Biochemical and Molecular Characterisation of FliI and FliH from *Helicobacter pylori*

A thesis presented in partial fulfilment
of
Doctor of Philosophy in Microbiology

at the Institute of Molecular BioSciences,
Massey University, Palmerston North,
New Zealand

Michael Lane

2006

Abstract

The bacterium *Helicobacter pylori* is a human pathogen that infects a large proportion of the world's population and is associated with serious diseases such as gastric ulcers and adenocarcinoma. The motility of this organism, by virtue of sheathed polar flagella is essential to colonisation and persistence in the human host.

The sequencing of the *H. pylori* genome in 1996 identified homologues of the majority of the flagellar genes found in *S. enterica* serovar *typhimurium.* These included genes encoding the flagellum ATPase, FliI and FliH a presumptive inhibitor, the primary focus of this study. Sequencing did not originally identify an *H. pylori* homologue of the flagellar chaperone FliJ, and this is also considered in this study.

Bioinformatic analysis and modeling suggests a structural and functional relationship between FliI and homologues such as F_1-ATPase α- and β-subunit. In particular, residues 2-91 of FliI resemble the N-terminal domain of the F_1-ATPase α- and β-subunits. Biochemical analyses reported in this thesis showed that a truncated FliI-(2- 91) protein was folded, although the N-terminal 18 residues were likely unstructured. Furthermore, deletion mutagenesis showed that this disordered segment of the protein mediates interaction with FliH and very likely forms an amphipathic α-helix upon forming of the FliI-FliH complex. The scanning mutagenesis of this interaction segment of FliI identified a cluster of conserved hydrophobic residues that was critical for the interaction with FliH. Thus, the interaction between FliI and FliH has similarities to the interaction between the N-terminal α-helix of the α-subunit and the globular domain of the δ-subunit of the F_1-ATPase. This similarity suggests that FliH, by analogy with the δ-subunit of the F_1-ATPase, may function as a molecular stator of the flagellum. The findings presented above have been published (96).

The function of a putative *H. pylori* FliJ homologue, HP0256, was also investigated by knock-out mutagenesis. Disruption of this gene does not abolish flagellar assembly, however further research continued beyond this thesis showed that the knock-out mutant results in impaired motility.

Acknowledgements

I would like to extend my deepest gratitude to Paul O'Toole and Stan Moore for being my supervisors during the course of this PhD, and providing the lab space and other resources necessary for this research. Thank you.

I would also like to thank and acknowledge Mike McManus and Jasna Rakonjac for their supervision and administration in New Zealand. I must also thank John Tweedie for his role in the administration of my PhD in New Zealand.

I am very grateful for the financial support I received from the award of a Massey University Doctoral Scholarship, and a travel grant received from the Institute of Molecular BioSciences. Thank you.

Electron microscopy was within this study was performed with the assistance of Doug Hopcroft at The Horticultural and Food Research Institute of New Zealand. I would like to thank Doug for his time and technical assistance and Hort. Research for the use of their facilities.

I also have to thank Ewa, George and Stacey for their support as technicians during my time as a PhD student.

During my time in Canada I carried out my research in the Department of Biochemistry, College of Medicine, University of Saskatchewan. I would like to thank the university and the department for accommodating me. Special thanks must also be given to the University of Saskatchewan Structural Sciences Centre and Jason Maley for the extensive use of their equipment, and their technical advice.

At times my circumstances in Saskatoon were difficult but I have to acknowledge that I have a fantastic group of friends there who made life a lot easier. In this regard I really have to thank Andrew, Kathryn, Mike, Curtis, Kris, Paul, Chauntelle, Scott, Radu, Dave, Jason, Nick, Ewa and George. I am glad I met you guys, without you my life would have been a lot less colourful.

Last but not least I really have to thank Mum and Dad for their continued love, understanding and support.

I have met a lot of people during the course of this study and if I have neglected to acknowledge anybody I apologise sincerely.

Table of Contents

List of Abbreviations

ADP adenosine diphosphate
AMP adenosine monophosphate
AMP-PNP adenylyl imidodiphosphate
Ap Ampicillin
APS ammonium persulphate
ATP adenosine triphosphate
BLAST Basic Local Alignment Search Tool
BSA bovine serum albumin
CAPS 3-(cyclohexylamino)-1-propanesulfonic acid
CD Circular dichroism
Cm Chloramphenicol
colE1 colicin E1
DLS Dynamic light scattering
DNA deoxyribonucleic acid
dNTP deoxynucleoside triphosphate
DTT dithiothreitol
EDTA ethylenediaminetetraacetic acid
FP forward primer
FPLC Fast protein liquid chromatography
FSB final sample buffer
GSP general secretory pathway
GEP general export pathway
GST Glutathione-S-transferase
HEPES N-(2-hydroxyethyl)-piperazine-N'-2-ethanesulfonic acid
HIC Hydrophobic interaction chromatography
IEC Ion exchange chromatography
IPTG isopropyl-β-D-galactoside
Kan Kanamycin
LB Luria-Bertani broth
LBA Luria-Bertani broth agar
MALT mucosa-associated lymphoid tissue
MCS multiple cloning site
na not applicable
NCBI National Centre for Biotechnology Information
OD Optical density
PBS Phosphate buffered saline
PCR polymerase chain reaction
PHI-BLAST Pattern Hit Initiated BLAST
pI isoelectric point
RP reverse primer
Rpm revolutions per minute
PSI-BLAST Position Specific Iterative BLAST
RT room temperature

SDS sodium dodecyl sulphate
SDS-PAGE Sodium dodecyl sulphate-Polyacrylamide gel electrophoresis
SEC Size exclusion chromatography
T3SSs Type III secretion systems
T4SSs Type IV secretion systems
TEMED N,N,N',N'-Tetramethylethylenediamine
TIGR The Institute for Genomic Research
Tris Tris(hydroxymethyl)aminomethane hydrochloride
TSB tryptic soy broth
U unit
w/v weight per volume

Amino acid abbreviations used:
A, Ala, Alanine
C, Cys, Cysteine
D, Asp, Aspartic acid
E, Glu, Glutamic acid
F, Phe, Penylalanine
G, Gly, Glycine
H, His, Histidine
I, Ile, Isoleucine
K, Lys, Lysine
L, Leu, Leucine
M, Met, Methionine
N, Asn, Asparagine
P, Pro, Proline
Q, Gln, Glutamine
R, Arg, Arginine
S, Ser, Serine
T, Thr, Threonine
V, Val, Valine
W, Trp, Tryptophan
Y, Tyr, Tyrosine

Deoxyribonucleosides: A, deoxyadenylate, C, deoxycytidylate, G, deoxyguanylate, T, deoxythymidylate

List of Figures

List of Tables

1. Introduction

1.1 Characteristics of the human gastric pathogen, *Helicobacter pylori*

Helicobacter pylori is a human gastric pathogen associated with chronic gastritis, duodenal and gastric ulcer disease, and two types of malignancies: lymphoma of the mucosa-associated lymphoid tissue (MALT) and adenocarcinoma. *H. pylori* infection occurs during childhood and persists for decades. Transmission has been proposed to occur within families, but while related strains can be identified within infected families this is not always the case (138). The current route of transmission is proposed to be faecal-oral and this has been lent indirect support by the association of infection with lower socio-economic status and associated poor hygiene, and the high occurrence of infection in the developing world. For example in West Africa infection amongst children can be as high as 80% whereas in the developed world infection rates are 5-20% amongst people younger than 40 (110)

1.1.1 General metabolism

H. pylori is a spiral-shaped Gram-negative bacterium, 2.5-5.0 μm in length and 0.5-1.0 μm wide, with 4-6 polar flagella (111). The niche of the bacterium is the mucous layer overlaying the human gastric epithelium. The gastric mucosa has a pH of 7.0, an oxygen tension of 12-16kPa and human body temperature is 37 °C. These characteristics of *H. pylori's* niche are reflected in the *in vitro* growth conditions of the organism. The bacterium is a microaerophile and thus requires an oxygen tension less than air. Consequently, it is grown in a CO_2 incubator in the presence of 5% CO_2, at 37 °C, and optimal growth occurs in a pH range of 6.8-7.5.

With a few variations there are broad similarities between the metabolism of *H. pylori* and other bacteria (41). Like other organisms *H. pylori* has the enzymes of a

glycolytic/gluconeogenesis pathway and a TCA cycle. While all the genes required for gluconeogenesis are present, pyruvate kinase and phosphofructokinase are absent from the glycolytic pathway. Consequently, *H. pylori* uses the Entner-Doudoroff (ED) pathway for the catabolism of glucose. Furthermore, only one sugar transporter, a glucose/galactose transporter has been identified in *H. pylori* suggesting a limited capacity for sugar catabolism. However, *H. pylori* does have a number of amino acid transporters and may use amino acids as an energy source.

Despite being a microaerophile *H. pylori* can perform aerobic respiration with oxygen as a terminal electron acceptor. It can also perform anaerobic respiration using N-oxides and fumarate as terminal electron acceptors.

To protect itself against reactive oxygen species *H. pylori* contains a catalase, superoxide dismutase and a peroxidase. Therefore, *H. pylori* microaerophily is not a consequence of the absence of these enzymes.

1.1.2 Virulence Factors

H. pylori has a number of virulence factors that contribute to this organism's ability to colonise and survive in the harsh gastric environment. The most significant of these are the urease, adhesins BabA and SabA, the vacuolating cytotoxin VacA, the *cag* pathogenicity island (cag-PAI) that encodes a type IV secretion system and a secreted protein CagA. Motility can also be considered a virulence factor, as it is required for colonisation and persistence in the host (48) (49) (50).

1.1.2.1 Urease

H. pylori is a neutralophile but it can colonise and survive in the stomach by acid acclimation involving an urea-importing channel and two enzymes, urease and a carbonic anhydrase. All three of these proteins are required to colonise animal models (47) (37) (165).

Urease is a cytoplasmic enzyme composed of two subunits, UreA and UreB. It buffers the cyotoplasm by converting urea to ammonia. UreI is a proton-gated channel for uptake of the urea into the bacterial cytoplasm. *H. pylori* expresses the urease at higher levels than any other known bacterial species (123). Urease can account for up to 15% of the total *H. pylori* protein content. Although the UreA/B urease exhibits optimal activity at neutral pH, urease activity inside the bacterial cell increases dramatically as pH decreases. This increase is a result of opening of the UreI channel (particularly at pH 5.0 and below) to increase the uptake of urea (163) (194) and of UreA/UreB association with UreI at the inner membrane (189) (72) to increase the turnover of the ammonia-producing urease activity. The rapid production of ammonia buffers the pH of the cytoplasm.

The periplasm is buffered by the second enzyme, carbonic anhydrase, to around neutral pH by converting CO_2 to HCO_3^-, with a pKa of 6.1 (108). Thus the urease has dual roles of producing ammonia to buffer the cytoplasm and CO_2 such that carbonic anhydrase can buffer the periplasm.

1.1.2.2 Adhesins

H. pylori colonises the gastric mucosa and epithelium by binding to the mucus and epithelial cells of the gastric mucosa. Thus far two adhesins have been shown to be significant for this binding activity, blood group antigen-binding adhesin (BabA) and sialic acid-binding adhesin (SabA) (74) (106).

Adherence to the gastric lining of healthy individuals occurs via the binding of BabA to the fucosylated blood group antigens H1 and Lewis b (Leb). The H1 antigen is the carbohydrate that defines the O blood group of the ABO blood group system and the Lewis antigen is formed by the addition of a branched fucose residue to H1. Amongst *H. pylori* strains with fucosylated blood group antigen binding activity, 95% can also bind the A, B and O antigens (13).

SabA adhesin binds to sialyl-Lewis a (sLea) and sialyl-Lewis x (sLex) glycosphingolipid. Such adhesion is proposed to have a role in chronic inflammation and possibly cancer as these lipids are expressed in inflamed tissue and tumors, respectively(106).

1.1.2.3 VacA and CagA

H. pylori strains with the *vacA*$^+$, cag-*PAI*$^+$ and *babA*$^+$ genotype are associated with more severe disease in comparison to the strains negative for these loci (136). Both the VacA and the *cag* pathogenicity island (*cag*-PAI) secreted protein CagA cause changes in pathology, histology and also modify (attenuate) the host immune response to allow persistent infection (146) (24). The effects of these proteins on the host epithelium are summarised in Figure 1.1.

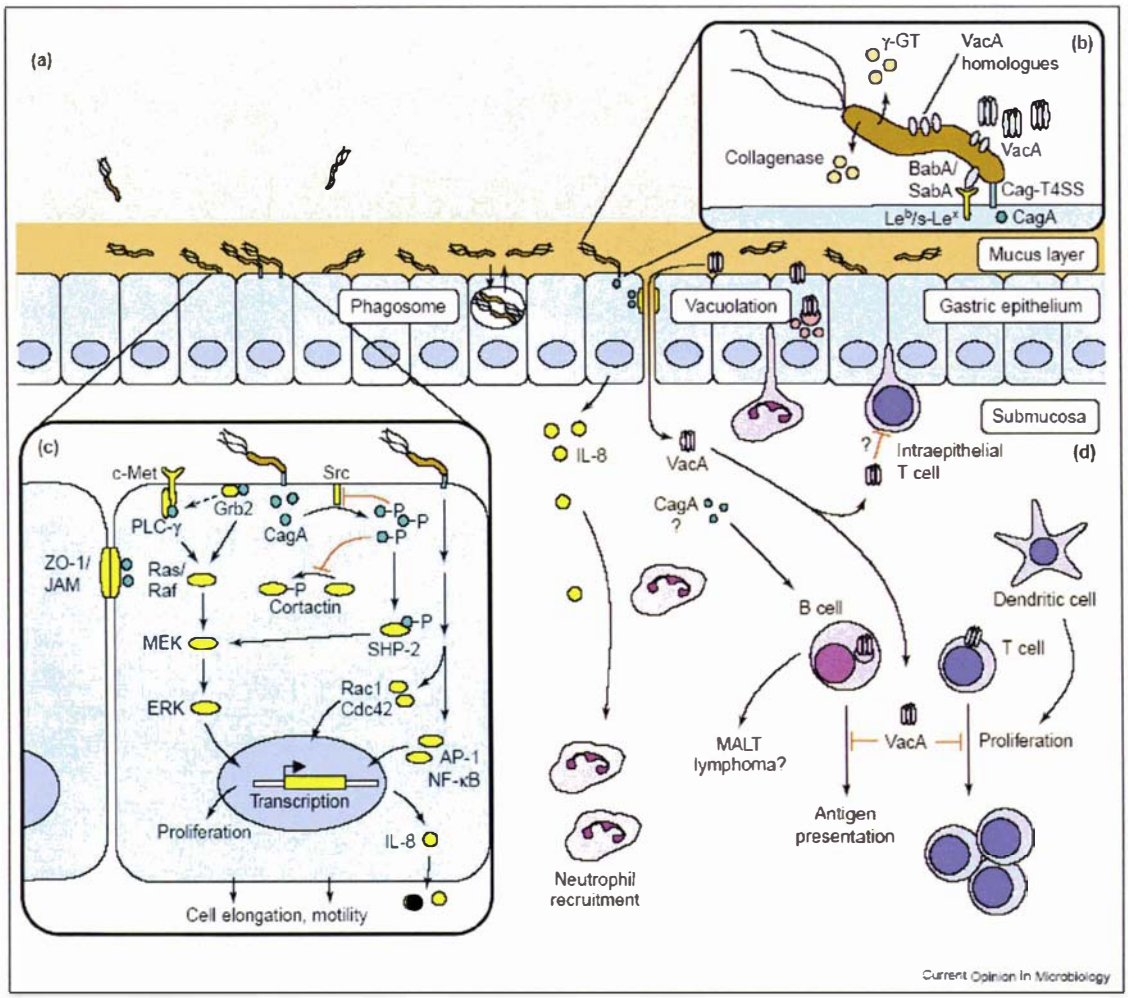


Figure 1.1. A schematic representation of the important events during *H. pylori* interaction with the host. Taken from Rieder *et al* (146). **(a).** *H. pylori* enters the mucous layer overlaying the gastric epithelial cells and adheres to the epithelial cells. **(b).** Close-up of *H. pylori* secreting virulence factors and adhering to cells using surface exposed adhesins. **(c).** Close-up of the signal transduction cascade modulated by translocated CagA. **(d).** Influence of VacA on immune cells.

Upon contact with host cells, VacA secreted from *H. pylori* inserts into the host membrane to form anion-selective hexameric channels (183) (179). These subsequently become endocytosed and cause the formation of membrane-bounded vacuoles from fusion events involving lysosomes and endosomes (124). As a consequence, VacA partially neutralises

these acidic intracellular compartments. This probably explains the observed inhibition of antigen processing by B-cells exposed to VacA and poor presentation to T-cells (125). In this manner VacA can suppress the development of an adaptive immune response and contribute to the establishment of a chronic infection. VacA can also loosen cell junctions and induce apoptosis. The molecular mechanisms of VacA have been reviewed by Montecucco *et al* (126).

The *cag*-PAI is associated with the stimulation of proinflammatory cytokine release following infection particularly IL-8 (57). The *cag*-PAI encodes a type IV secretion system that forms an extracellular needle complex which mediates translocation of CagA into the host cells (149). CagA is thus far the only known *cag*-PAI secretion substrate.

Following translocation into host cells, CagA is tyrosine-phosphorylated (171) and disrupts a number of signaling pathways leading to modification of the host immune response by inhibiting the proliferation of B-lymphocytes (184). CagA can also disrupt cell-cell adhesion independently of phosphorylation (7).

1.1.2.4 Motility

Although motility is a trait that is not attributable to a single factor it is well established as a virulence factor that is necessary for both colonisation and persistence of an *H. pylori* infection in animal models (48) (49) (50) . Furthermore, a recent animal infection study showed that chemotactic motility is necessary to colonise all the available niches i.e. to establish an infection, to maintain high numbers of bacteria and to compete against other *H. pylori* strains (180).

Further illustrating the importance of motility, *H. pylori* has also adapted the flagellum so that it is not a target for the innate immune response. Whereas the flagellin of *S. typhimurium* contains a conserved N-terminal Dl domain that is recognised by the TLR5 receptor and stimulates an innate immune response, in *H. pylori,* this region of flagellin is different from other bacteria and it is not recognised by the TLR5 receptor (8). This allows *H. pylori* to evade a TLR5-mediated initiation of the innate immune response.

1.2 Protein Export in Gram-Negative Bacteria

Protein export is essential for bacterial membrane biogenesis and communication with the extracellular environment. The export pathways of Gram-negative bacteria can be broadly classified into three groups according to the membranes crossed. The first class of export pathways facilitate the passage of proteins across the cytoplasmic membrane and includes the SecYEG/YidC pathway and the TAT ABC pathway. The second class of export pathways allows the passage of proteins from the periplasm across the outer membrane. This group includes the autotransporters, the chaperone-usher pathway, and the Type II pathway. The third class of export pathways initially allows the passage of proteins which assemble into a structure that forms a continuous passage that traverses both membranes, and then subsequent export of proteins occurs across both membranes in a single step. This group includes the Type I, III and IV secretion pathways.

In all of these export pathways there are three common elements: recognition of protein substrates, maintenance or conversion of substrates to an appropriate conformation for export, and energisation of the translocation process. It is the variation of these three elements that distinguishes the export pathways. Schematic representations of some of these export pathways are shown in Figures 1.2 and 1.3.

1.2.1 Class 1: Export pathways across the cytoplasmic membrane

1.2.1.1 The SecYEG/YidC pathway

The Sec pathway is the principal route for the insertion or translocation of proteins across the cytoplasmic membrane of bacteria, and this pathway has been shown to be essential for the viability of *E. coli* (186). Proteins recognised by this export system contain an N-terminal signal sequence consisting of a hydrophilic segment with one or more positively-

charged amino acids, followed by a hydrophobic H domain (33) (62). The export system is comprised of three main components, the SecYEG inner membrane channel, the SecB molecular chaperone and the SecA ATPase (186).

There is also an accessory heterotrimeric membrane complex associated with SecYEG, SecDF-YajC. The exact role of this complex is unknown but it may mediate the interaction of YidC with SecYEG. YidC is involved in the biogenesis of both sec-dependent and independent membrane proteins (186).

