crabtree, J. E., Ghiara, P., Borodovsky, M., Rappuoli, R. & Covacci, A. (1996). *Cag*, a pathogenicity island of *Helicobacter pylori*, encodes type I-specific and disease-associated virulence factors. *Proc Natl Acad Sci U S A* 93(25): 14648-14653.

Chan, W. Y., Hui, P. K., Leung, K. M., Chow, J., Kwok, F. & Ng, C. S. (1994). Coccoid forms of *Helicobacter pylori* in the human stomach. *Am J Clin Pathol* 102(4): 503-507.

Charkowski, A. O., Huang, H. C. & Collmer, A. (1997). Altered localization of HrpZ in *Pseudomonas syringae* pv. syringae *hrp* mutants suggests that different components of the type III secretion pathway control protein translocation across the inner and outer membranes of gram-negative bacteria. *J Bacteriol* 179(12): 3866-3874.

Chen, M. H., Lee, A. & Hazell, S. (1992). Immunisation against gastric *Helicobacter* infection in a mouse/*Helicobacter felis* model. *Lancet* 339: 1120-1121.

Christie, P. J. (1997a). *Agrobacterium tumefaciens* T-complex transport apparatus: a paradigm for a new family of multifunctional transporters in eubacteria. *J Bacteriol* 179(10): 3085-3094.

Christie, P. J. (1997b). The *cag* pathogenicity island: mechanistic insights. *Trends Microbiol* 5(7): 264-265.

Christie, P. J., Ward, J. E., Jr., Gordon, M. P. & Nester, E. W. (1989). A gene required for transfer of T-DNA to plants encodes an ATPase with autophosphorylating activity. *Proc Natl Acad Sci U S A* 86(24): 9677-9681.

Cover, T. L., Glupczynski, Y., Lage, A. P., Burette, A., Tummuru, M. K., Perez-Perez, G. I. & Blaser, M. J. (1995). Serologic detection of infection with *cagA*$^+$ *Helicobacter pylori* strains. *J Clin Microbiol* 33(6): 1496-1500.

Cover, T. L., Tummuru, M. K. R., Cao, P., Thompson, S. A. & Blaser, M. J. (1994). Divergence of genetic sequences for the vacuolating cytotoxin among *Helicobacter pylori* strains. *J Biol Chem* 269(14): 10566-10573.

Crabtree, J. E., Farmery, S. M., Lindley, I. J., Figura, N., Peichl, P. & Tompkins, D. S. (1994). CagA/cytotoxic strains of *Helicobacter pylori* and interleukin-8 in gastric epithelial cell lines. *J Clin Pathol* 47(10): 945-950.

Crabtree, J. E., Shallcross, T. M., Heatley, R. V. & Wyatt, J. I. (1991). Mucosal tumour necrosis factor alpha and interleukin-6 in patients with *Helicobacter pylori* associated gastritis. *Gut* 32(12): 1473-1477.

Cuenca, R., Blanchard, T. G., Czinn, S. J., Nedrud, J. G., Monath, T. P., Lee, C. K. & Redline, R. W. (1996). Therapeutic immunization against *Helicobacter mustelae* in naturally infected ferrets. *Gastroenterology* 110(6): 1770-1775.

D'Elios, M. M., Manghetti, M., Almerigogna, F., Amedei, A., Costa, F., Burroni, D., Baldari, C. T., Romagnani, S., Telford, J. L. & Del Prete, G. (1997). Different cytokine profile and antigen-specificity repertoire in *Helicobacter pylori*-specific T cell clones from the antrum of chronic gastritis patients with or without peptic ulcer. *Eur J Immunol* 27(7): 1751-1755.

Dingwall, A., Zhuang, W. Y., Quon, K. & Shapiro, L. (1992). Expression of an early gene in the flagellar regulatory hierarchy is sensitive to an interruption in DNA replication. *J Bacteriol* 174(6): 1760-1768.

Doolittle, R. F. (1997). A bug with excess gastric avidity. *Nature* 388(6642): 515-516.

Dorrell, N., Winzeler, E., Stabler, R. A., Allan, E., Foynes, S., Davis, R. W. & Wren, B. W. (1998). Simultaneous analysis of *Helicobacter pylori* mutants using unique DNA tags and the Affymetrix biochip. *Gut* 43 (Suppl 2), A13, 02/33.

Dreyfus, G., Williams, A. W., Kawagishi, I. & Macnab, R. M. (1993). Genetic and biochemical analysis of *Salmonella typhimurium* FliI, a flagellar protein related to the catalytic subunit of the F_0F_1 ATPase and to virulence proteins of mammalian and plant pathogens. *J Bacteriol* 175(10): 3131-3138.

Dubois, A., Fiala, N., Heman-Ackah, L. M., Drazek, E. S., Tarnawski, A., Fishbein, W. N., Perez-Perez, G. I. & Blaser, M. J. (1994). Natural gastric infection with *Helicobacter pylori* in monkeys: a model for spiral bacteria infection in humans. *Gastroenterology* 106(6): 1405-1417.

Dufour, A., Furness, R. B. & Hughes, C. (1998). Novel genes that upregulate the *Proteus mirabilis flhDC* master operon controlling flagellar biogenesis and swarming. *Mol Microbiol* 29(3): 741-751.

Dunn, B. E., Cohen, H. & Blaser, M. J. (1997). *Helicobacter pylori. Clin Microbiol Rev* 10(4): 720-741.

Eaton, K. A., Brooks, C. L., Morgan, D. R. & Krakowka, S. (1991). Essential role of urease in pathogenesis of gastritis induced by *Helicobacter pylori* in gnotobiotic piglets. *Infect. Immun.* 59: 2470-2475.

Eaton, K. A., Cover, T. L., Tummuru, M. K., Blaser, M. J. & Krakowka, S. (1997). Role of vacuolating cytotoxin in gastritis due to *Helicobacter pylori* in gnotobiotic piglets. *Infect Immun* 65(8): 3462-3464.

Eaton, K. A. & Krakowka, S. (1994). Effect of gastric pH on urease-dependent colonization of gnotobiotic piglets by *Helicobacter pylori. Infect Immun* 62(9): 3604-3607.

Eaton, K. A., Morgan, D. R. & Krakowka, S. (1989). *Campylobacter pylori* virulence factors in gnotobiotic piglets. *Infect Immun* 57(4): 1119-1125.

Eaton, K.A., Morgan, D. R. & Krakowka, S. (1992). Motility as a factor in the colonisation of gnotobiotic piglets by *Helicobacter pylori. J Med Microbiol* 37: 123-127.

Eaton, K. A., Suerbaum, S., Josenhans, C. & Krakowka, S. (1996). Colonization of gnotobiotic piglets by *Helicobacter pylori* deficient in two flagellin genes. *Infect Immun* 64(7): 2445-2448.

Evans, D. G., Karjalainen, T. K., Evans, D. J., Graham, D. Y. & Lee, C. H. (1993). Cloning, nucleotide sequence and expression of a gene encoding an adhesin subunit protein of *Helicobacter pylori*. *J Bacteriol* 175: 674-683.

Exner, M. M., Doig, P., Trust, T. J. & Hancock, R. E. (1995). Isolation and characterization of a family of porin proteins from *Helicobacter pylori*. *Infect Immun* 63(4): 1567-1572.