Following translation the N-terminal signal of proteins destined for secretion is recognised by SecB and trigger factor (TF) and targeted to the SecYEG translocon. Alternatively, proteins with a more hydrophobic signal sequence are recognised by the Signal Recognition Particle (SRP) chaperone and the SRP receptor, FtsY, and targeted to SecYEG translocon co-translationally. The SRP pathway was reviewed by Luirink *et al* (102) and it will not be considered further here.

SecA

SecA energises protein translocation by the hydrolysis of ATP, but the structure of this enzyme, and the mechanism by which energy release is coupled to translocation is different from other export ATPases. Although SecA has been visualised as tetrameric and higher order structures, it is currently thought to function as a dimer (44) (200) (80) (45) (191). This is in contrast to other ATPases such as the type IV ATPase VirB11 and the flagellar ATPase FliI, both functioning as hexamers (31) (206). Each SecA monomer is a 102 kDa protein consisting of three domains. An N-terminal motor domain, a Substrate Specificity Domain (SSD) and a C-terminal domain. The N-terminal motor domain is referred to as the DEAD motor due to the conservation of the DEAD amino acid motif also known as Walker box B, and this domain is homologous to the ATPase domains of the DEAD box nucleic acid helicases. The DEAD motor consists of two subdomains that have RecA-like folds, NBD (nucleotide binding domain) and IRA2 (intramolecular regulator of ATPase). A nucleotide binding cleft forms between these two subdomains. ATP is bound in a cleft

between the two RecA-like subdomains within the motor domain. The structure of SecA has been reviewed by Vrontou *et al* (190). While ATP binding involves the Walker A and B motifs of the NBD and occurs between the two RecA-like domains, similar to other ATPases such as F_1-ATPase, the tertiary structure of these enzymes is quite different. Within the F_1-ATPase each F_1 α- and β-subunit only contains one RecA-like domain as opposed to two, and the N- and C-terminally attached domains of F_1 bear no resemblance to the structure of the SSD and C-terminal domains of SecA (1). The DEAD motor domain also contains a binding site for the signal sequence of preproteins and the SecYEG translocase. The C-terminal 22 amino acids contain the SecB binding site.

SecA-mediated pre-protein translocation

Following the formation of a translocation pre-initiation complex at the inner membrane, consisting of SecYEG, SecA and SecB-preprotein, SecA binds ATP. ATP binding causes SecA to adopt a more extended conformation that drives insertion of SecA into the membrane. As this insertion begins SecB passes the preprotein to SecA and co-insertion of 20-30 amino acids of the preprotein occurs. Subsequent ATP hydrolysis then partially dissociates the preprotein from SecA and leads to SecA deinsertion from the membrane. Repeated cycles of SecA binding ATP, inserting into the membrane, hydrolysing ATP and deinsertion from the membrane eventually lead to threading of the preprotein through the SecYEG translocase. If SecA becomes dissociated from SecYEG, proton motif force (PMF) is capable of energising the final steps of the preprotein translocation (190).

After translocation is complete, the signal sequence is removed by signal peptidase and the protein folds into its native conformation and remains in the periplasm. Alternately it may be subsequently exported further by insertion into or by translocation across the outer membrane (190).

The linear threading of the preprotein through the inner membrane by cycles of insertion and deinsertion of SecA is in stark contrast to the rotary mechanism of type III and FliI hexameric ATPases being proposed in this study.

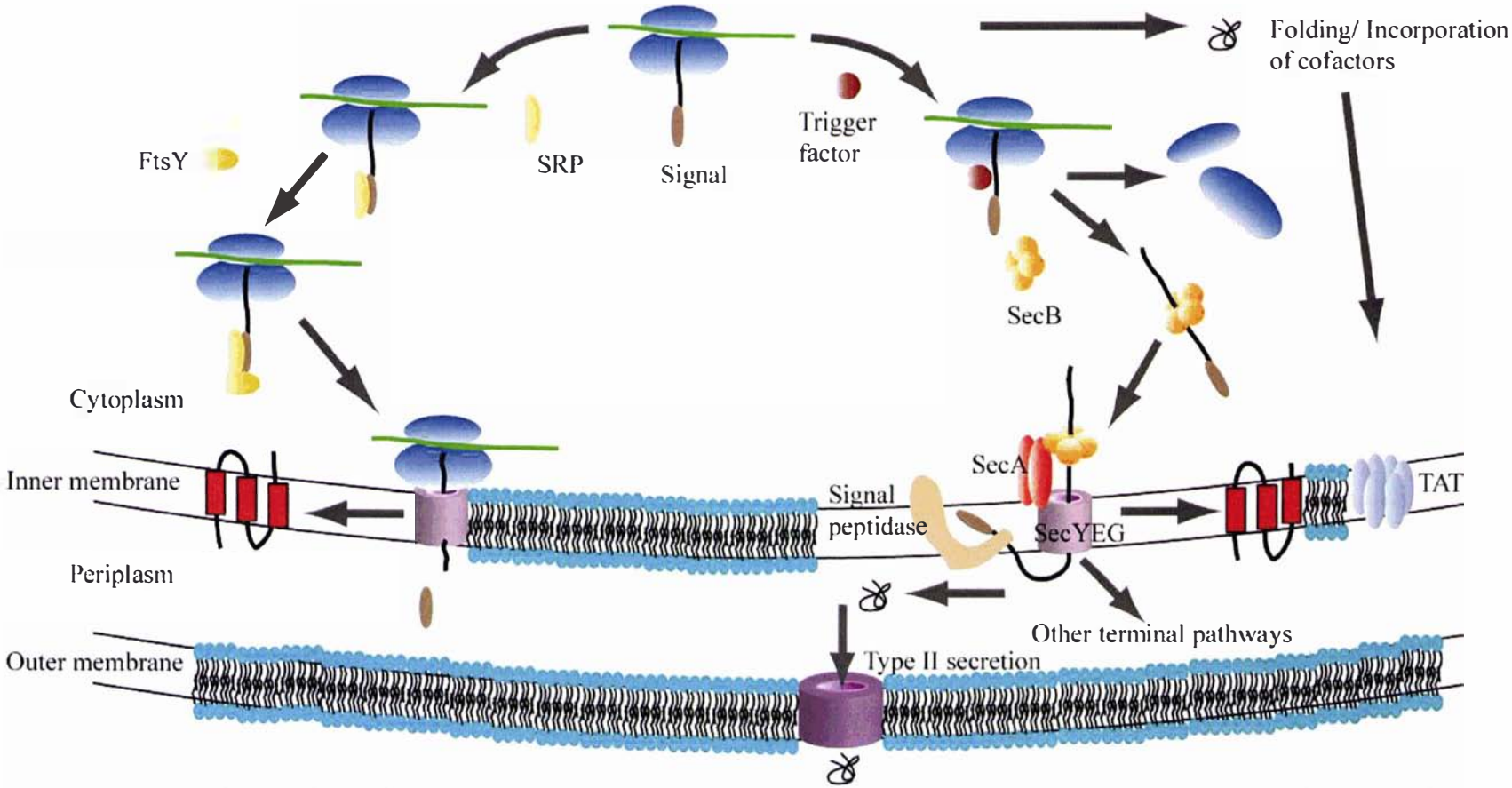


Figure 1.2 Sec pathway of translocation across the inner membrane. Based on a figure from Wickner *et al*, 2005 (198), with some modifications. Membrane proteins are targeted to SecYEG co-translationally, other proteins are recognised by SecB post-translationally and targeted to the SecA pre-initiation complex. The alternative TAT export pathway is also shown.

1.2.1.2 The TAT Export Pathway

The TAT (twin arginine translocation) export pathway functions to translocate folded preproteins across the inner membrane of Gram-negative bacteria (148). The proteins translocated are often those that require the insertion of heavy metal cofactors in the cytoplasm and therefore must be folded prior to export. An example is *E. coli* NapG, a member of the MauM family of ferredoxins (20). Some TAT pathway substrates do not require metal cofactor insertion, and these may simply fold too rapidly for export via other pathways.

Export substrates of the TAT pathway have traditional signal sequence similar to the one used by the Sec translocon (62). The distinguishing features of TAT signal peptides are an H region of lower hydrophobicity and an N-region sequence motif SRRXFLK (20) (35). The arginines of the motif give the pathway its name as they are invariant; the other residues occur in more than 50% of the TAT signal peptides (20). If either arginine is substituted by lysine, the subsequent export of TAT substrates is less efficient (170).

The TAT proteins

The TAT translocon is formed by three proteins encoded by the *tatABC* operon (197). The TatBC complex is currently thought to be a recognition complex. TatC contacts a discrete area around the consensus sequence (SRRXFLK) whereas TatB contacts the entire TAT signal sequence and the adjacent parts of the mature protein (3). TatA is assumed to form the pore of the TatABC translocase (155) (141).

The mechanism of protein translocation via the TAT pathway is still very poorly understood. As the TAT substrates come in a range of sizes it is hypothesised that the translocon is adaptive and the pore size can be adjusted to accommodate different size substrates. The energy for translocation is assumed to come from proton motive force.

1.2.2 Class 2: Export pathways across the outer membrane

1.2.2.1 Type II or the Main terminal branch of the General Secretory Pathway (GSP)

The aforementioned Sec system by which proteins containing an N-terminal signal peptide are translocated across the bacterial inner membrane is also known as the General Export Pathway. This is the first step of the General Secretory Pathway for proteins that are destined to be secreted from the cell. Following the cleavage of the signal peptides, proteins entering the periplasm via the Sec pathway often enter the Type II secretion pathway (reviewed by Filloux (56)). This is the main terminal branch of the GSP, but additional terminal branches have been identified such as the chaperone-usher pathway, the autotransporter (AT) and Two-Partner Secretion (TPS). Proteins secreted by the GSP and other terminal branch pathways are folded, and the secretion process is specific, but the secretion signal that targets these proteins to the outer membrane is not well defined.

The machinery of the Type II secretion system consists of three substructures, an inner membrane platform, a pseudopilus that spans the periplasm and a secretin in the outer membrane (56). The pseudopilus is so named as it is composed of pseudopilins which contain a typical pilin N-terminal leader peptide that is processed before being helically packed into a pilus structure. Pili are fibrous organelles that can mediate attachment of bacteria to host tissue. Each pilus type differs slightly in structure (53). Unlike the pili, pseudopili are not extruded from the cell surface in wild-type cells. A schematic of the proposed structure of the type II secretion machinery is shown in Figure 1.3 below.

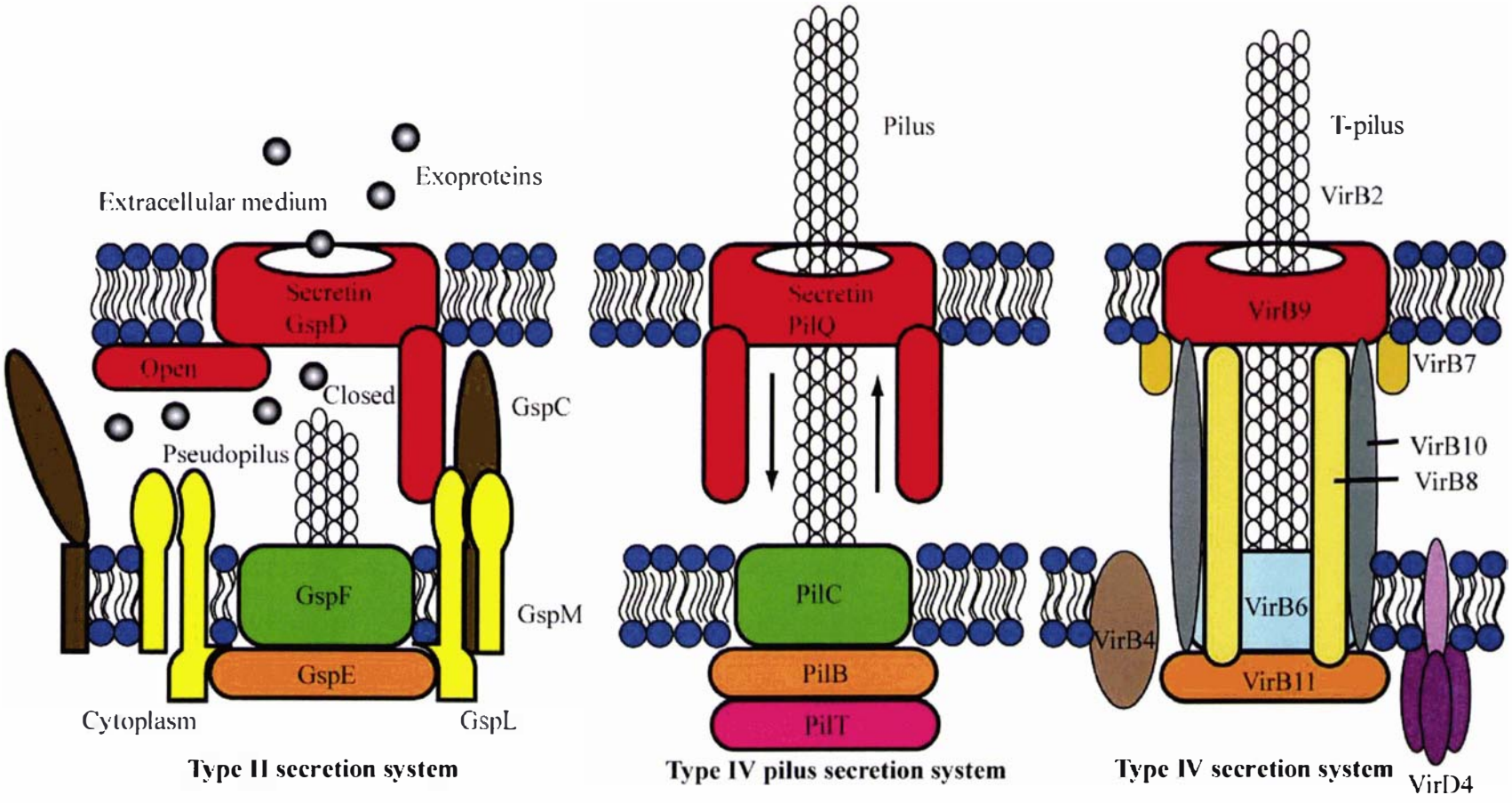
The inner membrane platform consists of three inner membrane proteins GspF, L and M. The traffic ATPase, GspE, associates with the membrane via interactions with GspL and GspF, and energises the translocation of pseudopilins across the cytoplasmic membrane. The manner in which GspE couples ATP hydrolysis to psuedopilin translocation remains to

be determined, but the crystal structure of *V. cholerae* type II ATPase, EpsE, has offered some insight (147). The protein consists of an N-terminal and a C-terminal domain. The closest structural homologue of these domains are the N- and C-terminal domains of the type IV ATPase, VirB11 of *H. pylori* (HP0525) (206). Based on this homology EpsE is expected to assemble as a hexamer, with the N- and C-terminal domains forming rings separated by a flexible linker. This assembly was proposed to energise and gate the translocation of pseudopilins across the inner membrane (see the VirB11 discussion below).

The pseudopilus

The translocated pseudopilins assemble into the pseudopilus in the periplasm, but these pseudopili are not extruded onto the surface of wild-type cells. However, when the pseudopilus subunit GspG is overexpressed in *P. aeruginosa*, thick pseudopili can be observed on the surface of cells (46) (187). It has been proposed that the extension of the pseudopilus through the outer membrane pore formed by GspD pushes exoproteins through the secretin into the extracellular environment.

Figure 1.3. Models of the type II, type IV pilus (TFP), and type IV secretion machinery. Pili extruded by the type IV pilus and the type IV secretion systems have been observed. The type II pili are only observed in mutants. The models are based on interaction and localisation studies. This figure is based on figures presented in: (30, 40, 56).



1.2.2.2 The Chaperone-Usher Pathway

Gram-negative bacteria use the chaperone-usher pathway as another terminal branch of the general secretion pathway, following Sec-dependent export, to secrete proteins across the outer membrane. This pathway is best characterised in the context of Type I (*fim* gene cluster) and P pilus (*pap* gene cluster) assembly in *E. coli* (Reviewed by Sauer and colleagues (156). Pilus assembly requires a periplasmic chaperone, PapD/FimC, and an outer membrane usher, PapC/FimD. The chaperone assists the folding of the pilins and then prevents their premature association while they are targeted to the usher (16) (157). Folded pilin subunits are assembled into the pilus and secreted through the usher, which forms a pore in the outer membrane (98) (129) (158). The secretion of the pilus is thought to be energised by conformational changes during pilus assembly.

1.2.2.3 The Autotransporter and Two-partner secretion pathways

Following translocation across the inner membrane via the Sec pathway, both the Autotransporter (AT) and Two-partner secretion (TPS) pathways facilitate the secretion of substrates across the bacterial outer membrane. Preceding the typical Sec secretion signal, autotransporter and component A of the TPS pathway contain an extended signal consisting of a sequence rich in aromatic and hydrophobic amino acids of approximately 25 residues that terminates with the consensus motif LIAVSELAR (67). The function of this additional sequence is unknown. Unlike the other export pathways such as the type II, III and IV pathways, there is no complex secretion machinery assembled in these pathways. Once in the periplasm, AT protein passenger domains are secreted via their own transporter domain in the C-terminus of the protein. The transporter domain forms a pore in the outer membrane through which the N-terminal substrate can pass (137). TPS systems consist of two components, the secretion substrate TpsA and the partner protein TpsB which forms a pore in the outer membrane for secretion and in some cases its role is the activation of TpsA (85) (203).

It is unknown how the secretion process is energised in these pathways, but as the passenger domain and TpsA are thought to fold during secretion into the extracellular environment, this may provide energy to power secretion. Following secretion, many passenger domains and TpsA proteins undergo proteolytic maturation which, in the case of the AT pathway, releases the passenger domain from the translocator domain. The mature protein may then be released into the extracellular milieu or remain associated with the membrane.

Proteins secreted by the AT and TPS pathways have diverse functions. AT passenger proteins include the vacuolating cytotoxin VacA of *H. pylori* and *Yersinia enterocolitica* adhesin YadA. TpsA proteins include the HMW1 adhesin from *Haemophilus influenzae* and the cytolysin of *Proteus mirabilis*. The AT and Tps pathways were reviewd by Jacob-Dubuisson et al (76).

1.2.3 Class 3: Export pathways that cross both membranes in a single step

1.2.3.1 The Type IV secretion pathway

The type IV secretion systems translocate DNA and protein substrates across the bacterial cell envelope. The archetype type IV secretion system is the virB/virD4 system found in the plant pathogen *Agrobacterium tumefaciens* which uses this secretion system to translocate a T-DNA protein complex into host cells causing the crown galls seen on infected plants. Homologous type IV secretion systems are also found in bacterial conjugation systems and human pathogens such as the causative agent of whooping cough *Bordetella pertussis*, the intracellular pathogens *Brucella suis*, *Legionella pneumophila* and *H. pylori*. In these pathogens the type IV secretion system translocates effector proteins rather than DNA into eukaryote host cells. The Type IV secretion pathway has been reviewed by Yeo et al, Ding et al and Christie (207) (40) (30).

The structure of the type IV secretion machinery is poorly characterised. The hypothesised

architecture is based on the prediction of transmembrane spans, interaction studies, and fractionation studies that localise the various Vir proteins to the outer membrane, periplasm, inner membrane or cytoplasm. Some Vir proteins co-partition with both the inner and outer membranes. This suggests that a transenvelope structure is formed similar to the type III needle complex or the flagellum basal body (207). However, such a structure has not been observed yet. A proposed assembly of the secretion machinery is presented in Figure 1.3, for comparison with the type II and TFP (not considered here) secretion machinery.

VirB11, a cytoplasmic protein peripherally associated with the inner membrane

Mutagenesis studies suggest that VirB11 energises pilus assembly and the assembly or function of the type IV secretion machinery (Figure 1.3; (153)). VirB11 forms a homohexameric ring structure with a central 50 Å cavity similar to type II ATPases, but quite obviously different from SecA. The VirB11 hexamer appears to be made of two stacked rings in electron micrographs (88). In fact the first of the two rings is formed by the N-terminal domain which has a unique fold. The other ring is formed by the C-terminal domain which has a RecA-like fold. Nucleotide is bound at the interface of the two domains. The structure of the protein has been determined in the absence of nucleotide and in the presence of ADP and non-hydrolysable ATPγS (206) (159). In these studies nucleotide is bound at the interface of the N- and C-domains, and binding of nucleotide is associated with a swiveling of the N-terminal domain which converts the complex from an open asymmetric hexamer to a more closed and compact symmetrical hexamer. Furthermore, ATP binding and hydrolysis influences the association of VirB11 with the inner membrane as a Walker A mutant of VirB11 binds the inner membrane more tightly (144).