Fahrner, K. A., Block, S. M., Krishnaswamy, S., Parkinson, J. S. & Berg, H. C. (1994). A mutant hook-associated protein (HAP3) facilitates torsionally induced transformations of the flagellar filament of *Escherichia coli*. *J Mol Biol* 238(2): 173-186.

Fan, F. & Macnab, R. M. (1996). Enzymatic Characterization Of FliI - an ATPase involved in flagellar assembly in *Salmonella typhimurium*. *J Biol Chem* 271(50): 31981-31988.

Fan, F., Ohnishi, K., Francis, N. R. & Macnab, R. M. (1997). The FliP and FliR proteins of *Salmonella typhimurium*, putative components of the type III flagellar export apparatus, are located in the flagellar basal body. *Mol Microbiol* 26(5): 1035-1046.

Fath, M. J. & Kolter, R. (1993). ABC transporters: bacterial exporters. *Microbiol Rev* 57(4): 995-1017.

Feldman, R. A., Eccersley, J. P. & Hardie, J. M. (1997). Transmission of *Helicobacter pylori*. *Curr Opin Gastroenterol* 13 (Suppl 1): 8-12.

Fenselau, S., Balbo, I. & Bonas, U. (1992). Determinants of pathogenicity in *Xanthomonas campestris* pv. vesicatoria are related to proteins involved in secretion in bacterial pathogens of animals. *Mol Plant Microbe Interact* 5(5): 390-396.

Ferrero, R. L., Thiberge, J. M., Huerre, M. & Labigne, A. (1994). Recombinant antigens prepared from the urease subunits of *Helicobacter* spp.: evidence of protection in a mouse model of gastric infection. *Infect Immun* 62(11): 4981-4989.

Ferrero, R. L., Thiberge, J. M., Kansau, I., Wuscher, N., Huerre, M. & Labigne, A. (1995). The GroES homolog of *Helicobacter pylori* confers protective immunity against mucosal infection in mice. *Proc Natl Acad Sci U S A* 92(14): 6499-6503.

Figura, N. (1996). Mouth-to-mouth resuscitation and *Helicobacter pylori* infection. *Lancet* 347(9011): 1342.

Finlay, B. B. & Falkow, S. (1997). Common themes in microbial pathogenicity revisited. *Microbiol Mol Biol Rev* 61(2): 136-169.

Fox, J. G., Correa, P., Taylor, N. S., Lee, A., Otto, G., Murphy, J. C. & Rose, R. (1990). *Helicobacter mustelae*-associated gastritis in ferrets. An animal model of *Helicobacter pylori* gastritis in humans. *Gastroenterology* 99(2): 352-361.

Francis, N. R., Irikura, V. M., Yamaguchi, S., DeRosier, D. J. & Macnab, R. M. (1992). Localization of the *Salmonella typhimurium* flagellar switch protein FliG to the cytoplasmic M-ring face of the basal body. *Proc Natl Acad Sci U S A* 89(14): 6304-6308.

Freedburg, A. S. (1940). The presence of spirochaetes in human gastric mucosa. *Am J Dig Dis* 7: 443-445.

Galan, J. E. (1996). Molecular genetic basis of *Salmonella* entry into host cells. *Mol Microbiol* 20(2): 263-271.

Galan, J. E., Ginocchio, C. & Costeas, P. (1992). Molecular and functional characterization of the *Salmonella* invasion gene *invA*: homology of InvA to members of a new protein family. *J Bacteriol* 174(13): 4338-4349.

Ge, Y. G. & Charon, N. W. (1997). Identification of a large motility operon in *Borrelia burgdorferi* by semi-random PCR chromosome walking. *Gene* 189(2): 195-201.

Ge, Y. G., Old, I., Saintgirons, I., Yelton, D. B. & Charon, N. W. (1996). FliH and FliI of *Borrelia burgdorferi* are similar to flagellar and virulence factor export proteins of other bacteria. *Gene* 168(1): 73-75.

Ge, Y. G., Old, I. G., Girons, I. S. & Charon, N. W. (1997a). Molecular characterization of a large *Borrelia burgdorferi* motility operon which is initiated by a consensus sigma(70) promoter. *J Bacteriol* 179(7): 2289-2299.

Ge, Y. G., Old, I. G., Saintgirons, I. & Charon, N. W. (1997b). The Flgk motility operon of *Borrelia burgdorferi* is initiated by a sigma(70)-like promoter. *Microbiology* 143(5): 1681-1690.

Ghiara, P., Marchetti, M., Blaser, M. J., Tummuru, M. K., Cover, T. L., Segal, E. D., Tompkins, L. S. & Rappuoli, R. (1995). Role of the *Helicobacter pylori* virulence factors vacuolating cytotoxin, CagA, and urease in a mouse model of disease. *Infect Immun* 63(10): 4154-4160.

Gilbert, J. V., Ramakrishna, J., Sunderman, F. W., Jr., Wright, A. & Plaut, A. G. (1995). Protein Hpn: cloning and characterization of a histidine-rich metal-binding polypeptide in *Helicobacter pylori* and *Helicobacter mustelae*. *Infect Immun* 63(7): 2682-2688.

Gillen, K. L. & Hughes, K. T. (1993). Transcription from two promoters and autoregulation contribute to the control of expression of the *Salmonella typhimurium* flagellar regulatory gene *flgM*. *J Bacteriol* 175(21): 7006-7015.

Glisin, V., Crkvenjakov, R. & Byus, C. (1974). Ribonucleic acid isolated by cesium chloride centrifugation. *Biochemistry* 13(12): 2633-2637.

Go, M. F., Kapur, V., Graham, D. Y. & Musser, J. M. (1996). Population genetic analysis of *Helicobacter pylori* by multilocus enzyme electrophoresis: extensive allelic diversity and recombinational population structure. *J Bacteriol* 178(13): 3934-3938.

Guilford, P., Hopkins, J., Harraway, J., McLeod, M., McLeod, N., Harawira, P., Taite, H., Scoular, R., Miller, A. & Reeve, A. E. (1998). E-cadherin germline mutations in familial gastric cancer. *Nature* 392(6674): 402-405.

Gygi, D., Bailey, M. J., Allison, C. & Hughes, C. (1995). Requirement for FlhA in flagella assembly and swarm-cell differentiation by *Proteus mirabilis*. *Mol Microbiol* 15(4): 761-769.

Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol* 166(4): 557-580.

Hancock, R. E., Alm, R., Bina, J. & Trust, T. (1998). *Helicobacter pylori*: a surprisingly conserved bacterium. *Nat Biotechnol* 16(3): 216-217.

Handt, L. K., Fox, J. G., Dewhirst, F. E., Fraser, G. J., Paster, B. J., Yan, L. L., Rozmiarek, H., Rufo, R. & Stalis, I. H. (1994). *Helicobacter pylori* isolated from the domestic cat: public health implications. *Infect Immun* 62(6): 2367-2374.

Hardham, J. M., Frye, J. G. & Stamm, L. V. (1995). Identification and sequences of the *Treponema pallidum fliM', fliY, fliP, fliQ, fliR* and *flhB'* genes. *Gene* 166(1): 57-64.

Harley, C. B. & Reynolds, R. P. (1987). Analysis of *E. coli* promoter sequences. *Nucleic Acids Res* 15(5): 2343-2361.