There are two other ATPases involved in type IV secretion, VirB4 and VirD4. The exact role of these proteins in secretion is unknown. VirD4 may be involved in the recognition

and delivery of substrates to the secretion machinery. Unlike VirB11, VirD4 and VirB4 are inner membrane proteins rather than being peripherally associated with the membrane (30).

Secretion by the Type IV Secretion System machinery

Secretion substrates such as VirE2 contain a C-terminus which is rich in positively charged amino acids. Furthermore, the C-terminal domain of VirE2 can drive the secretion of chimeric fusion proteins. This indicates that the Type IV secretion signal is formed by C-terminal positive charges and VirD4 can recognise the C-terminus of proteins to be secreted. Interestingly, the Type IV secretion system of *B. pertussis* does not contain a VirD4 homologue and the secretion substrate of this system, the pertussis toxin, contains a typical general export pathway (GEP) N-terminal signal peptide. This suggests that the *B. pertussis* Type IV secretion system operates by a different mechanism from that of *A. tumifaciens* (40).

Following the recognition of the Type IV secretion substrate by VirD4 and the entry of this protein into the secretion channel, the Type IV pilus is assembled and extruded. It does not function as a conduit (19), and is currently assumed to draw the target cell closer. The assembly of the pilus is energised by VirB11. The crystal structure of VirB11 bound to ATPγS, ADP and free nucleotide has allowed a model to be generated to explain the coupling of the ATP hydrolysis cycle to protein translocation. This is summarised in Figure 1.4 below (159). Each N-terminal domain (NTD) of the hexamer can move about a flexible linker between the NTD and C-terminal domain (CTD), to form an opening 50 Å in diameter. In contrast the CTDs combine to form a closed hexameric ring with a "six-clawed" grapple around a central pore of approximately 10 Å. Following ATP binding, the NTDs adopt a more rigid closed conformation. Then by an unknown mechanism ATP hydrolysis is assumed to generate the necessary mechanical force to drive translocation of an export substrate (159) (206).

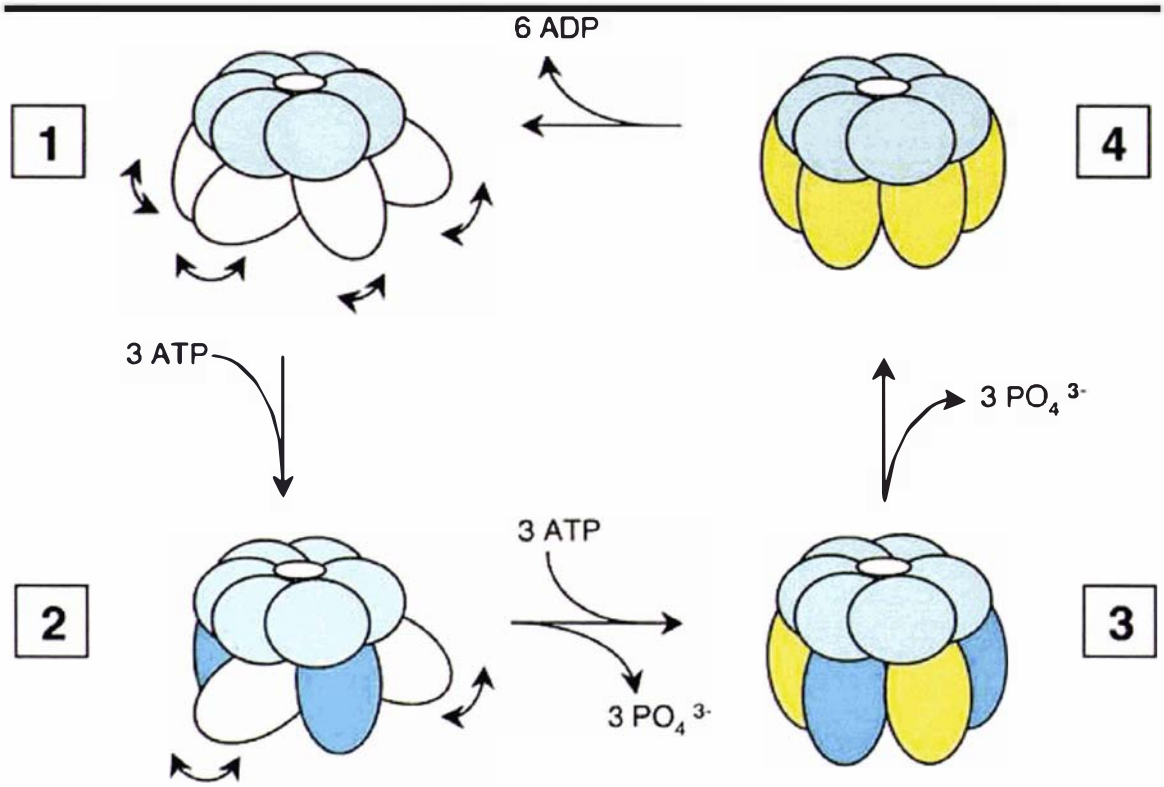


Figure 1.4. Model for the mode of action of VirB11 ATPases. Taken from Figure 6 of Savvides, et al (159), the N-terminal domains are pink in panel one the C-terminal domains are blue. The N-terminal domains become locked into a rigid conformation by binding ATP.

Following assembly of the pilus, the substrate is secreted and it enters the target cell, this may also be energised by VirB11. However, the pilus is not used as a conduit to transfer the secretion substrate to the target cell. This is deduced from the phenotype of mutants that do not produce pili, but still secrete substrates (19).

1.2.3.2 Type I Secretion

Type I secretion systems secrete cognate substrates to the extracellular environment in a single step without a periplasmic intermediate. The secretion signal is C-terminal, however it varies and it is hence poorly understood. The substrates vary in function and include toxins, proteases and phosphatases. Despite unrelated primary structures and functions, these substrates have some common characteristics. Many contain glycine-rich repeats, few or no cysteines and with the exception of one substrate, they all have an acidic pI around 4 (39).

The secretion machinery consists of three components; an ATP-binding cassette (ABC) protein associated with the inner membrane, a membrane fusion protein (MFP) that is anchored in the cytoplasmic membrane and traverses the periplasm, and an outer membrane channel of the TolC class (9) (87) (10). The ABC protein recognises export substrates and energises export (86) (21) while the MFP connects the ABC protein and TolC (181). The MFP may also open TolC following recognition of the substrate. Although substrates are exported across the cytoplasmic membrane in an unfolded state (38), experiments suggest that proteins fold within the TolC channel (185).

The ABC protein functions as a dimer and consist of two nucleotide binding domains (NBDs) and two transmembrane domains (TMDs). Thus, unlike other export ATPases considered here it is an integral membrane protein rather than being peripherally associated with the inner membrane (69) (161).

1.3 Type III virulence export system

A number of Gram-negative bacterial pathogens use type III secretion systems (TTSSs) to transfer effector proteins to the eukaryotic cells of mammalian hosts, where they have a variety of effects. Intracellular pathogens such *Salmonella* and *Shigella* use type III secretion systems to invade and multiply within host cells (60) (113), whereas the extracellular pathogen *Yersinia* causes changes in the cytoskeleton, apoptosis of host cells and ultimately avoids phagocytosis (34). There are a number of similarities between the type III virulence secretion machinery and the bacterial flagellum. These include the ultrastructure of the transenvelope secretion machinery, the ordered assembly of this machinery and the components of the export apparatus. These similarities have led to the classification of the flagellar export system as a type III pathway.

The TTSSs of *Shigella* and *Salmonella* form the so-called needle complex that traverses both bacterial membranes. In this introduction, the emphasis will be on the *Shigella* or *Salmonella* TTSS needle complex (89). As viewed by electron microscopy the ultrastructure of this needle complex is remarkably similar to the basal body complex of the flagellum (109) (Figure 1.5). Furthermore, recent studies of the *Shigella* needle have shown that it bears a close resemblance to the flagellum filament, with the MxiH subunit proteins packing in a helical manner, 5.6 units/turn. The needle complex consists of four distinct substructures. Two inner membrane rings 40 nm in diameter formed by PrgH and PrgK, two outer membrane rings formed by InvG a protein of the secretin family, a central rod formed by PrgJ that connects the two pairs of rings, and an external needle extension formed by PrgI that is 80 nm in length and 8 nm in diameter (89) (91) (109). A pore of approximately 2 nm runs through the centre of the structure.

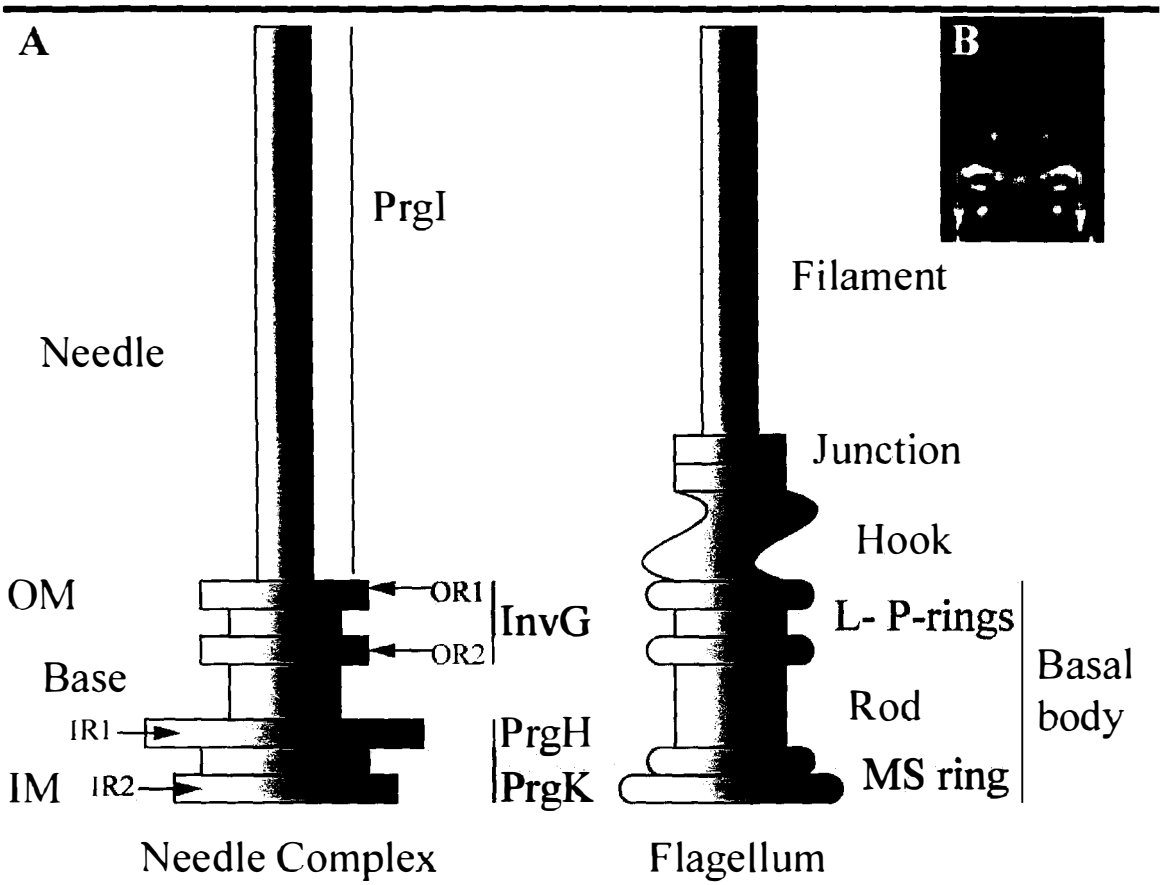


Figure 1.5. A. Schematic comparison of the *S. typhimurium* TTSS needle complex, and the *S. typhimurium* flagellum. The schematic representation of the needle complex is from Marlovits *et al* (109). OR1 and 2 are rings associated with the outer membrane and are composed of InvG. IR1 and IR2 are rings associated with the inner membrane and they are composed of proteins PrgH and PrgK. **B.** Electron micrograph of the needle complex (109).

Mutation of the structural counterparts in these two systems block their assembly at the equivalent steps, further emphasising the similarities of the type III virulence and flagellum export pathways. These observations indicate the assembly pathway of the needle complex proceeds with the formation of defined substructures like the flagellum (90) (176). The first components of the needle-complex to be assembled are the base substructure consisting of the inner membrane PrgH-K rings, and the outer membrane InvG ring. These proteins are exported in a sec-dependent manner. The PrgH-K rings act as a mounting plate for the assembly of other proteins to build the complete type III export apparatus. Subsequent protein export to assemble the inner rod and needle and the secretion of effectors from the needle, proceeds via a sec-independent type III mechanism. Export via this pathway is energised by an ATPase (InvC) of the export apparatus.

Perhaps most significantly, a number of the export apparatus components of the virulence associated TTSSs have significant sequence similarities with the flagellar export system, including the ATPase InvC (Table 1.1). This led to the assumption that virulence TTSS evolved from the flagellum TTSS. However, phylogenetic analysis of four of the core components of the Type III and flagellum export apparatus, (YscN/FliI, LcrD/FlhA, YscR/FliP, YscS/FliQ), placed doubt on this assumption. In contrast,it suggested that they both evolved from a common ancestral structure (65).

Table 1.1. Proteins of the type III virulence secretion systems of *S. flexneri*, *S. typhimurium* and the *S. typhimurium* flagellar secretion system that share sequence similarity.[a]

Shigella flexneri	*Salmonella* SPI-1	Flagellar proteins	Localisation
MxiA	InvA	FlhA	Inner membrane
Spa47	InvC	FliI	Cytoplasm/inner membrane
Spa33	SpaO	FliN	Inner membrane (SpaO is secreted)
Spa24	SpaP	FliP	Inner membrane
Spa9	SpaQ	FliQ	Inner membrane
Spa29	SpaR	FliR	Inner membrane
Spa40	SpaS	FlhB	Inner membrane
MxiJ	PrgK	FliF	Inner membrane /needle base
MxiD	InvG		Outer membrane /needle base
MxiH	PrgI	FlgE	Needle
Spa32	InvJ	FliK	Secreted, needle length control
MxiG	PrgH		Needle base
MxiI	PrgJ		Needle, cap protein

[a] This table is an abbreviated version of Table 1 from He *et al* (66).

1.3.1 The Type III virulence secretion signal

There is specific recognition of the type III export substrates by the export apparatus. The type III secretion signal has been most extensively studied in the virulence export systems, but parallels with putative flagellar signal sequences are emerging. There is no clear consensus sequence motif and there is some controversy regarding the nature of the signal. It has been suggested that the signal may be within the mRNA, because if a single synonymous change is made in codon three of *Yersinia yopQ*, it destroys the capacity of the first 10 codons to direct secretion via the type III pathway when fused to Npt (143). At the same time there are examples of dramatic changes in the mRNA sequence that leave protein sequence intact and secretion still occurs. For example, when 17 of the 27 nucleotides in codons 2-10 of the YopE signal are changed export of this protein is not affected (100). Regardless of its nature, the signal has been localised to approximately the first 15 amino acids or codons of type III virulence effectors (11) (160) (168).

1.3.2 Substrate-specific chaperones of the Type III virulence secretion system

An additional type III secretion signal may be present in the type III substrate specific chaperones. For example, a mutant of the secreted protein YopE lacking the N-terminal 15 residues is exported in a SycE (chaperone) dependent manner (29). Targeting to the export apparatus is just one of the roles suggested for substrate specific chaperones, others include passive protection, secretion competence and unfolding. Holding the protein in an unfolded state is particularly important as the pore in the centre of the needle structure is only 2 nm in diameter.

1.4 The Type III Flagellum Export System

The flagellum is a cell surface organelle that mediates bacterial motility. Proton motive force or sodium motif force provide the energy for motor rotation. Rotation of the motor leads to the rotation of a rod which functions as drive shaft, then to rotation of a flexible universal joint termed the hook. This in turn leads to the rotation of the helical filament that results in thrust to propel the cell. The motor rotation and consequently the filament rotation will occur in a counter-clockwise direction if the chemoreceptors for attractants in the plasma membrane are occupied, resulting in directional bacterial movement or "swimming". When the occupancy of these receptors drops the motor rotation changes to a clockwise direction and causes the cell to tumble randomly rather than swim.

The model organism for the study of bacterial motility and flagellum protein export is *Salmonella enterica* serovar *typhimurium*, commonly referred to as *S. typhimurium* in the literature. As can be seen in Figure 3.1 of the results chapter, the majority of the structural genes of the *S. typhimurium* flagellum have been annotated in the genome of *H. pylori*. There are some significant differences in *H. pylori* flagellar gene regulation that will be discussed later.

1.4.1 Morphological assembly pathway and structure of the bacterial flagellum

The flagellum organelle consists of several substructures. A basal body, consisting of the MS ring in the cytoplasmic membrane; the attached rod that traverses the periplasm, and L-(LPS) and P-(peptidoglycan) rings that assemble around the rod in the outer membrane; the hook containing the hook protein and the hook-associated proteins; and the filament consisting of flagellin and the filament capping protein.

Like the assembly of the type III virulence needle complex, the first step of flagellum assembly involves the Sec-dependent export and assembly of two inner membrane rings.

Twenty six copies of FliF oligomerise in the cytoplasmic membrane to form the M and S rings with a central pore 10 nm in diameter (82). The MS ring was thought to be a passive mounting plate, but recent structural studies have revealed that the FliF pore is tightly closed by domains contributed by FliF (177) (178). Therefore, it seems that FliF also has a gating function. Unlike the needle complex, the flagellum rotates. This requires the mounting of the switch complex or C-ring on the MS ring in addition to the export apparatus.

The rotor/switch complex is formed by FliG, FliM and FliN. This complex associates with the cytoplasmic membrane via an interaction with the MS ring mediated by FliG and this interaction is independent of other flagellar proteins (92) (103). The interaction involves the N-terminal domain of FliG. The C-terminal domain of FliG is essential for flagellar rotation (101) and it contains some conserved charged residues that are proposed to interact with the cytoplasmic loop of the stator protein MotA, an inner membrane protein peripherally associated with FliG/M/N (Figure 3.1) (210). FliG also interacts with FliM (112), a protein important for control of CW/CCW switching of rotation of the flagellum (167) following the binding of chemotaxis response regulator CheY-P in a phosphorylation dependent manner (152). FliM in turn interacts with FliN. The function of FliN in switching is unknown (75) but it is essential for flagellum assembly (188).

The export apparatus is formed by the integral membrane proteins FlhA, FlhB, FliO, FliP, FliQ, FliR and the soluble components FliI, FliH, FliJ, FliS, FliT, and FlgN (90) (117) (134). The integral membrane components are assumed to assemble in a patch of specialised membrane within the pore of the MS ring. This is supported by the experiments that localise FliP and FliR to the basal body (51) and mutations in the FlhA integral membrane domain that suppress mutations in FliF (83). In spite of this evidence, an export apparatus complex has not been isolated and a recent structure of the basal body suggests that the FliF pore is closed or very narrow and it is difficult to see how the export apparatus could be accommodated (178). Following completion of the export apparatus, assembly of the rod, hook and filament proceed via the type III pathway, in a proximal to distal fashion. The structures of the hook and filament have been determined. Although the hook is

flexible and the filament is rigid, both are tubular structures consisting of 11 helical protofilaments formed from the hook and flagellin subunits respectively (209) (154). The rod, hook and flagellin subunits are assembled following the translocation and diffusion of unfolded proteins through the central 2 nm pore in the central channel. The translocation of these proteins is energised by FliI.