Hawtin, P. R., Stacey, A. R. & Newell, D. G. (1990). Investigation of the structure and localization of the urease of *Helicobacter pylori* using monoclonal antibodies. *J Gen Microbiol* 136(10): 1995-2000.

Hazell, S. L., Lee, A., Brady, L. & Hennessy, W. (1986). *Campylobacter pyloridis* and gastritis: association with intercellular spaces and adaptation to an environment of mucus as important factors in colonization of the gastric epithelium. *J Infect Dis* 153(4): 658-663.

He, S. Y. (1997). Hrp-controlled interkingdom protein transport: learning from flagellar assembly? *Trends Microbiol* 5(12): 489-495.

Helmann, J. D. (1991). Alternative sigma factors and the regulation of flagellar gene expression. *Mol Microbiol* 5(12): 2875-2882.

Henrichsen, J. (1983). Twitching motility. *Annu Rev Microbiol* 37: 81-93.

Higgins, C. F., Hiles, I. D., Salmond, G. P., Gill, D. R., Downie, J. A., Evans, I. J., Holland, I. B., Gray, L., Buckel, S. D., Bell, A. W. & et al. (1986). A family of related ATP-binding subunits coupled to many distinct biological processes in bacteria. *Nature* 323(6087): 448-450.

Higgins, D. G. & Sharp, P. M. (1989). Fast and sensitive multiple sequence alignments on a microcomputer. *CABIOS Commun* 5: 151-153.

Hirano, T., Yamaguchi, S., Oosawa, K. & Aizawa, S. (1994). Roles of FliK and FlhB in determination of flagellar hook length in *Salmonella typhimurium*. *J Bacteriol* 176(17): 5439-5449.

Hobbs, M. & Mattick, J. S. (1993). Common components in the assembly of type 4 fimbriae, DNA transfer systems, filamentous phage and protein-secretion apparatus: a general system for the formation of surface-associated protein complexes. *Mol Microbiol* 10(2): 233-243.

Homma, M. & Iino, T. (1985). Locations of hook-associated proteins in flagellar structures of *Salmonella typhimurium*. *J Bacteriol* 162(1): 183-189.

Hosking, S. W., Ling, T. K., Chung, S. C., Yung, M. Y., Cheng, A. F., Sung, J. J. & Li, A. K. (1994). Duodenal ulcer healing by eradication of *Helicobacter pylori* without anti-acid treatment: randomised controlled trial. *Lancet* 343(8896): 508-510.

Hu, L. T. & Mobley, H. L. T. (1990). Purification and N-terminal analysis of urease from *Helicobacter pylori*. *Infect. Immun.* 58(4): 992-998.

Huang, Z., Fasco, M. J. & Kaminsky, L. S. (1996). Optimization of Dnase I removal of contaminating DNA from RNA for use in quantitative RNA-PCR. *Biotechniques* 20(6): 1012-4, 1016, 1018-20.

Hueck, C. J. (1998). Type III protein secretion systems in bacterial pathogens of animals and plants. *Microbiol Mol Biol Rev* 62(2): 379-433.

Hughes, K. T., Gillen, K. L., Semon, M. J. & Karlinsey, J. E. (1993). Sensing structural intermediates in bacterial flagellar assembly by export of a negative regulator. *Science* 262(5137): 1277-1280.

Ikeda, T., Yamaguchi, S. & Hotani, H. (1993). Flagellar growth in a filament-less *Salmonella fliD* mutant supplemented with purified hook-associated protein 2. *J Biochem* 114(1): 39-44.

Ilver, D., Arnqvist, A., Ogren, J., Frick, I. M., Kersulyte, D., Incecik, E. T., Berg, D. E., Covacci, A., Engstrand, L. & Boren, T. (1998). *Helicobacter pylori* adhesin binding fucosylated histo-blood group antigens revealed by retagging. *Science* 279(5349): 373-377.

Jalava, K., Kaartinen, M., Utriainen, M., Happonen, I. & Hanninen, M. L. (1997). *Helicobacter salomonis* sp. *nov.*, a canine gastric *Helicobacter* sp. related to *Helicobacter felis* and *Helicobacter bizzozeronii. Int J Syst Bacteriol* 47(4): 975-82.

Jenks, P. J., Foynes, S., Ward, S. J., Constantinidou, C., Penn, C. W. & Wren, B. W. (1997). A flagellar-specific ATPase (FliI) is necessary for flagellar export *in Helicobacter pylori. FEMS Microbiol Lett* 152(2): 205-211.

Jiang, Q., Hiratsuka, K. & Taylor, D. E. (1996). Variability of gene order in different *Helicobacter pylori* strains contributes to genome diversity. *Mol Microbiol* 20(4): 833-842.

Jones, A. C., Logan, R. P., Foynes, S., Cockayne, A., Wren, B. W. & Penn, C. W. (1997). A flagellar sheath protein of *Helicobacter pylori* is identical to HpaA, a putative N-acetylneuraminyllactose-binding hemagglutinin, but is not an adhesin for AGS cells. *J Bacteriol* 179(17): 5643-5647.

Josenhans, C., Labigne, A. & Suerbaum, S. (1995). Comparative ultrastructural and functional studies of *Helicobacter pylori* and *Helicobacter mustelae* flagellin mutants: both flagellin subunits, FlaA and FlaB, are necessary for full motility in *Helicobacter* species. *J Bacteriol* 177(11): 3010-3020.

Karlinsey, J. E., Pease, A. J., Winkler, M. E., Bailey, J. L. & Hughes, K. T. (1997). The *flk* gene of *Salmonella typhimurium* couples flagellar P- and L-ring assembly to flagellar morphogenesis. *J Bacteriol* 179(7): 2389-2400.

Karlinsey, J. E., Tsui, H. C. T., Winkler, M. E. & Hughes, K. T. (1998). Flk couples *flgM* translation to flagellar ring assembly in *Salmonella typhimurium*. *J Bacteriol* 180(20): 5384-5397.

Khan, S., Khan, I. H. & Reese, T. S. (1991). New structural features of the flagellar base in *Salmonella typhimurium* revealed by rapid-freeze electron microscopy. *J Bacteriol* 173(9): 2888-2896.

Kihara, M., Homma, M., Kutsukake, K. & Macnab, R. M. (1989). Flagellar switch of *Salmonella typhimurium*: gene sequences and deduced protein sequences. *J Bacteriol* 171(6): 3247-3257.

Klose, K. E. & Mekalanos, J. J. (1998). Distinct roles of an alternative sigma factor during both free-swimming and colonizing phases of the *Vibrio cholerae* pathogenic cycle. *Mol Microbiol* 28(3): 501-520.

Komeda, Y., Suzuki, H., Ishidsu, J. I. & Iino, T. (1976). The role of cAMP in flagellation of *Salmonella typhimurium*. *Mol Gen Genet* 142(4): 289-298.

Kostrzynska, M., Betts, J. D., Austin, J. W. & Trust, T. J. (1991). Identification, characterization, and spatial localization of two flagellin species in *Helicobacter pylori* flagella. *J Bacteriol* 173(3): 937-946.

Kostrzynska, M., O'Toole, P. W., Taylor, D. E. & Trust, T. J. (1994). Molecular characterization of a conserved 20-kilodalton membrane- associated lipoprotein antigen of *Helicobacter pylori*. *J Bacteriol* 176(19): 5938-5948.