1.4.2 The Flagellum Export Pathway

As can be seen in the previous section, the flagellum is a complex structure involving the coordinated export and assembly of 12 axial proteins: FliE (rod adaptor), FlgB (rod), FlgC (rod), FlgF (rod), FlgG (rod), FlgJ (muramidase), FlgD (hook cap), FlgE (hook), FlgK (hook-filament junction protein), FlgL (hook-filament junction protein), FliD (filament cap), FliC (flagellin, FlaA and B in *H. pylori* Figure 3.1). As shown in the mutagensis studies that defined the assembly pathway, these proteins are assembled into substructures in a defined order. This presents a logistical problem for the export apparatus as there are multiple different proteins to be exported in various copy numbers, and at defined times. For example, export of 120 copies of the hook subunit precedes the massive export of 20,000 copies of the flagellin to form the filament (104). Furthermore, these proteins must be in a partially-unfolded secretion competent state as the internal diameter of the hook, filament and presumably the rod through which the axial proteins are exported is only 20 Å. Finally, this export must be energised.

To deal with this logistical problem the flagellum has multiple levels of regulation that include transcriptional control of export substrates, and translational control and recognition of substrates by the export apparatus. The transcriptional control coordinates gene expression with the stage of assembly; this ensures that only some of the flagellar proteins are present in the cytoplasm competing for export prior to the next morphological check-point. This gene regulation is different in *H. pylori* as compared to *S. typhimurium*. Following transcription of an export substrate gene, flagellum-specific chaperones are involved in both the regulation of translation and stabilisation of partially folded flagellum proteins in a secretion competent state prior to docking with the export apparatus. At the export apparatus, recognition presumably involves a poorly defined N-terminal signal and possibly a chaperone or a substrate-chaperone complex signal. The interaction of the export

apparatus with the rod/hook subunit-chaperone complexes changes the recognition interface of the apparatus allowing it to recognise and export the late filament-type protein-chaperone complexes, and consequently change the specificity of the export process following hook completion.

1.4.3 Flagellar Gene Regulation

The regulation of bacterial flagellar gene expression is best characterised in *S. typhimurium*. The flagellar genes of *S. typhimurium* are clustered in three regions of the chromosome and these clusters represent three levels of an expression hierarchy (105). At the first level of the hierarchy class 1 genes *flhC* and *flhD* of the master operon are expressed. The heterotetrameric complex $FlhD_2C_2$ binds to an imperfect palindrome that contains two inverted repeats within the promoters of class 2 flagellar genes and activates their transcription (99) (32). The class 2 genes encode the proteins of the flagellum export apparatus and the basal body in addition to two regulatory proteins, FlgM and σ^{28} (RpoF). σ^{28}, encoded by *fliA*, is necessary for the transcription of class 3 genes in the subsequent level of the hierarchy and it is negatively regulated by FlgM, an anti-sigma factor which binds σ^{28} and prevents the interaction of the RNA-polymerase-σ^{28} holoenzyme with class 3 promoters (135) (94) (133). Following expression of the class 2 genes and export and assembly of the basal body and hook, FlgM is exported through the BBH (basal-body hook) to the extracellular medium (73). This relieves the inhibition of σ^{28}-dependent transcription of the class 3 flagellar genes. The class 3 genes encode the so-called "late" flagellar genes, the HAPs (hook-associated proteins), the filament, the hook capping protein and the chemotaxis and motor proteins (105).

The regulation of flagellar gene expression of *C. crescentus* has also been well characterised. As opposed to *S. typhimurium* in which the flagella are peritrichous (distributed around the periphery of the cell) the single flagellum of *C. crescentus* is polar (located at one of the poles of the cell). Different regulation in the two bacterial species may be representative of differences between the gene expression in polar vs. peritrichous flagellated micro-organisms. The gene expression hierarchy of *C. crescentus* consists of four levels. Class 1 of the hierarchy consists solely of *ctrA*, a regulator that controls the σ^{70}

dependent transcription of class 2 genes in response to signals associated with the progression of the cell cycle. Class 2 genes encode the MS ring complex and the export apparatus in addition to FlbD, an activator of σ^{54}-associated transcription of class 3 and 4 genes. The class 3 genes encode the hook and basal body proteins and class 4 genes encode the three flagellin genes of the flagellar filament (201) (202).

H. pylori has 4-6 polar flagella. Interestingly, the regulation of *H. pylori* flagellar gene expression is not alike that of *C. crescentus*, but appears to be unique. The *H. pylori* regulatory pathways are summarised below in Figure 1.6 taken from a genome-wide analysis of flagellar gene expression by Niehus and colleagues (130). Although there is a clear three-tier expression hierarchy differentiated by different sigma factor requrements, there is also an intermediate class of genes that are transcribed by both sigma factor σ^{28} and σ^{54}. Furthermore, *H. pylori* regulatory cascade contains some regulatory proteins homologous to those of *S. typhimurium* and *C. crescentus*, such as a σ^{54}-associated transcription activator, FlgR of the NtrC family and the anti-sigma factor FlgM. However, the place of these proteins in *H. pylori* regulatory cascade and the manner in which the gene expression hierarchy is regulated appears to be unique.

At the first level of the hierarchy, class 1 genes are expressed from a σ^{80} promoter in response to an unknown signal. The class 1 genes encode chemotaxis and motor proteins, the MS ring protein FliF, the L- and P-ring proteins FlgH and I respectively, the rod proteins, proteins of the export apparatus such as FliH, FliI, and FlhB, and regulatory proteins such as sigma factor σ^{54}, FlgR, FlgS, and FlhA. FlhA is a part of the export apparatus, but it also appears to be a master regulator. A knock-out mutant of *flhA* exhibits decreased transcription of class 2, 3 and intermediate *H. pylori* flagellar genes. Surprisingly, in an *flhA/flgM* double mutant the transcription of these genes is upregulated, suggesting the involvement of FlgM in a negative feedback mechanism of regulation when the flagellum structure is compromised by the absence of FlhA. The transcription profile of an *flhF* and an *flhF/flgM* double mutant is similar to that of the *flhA* and *flhA/flhM* mutants (130).

The class 2 genes include the *flgK* and *L* (hook-associated proteins or HAPs), the hook protein *flgE*, the flagellin *flaB*, a *flgJ*-like gene (muramidase) and proteins that may be involved in the synthesis of the lipid sheath of the flagellum. Intermediate genes include the rod proteins, the flagellar cap protein *fliD*, and the flagellum-specific chaperones *fliS* and *fliT*, *flhF* and *flgM*. Class 2 and intermediate gene expression is activated by the phosphorylation of FlgR of the NtrC family by an NtrB-like histidine kinase FlgS. While σ^{54}-associated transcription occurs in other bacteria with polar flagella, both FlgR and FlgS are atypical. FlgS lacks a transmembrane domain and the signal that stimulates its activity is unknown. FlgR does not contain a known DNA binding domain and it engages in DNA-independent activation of transcription by σ^{54} and RNA polymerase. FlgR also suppresses transcription from the σ^{28} promoter of the flagellin gene, *flaA*. As novobiocin, a repressor of bacterial gyrases, suppresses expression from the σ^{54} promoter of *flaB* but not the σ^{28} promoter of *flaA*, DNA supercoiling may regulate expression from the *flaB* promoter (169) (17) (26).

The only class 3 *H. pylori* flagellar gene is the major flagellin, *flaA*. The transition from class 2 to class 3 flagellar gene expression is particularly interesting in *H. pylori*. In *S. typhimurium* this transition involves the completion of the flagellar hook which in turn allows the export of the σ^{28} anti-sigma factor FlgM to allow expression of late flagellar genes from σ^{28} promoters. Recent experiments with the *H. pylori* hook-length control protein, *fliK*, suggest that this protein is necessary to turn off expression from the class 2 regulon (151), but the mechanism by which σ^{28}-based flagellum gene expression is upregulated is less clear. As the *H. pylori* flagellum is sheathed it may not be possible to export FlgM via the flagellum. Furthermore, FlgM lacks the N-terminal 20 amino acids present in the *S. typhimurium* FlgM, required for interaction with the FlgM chaperone, FlgN.

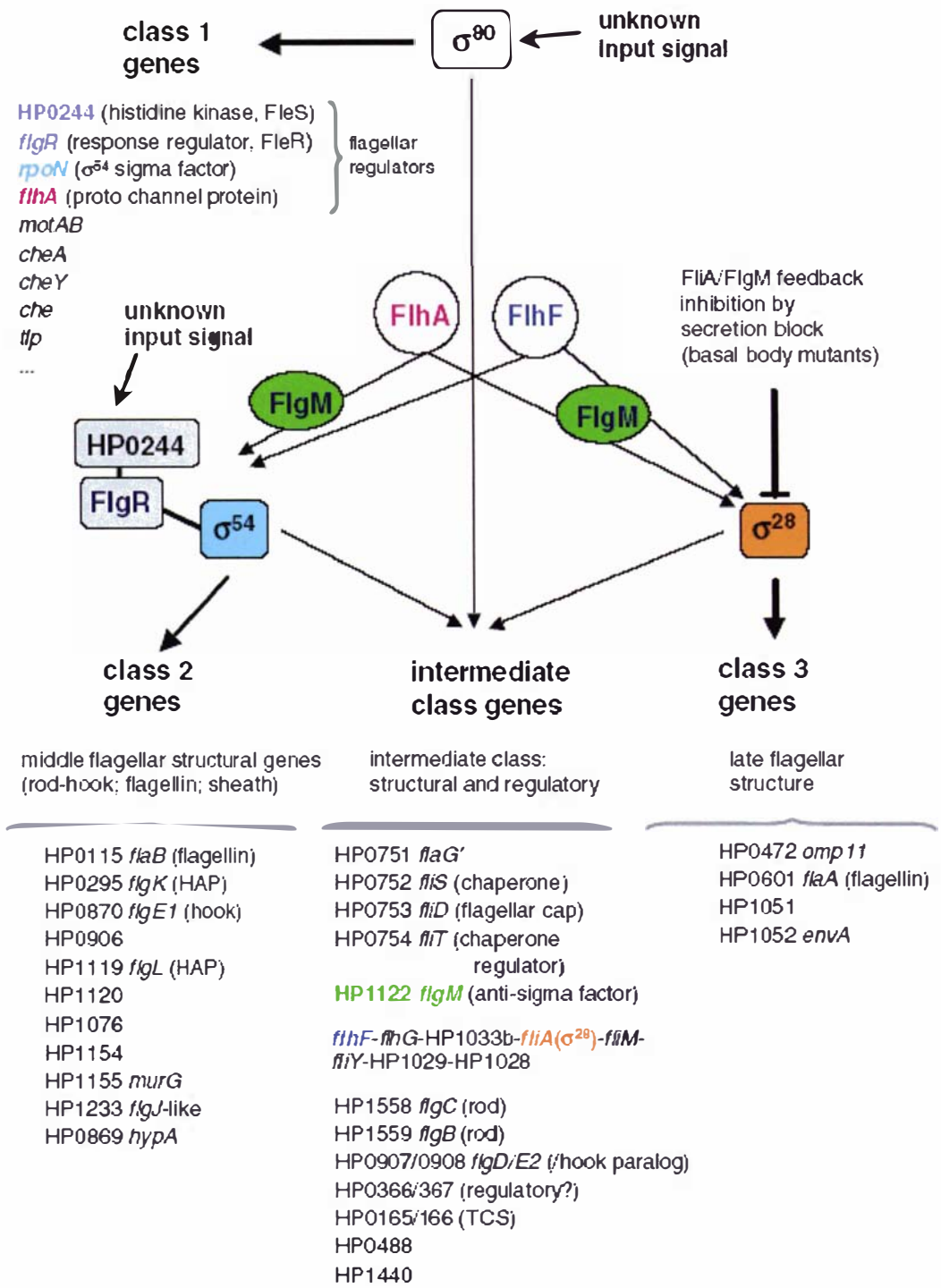


Figure 1.6. Current model of gene regulation pathways of flagellum biogenesis in *H. pylori*. Taken from Niehus *et al* (130).

1.4.4 Flagellum-specific chaperones

Four flagellum-specific chaperones, FlgN, FliS, FliT and FliJ, have been identified in *S. typhimurium*. Recently another role for flagellar chaperones in addition to post-translational functions similar to Type III virulence chaperones, has begun to emerge. It appears chaperones can also have a role in regulating gene expression at both the transcriptional and translational levels. For example, FlgN is required for regulation of FlgM as a class 3 transcript, and FliT has been shown to be a negative regulator of transcription from class 2 promoters (93) (81) (4) .

The post-translational role of flagellar chaperones in flagellar assembly involves stabilising and protecting export substrates from aggregation/polymerisation and degradation. Null mutants of *fliT* exhibit a motility defect due to decreased export of FliD. FliD does not accumulate intracellularly in this mutant strain suggesting that FliT has a role in protecting FliD from degradation. Consistent with this role, the C-termini of FliT homodimers bind FliD *in vitro* (18). Similar genetic and biochemical observations were made for the chaperone FlgN, which is essential for motility, binds FlgK and FlgL *in vitro*, and inhibits FlgK aggregation *in vitro* (18). Mutants of the chaperone FliS produce short flagella and therefore also display defective motility (208). It has subsequently been shown that FliS binds to the disordered D0 domain at the C-terminus of flagellin. Such binding inhibits polymerisation of FliC *in vitro* and protects the C-terminus from degradation *in vivo*. Thus, the role of FliS is to prevent premature polymerisation of FliC prior to export and to protect it from degradation (15) (139). FliJ is also classified as a flagellar chaperone as mutants have a defective motility phenotype and co-overexpression with FliE or FlgG hindered the aggregation of these proteins. But unlike other flagellar chaperones FliJ is more promiscuous as it interacts with flagellin, and both hook and rod proteins. In addition FliJ interacts with the soluble components of the export apparatus, the flagellum-specific ATPase FliI, FliH, and the cytoplasmic domain of FlhA (119) (114) (58).

1.4.5 Flagellum N-terminal secretion signal

Although the N-terminal segments of a number of the exported flagellar proteins can be aligned, no clear signal sequence has been identified for flagellar proteins (71) (70). Furthermore, most of the axial proteins contain heptad repeats within their N-termini indicative of amphipathic helix formation. However, all of the axial proteins that are exported by the flagellum are disordered at their respective N-termini and have no observable secondary/tertiary structure in this region. This has led some researchers to propose that the terminal disorder is the secretion signal. The heptad repeats probably indicate the helical structure of termini following the assembly of the axial proteins into the flagellum (71) (70) (36). Intriguingly, *Salmonella* type III virulence secretion substrates can be secreted by the flagellum suggesting that both systems use a common N-terminal secretion signal (97).

1.4.6 Substrate-specificity switching by the export apparatus

Export substrates are recognised by an N-terminal signal. This signal is still poorly characterised, and it is not known if it can allow differentiation of different export substrates. But it has been determined that the export apparatus can switch between exporting class 2 substrates and class 3 substrates or late flagellar genes, because hook-length is controlled. This change in export specificity involves the export of a hook-length control protein and a change in the cytoplasmic domain of FlhB.

FliK is a sensor of flagellum hook length and it relays this information by interacting with FlhB to induce a conformational change (120). This change is associated with the autoproteolytic cleavage of the cytoplasmic domain of FlhB into two fragments that remain associated, FlhBn and FlhBc (59) (54). This cleavage results in a switch from exporting

rod/hook-type substrates to filament substrates. Consequently, *fliK* null mutants produce polyhooks and filaments do not assemble (140).

The assembly of the hook is also influenced by the capacity of the export apparatus, and this is currently assumed to be defined by the C-ring, as mutations in any of the switch proteins lead to short hooks (107). Furthermore, the hook assembly is also influenced by the concentration of hook subunits (FlgE) as superpolyhooks are formed if FlgE is overexpressed (128). This has led to a model of substrate specificity switching in which FlgE fills the C-ring and is exported allowing FliK to interact with FlhB, change the export specificity and be exported itself.

1.5 Function of FliI and FliH in flagellar export

1.5.1 FliI

As previously discussed, flagellum assembly is regulated at the level of transcription and post-transcriptionally. The studies of hook-length control show that the export apparatus can change. Perhaps it is a dynamic structure involved in the post-translational regulation of export, by changing in response to different N-terminal secretion signals. However, this remains to be determined. Very little is currently known about the structure of the export apparatus or the translocation events following recognition of the substrate. The C-rod extending from the C-ring has been visualised and this is thought to constitute the export apparatus. The translocation process is expected to be energised by the flagellum-specific ATPase FliI. To perform this role, FliI is thought to interact with both components of the export apparatus and export substrates (119), and ATPase activity is predicted to be regulated to ensure ATP hydrolysis is coupled to export. This regulatory function is currently thought to be performed by FliH (118).

Both *S. typhimurium* and *H. pylori* FliI share 26% sequence identity with the β-subunit of Bovine F_1-ATPase, and contain the Walker A and B motifs. When the following mutations were made in *S. typhimurium* FliI: K188 to I (Walker A), D272 to N (Walker B) and Y363 to S (adenine binding pocket), the ATPase activity dropped 100-fold. All three mutations

also rendered *S. typhimurium* aflagellate and non-motile. Similarly, mutations of the Walker A and B residues of *H. pylori* FliI produced mutant strains that were non-motile and aflagellate (77). This mutagenesis data supports a role for FliI energising flagellar protein export.

However, the Walker A and B motifs are very common in ATP-utilising enzymes and therefore to verify that FliI has a homologous structure to the F_1 ATPase catalytic subunits and a similar catalytic mechanism, further structural studies are required. These structural studies are in progress. A recent structural study of *S. typhimurium* FliI has shown that like the F_1-ATPase, FliI assembles as a hexamer, and FliI demonstrates positive catalytic cooperativity (V_{max} = 0.28 μmol ATP hydrolysed min^{-1} mg^{-1}).

In addition to the catalytic residues, a temperature-sensitive L12P mutation of *S. typhimurium* FliI has also been identified, that shows that the residues of the N-terminus are also important for the function of FliI. This mutant has a defect in flagellation and motility that becomes more prominent when cells are shifted from 30 °C to 42 °C. An R7C/L12P mutation was also created that has an even greater defect in motility at the permissive temperature. Despite the defective motility these proteins exhibited wild-type ATPase activity (52). This suggested that their aberrant function may be due to the disruption of an interaction with another protein of the export apparatus. The L12P and R7C/L12P mutations were not dominant *in trans* (52), but interestingly when mutant proteins were created that contained both the R7C/L12P mutation and the Y363S or K188I catalytic mutations the dominance of the catalytic mutations was lost (118). It was subsequently demonstrated that these N-terminal mutants and N-terminally truncated FliI cannot interact with FliH (118) (121).

1.5.2 FliH

S. typhimurium FliH has been classified as a soluble component of the export apparatus because it interacts with other cytoplasmic components, FliI and the cytoplasmic domains of FlhA and FlhB. Furthermore, *fliH* null mutants fail to export flagellar proteins and consequently they are non-motile.

In vitro characterisation of *S. typhimurium* FliH has revealed that it forms a dimer in solution. The dimer has an elongated shape as it elutes with an estimated molecular mass of 232 kDa by gel filtration despite having a theoretical mass of 56 kDa (116). Deletions of FliH were also made to functionally dissect the protein. This identified three domains, an N-terminal domain, a dimerisation domain or interface and a C-terminal domain. The N-terminal domain consisting of residues 2-100 was found to be primarily responsible for the elongated shape of FliH (64), whereas deletion of residues 100-140 defined the dimerisation interface by abolishing dimerisation. Furthermore, FliH lacking residues 100-140 was unable to complement a *fliH* null mutant, indicating that the FliH dimer is the functionally relevant form of the protein *in vivo*. The C-terminal domain is critical for the interaction with FliI and consists of residues 100-235.

At the beginning of this project no data on *H. pylori* FliH had been published.

1.5.3 The FliH-FliI complex

FliH and FliI of *S. typhimurium* form an elongated $(FliH)_2FliI$ complex in solution that involves the C-termini of both FliH subunits and protects the N-terminal 26 amino acids of FliI in limited proteolysis experiments (118) (121) (116). FliI in this complex exhibits a 10-fold decrease in ATPase activity (118), but this result is hard to interpret as FliI was not a hexamer in this complex. Interestingly, following the binding of ATP, AMP-PNP or ADP to FliI, the N-terminal 26 residues become susceptible to proteolysis as they are in the absence of FliH (121), suggesting that conformational changes are coupled to ATP binding.