Krakowka, S., Eaton, K. A., Rings, D. M. & Morgan, D. R. (1991). Gastritis induced by *Helicobacter pylori* in gnotobiotic piglets. *Rev Infect Dis* 13: S681-S685.

Kreiss, C., Buclin, T., Cosma, M., Corthesy-Theulaz, I. & Michetti, P. (1996). Safety of oral immunisation with recombinant urease in patients with *Helicobacter pylori* infection. *Lancet* 347(9015): 1630-1631.

Kubori, T., Shimamoto, N., Yamaguchi, S., Namba, K. & Aizawa, S. (1992). Morphological pathway of flagellar assembly in *Salmonella typhimurium*. *J Mol Biol* 226(2): 433-446.

Kubori, T., Yamaguchi, S. & Aizawa, S. I. (1997). Assembly of the switch complex onto the MS ring complex of *Salmonella typhimurium* does not require any other flagellar proteins. *J Bacteriol* 179(3): 813-817.

Kunst, F., Ogasawara, N., Moszer, I., Albertini, A. M., Alloni, G., Azevedo, V., Bertero, M. G., Bessieres, P., Bolotin, A., Borchert, S., Borriss, R., Boursier, L., Brans, A., Braun, M., Brignell, S. C., Bron, S., Brouillet, S., Bruschi, C. V., Caldwell, B., Capuano, V., Carter, N. M., Choi, S. K., Codani, J. J., Connerton, I. F., Danchin, A. & *et al.* (1997). The complete genome sequence of the gram-positive bacterium *Bacillus subtilis. Nature* 390(6657): 249-256.

Kusters, J. G., Gerrits, M. M., Van Strijp, J. A. & Vandenbroucke-Grauls, C. M. (1997). Coccoid forms of *Helicobacter pylori* are the morphologic manifestation of cell death. *Infect Immun* 65(9): 3672-3679.

Kutsukake, K. (1994). Excretion of the anti-sigma factor through a flagellar substructure couples flagellar gene expression with flagellar assembly in *Salmonella typhimurium. Mol Gen Genet* 243(6): 605-612.

Kutsukake, K. (1997a). Autogenous and global control of the flagellar master operon, *flhD*, in *Salmonella typhimurium. Mol Gen Genet* 254(4): 440-448.

Kutsukake, K. (1997b). Hook-length control of the export-switching machinery involves a double-locked gate in *Salmonella typhimurium* flagellar morphogenesis. *J Bacteriol* 179(4): 1268-1273.

Kutsukake, K. & Ide, N. (1995). Transcriptional analysis of the *flgK* and *fliD* operons of *Salmonella typhimurium* which encode flagellar hook-associated proteins. *Mol Gen Genet* 247(3): 275-281.

Kutsukake, K. & Iino, T. (1994). Role of the FliA-FlgM regulatory system on the transcriptional control of the flagellar regulon and flagellar formation in *Salmonella typhimurium. J Bacteriol* 176(12): 3598-3605.

Kutsukake, K., Minamino, T. & Yokoseki, T. (1994). Isolation and characterization of FliK-independent flagellation mutants from *Salmonella typhimurium. J Bacteriol* 176(24): 7625-7629.

Kutsukake, K., Ohya, Y. & Iino, T. (1990). Transcriptional analysis of the flagellar regulon of *Salmonella typhimurium. J Bacteriol* 172(2): 741-747.

Labigne, A., Courcoux, P. & Tompkins, L. (1992). Cloning of *Campylobacter jejuni* genes required for leucine biosynthesis, and construction of a *leu*-negative

mutant of *C. jejuni* by shuttle transposon mutagenesis. *Res Microbiol* 143(1): 15-26.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227: 680-685.

Lam, S. K., Ching, C. K., Lai, K. C., Wong, B. C., Lai, C. L., Chan, C. K. & Ong, L. (1997). Does treatment of *Helicobacter pylori* with antibiotics alone heal duodenal ulcer? A randomised double blind placebo controlled study. *Gut* 41(1): 43-48.

Leclerc, G., Wang, S. P. & Ely, B. (1998). A new class of *Caulobacter crescentus* flagellar genes. *J Bacteriol* 180(19): 5010-5019.

Lee, A. (1998). The *Helicobacter pylori* genome: the bacteriologist's viewpoint. *Curr Opin Gastroenterol* 14 (Suppl 1): S5-S6.

Lee, A., Fox, J. G., Otto, G. & Murphy, J. (1990). A small animal model of human *Helicobacter pylori* active chronic gastritis. *Gastroenterology* 99: 1315-1323.

Lee, A., O'Rourke, J., Deungria, M. C., Robertson, B., Daskalopoulos, G. & Dixon, M. F. (1997a). A standardized mouse model of *Helicobacter pylori* infection - introducing the Sydney strain. *Gastroenterology* 112(4): 1386-1397.

Lee, C. A. (1997). Type III secretion systems: machines to deliver bacterial proteins into eukaryotic cells? *Trends Microbiol* 5(4): 148-156.

Lee, W. K., An, Y. S., Kim, K. H., Kim, S. H., Song, J. Y., Ryu, B. D., Choi, Y. J., Yoon, Y. H., Baik, S. C., Rhee, K. H. & Cho, M. J. (1997b). Construction of a *Helicobacter pylori-Escherichia coli* shuttle vector for gene transfer in *Helicobacter pylori*. *Appl Environ Microbiol* 63(12): 4866-4871.

Lessl, M., Balzer, D., Pansegrau, W. & Lanka, E. (1992). Sequence similarities between the RP4 Tra2 and the Ti VirB region strongly support the conjugation model for T-DNA transfer. *J Biol Chem* 267(28): 20471-20480.

Levin, P. A. & Losick, R. (1994). Characterization of a cell division gene from *Bacillus subtilis* that is required for vegetative and sporulation septum. *J Bacteriol* 176(5): 1451-1459.

Leying, H., Suerbaum, S., Geis, G. & Haas, R. (1992). Cloning and genetic characterization of a *Helicobacter pylori* flagellin gene. *Mol Microbiol* 6(19): 2863-2874.

Lin, T. S. & Kado, C. I. (1993). The *virD4* gene is required for virulence while *virD3* and *orf5* are not required for virulence of *Agrobacterium tumefaciens*. *Mol Microbiol* 9(4): 803-812.

Liu, X. & Matsumura, P. (1994). The FlhD/FlhC complex, a transcriptional activator of the *Escherichia coli* flagellar class II operons. *J Bacteriol* 176(23): 7345-7351.

Logan, R. P. & Berg, D. E. (1996). Genetic diversity of Helicobacter pylori. *Lancet* 348(9040): 1462-1463.

Logan, R. P. H. (1996). Adherence of *Helicobacter pylori*. *Aliment Pharmacol Ther* 10 (Suppl 1): 3-15.

Macnab, R. M. (1987). Motility and chemotaxis. In *Escherichia coli and Salmonella typhimurium* (Neidhardt, F. C., III, Ingraham, J. L., Low, K. B., Magasanik, B., Schaechter, M. & Umbarger, H. E., eds.), pp. 732-759. American Society for Microbiology Press, Washington D.C.

Macnab, R. M. (1992). Genetics and biogenesis of bacterial flagella. *Annu Rev Genet* 26: 131-158.

Macnab, R. M. (1996). Flagella and motility. In *Escherichia coli and Salmonella: cellular and molecular biology* 2nd ed. edit. (Neidhardt, F. C., III, R. C., Ingraham, J. L., Lin, E. C. C., Low, K. B., Magasanik, B., Reznikoff, W. S., Riley, M., Schaechter, M. & Umbarger, H. E., eds.), pp. 123-145. American Society for Microbiology Press, Washington, D.C.

Malakooti, J., Ely, B. & Matsumura, P. (1994). Molecular characterization, nucleotide sequence, and expression of the *fliO, fliP, fliQ*, and *fliR* genes of *Escherichia coli*. *J Bacteriol* 176(1): 189-197.

Mangan, E. K., Bartamian, M. & Gober, J. W. (1995). A mutation that uncouples flagellum assembly from transcription alters the temporal pattern of flagellar gene expression in *Caulobacter crescentus*. *J Bacteriol* 177(11): 3176-3184.

Manson, M. D., Armitage, J. P., Hoch, J. A. & Macnab, R. M. (1998). Bacterial locomotion and signal transduction. *J Bacteriol* 180(5): 1009-1022.

Marchetti, M., Arico, B., Burroni, D., Figura, N., Rappuoli, R. & Ghiara, P. (1995). Development of a mouse model of *Helicobacter pylori* infection that mimics human disease. *Science* 267(5204): 1655-1658.

Marshall, B. J. (1998). *Helicobacter pylori*: past, present and future (lecture). In *XIth International Workshop Gastroduodenal pathology and Helicobacter pylori*, Budapest.

Marshall, B. J., Armstrong, J. A., McGechie, D. B. & Glancy, R. J. (1985). Attempt to fulfil Koch's postulates for pyloric *Campylobacter*. *Med J Aust* 142(8): 436-439.

Marshall, B. J. & Warren, J. R. (1984). Unidentified curved bacilli in the stomach of patients with gastritis and peptic ulceration. *Lancet* 1(8390): 1311-1315.

Mattick, J. S. & Alm, R. A. (1995). Common architecture of type-4 fimbriae and complexes involved in macromolecular traffic. *Trends in Microbiology* 3: 411-413.

Mecsas, J. & Strauss, E. J. (1996). Molecular mechanisms of bacterial virulence - type III secretion and pathogenicity islands. *Emerg Infect Dis* 2(4): 271-288.

Megraud, F. (1995). Transmission of *Helicobacter pylori*: faecal-oral versus oral-oral route. *Aliment Pharmacol Ther* 9 (Suppl 2): 85-91.

Midolo, P. D., Korman, M. G., Turnidge, J. D. & Lambert, J. R. (1996). *Helicobacter pylori* resistance to tetracycline. *Lancet* 347(9009): 1194-1195.

Miehlke, S., Hackelsberger, A., Meining, A., von Arnim, U., Muller, P., Ochsenkuhn, T., Lehn, N., Malfertheiner, P., Stolte, M. & Bayerdorffer, E. (1997). Histological diagnosis of *Helicobacter pylori* gastritis is predictive of a high risk of gastric carcinoma. *Int J Cancer* 73(6): 837-839.

Miller, S., Pesci, E. C. & Pickett, C. L. (1993). A *Campylobacter jejuni* homolog of the LcrD/FlbF family of proteins is necessary for flagellar biogenesis. *Infect Immun* 61(7): 2930-2936.

Mimori, Y., Yamashita, I., Murata, K., Fujiyoshi, Y., Yonekura, K., Toyoshima, C. & Namba, K. (1995). The structure of the R-type straight flagellar filament of *Salmonella* at 9 Å resolution by electron cryomicroscopy. *J Mol Biol* 249(1): 69-87.

Minamino, T., Iino, T. & Kutuskake, K. (1994). Molecular characterization of the *Salmonella typhimurium flhB* operon and its protein products. *J Bacteriol* 176(24): 7630-7637.

Misra, T. K. (1987). DNA sequencing: a new strategy to create ordered deletions, modified M13 vector, and improved reaction conditions for sequencing by dideoxy chain termination method. *Methods Enzymol* 155: 119-139.

Moore, S. (1999). Personal communication.

Moran, A. P., Lindner, B. & Walsh, E. J. (1997). Structural characterization of the lipid A component of *Helicobacter pylori* rough- and smooth-form lipopolysaccharides. *J Bacteriol* 179(20): 6453-6463.

Morgan, D. G., Owen, C., Melanson, L. A. & DeRosier, D. J. (1995). Structure of bacterial flagellar filaments at 11 Å resolution: packing of the alpha-helices. *J Mol Biol* 249(1): 88-110.

Morris, A. & Nicholson, G. (1987). Ingestion of *Campylobacter pyloridis* causes gastritis and raised fasting gastric pH. *Am J Gastroenterol* 82(3): 192-199.

Mulholland, V., Hinton, J. C., Sidebotham, J., Toth, I. K., Hyman, L. J., Perombelon, M. C., Reeves, P. J. & Salmond, G. P. (1993). A pleiotropic reduced virulence (Rvi-) mutant of *Erwinia carotovora* subspecies *atroseptica* is defective in flagella assembly proteins that are conserved in plant and animal bacterial pathogens. *Mol Microbiol* 9(2): 343-356.

Mullin, D., Minnich, S., Chen, L. S. & Newton, A. (1987). A set of positively regulated flagellar gene promoters in *Caulobacter crescentus* with sequence homology to the *nif* gene promoters of *Klebsiella pneumoniae*. *J Mol Biol* 195(4): 939-943.

Muramoto, K., Makishima, S., Aizawa, S. I. & Macnab, R. M. (1998). Effect of cellular level of FliK on flagellar hook and filament assembly in *Salmonella typhimurium*. *J Mol Biol* 277(4): 871-882.

Namavar, F., Sparrius, M., Veerman, E. C., Appelmelk, B. J. & Vandenbroucke-Grauls, C. M. (1998). Neutrophil-activating protein mediates adhesion of *Helicobacter pylori* to sulfated carbohydrates on high-molecular-weight salivary mucin. *Infect Immun* 66(2): 444-447.

Namba, K., Yamashita, I. & Vonderviszt, F. (1989). Structure of the core and central channel of bacterial flagella. *Nature* 342(6250): 648-654.

Nilius, M. & Malfertheiner, P. (1996). *Helicobacter pylori* enzymes. *Aliment Pharmacol Ther* 10 (Suppl 1): 65-71.

Noonan, B. (1998). unpublished data.

O'Toole, P. W., Janzon, L., Doig, P., Huang, J., Kostrzynska, M. & Trust, T. J. (1995). The putative neuraminyllactose-binding hemagglutinin HpaA of *Helicobacter pylori* CCUG 17874 is a lipoprotein. *J Bacteriol* 177(21): 6049-6057.

O'Toole, P. W., Kostrzynska, M. & Trust, T. J. (1994). Non-motile mutants of *Helicobacter pylori* and *Helicobacter mustelae* defective in flagellar hook production. *Mol Microbiol* 14(4): 691-703.

O'Toole, P. W., Logan, S. M., Kostrzynska, M., Wadström, T. & Trust, T. J. (1991). Isolation and biochemical and molecular analyses of a species-specific protein antigen from the gastric pathogen *Helicobacter pylori. J. Bacteriol.* 173: 505-513.