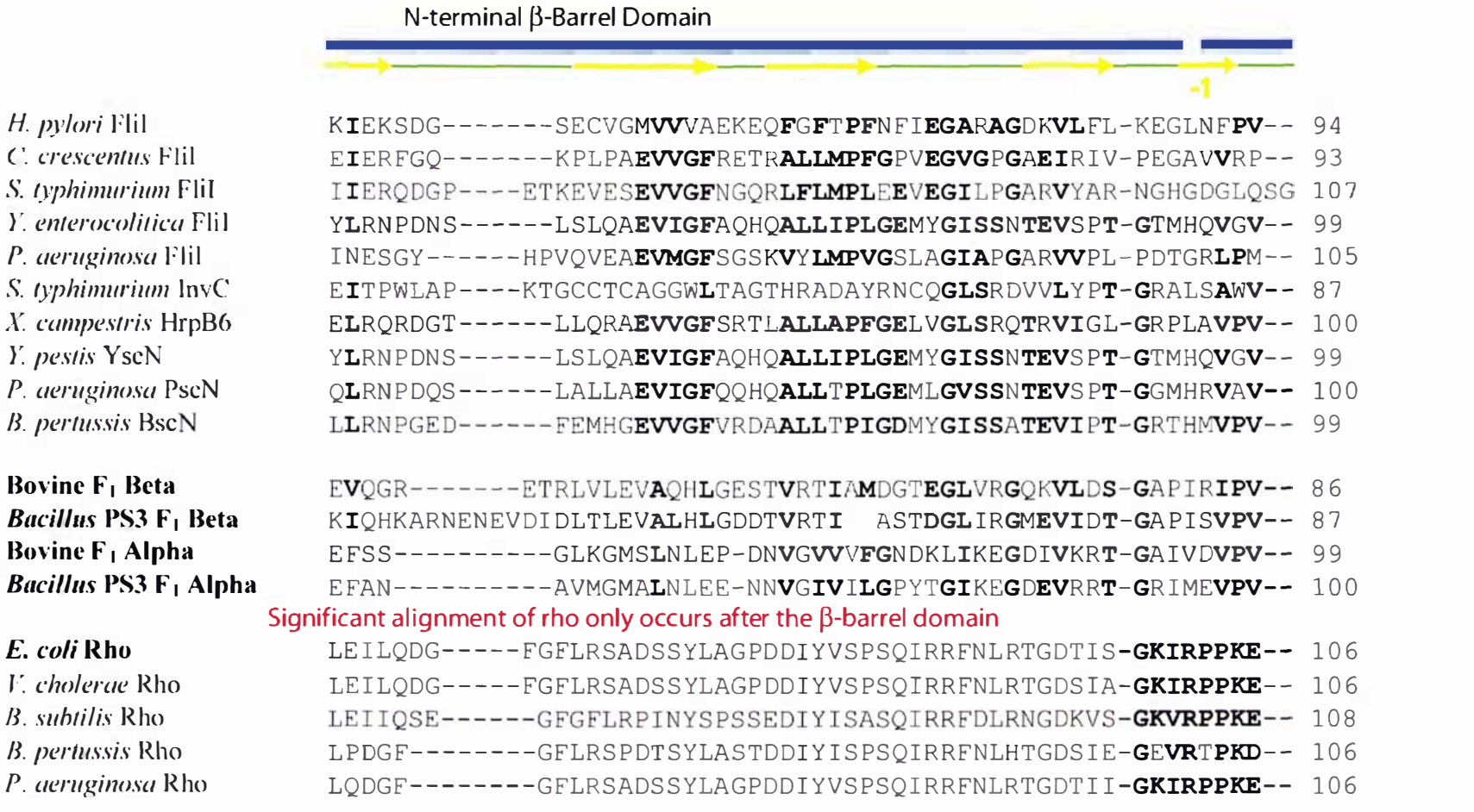
The molecular details of the role of FliI and FliH in the export apparatus are still being characterised. Both proteins have been shown to interact with the cytoplasmic domains of FlhA and FlhB, albeit weakly, and export substrates such as the hook FlgE and flagellin FliC which also require FliI for their export (117) (119) (211). These and the previous observations regarding FliI, FliH and the (FliH)2FliI complex allow a model to be proposed for the function of a soluble export complex that forms before the translocation of flagellar export substrates.

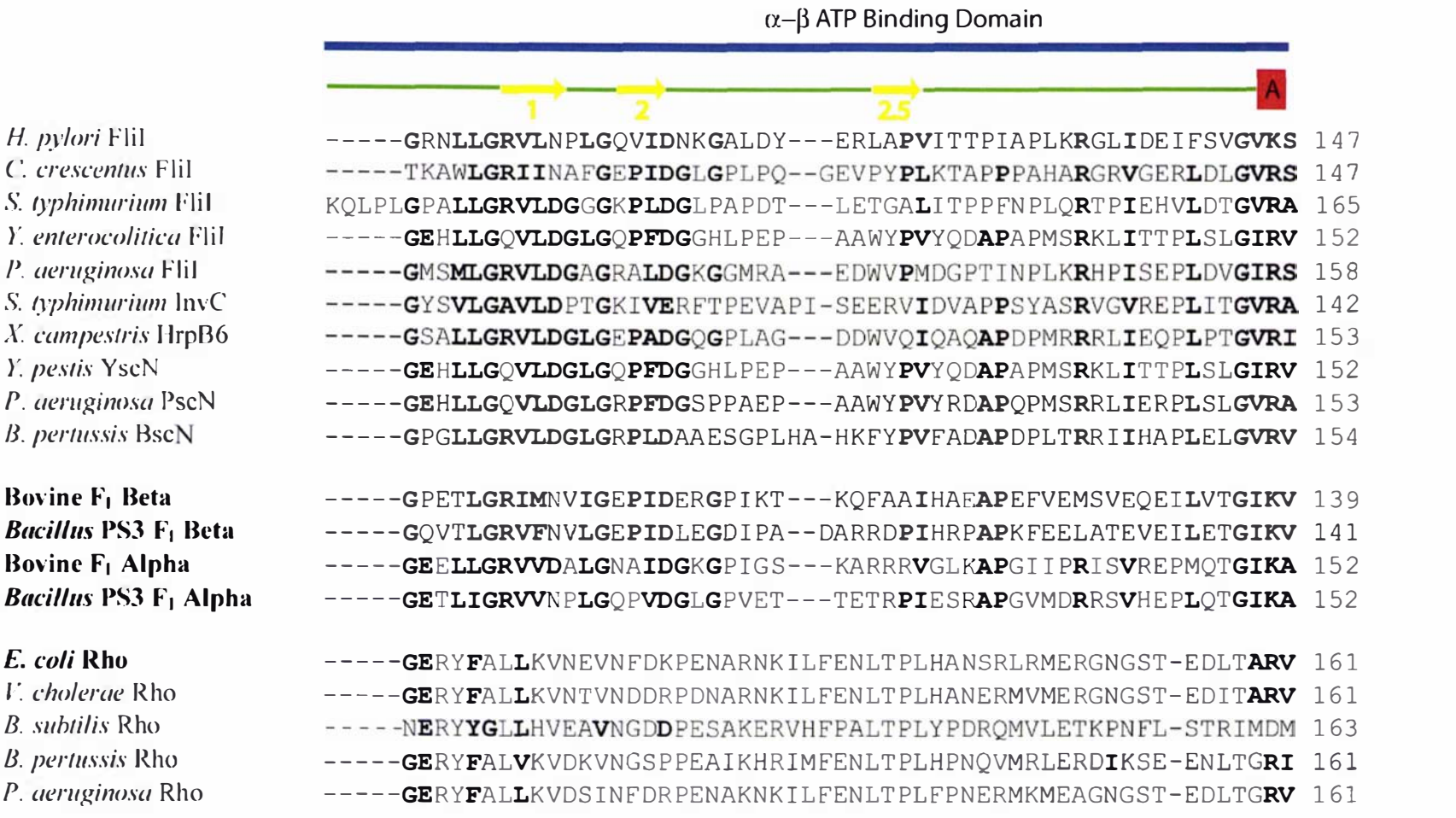
A $(FliH)_2FliI$ complex forms in the cytoplasm. In this complex FliH holds FliI in a conformation which allows substrate/chaperone complex binding but reduces ATP hydrolysis to prevent futile ATP hydrolysis when the ATPase is not coupled to substrate translocation. Both FliH and FliI have an intrinsic ability to associate with the membrane (14) (115) and this leads to stimulation of FliI oligomerisation and ATPase activity (31). The binding of ATP changes the conformation of the N-terminal domain of FliI (121) and allows either the binding of the substrate to FliI and/or the association of the complex with FlhA and B of the export apparatus. Association of the complex with FlhA and FlhB could then trigger the dissociation of FliH, and stimulated ATPase activity ensues to translocate a flagellar protein into the MS ring pore.

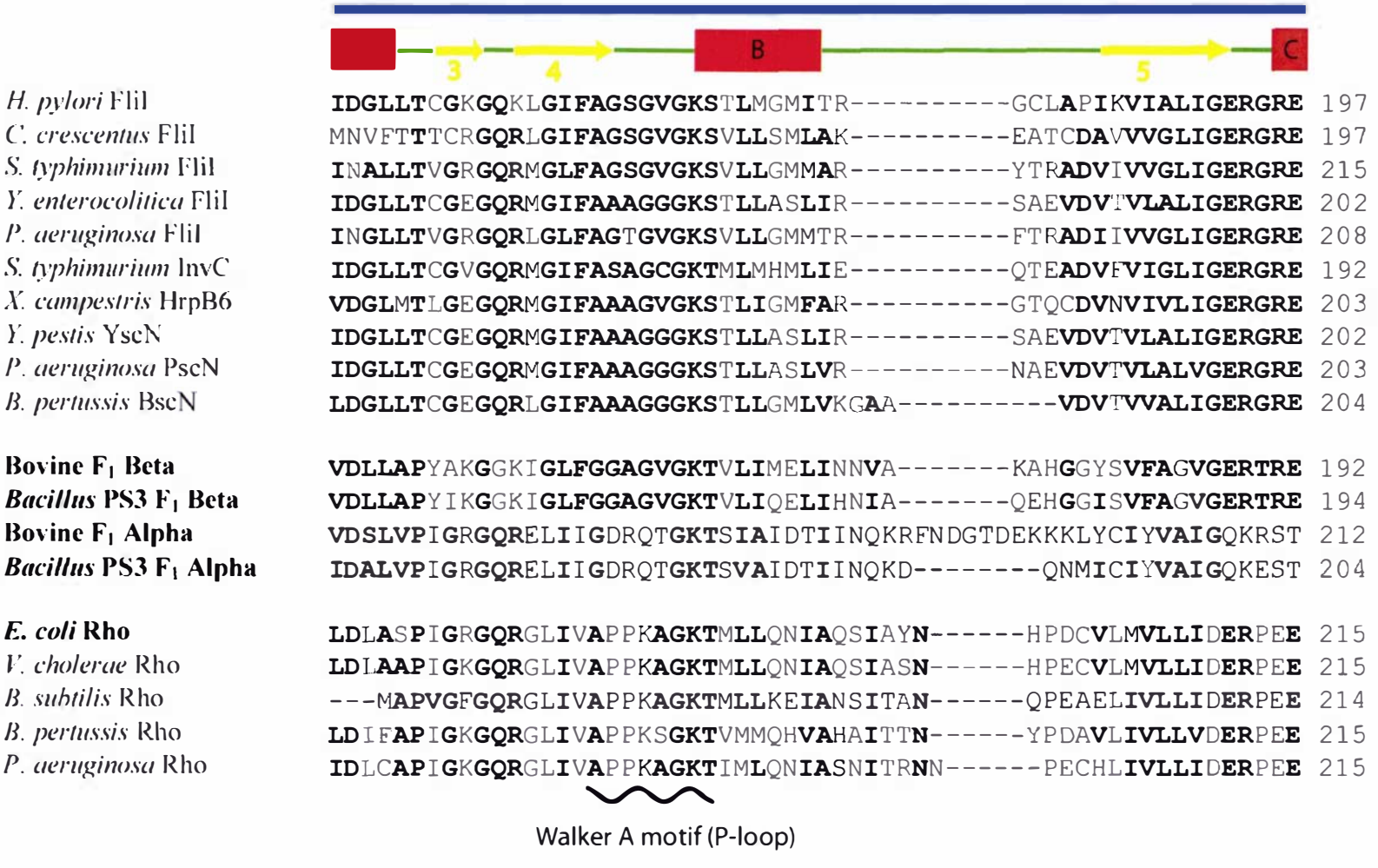
1.5.4 Structure of FliI

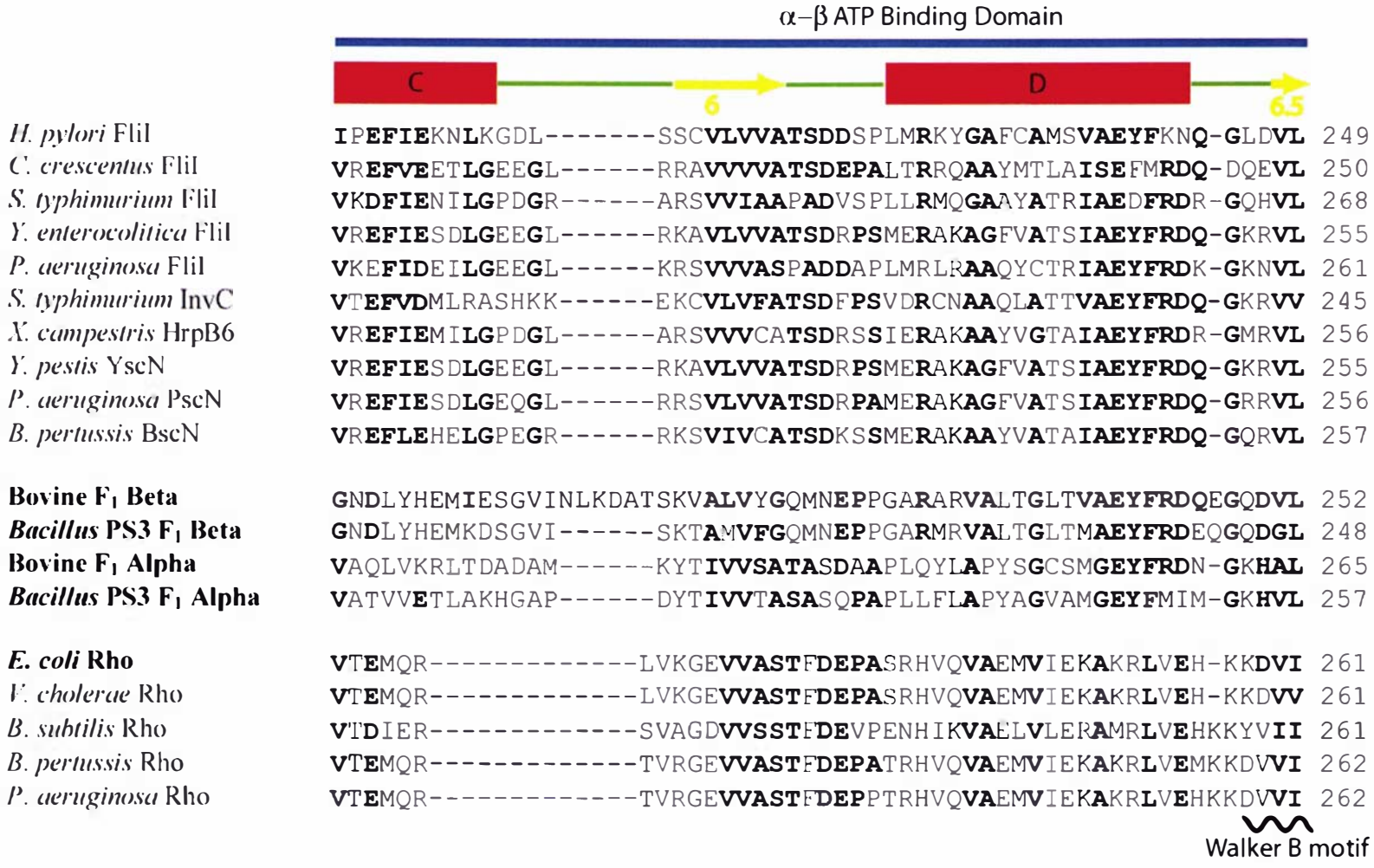
Significant alignments of FliI homologues and the known structure of the F_1 β subunit (Figure 1.7) allow the modeling of the three-dimensional structure of FliI. Such a model is presented in Figure 1.8.

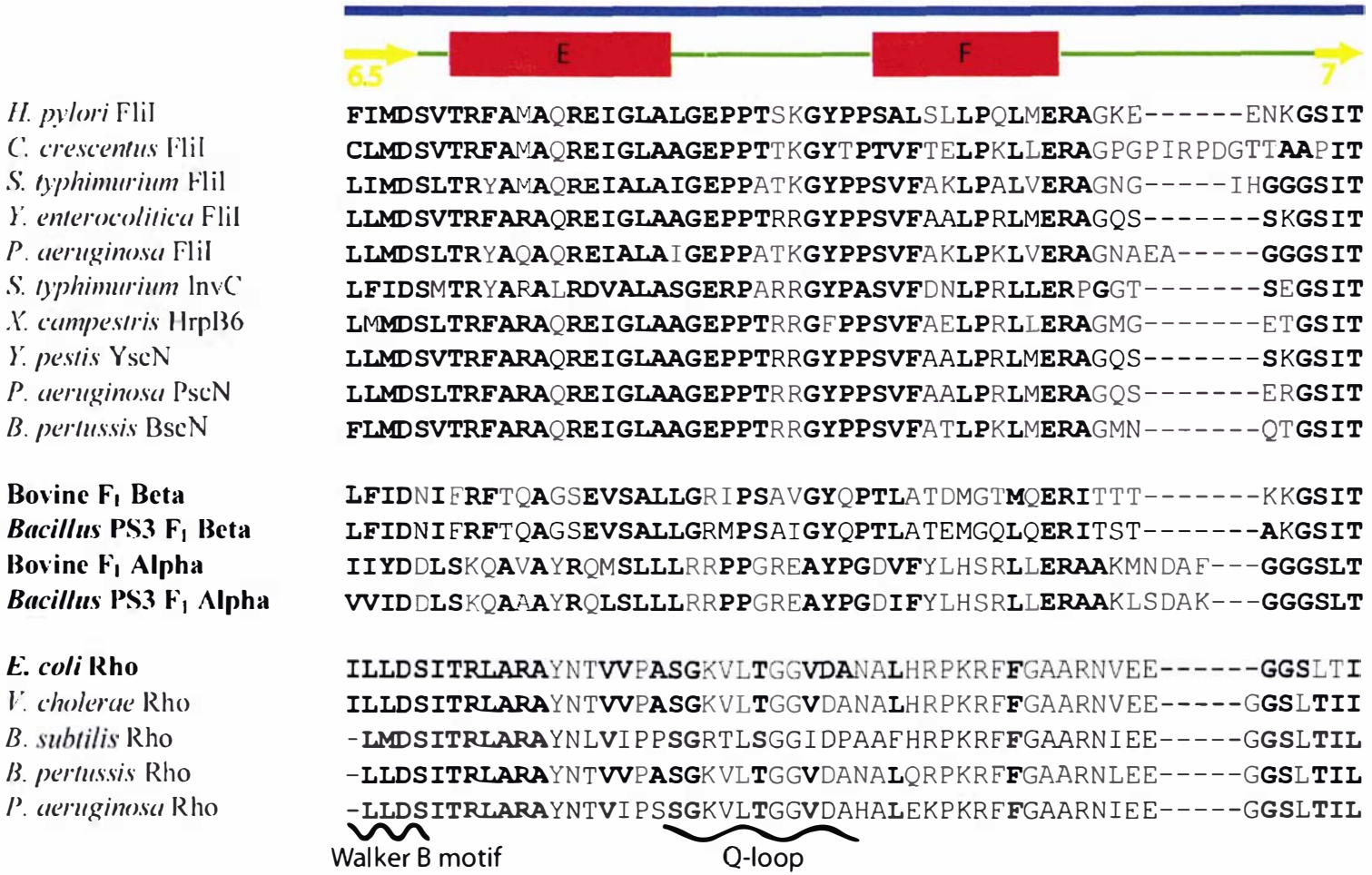
```
H. pylori FliI          ----------MPLKSLKNRLNQHFDLSPRYGSVKKIMPNIVYADG------FNPSVGDVV 44
C. crescentus FliI      ------------MRSLIAAVER-IDPLTIYGRVAAVNGLLIEVRGGL----TRLAVGARV 43
S. typhimurium FliI     -MTTRLTRWLTALDNFEAKMAL-LPAVRRYGRLTPATGLVLEATG------LQLPLGATC 52
Y. enterocolitica FliI  ------MLSLDQIPHHIRHGIVGSRLIQIRGRVTQVTGTLLKAVVP------GVRIGELC 48
P. aeruginosa FliI      --MRLERTSFARRLEGYTEAVSLPAQPVVEGRLLRMVGLTLEAEGLQ----AAVGSRCNV 54
S. typhimurium InvC     --------------------MKTPRLLQYLAYPQKITGPIIEAELR------DVAIGELC 34
X. campestris HrpB6     ---------MLAEMPLLQTTLERELAALAFGRRYGKVVEVIGTMLK--VAGVQVSLGEVC 49
Y. pestis YscN          ------MLSLDQIPHHIRHGIVGSRLIQIRGRVTQVTGTLLKAVVPG------VRIGELC 48
P. aeruginosa PscN      -----MPAPLSPLIVRMRHAIEGCRPIQIRGRVTQVTGTLLKAVVPG------VRIGELC 49
B. pertussis BscN       ------MRQYHYITEMMRVALQDLSTLRIKGRVVQVVGTIIKAVVPM------VKIGEVC 48

Bovine F1 Beta          --------------------SPSPKAGATTGRIVAVIGAVVDVQFDEG----LPPILNAL 36
Bacillus PS3 F1 Beta    ---------------------------MTRGRVIQVMGPVVDVKFENGH---LPAIYNAL 30
Bovine F1 Alpha         --QKTGTAEVSSILEERILGADTSVDLEETGRVLSIGDGIARVHGLRN-----VQAEEMV 53
Bacillus PS3 F1 Alpha   --MSIRAEEISALIKQQIENYESQIQVSDVGTVIQVGDGIARAHGLDN-----VMSGEAV 53

E. coli Rho             -----MNLTELKNTPVSELITLGENMGLENLARMRKQDIIFAILKQHAKS-GEDIFGDGV 54
V. cholerae Rho         -----MNLTELKNTPVSDLVKLGESLGLENLARLRKQDIIFAILKAHAKS-GEDIFGDGV 54
B. subtilis Rho         ---MKDVSISSLENMKLKELYELARHYKISYYSKLTKKELIFAILKANAEQEDLLFMEGV 57
B. pertussis Rho        ---MHLNELKALHVSQLLEMAAGLEIENANRLRKQELMFAIMKRRAKQGEQIFGDGVLEV 57
P. aeruginosa Rho       ---MNLTELKQKPIAELLEMSDAMGLENMARSRKQDIIFALLKKHAKSGEEISGDGVLEI 57
```

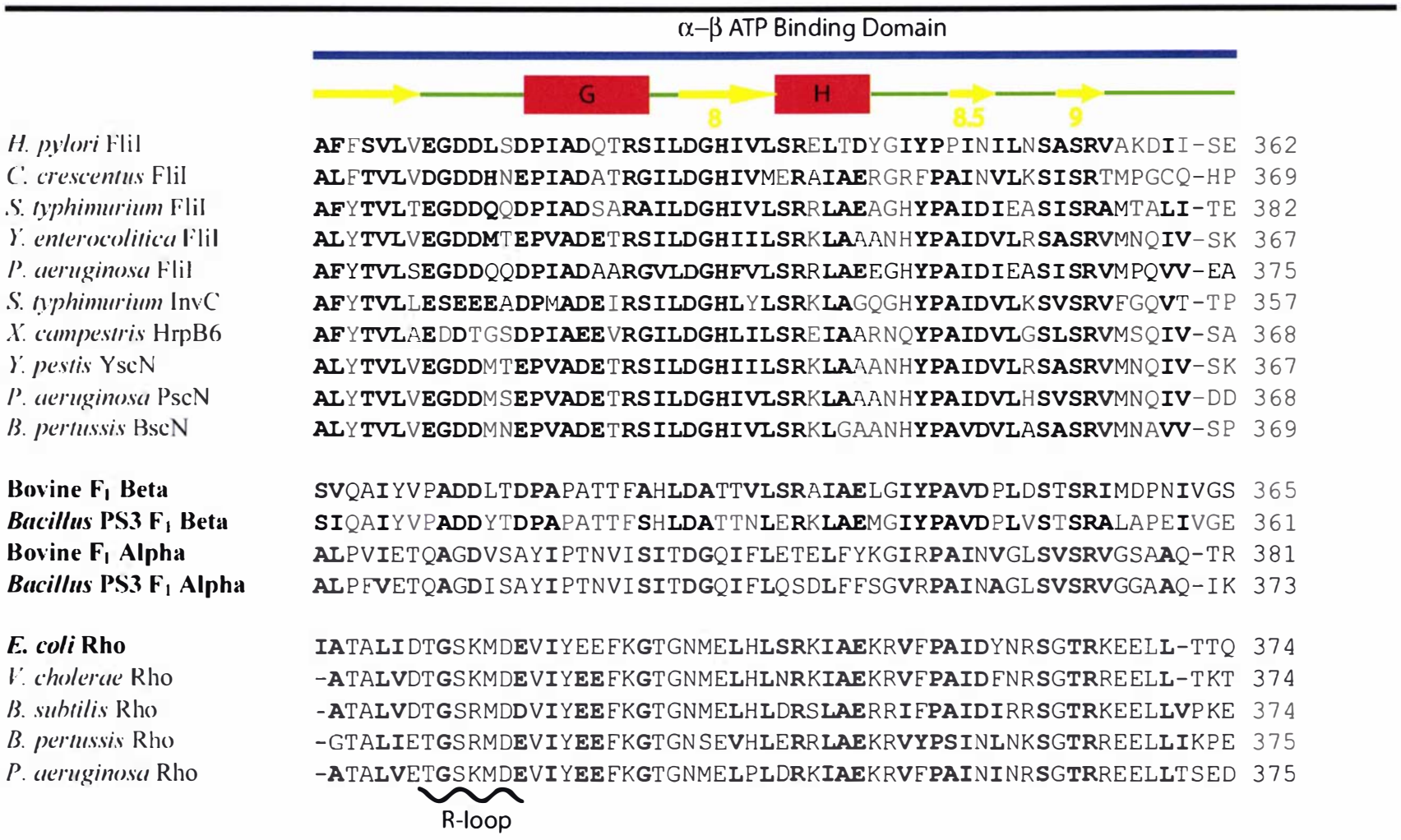

N-terminal β-Barrel Domain
-1
H. pylori FliI KIEKSDG-------SECVGMVVVAEKEQFGFTPFNFIEGARAGDKVLFL-KEGLNFPV-- 94
C. crescentus FliI EIERFGQ-------KPLPAEVVGFRETRALLMPFGPVEGVGPGAEIRIV-PEGAVVRP-- 93
S. typhimurium FliI IIERQDGP----ETKEVESEVVGFNGQRLFLMPLEEVEGILPGARVYAR-NGHGDGLQSG 107
Y. enterocolitica FliI YLRNPDNS------LSLQAEVIGFAQHQALLIPLGEMYGISSNTEVSPT-GTMHQVGV-- 99
P. aeruginosa FliI INESGY------HPVQVEAEVMGFSGSKVYLMPVGSLAGIAPGARVVPL-PDTGRLPM-- 105
S. typhimurium InvC EITPWLAP----KTGCCTCAGGWLTAGTHRADAYRNCQGLSRDVVLYPT-GRALSAWV-- 87
X. campestris HrpB6 ELRQRDGT------LLQRAEVVGFSRTLALLAPFGELVGLSRQTRVIGL-GRPLAVPV-- 100
Y. pestis YscN YLRNPDNS------LSLQAEVIGFAQHQALLIPLGEMYGISSNTEVSPT-GTMHQVGV-- 99
P. aeruginosa PscN QLRNPDQS------LALLAEVIGFQQHQALLTPLGEMLGVSSNTEVSPT-GGMHRVAV-- 100
B. pertussis BscN LLRNPGED------FEMHGEVVGFVRDAALLTPIGDMYGISSATEVIPT-GRTHMVPV-- 99
Bovine F1 Beta EVQGR-------ETRLVLEVAQHLGESTVRTIAMDGTEGLVRGQKVLDS-GAPIRIPV-- 86
Bacillus PS3 F1 Beta KIQHKARNENEVDIDLTLEVALHLGDDTVRTI ASTDGLIRGMEVIDT-GAPISVPV-- 87
Bovine F1 Alpha EFSS----------GLKGMSLNLEP-DNVGVVVFGNDKLIKEGDIVKRT-GAIVDVPV-- 99
Bacillus PS3 F1 Alpha EFAN----------AVMGMALNLEE-NNVGIVILGPYTGIKEGDEVRRT-GRIMEVPV-- 100
Significant alignment of rho only occurs after the β-barrel domain
E. coli Rho LEILQDG-----FGFLRSADSSYLAGPDDIYVSPSQIRRFNLRTGDTIS-GKIRPPKE-- 106
V. cholerae Rho LEILQDG-----FGFLRSADSSYLAGPDDIYVSPSQIRRFNLRTGDSIA-GKIRPPKE-- 106
B. subtilis Rho LEIIQSE------GFGFLRPINYSPSSEDIYISASQIRRFDLRNGDKVS-GKVRPPKE-- 108
B. pertussis Rho LPDGF--------GFLRSPDTSYLASTDDIYISPSQIRRFNLHTGDSIE-GEVRTPKD-- 106
P. aeruginosa Rho LQDGF--------GFLRSADSSYLAGPDDIYVSPSQIRRFNLRTGDTII-GKIRPPKE-- 106

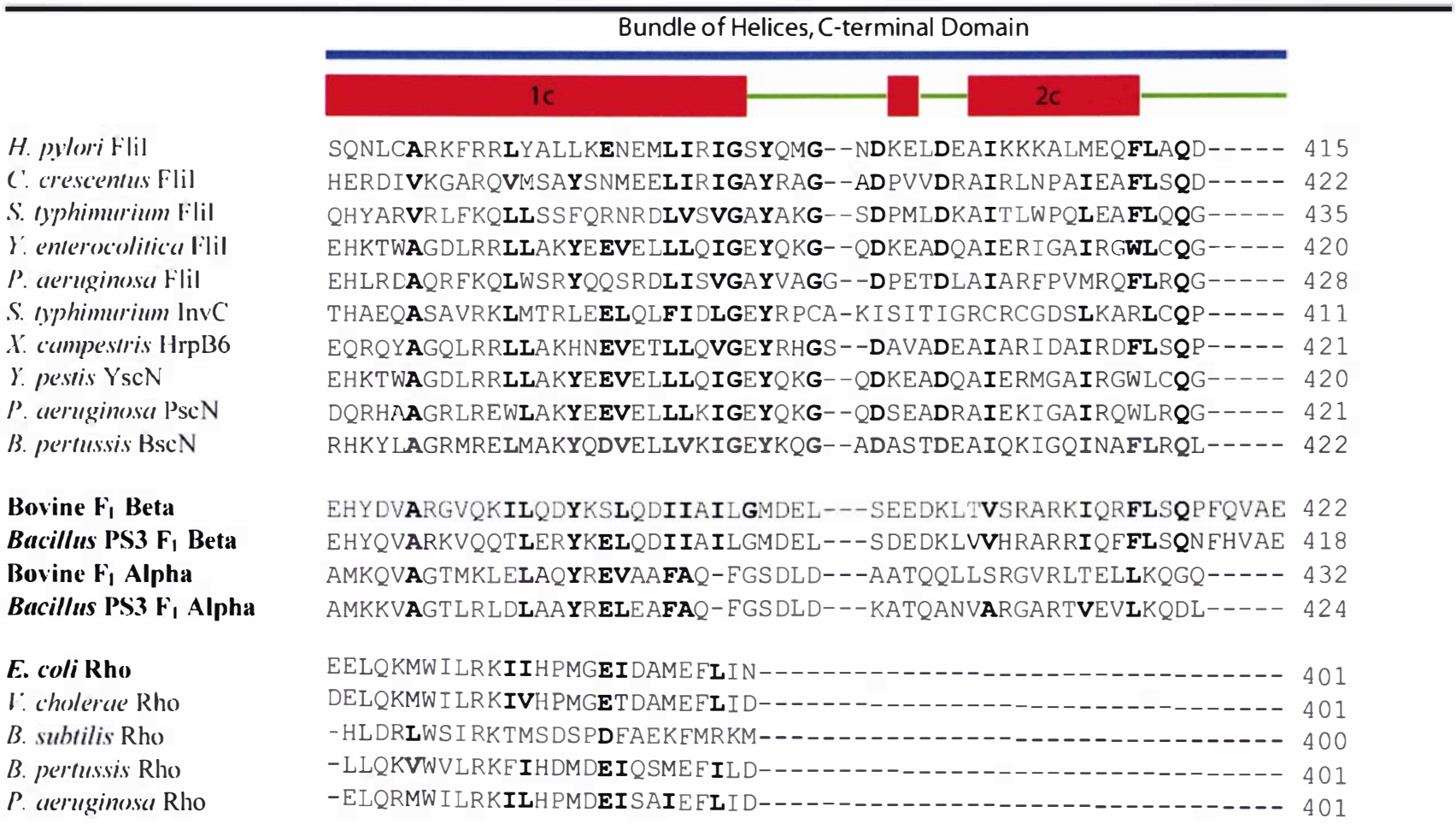

α–β ATP Binding Domain
1
2
2.5
A
H. pylori FliI -----GRNLLGRVLNPLGQVIDNKGALDY---ERLAPVITTPIAPLKRGLIDEIFSVGVKS 147
C. crescentus FliI -----TKAWLGRIINAFGEPIDGLGPLPQ--GEVPYPLKTAPPPAHARGRVGERLDLGVRS 147
S. typhimurium FliI KQLPLGPALLGRVLDGGGKPLDGLPAPDT---LETGALITPPFNPLQRTPIEHVLDTGVRA 165
Y. enterocolitica FliI -----GEHLLGQVLDGLGQPFDGGHLPEP---AAWYPVYQDAPAPMSRKLITTPLSLGIRV 152
P. aeruginosa FliI -----GMSMLGRVLDGAGRALDGKGGMRA---EDWVPMDGPTINPLKRHPISEPLDVGIRS 158
S. typhimurium InvC -----GYSVLGAVLDPTGKIVERFTPEVAPI-SEERVIDVAPPSYASRVGVREPLITGVRA 142
X. campestris HrpB6 -----GSALLGRVLDGLGEPADGQGPLAG---DDWVQIQAQAPDPMRRRLIEQPLPTGVRI 153
Y. pestis YscN -----GEHLLGQVLDGLGQPFDGGHLPEP---AAWYPVYQDAPAPMSRKLITTPLSLGIRV 152
P. aeruginosa PscN -----GEHLLGQVLDGLGRPFDGSPPAEP---AAWYPVYRDAPQPMSRRLIERPLSLGVRA 153
B. pertussis BscN -----GPGLLGRVLDGLGRPLDAAESGPLHA-HKFYPVFADAPDPLTRRIIHAPLELGVRV 154
Bovine F1 Beta -----GPETLGRIMNVIGEPIDERGPIKT---KQFAAIHAEAPEFVEMSVEQEILVTGIKV 139
Bacillus PS3 F1 Beta -----GQVTLGRVFNVLGEPIDLEGDIPA--DARRDPIHRPAPKFEELATEVEILETGIKV 141
Bovine F1 Alpha -----GEELLGRVVDALGNAIDGKGPIGS---KARRRVGLKAPGIIPRISVREPMQTGIKA 152
Bacillus PS3 F1 Alpha -----GETLIGRVVNPLGQPVDGLGPVET---TETRPIESRAPGVMDRRSVHEPLQTGIKA 152
E. coli Rho -----GERYFALLKVNEVNFDKPENARNKILFENLTPLHANSRLRMERGNGST-EDLTARV 161
V. cholerae Rho -----GERYFALLKVNTVNDDRPDNARNKILFENLTPLHANERMVMERGNGST-EDITARV 161
B. subtilis Rho -----NERYYGLLHVEAVNGDDPESAKERVHFPALTPLYPDRQMVLETKPNFL-STRIMDM 163
B. pertussis Rho -----GERYFALVKVDKVNGSPPEAIKHRIMFENLTPLHPNQVMRLERDIKSE-ENLTGRI 161
P. aeruginosa Rho -----GERYFALLKVDSINFDRPENAKNKILFENLTPLFPNERMKMEAGNGST-EDLTGRV 161


B
C
3
4
5
H. pylori FliI IDGLLTCGKGQKLGIFAGSGVGKSTLMGMITR----------GCLAPIKVIALIGERGRE 197
C. crescentus FliI MNVFTTTCRGQRLGIFAGSGVGKSVLLSMLAK----------EATCDAVVVGLIGERGRE 197
S. typhimurium FliI INALLTVGRGQRMGLFAGSGVGKSVLLGMMAR----------YTRADVIVVGLIGERGRE 215
Y. enterocolitica FliI IDGLLTCGEGQRMGIFAAAGGGKSTLLASLIR----------SAEVDVTVLALIGERGRE 202
P. aeruginosa FliI INGLLTVGRGQRLGLFAGTGVGKSVLLGMMTR----------FTRADIIVVGLIGERGRE 208
S. typhimurium InvC IDGLLTCGVGQRMGIFASAGCGKTMLMHMLIE----------QTEADVFVIGLIGERGRE 192
X. campestris HrpB6 VDGLMTLGEGQRMGIFAAAGVGKSTLIGMFAR----------GTQCDVNVIVLIGERGRE 203
Y. pestis YscN IDGLLTCGEGQRMGIFAAAGGGKSTLLASLIR----------SAEVDVTVLALIGERGRE 202
P. aeruginosa PscN IDGLLTCGEGQRMGIFAAAGGGKSTLLASLVR----------NAEVDVTVLALVGERGRE 203
B. pertussis BscN LDGLLTCGEGQRLGIFAAAGGGKSTLLGMLVKGAA----------VDVTVVALIGERGRE 204
Bovine F1 Beta VDLLAPYAKGGKIGLFGGAGVGKTVLIMELINNVA-------KAHGGYSVFAGVGERTRE 192
Bacillus PS3 F1 Beta VDLLAPYIKGGKIGLFGGAGVGKTVLIQELIHNIA-------QEHGGISVFAGVGERTRE 194
Bovine F1 Alpha VDSLVPIGRGQRELIIGDRQTGKTSIAIDTIINQKRFNDGTDEKKKLYCIYVAIGQKRST 212
Bacillus PS3 F1 Alpha IDALVPIGRGQRELIIGDRQTGKTSVAIDTIINQKD--------QNMICIYVAIGQKEST 204
E. coli Rho LDLASPIGRGQRGLIVAPPKAGKTMLLQNIAQSIAYN------HPDCVLMVLLIDERPEE 215
V. cholerae Rho LDLAAPIGKGQRGLIVAPPKAGKTMLLQNIAQSIASN------HPECVLMVLLIDERPEE 215
B. subtilis Rho ---MAPVGFGQRGLIVAPPKAGKTMLLKEIANSITAN------QPEAELIVLLIDERPEE 214
B. pertussis Rho LDIFAPIGKGQRGLIVAPPKSGKTVMMQHVAHAITTN------YPDAVLIVLLVDERPEE 215
P. aeruginosa Rho IDLCAPIGKGQRGLIVAPPKAGKTIMLQNIASNITRNN------PECHLIVLLIDERPEE 215
Walker A motif (P-loop)