Ohnishi, K., Fan, F., Schoenhals, G. J., Kihara, M. & Macnab, R. M. (1997). The FliO, FliP, FliQ, and FliR proteins of *Salmonella typhimurium*: putative components for flagellar assembly. *J Bacteriol* 179(19): 6092-6099.

Ohnishi, K., Kutsukake, K., Suzuki, H. & Lino, T. (1992). A novel transcriptional regulation mechanism in the flagellar regulon of *Salmonella typhimurium*: an antisigma factor inhibits the activity of the flagellum-specific sigma factor, sigma F. *Mol Microbiol* 6(21): 3149-3157.

Ohnishi, K., Ohto, Y., Aizawa, S., Macnab, R. M. & Iino, T. (1994). FlgD is a scaffolding protein needed for flagellar hook assembly in *Salmonella typhimurium. J Bacteriol* 176(8): 2272-2281.

Ottow, J. C. (1975). Ecology, physiology, and genetics of fimbriae and pili. *Annu Rev Microbiol* 29: 79-108.

Owen, R. J. (1998). The genome: viewpoint of a molecular biologist. *Curr Opin Gastroenterol* 14 (Suppl 1): S1-S3.

Palli, D. (1997). Gastric cancer and *Helicobacter pylori*: a critical evaluation of the epidemiological evidence. *Helicobacter* 2 (Suppl 1): S50-55.

Palmer, E. D. (1954). Investigation of the gastric spirochaetes of the human. *Gastroenterology* 27: 218-220.

Parsonnet, J., Friedman, G. D., Orentreich, N. & Vogelman, H. (1997). Risk for gastric cancer in people with CagA positive or CagA negative *Helicobacter pylori* infection. *Gut* 40(3): 297-301.

Peek, R. M., Jr., Moss, S. F., Tham, K. T., Perez-Perez, G. I., Wang, S., Miller, G. G., Atherton, J. C., Holt, P. R. & Blaser, M. J. (1997). Helicobacter pylori *cagA*$^{+}$ strains and dissociation of gastric epithelial cell proliferation from apoptosis. *J Natl Cancer Inst* 89(12): 863-868.

Perez-Casal, J., Caparon, M. G. & Scott, J. R. (1991). Mry, a trans-acting positive regulator of the M protein gene of Streptococcus pyogenes with similarity to the receptor proteins of two- component regulatory systems. *J Bacteriol* 173(8): 2617-2624.

Pettersson, J., Nordfelth, R., Dubinina, E., Bergman, T., Gustafsson, M., Magnusson, K. E. & Wolf-Watz, H. (1996). Modulation of virulence factor expression by pathogen target cell contact. *Science* 273(5279): 1231-1233.

Phadnis, S. H., Ilver, D., Janzon, L., Normark, S. & Westblom, T. U. (1994). Pathological significance and molecular characterization of the vacuolating toxin gene of *Helicobacter pylori*. *Infect Immun* 62(5): 1557-1565.

Phadnis, S. H., Parlow, M. H., Levy, M., Ilver, D., Caulkins, C. M., Connors, J. B. & Dunn, B. E. (1996). Surface localization Oof *Helicobacter pylori* urease and a heat shock protein homolog requires bacterial autolysis. *Infect Immun* 64(3): 905-912.

Pitcher, D. G., Saunders, N. A. & Owen, R. J. (1989). Rapid extraction of bacterial genomic DNA with guanidium thiocyanate. *Lett Appl Microbiol* 8: 151-156.

Plano, G. V., Barve, S. S. & Straley, S. C. (1991). LcrD, a membrane-bound regulator of the *Yersinia pestis* low-calcium response. *J Bacteriol* 173(22): 7293-7303.

Pohlman, R. F., Genetti, H. D. & Winans, S. C. (1994). Common ancestry between IncN conjugal transfer genes and macromolecular export systems of plant and animal pathogens. *Mol Microbiol* 14(4): 655-668.

Pugsley, A. P. (1993). The complete general secretory pathway in gram-negative bacteria. *Microbiol Rev* 57(1): 50-108.

Pugsley, A. P., Francetic, O., Possot, O. M., Sauvonnet, N. & Hardie, K. R. (1997). Recent progress and future directions in studies of the main terminal branch of the general secretory pathway in Gram-negative bacteria - a review. *Gene* 192(1): 13-19.

Quon, K. C., Marczynski, G. T. & Shapiro, L. (1996). Cell cycle control by an essential bacterial two-component signal transduction protein. *Cell* 84(1): 83-93.

Radcliff, F. J., Hazell, S. L., Kolesnikow, T., Doidge, C. & Lee, A. (1997). Catalase, a novel antigen for *Helicobacter pylori* vaccination. *Infect Immun* 65(11): 4668-4674.

Ramakrishnan, G., Zhao, J. L. & Newton, A. (1991). The cell cycle-regulated flagellar gene *flbF* of *Caulobacter crescentus* is homologous to a virulence locus (*lcrD*) of *Yersinia pestis*. *J Bacteriol* 173(22): 7283-7292.

Rieder, G., Hatz, R. A., Moran, A. P., Walz, A., Stolte, M. & Enders, G. (1997). Role of adherence in interleukin-8 induction in *Helicobacter pylori*- associated gastritis. *Infect Immun* 65(9): 3622-3630.

Romaniuk, P. J., Zoltowska, B., Trust, T. J., Lane, D. J., Olsen, G. J., Pace, N. R. & Stahl, D. A. (1987). *Campylobacter pylori*, the spiral bacterium associated with human gastritis, is not a true *Campylobacter* sp. *J Bacteriol* 169(5): 2137-2141.

Rosqvist, R., Magnusson, K. E. & Wolf-Watz, H. (1994). Target cell contact triggers expression and polarized transfer of *Yersinia* YopE cytotoxin into mammalian cells. *EMBO J* 13(4): 964-972.

Rowland, S. L., Errington, J. & Wake, R. G. (1995). The *Bacillus subtilis* cell-division 135-137 degrees region contains an essential orf with significant similarity to *murB* and a dispensable *sbp* gene. *Gene* 164(1): 113-116.

Rudmann, D. G., Eaton, K. A. & Krakowka, S. (1992). Ultrastructural study of *Helicobacter pylori* adherence properties in gnotobiotic piglets. *Infect Immun* 60: 2121-2124.

Rüssmann, H., Shams, H., Poblete, F., Fu, Y., Galan, J. E. & Donis, R. O. (1998). Delivery of epitopes by the *Salmonella* type III secretion system for vaccine development. *Science* 281(5376): 565-568.

Salmond, G. P. (1994). Secretion of extracellular virulence factors by plant pathogenic bacteria. *Ann Rev Phytopathol* 32: 181-200.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular cloning: a laboratory manual. 2nd Ed.*, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Sanger, F., Donelson, J. E., Coulson, A. R., Kossel, H. & Fischer, D. (1974). Determination of a nucleotide sequence in bacteriophage f1 DNA by primed synthesis with DNA polymerase. *J Mol Biol* 90(2): 315-333.

Schmitt, W. & Haas, R. (1994). Genetic analysis of the *Helicobacter pylori* vacuolating cytotoxin: structural similarities with the IgA protease type of exported protein. *Mol Microbiol* 12: 307-319.

Schmitz, A., Josenhans, C. & Suerbaum, S. (1997). Cloning and characterization of the *Helicobacter pylori flba* gene, which codes for a membrane protein involved in coordinated expression of flagellar genes. *J Bacteriol* 179(4): 987-997.