α–β ATP Binding Domain
C
D
6
6.5
H. pylori FliI IPEFIEKNLKGDL-------SSCVLVVATSDDSPLMRKYGAFCAMSVAEYFKNQ-GLDVL 249
C. crescentus FliI VREFVEETLGEEGL------RRAVVVVATSDEPALTRRQAAYMTLAISEFMRDQ-DQEVL 250
S. typhimurium FliI VKDFIENILGPDGR------ARSVVIAAPADVSPLLRMQGAAYATRIAEDFRDR-GQHVL 268
Y. enterocolitica FliI VREFIESDLGEEGL------RKAVLVVATSDRPSMERAKAGFVATSIAEYFRDQ-GKRVL 255
P. aeruginosa FliI VKEFIDEILGEEGL------KRSVVVASPADDAPLMRLRAAQYCTRIAEYFRDK-GKNVL 261
S. typhimurium InvC VTEFVDMLRASHKK------EKCVLVFATSDFPSVDRCNAAQLATTVAEYFRDQ-GKRVV 245
X. campestris HrpB6 VREFIEMILGPDGL------ARSVVVCATSDRSSIERAKAAYVGTAIAEYFRDR-GMRVL 256
Y. pestis YscN VREFIESDLGEEGL------RKAVLVVATSDRPSMERAKAGFVATSIAEYFRDQ-GKRVL 255
P. aeruginosa PscN VREFIESDLGEQGL------RRSVLVVATSDRPAMERAKAGFVATSIAEYFRDQ-GRRVL 256
B. pertussis BscN VREFLEHELGPEGR------RKSVIVCATSDKSSMERAKAAYVATAIAEYFRDQ-GQRVL 257
Bovine F1 Beta GNDLYHEMIESGVINLKDATSKVALVYGQMNEPPGARARVALTGLTVAEYFRDQEGQDVL 252
Bacillus PS3 F1 Beta GNDLYHEMKDSGVI------SKTAMVFGQMNEPPGARMRVALTGLTMAEYFRDEQGQDGL 248
Bovine F1 Alpha VAQLVKRLTDADAM------KYTIVVSATASDAAPLQYLAPYSGCSMGEYFRDN-GKHAL 265
Bacillus PS3 F1 Alpha VATVVETLAKHGAP------DYTIVVTASASQPAPLLFLAPYAGVAMGEYFMIM-GKHVL 257
E. coli Rho VTEMQR-------------LVKGEVVASTFDEPASRHVQVAEMVIEKAKRLVEH-KKDVI 261
V. cholerae Rho VTEMQR-------------LVKGEVVASTFDEPASRHVQVAEMVIEKAKRLVEH-KKDVV 261
B. subtilis Rho VTDIER-------------SVAGDVVSSTFDEVPENHIKVAELVLERAMRLVEHKKYVII 261
B. pertussis Rho VTEMQR-------------TVRGEVVASTFDEPATRHVQVAEMVIEKAKRLVEMKKDVVI 262
P. aeruginosa Rho VTEMQR-------------TVRGEVVASTFDEPPTRHVQVAEMVIEKAKRLVEHKKDVVI 262
Walker B motif


E
F
6.5
7
H. pylori FliI FIMDSVTRFAMAQREIGLALGEPPTSKGYPPSALSLLPQLMERAGKE------ENKGSIT 303
C. crescentus FliI CLMDSVTRFAMAQREIGLAAGEPPTTKGYTPTVFTELPKLLERAGPGPIRPDGTTAAPIT 310
S. typhimurium FliI LIMDSLTRYAMAQREIALAIGEPPATKGYPPSVFAKLPALVERAGNG-----IHGGGSIT 323
Y. enterocolitica FliI LLMDSVTRFARAQREIGLAAGEPPTRRGYPPSVFAALPRLMERAGQS-------SKGSIT 308
P. aeruginosa FliI LLMDSLTRYAQAQREIALAIGEPPATKGYPPSVFAKLPKLVERAGNAEA-----GGGSIT 316
S. typhimurium InvC LFIDSMTRYARALRDVALASGERPARRGYPASVFDNLPRLLERPGGT-------SEGSIT 298
X. campestris HrpB6 LMMDSLTRFARAQREIGLAAGEPPTRRGFPPSVFAELPRLLERAGMG-------ETGSIT 309
Y. pestis YscN LLMDSVTRFARAQREIGLAAGEPPTRRGYPPSVFAALPRLMERAGQS-------SKGSIT 308
P. aeruginosa PscN LLMDSLTRFARAQREIGLAAGEPPTRRGYPPSVFAALPRLMERAGQS-------ERGSIT 309
B. pertussis BscN FLMDSVTRFARAQREIGLAAGEPPTRRGYPPSVFATLPKLMERAGMN-------QTGSIT 310
Bovine F1 Beta LFIDNIFRFTQAGSEVSALLGRIPSAVGYQPTLATDMGTMQERITTT-------KKGSIT 305
Bacillus PS3 F1 Beta LFIDNIFRFTQAGSEVSALLGRMPSAIGYQPTLATEMGQLQERITST-------AKGSIT 301
Bovine F1 Alpha IIYDDLSKQAVAYRQMSLLLRRPPGREAYPGDVFYLHSRLLERAAKMNDAF---GGGSLT 322
Bacillus PS3 F1 Alpha VVIDDLSKQAAAYRQLSLLLRRPPGREAYPGDIFYLHSRLLERAAKLSDAK---GGGSLT 314
E. coli Rho ILLDSITRLARAYNTVVPASGKVLTGGVDANALHRPKRFFGAARNVEE------GGSLTI 315
V. cholerae Rho ILLDSITRLARAYNTVVPASGKVLTGGVDANALHRPKRFFGAARNVEE-----GGSLTII 316
B. subtilis Rho -LMDSITRLARAYNLVIPPSGRTLSGGIDPAAFHRPKRFFGAARNIEE-----GGSLTIL 315
B. pertussis Rho -LLDSITRLARAYNTVVPASGKVLTGGVDANALQRPKRFFGAARNLEE-----GGSLTIL 316
P. aeruginosa Rho -LLDSITRLARAYNTVIPSSGKVLTGGVDAHALEKPKRFFGAARNIEE-----GGSLTIL 316
Walker B motif
Q-loop


α–β ATP Binding Domain
G
H
8
8.5
9
H. pylori FliI AFFSVLVEGDDLSDPIADQTRSILDGHIVLSRELTDYGIYPPINILNSASRVAKDII-SE 362
C. crescentus FliI ALFTVLVDGDDHNEPIADATRGILDGHIVMERAIAERGRFPAINVLKSISRTMPGCQ-HP 369
S. typhimurium FliI AFYTVLTEGDDQQDPIADSARAILDGHIVLSRRLAEAGHYPAIDIEASISRAMTALI-TE 382
Y. enterocolitica FliI ALYTVLVEGDDMTEPVADETRSILDGHIILSRKLAAANHYPAIDVLRSASRVMNQIV-SK 367
P. aeruginosa FliI AFYTVLSEGDDQQDPIADAARGVLDGHFVLSRRLAEEGHYPAIDIEASISRVMPQVV-EA 375
S. typhimurium InvC AFYTVLLESEEEADPMADEIRSILDGHLYLSRKLAGQGHYPAIDVLKSVSRVFGQVT-TP 357
X. campestris HrpB6 AFYTVLAEDDTGSDPIAEEVRGILDGHLILSREIAARNQYPAIDVLGSLSRVMSQIV-SA 368
Y. pestis YscN ALYTVLVEGDDMTEPVADETRSILDGHIILSRKLAAANHYPAIDVLRSASRVMNQIV-SK 367
P. aeruginosa PscN ALYTVLVEGDDMSEPVADETRSILDGHIVLSRKLAAANHYPAIDVLHSVSRVMNQIV-DD 368
B. pertussis BscN ALYTVLVEGDDMNEPVADETRSILDGHIVLSRKLGAANHYPAVDVLASASRVMNAVV-SP 369
Bovine F1 Beta SVQAIYVPADDLTDPAPATTFAHLDATTVLSRAIAELGIYPAVDPLDSTSRIMDPNIVGS 365
Bacillus PS3 F1 Beta SIQAIYVPADDYTDPAPATTFSHLDATTNLERKLAEMGIYPAVDPLVSTSRALAPEIVGE 361
Bovine F1 Alpha ALPVIETQAGDVSAYIPTNVISITDGQIFLETELFYKGIRPAINVGLSVSRVGSAAQ-TR 381
Bacillus PS3 F1 Alpha ALPFVETQAGDISAYIPTNVISITDGQIFLQSDLFFSGVRPAINAGLSVSRVGGAAQ-IK 373
E. coli Rho IATALIDTGSKMDEVIYEEFKGTGNMELHLSRKIAEKRVFPAIDYNRSGTRKEELL-TTQ 374
V. cholerae Rho -ATALVDTGSKMDEVIYEEFKGTGNMELHLNRKIAEKRVFPAIDFNRSGTRREELL-TKT 374
B. subtilis Rho -ATALVDTGSRMDDVIYEEFKGTGNMELHLDRSLAERRIFPAIDIRRSGTRKEELLVPKE 374
B. pertussis Rho -GTALIETGSRMDEVIYEEFKGTGNSEVHLERRLAEKRVYPSINLNKSGTRREELLIKPE 375
P. aeruginosa Rho -ATALVETGSKMDEVIYEEFKGTGNMELPLDRKIAEKRVFPAININRSGTRREELLTSED 375
R-loop


Bundle of Helices, C-terminal Domain
1c
2c
H. pylori FliI SQNLCARKFRRLYALLKENEMLIRIGSYQMG--NDKELDEAIKKKALMEQFLAQD----- 415
C. crescentus FliI HERDIVKGARQVMSAYSNMEELIRIGAYRAG--ADPVVDRAIRLNPAIEAFLSQD----- 422
S. typhimurium FliI QHYARVRLFKQLLSSFQRNRDLVSVGAYAKG--SDPMLDKAITLWPQLEAFLQQG----- 435
Y. enterocolitica FliI EHKTWAGDLRRLLAKYEEVELLLQIGEYQKG--QDKEADQAIERIGAIRGWLCQG----- 420
P. aeruginosa FliI EHLRDAQRFKQLWSRYQQSRDLISVGAYVAGG--DPETDLAIARFPVMRQFLRQG----- 428
S. typhimurium InvC THAEQASAVRKLMTRLEELQLFIDLGEYRPCA-KISITIGRCRCGDSLKARLCQP----- 411
X. campestris HrpB6 EQRQYAGQLRRLLAKHNEVETLLQVGEYRHGS--DAVADEAIARIDAIRDFLSQP----- 421
Y. pestis YscN EHKTWAGDLRRLLAKYEEVELLLQIGEYQKG--QDKEADQAIERMGAIRGWLCQG----- 420
P. aeruginosa PscN DQRHAAGRLREWLAKYEEVELLLKIGEYQKG--QDSEADRAIEKIGAIRQWLRQG----- 421
B. pertussis BscN RHKYLAGRMRELMAKYQDVELLVKIGEYKQG--ADASTDEAIQKIGQINAFLRQL----- 422
Bovine F1 Beta EHYDVARGVQKILQDYKSLQDIIAILGMDEL---SEEDKLTVSRARKIQRFLSQPFQVAE 422
Bacillus PS3 F1 Beta EHYQVARKVQQTLERYKELQDIIAILGMDEL---SDEDKLVVHRARRIQFFLSQNFHVAE 418
Bovine F1 Alpha AMKQVAGTMKLELAQYREVAAFAQ-FGSDLD---AATQQLLSRGVRLTELLKQGQ----- 432
Bacillus PS3 F1 Alpha AMKKVAGTLRLDLAAYRELEAFAQ-FGSDLD---KATQANVARGARTVEVLKQDL----- 424
E. coli Rho EELQKMWILRKIIHPMGEIDAMEFLIN---------- 401
V. cholerae Rho DELQKMWILRKIVHPMGETDAMEFLID---------- 401
B. subtilis Rho -HLDRLWSIRKTMSDSPDFAEKFMRKM---------- 400
B. pertussis Rho -LLQKVWVLRKFIHDMDEIQSMEFILD---------- 401
P. aeruginosa Rho -ELQRMWILRKILHPMDEISAIEFLID---------- 401

H. pylori FliI	-----ENALQPFETSFQQLEEILR----------	434
C. crescentus FliI	-----KEEATSLDDSFGMLGQILQSEY----------	444
S. typhimurium FliI	-----IFERADWEDSLQALDLIFPTV----------	456
Y. enterocolitica FliI	-----THELSHFNETLNLLETLTQ----------	439
P. aeruginosa FliI	-----LDESESLAESRARLASLLAGGQA----------	451
S. typhimurium InvC	-----VAQYSSFDDTLSGMNAFADQN----------	432
X. campestris HrpB6	-----TDQLSDYDTILEQLAGVIDDA----------	442
Y. pestis YscN	-----THELSHFNETLNLLETLTQ----------	439
P. aeruginosa PscN	-----THETSDYAQACAQLRSLCA----------	440
B. pertussis BscN	-----TDEREAFEDTVLRMAEIIGPES----------	444
Bovine F_1 Beta	VFTGHLGKLVPLKETIKGFQQILAGEYDHLPEQA---FYMVGPIEEAVAKADKLAEEHS-	478
***Bacillus* PS3 F_1 Beta**	QFTGQPGSYVPVKETVRGFKEILEGKYDHLPEDR---FRLVGRIEEVVEKAKAMGVEV--	473
Bovine F_1 Alpha	------YSPMAIEEQVAVIYAGVRGYLDKLEPSK-------ITKFENAFLSHVISQHQAL	460
***Bacillus* PS3 F_1 Alpha**	------HQPIPVEKQVLIIYALTRGFLDDIPVED-------VRRFEKEFYLWLDQNGQHL	452
***E. coli* Rho**	----KLAMTKTNDDFFEMMKRS----------	419
V. cholerae Rho	----KLAMTKTNDEFFDAMRRQ----------	419
B. subtilis Rho	----KKTKTNQEFFDILNQEWKQANLSSARR----------	427
B. pertussis Rho	----KMRATKTNAEFFDMMKK----------	418
P. aeruginosa Rho	----KLKQTKTNDEFFDSMKRK----------	419

Figure 1.7. Multiple alignment of the primary sequence and secondary structure elements of homologues of FliI, the Type III export ATPases, the Rho transcription terminator and the F_1-ATPase α- and β-subunits. The NCBI gi numbers for these protein sequences are: 15646029, *H. pylori* FliI, 16127270, *C. crescentus* FliI, 120332, *S. typhimurium* FliI; 732261, *Y. enterocolitica* FliI, 730791, *P. aeruginosa* FliI; the Type III export system ATPases, 497222 *S. typhimurium* InvC, 21112282 *X. campestris* HrpB6, 16082728 *Y. pestis* YscN, 9947670 *P. aeruginosa* PscN, 29603230 *B. pertussis* BscN; 1943080, 114543, Bovine F_1-ATPase α- and β-subunits respectively; 114531, 114571, *B. subtilis* PS3 F_1-ATPase α- and β-subunits respectively; 26110951 *E. coli* Rho, 9654720 *V. cholerae* Rho, 2636233 *B. subtilis* Rho, 33602207 *B. pertussis* Rho, 15600432 *P. aeruginosa* Rho. The initial alignment of the sequences was made using Clustal W. The boundaries of the secondary structure elements were determined by reference to the reported structure of Bovine F_1-ATPase as reported by Abrahams *et al*, 1994. The alignment accommodates poorly aligning sequence by introducing gaps, and the alignment could be adjusted such that these gaps lie in the loops between secondary structure elements. The blue lines indicate the predicted domain boundaries, the red boxes represent helices, the yellow arrows represent β-strands, and the green lines indicate loops. The red boxes shaded with a red gradient and the pale green lines indicate helices and loops respectively that are present in the F_1-ATPase β-subunit structure but are not predicted in the FliI structure. Within the ATP binding domain the strands are numbered 1-9, and the helices are labeled A-H, in the C-terminal domain the helices are numbered 1c-3c. These are the same designations used by Abrahams *et al*, 1994 for the Bovine F_1-ATPase structure and by Miwa *et al*, for a model of the *E. coli* rho structure. Helix I appears to be absent in the FliI structure prediction, it should be short helix following strand 9. The sequence names in bold type indicate proteins of known tertiary structure. Note: The *H. pylori* FliI sequence is from strain 26695. The DNA used in the cloning and mutagenesis of *fliI* was from strain 17874. Strain 17874 contains the following point mutations I_{140}-V and T_{323}-A (142). In addition three mutations were introduced during cloning, S_{52}-T, I_{360}-T and T_{424}-Q.

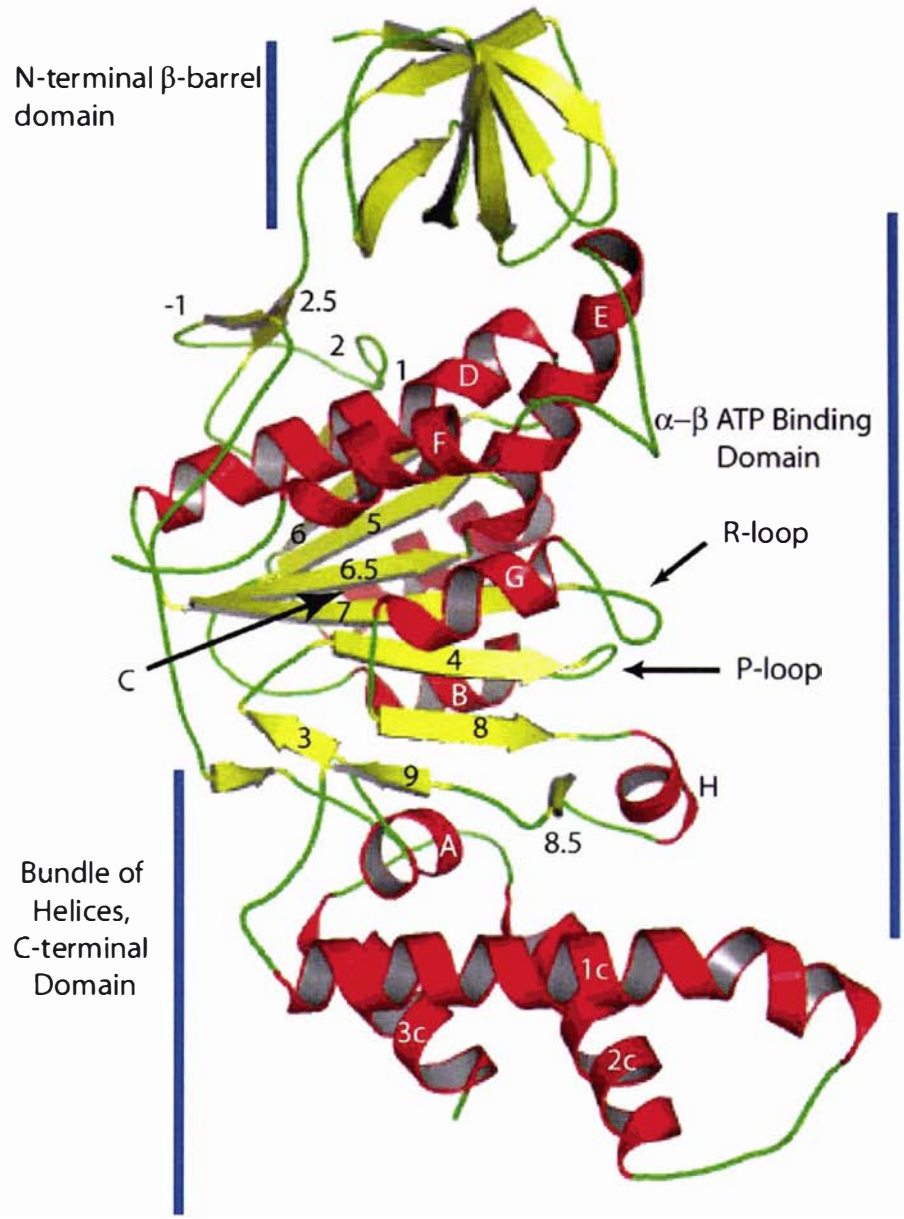


Figure 1.8 Model of the three dimensional structure of *H. pylori* FliI. This model is based on the structure of the Bovine F_1-ATPase β-subunit as published by Abrahams *et al*, 1994. Following alignment of the primary sequences of a number of FliI homologues and the known secondary structural elements of Bovine F_1-ATPase β-subunnit in Figure 1.7, the amino acids of Bovine F_1-ATPase β-subunit were substituted for those of FliI or deleted where appropriate, and the torsion angles were adjusted where necessary to accommodate these changes without steric clashes. The β-strands are coloured yellow and numbered 1-9, the helices are red and labeled A-H and 1c-3c, and the loops are green. Some elements of the structure such as helix C are difficult to see as they lie behind other elements of the three dimensional structure.

Like the β-subunit of the F_1-ATPase, the three-dimensional structure of *H. pylori* FliI is predicted to consist of three domains: an N-terminal β-barrel of six strands, a central nucleotide binding domain containing a nine-stranded β-sheet with nine associated helices, and a C-terminal bundle of helices. However, the C-terminal domain of FliI consists of four helices rather than the five seen in the C-terminal domain of F_1-ATPase α- and β-subunits. The model also suggests that the catalytic site is conserved with the residues of the conserved P-loop (Walker A motif) positioned to bind the phosphates of ATP, and Glutamate 193 is appropriately positioned to activate a water molecule for a nucleophilic attack on the terminal phosphate leading to ATP hydrolysis.