Scott, D. R., Weeks, D., Hong, C., Postius, S., Melchers, K. & Sachs, G. (1998). The role of internal urease in acid resistance of *Helicobacter pylori. Gastroenterology* 114(1): 58-70.

Segal, E. D. & Tompkins, L. S. (1993). Transformation of *Helicobacter pylori* by electroporation. *Biotechniques* 14: 225-226.

Shah, D. S. & Sockett, R. E. (1995). Analysis of the *motA* flagellar motor gene from *Rhodobacter sphaeroides*, a bacterium with a unidirectional, stop-start flagellum. *Mol Microbiol* 17(5): 961-969.

Shirakihara, Y., Leslie, A. G., Abrahams, J. P., Walker, J. E., Ueda, T., Sekimoto, Y., Kambara, M., Saika, K., Kagawa, Y. & Yoshida, M. (1997). The crystal structure of the nucleotide-free alpha 3 beta 3 subcomplex of F_1-ATPase from the thermophilic Bacillus PS3 is a symmetric trimer. *Structure* 5(6): 825-836.

Shirasu, K., Koukolikova-Nicola, Z., Hohn, B. & Kado, C. I. (1994). An inner-membrane-associated virulence protein essential for T-DNA transfer from *Agrobacterium tumefaciens* to plants exhibits ATPase activity and similarities to conjugative transfer genes. *Mol Microbiol* 11(3): 581-588.

Southern, E. M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J Mol Biol* 98(3): 503-517.

Stachel, S. E., Timmerman, B. & Zambryski, P. (1987). Activation of *Agrobacterium tumefaciens vir* gene expression generates multiple single-stranded T-strand molecules from the pTiA6 T-region: requirement for 5' *virD* gene products. *EMBO J* 6(4): 857-863.

Stallmeyer, M. J., Hahnenberger, K. M., Sosinsky, G. E., Shapiro, L. & DeRosier, D. J. (1989). Image reconstruction of the flagellar basal body of *Caulobacter crescentus. J Mol Biol* 205(3): 511-518.

Stark, R. M., Greenman, J. & Millar, M.R. (1995). Physiology and biochemistry of *Helicobacter pylori. Brit J Biomed Sc* 52: 282-290.

Stephens, C., Mohr, C., Boyd, C., Maddock, J., Gober, J. & Shapiro, L. (1997). Identification of the *fliI* and *fliJ* components of the *Caulobacter* flagellar type III protein secretion system. *J Bacteriol* 179(17): 5355-5365.

Stephens, C. M. & Shapiro, L. (1993). An unusual promoter controls cell-cycle regulation and dependence on DNA replication of the *Caulobacter fliLM* early flagellar operon. *Mol Microbiol* 9(6): 1169-1179.

Stryer, L. (1988). *Biochemistry*. 3rd edit, W.H.Freeman & Co., New York.

Suerbaum, S. (1995). The complex flagella of gastric *Helicobacter* species. *Trends Microbiol* 3(5): 168-170.

Suerbaum, S., Josenhans, C. & Labigne, A. (1993). Cloning and genetic characterization of the *Helicobacter pylori* and *Helicobacter mustelae flaB* flagellin genes and construction of *H. pylori flaA*- and *flaB*-negative mutants by electroporation-mediated allelic exchange. *J Bacteriol* 175(11): 3278-3288.

Suerbaum, S., Smith, J. M., Bapumia, K., Morelli, G., Smith, N. H., Kunstmann, E., Dyrek, I. & Achtman, M. (1998). Free recombination within *Helicobacter pylori*. *Proc Natl Acad Sci U S A* 95(21): 12619-12624.

Suerbaum, S., Thiberge, J. M., Kansau, I., Ferrero, R. L. & Labigne, A. (1994). *Helicobacter pylori hspA-hspB* heat-shock gene cluster: nucleotide sequence, expression, putative function and immunogenicity. *Mol Microbiol* 14(5): 959-974.

Takahashi, S., Igarashi, H., Nakamura, K., Masubuchi, N., Saito, S., Aoyagi, T., Itoh, T. & Hirata, I. (1993). *Helicobacter pylori* and urease activity--comparative study between urease positive and negative mutant strains. *Jap J Clin Med* 51(12): 3149-3153.

Taku, A., Gunetileke, K. G. & Anwar, R. A. (1970). Biosynthesis of uridine diphospho-N-acetylmuramic acid. 3. Purification and properties of uridine diphospho-N-acetylenolpyruvyl-glucosamine reductase. *J Biol Chem* 245(19): 5012-5016.

Taylor, D. E., Eaton, M., Chang, N. & Salama, S. M. (1992). Construction of a *Helicobacter pylori* genome map and demonstration of diversity at the genome level. *J Bacteriol* 174: 6800-6806.

Telford, J. L., Ghiara, P., Dell'Orco, M., Comanducci, M., Burroni, D., Bugnoli, M., Tecce, M. F., Censini, S., Covacci, A., Xiang, Z. & et al. (1994). Gene

structure of the *Helicobacter pylori* cytotoxin and evidence of its key role in gastric disease. *J Exp Med* 179(5): 1653-1658.

Terres, A. M., Pajares, J. M., O'Toole, D., Ahern, S. & Kelleher, D. (1998). *H. pylori* infection is associated with downregulation of E-cadherin, a molecule involved in epithelial cell adhesion and proliferation control. *J Clin Pathol* 51(5): 410-412.

Thöny, B. & Hennecke, H. (1989). The -24/-12 promoter comes of age. *FEMS Microbiol Rev* 5(4): 341-357.

Thorstenson, Y. R., Kuldau, G. A. & Zambryski, P. C. (1993). Subcellular localization of seven VirB proteins of *Agrobacterium tumefaciens*: implications for the formation of a T-DNA transport structure. *J Bacteriol* 175(16): 5233-5241.

Tomb, J. F., White, O., Kerlavage, A. R., Clayton, R. A., Sutton, G. G., Fleischmann, R. D., Ketchum, K. A., Klenk, H. P., Gill, S., Dougherty, B. A., Nelson, K., Quackenbush, J., Zhou, L. X., Kirkness, E. F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H. G., Glodek, A., McKenney, K., Fitzegerald, L. M., Lee, N., Adams, M. D., Hickey, E. K. & et al. (1997). The complete genome sequence of the gastric pathogen *Helicobacter pylori*. *Nature* 388(6642): 539-547.

Towbin, H., Staehelin, T. & Gordon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: Procedure and some applications. *Proc. Natl. Acad. Sci. USA* 74: 4350-4354.

Tsuda, M., Karita, M., Morshed, M. G., Okita, K. & Nakazawa, T. (1994). A urease-negative mutant of *Helicobacter pylori* constructed by allelic exchange mutagenesis lacks the ability to colonize the nude mouse stomach. *Infect Immun* 62(8): 3586-3589.

Tsujii, M., Kawano, S., Tsuji, S., Ito, T., Nagano, K., Sasaki, Y., Hayashi, N., Fusamoto, H. & Kamada, T. (1993). Cell kinetics of mucosal atrophy in rat stomach induced by long-term administration of ammonia. *Gastroenterology* 104(3): 796-801.

Tummuru, M. K., Cover, T. L. & Blaser, M. J. (1993). Cloning and expression of a high-molecular-mass major antigen of *Helicobacter pylori*: evidence of linkage to cytotoxin production. *Infect Immun* 61(5): 1799-1809.

Ueno, T., Oosawa, K. & Aizawa, S. (1994). Domain structures of the MS ring component protein (FliF) of the flagellar basal body of *Salmonella typhimurium*. *J Mol Biol* 236(2): 546-555.

Vaira, D., Ali', A., Gatta, L. & O'Morain, C. (1998). Treatment of *Helicobacter pylori*. *Curr Opin Gastroenterol* 14 (Suppl 1): S71-S78.

Valkonen, K. H., Wadstrom, T. & Moran, A. P. (1997). Identification of the N-acetylneuraminyllactose-specific laminin- binding protein of *Helicobacter pylori*. *Infect Immun* 65(3): 916-923.

Valle, J., Seppala, K., Sipponen, P. & Kosunen, T. (1991). Disappearance of gastritis after eradication of *Helicobacter pylori*. A morphometric study. *Scand J Gastroenterol* 26(10): 1057-1065.

van Doorn, L. J., Figueiredo, C., Sanna, R., Plaisier, A., Schneeberger, P., de Boer, W. & Quint, W. (1998). Clinical relevance of the *cagA*, *vacA*, and *iceA* status of *Helicobacter pylori*. *Gastroenterology* 115(1): 58-66.

Vanet, A. & Labigne, A. (1998). Evidence for specific secretion rather than autolysis in the release of some *Helicobacter pylori* proteins. *Infect Immun* 66(3): 1023-1027.

Vasquez, A., Valdez, Y., Gilman, R. H., McDonald, J. J., Westblom, T. U., Berg, D., Mayta, H. & Gutierrez, V. (1996). Metronidazole and clarithromycin resistance in *Helicobacter pylori* determined by measuring MICs of antimicrobial agents in color indicator egg yolk agar in a miniwell format. *J Clin Microbiol* 34(5): 1232-1234.

Venkatesan, M. M., Buysse, J. M. & Oaks, E. V. (1992). Surface presentation of *Shigella flexneri* invasion plasmid antigens requires the products of the *spa* locus. *J Bacteriol* 174(6): 1990-2001.

Vogel, J. P., Andrews, H. L., Wong, S. K. & Isberg, R. R. (1998). Conjugative transfer by the virulence system of *Legionella pneumophila*. *Science* 279(5352): 873-876.

Vogler, A. P., Homma, M., Irikura, V. M. & Macnab, R. M. (1991). *Salmonella typhimurium* mutants defective in flagellar filament regrowth and sequence similarity of FliI to F_0F_1, vacuolar, and archaebacterial ATPase subunits. *J Bacteriol* 173(11): 3564-3572.

Wadström, T., Hirmo, S. & Boren, T. (1996). Biochemical aspects of *Helicobacter pylori* colonization of the human gastric mucosa. *Aliment Pharmacol Ther* 10 (Suppl 1): 17-27.

Walker, J. E., Saraste, M., Runswick, M. J. & Gay, N. J. (1982). Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and

other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J* 1(8): 945-951.

Ward, J. E., Jr., Dale, E. M., Christie, P. J., Nester, E. W. & Binns, A. N. (1990). Complementation analysis of *Agrobacterium tumefaciens* Ti plasmid *virB* genes by use of a *vir* promoter expression vector: *virB9, virB10* and *virB11* are essential virulence genes. *J Bacteriol* 172(9): 5187-5199.

Watanabe, T., Tada, M., Nagai, H., Sasaki, S. & Nakao, M. (1998). *Helicobacter pylori* infection induces gastric cancer in mongolian gerbils. *Gastroenterology* 115(3): 642-648.

Watarai, M., Tobe, T., Yoshikawa, M. & Sasakawa, C. (1995). Contact of *Shigella* with host cells triggers release of Ipa invasins and is an essential function of invasiveness. *EMBO J* 14(11): 2461-2470.

Wattiau, P., Woestyn, S. & Cornelis, G. R. (1996). Customized secretion chaperones in pathogenic bacteria. *Mol Microbiol* 20(2): 255-262.

Weiss, A. A., Johnson, F. D. & Burns, D. L. (1993). Molecular characterization of an operon required for pertussis toxin secretion. *Proc Natl Acad Sci U S A* 90(7): 2970-2974.

Winans, S. C., Burns, D. L. & Christie, P. J. (1996). Adaptation of a conjugal transfer system for the export of pathogenic macromolecules. *Trends Microbiol* 4(2): 64-68.

Wirth, H. P., Yang, M., Peek, R. M., Jr., Tham, K. T. & Blaser, M. J. (1997). *Helicobacter pylori* Lewis expression is related to the host Lewis phenotype. *Gastroenterology* 113(4): 1091-1098.

World Health Organization. (1994). Schistosomes, liver flukes and *Helicobacter pylori*. In *IARC monographs on the evaluation of carcinogenic risks to humans*, Vol. 61, pp. 177-240. World Health Organization, Lyon.

Wotherspoon, A. C. (1998). Gastric lymphoma of mucosa-associated lymphoid tissue and *Helicobacter pylori*. *Annu Rev Med* 49: 289-299.

Wu, J., Benson, A. K. & Newton, A. (1995). Global regulation of a sigma 54-dependent flagellar gene family in *Caulobacter crescentus* by the transcriptional activator FlbD. *J Bacteriol* 177(11): 3241-3250.

Wu, J. G. & Newton, A. (1997). Regulation of the *Caulobacter* flagellar gene hierarchy - not just for motility. *Mol Microbiol* 24(2): 233-239.

Xiang, Z., Censini, S., Bayeli, P. F., Telford, J. L., Figura, N., Rappuoli, R. & Covacci, A. (1995). Analysis of expression of CagA and VacA virulence factors in 43 strains of *Helicobacter pylori* reveals that clinical isolates can be divided into two major types and that CagA is not necessary for expression of the vacuolating cytotoxin. *Infect Immun* 63(1): 94-98.

Yamaguchi, H., Osaki, T., Taguchi, H., Hanawa, T., Yamamoto, T. & Kamiya, S. (1997). Production and characterisation of monoclonal antibodies to heat-shock protein 60 of *Helicobacter pylori*. *J Med Microbiol* 46(10): 819-824.

Yamaoka, Y., Kita, M., Kodama, T., Sawai, N., Kashima, K. & Imanishi, J. (1997). Induction of various cytokines and development of severe mucosal inflammation by *cagA* gene positive *Helicobacter pylori* strains. *Gut* 41(4): 442-451.

Yanisch-Perron, C., Vieira, J. & Messing, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* 33(1): 103-119.

Yao, R., Alm, R. A., Trust, T. J. & Guerry, P. (1993). Construction of new *Campylobacter* cloning vectors and a new mutational *cat* cassette. *Gene* 130(1): 127-130.

Zhuang, W. Y. & Shapiro, L. (1995). *Caulobacter* FliQ and FliR membrane proteins, required for flagellar biogenesis and cell division, belong to a family of virulence factor export proteins. *J Bacteriol* 177(2): 343-356.

Zupan, J. R., Ward, D. & Zambryski, P. (1998). Assembly of the VirB transport complex for DNA transfer from *Agrobacterium tumefaciens* to plant cells. *Curr Opin Microbiol* 1(6): 649-655.