1.6 FliI Homologues

1.6.1 F_0F_1 ATP-synthase

Given the predicted structural homology of the F_1-ATPase β-subunit to FliI (and to a lesser extent the α-subunit), it is worthwhile considering the catalytic cycle of this enzyme, the structural transitions involved and the coupling of these to enzyme function.

Oxidative phosphorylation establishes transmembrane proton gradients. When protons move down this concentration gradient via the F_0 subunit, releasing free energy, the F_0F_1 ATP-synthase couples this free energy release to ATP synthesis (Figure 1.9). Disruption of the stalk connecting the F_0 and F_1 releases the F_1 catalytic subunit of the enzyme from the membrane bound F_0 component responsible for proton conductance. The F_1 subunit consists of five different proteins present in the following stoichiometry: $3\alpha:3\beta:1\gamma:1\delta:1\varepsilon$. The α- and β-subunits alternate to form a roughly spherical complex with γ helices at the centre (Figure 1.10).

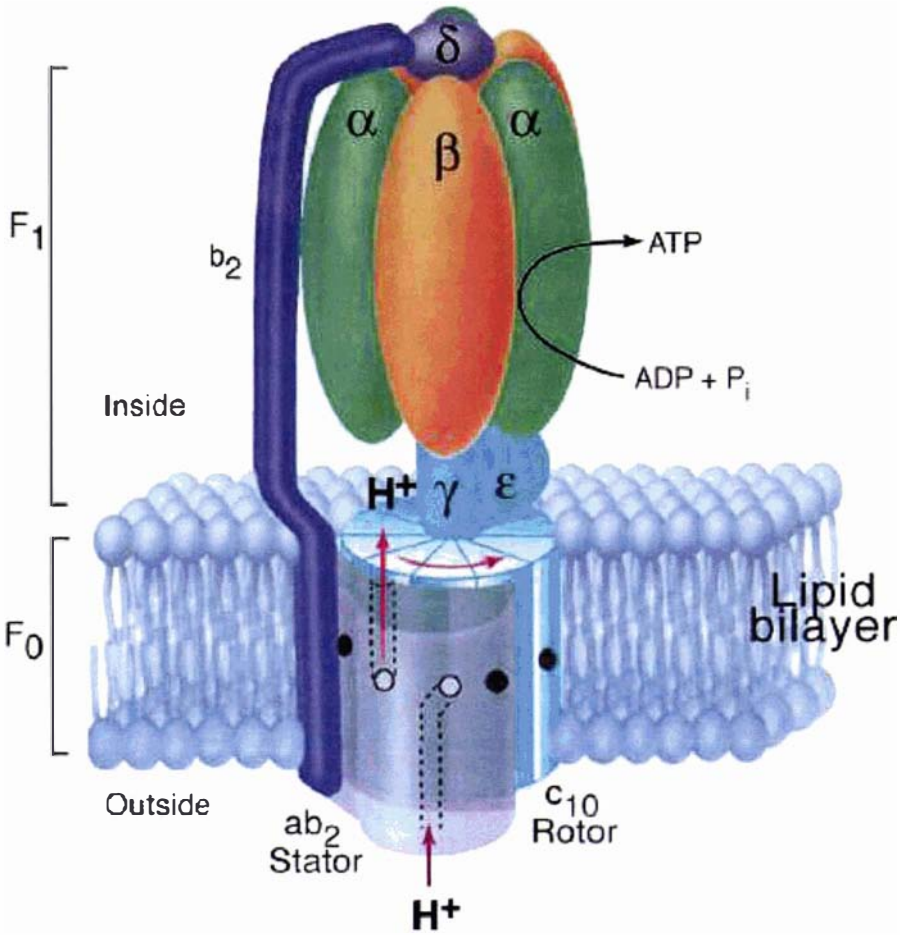


Figure 1.9. Model of the synthesis of ATP by the *E. coli* F_0F_1 ATP-synthase. From Fillingame *et al.* (55).

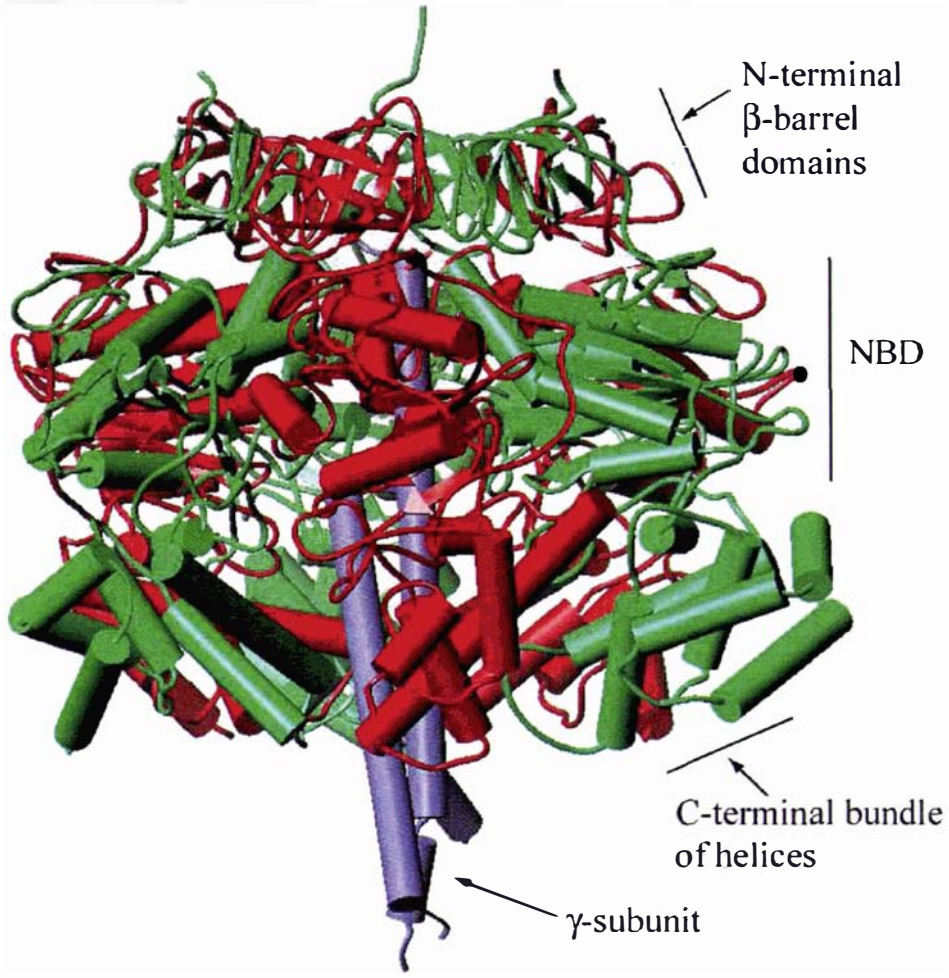


Figure 1.10. The structure of the αβ-hexamer of the F_1-ATPase. In this figure the α-subunits are shown in green, β-subunits in red, the central γ-subunit is shown in blue. This structure was generated with the PDB file of Abrahams *et al* (1), that contains the coordinates for the crystal structure of the Bovine αβγ-complex of the F_1-ATPase.

1.6.1.1 Structure of the F_1-ATPase

The Bovine F_1-ATPase consists of three α- and three β-subunits arranged in a pseudohexameric fashion around a γ-subunit in the core of the structure. The α- and β-subunits only share 20% identity but they have almost identical folds consisting of three domains. The N-terminal domain consists of a six stranded β-barrel which together with the other α- and β-subunits form a continuous 24-stranded β-sheet. The first 18-24 residues of the N-terminus are disordered and not visible in the crystal structure of F_1-ATPase, but the

22 N-terminal residues of the α-subunit have been shown to form an α-helix that interacts with the δ-subunit in the complete F_0F_1 complex (193) (1). Following the N-terminal domain is a central nucleotide binding domain (NBD) containing a nine-stranded β-sheet with nine associated helices. This is followed by a C-terminal bundle of helices consisting of seven and six helices in the α- and β-subunits respectively.

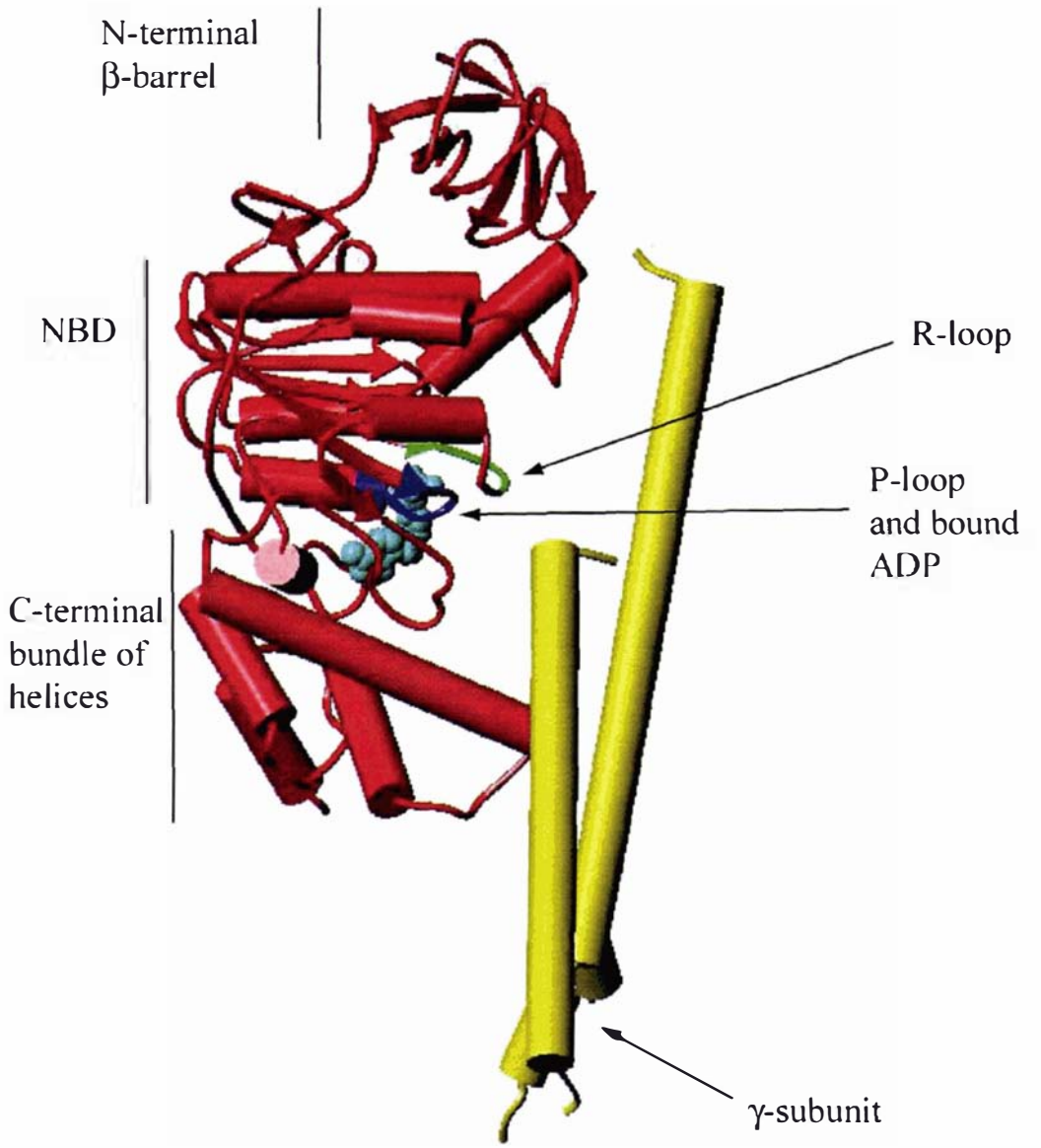


Figure 1.11. An F_1-ATPase β-subunit and the γ-subunit. The β-subunit is shown in red and ADP is bound in the catalytic site. This structure was generated with the PDB file of Abrahams *et al* (1), that contains the coordinates for the crystal structure of the αβγ-complex of the F_1-ATPase.

1.6.1.2 Catalytic sites of the F_1-ATPase

The catalytic sites of the F_1-ATPase are formed primarily by residues of the β-subunit and lie at the interface of the α- and β-subunits in the NBD. The α-subunits do not contain active catalytic sites. The residues involved in binding the phosphates of ATP are contributed by the conserved P-loop (Walker A motif) with the consensus sequence GXXXXGKT/S, while the adenine moiety binds in a hydrophobic pocket formed by α-Y345, α-F424 and α-F418. The octahedral coordination of the Mg^{2+} in the active site involves the hydroxyl oxygen of β-T163 (Walker A), a water molecule hydrogen bonded to β-D256 (Walker B), two water molecules hydrogen bonded to β-E192 and the β- and γ-phosphate oxygens of ATP. The only significant contribution from the α-subunit to the catalytic sites is α-R373. The guanidinium group of α-R373 serves to stabilise the negative charge on the terminal phosphate following nucleophilic attack by a water molecule that is activated by the carboxylate group of β-E188. The absence of a functional equivalent of β-E188 in the nucleotide binding sites of the α-subunit renders these sites inactive (1).

1.6.1.3 Rotary Catalysis

The three active catalytic site of the F_1-ATPase are in different conformations despite the similarities of the α- and β-subunits because the quarternary structure of the F_1-ATPase is asymmetric. The asymmetry is because the subunits differ in their interactions with each other and make unique contacts with the central γ-subunit (Figure 1.11). For example in the crystal structure of Abrahams and colleagues (1), the R-loop of the β_E-subunit (empty catalytic site) forms hydrogen bonds with γ-R254 and γ-Q255. This interaction disrupts the interaction between strands three and seven of the NBD, which in turn dislocates the P-loop from the active site. Experiments have shown that rotation of the γ-subunit occurs in response to ATP hydrolysis (131). As the γ-subunit rotates during hydrolysis the interaction with the R-loop will be broken, allowing the β_E active site to adopt a different conformation, and in this manner the R-loop is thought to couple ATP hydrolysis to the rotation of the γ-subunit.

1.6.2 Rho transcriptional terminator

Transcription termination in bacteria occurs by two mechanisms (Richardson and Greenblatt, 1996). One of these involves the formation of a stem loop in the mRNA which dissociates the RNA polymerase from the DNA template, whereas the other mechanism relies on Rho, an RNA-DNA helicase. Rho binds to mRNA at a cytidine rich Rho utilisation sequence (*rut*) (127) (145) and translocates along the RNA in a 5'-3' direction until it dissociates the DNA-RNA hybrid.

Structurally the Rho transcription termination factor is similar to the α- and β- subunits of the F_1-ATPase (Figure 1.11). It shares 17% and 21% sequence identity with these subunits respectively, and adopts a similar tertiary structure. This consists of an N-terminal RNA binding domain and a C-terminal nucleotide binding domain (Figure 1.12). However, the N-domain of Rho is significantly different from the N-domain of the F_1-ATPase catalytic subunits. It contains two sub-domains: an N-terminal bundle of three α-helices and a C-terminal β-barrel domain. This five-stranded β-barrel has an oligosaccharide/olignucleotide (OB) fold and it contains an RNP 1-like nucleotide binding motif that forms the primary, high affinity, binding site (27) (5). The C-terminal nucleotide binding domain of Rho is similar to the nucleotide binding domain of the α- and β- subunits of the F_1-ATPase, and the FliI model already mentioned, it is comprised of seven parallel β-strands sandwiched between seven α-helices. This domain contains the Walker A and B motifs important for the function of the catalytic sites, and as such mutations within these motifs cause a drastic decrease in ATP hydrolysis and transcription termination by Rho (42). Like F_1-ATPase, Rho contains three catalytic sites in different conformations that lie at the interface of two subunits (164), and these catalytic sites act sequentially in a cooperative manner (174) (25).

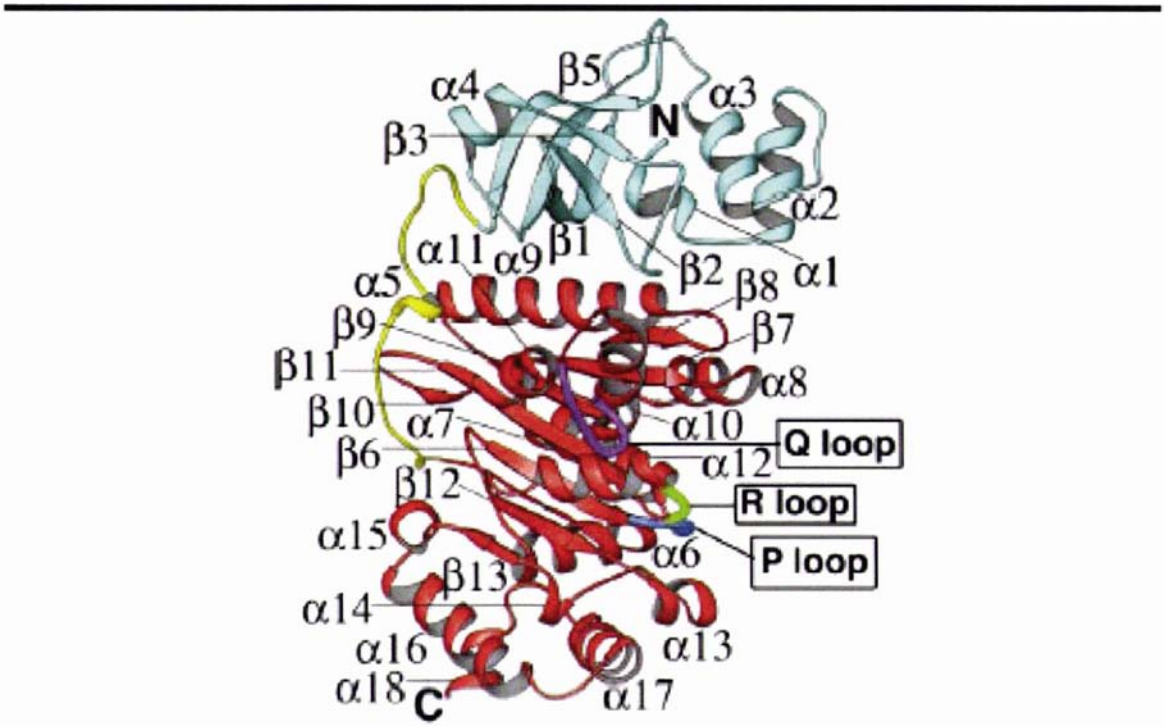


Figure 1.12. The structure of Rho transcription termination factor. From Skordalakes *et al* (164).

The crystal structure of Rho also shows that the subunits form a hexameric ring 120 Å wide, with an internal hole that varies in diameter from 20-35 Å. The R- and Q-loops line the internal hole of the Rho hexamer. The Q-loops form a constriction in the hole in the centre of the Rho hexamer 20 Å in diameter, and are thought to be a secondary low affinity RNA binding site. The R-loop is important for both ATP and RNA binding and given the proximity of the R- and P-loops, the R-loop may couple ATP hydrolysis to RNA translocation.

The structure of Rho has given considerable insight into the mechanism of RNA binding and translocation (Figures 1.12, 1.13). Firstly, the N-terminal primary RNA binding sites face inwards and down such that the 3' end is oriented towards the interior of the hexamer channel where RNA is known to bind (164) (84). Secondly, the hexameric ring was crystallised in the open "lock washer" conformation that is often seen by electron microscopy, and this is thought to allow RNA entry to the Rho channel from the side

following binding to the primary binding site. The 12 Å opening is formed by the upward rotation of one subunit with respect to another by 15 degrees, and the perpetuation of this rotation around the ring. Thirdly, the three catalytic sites each have different conformations and binding involves the residues of the Walker A and B motifs, but in the open conformation these residues do not make contact with the γ- and β- phosphates. These structural observations and other data lead to the following mechanistic model. The transcript mRNA binds to the N-terminal high affinity primary binding sites (22), and the 3'end of the RNA is directed towards the Rho channel (164). This binding leads to the opening of the hexameric ring, and the RNA inserts through the side of the ring and binds to the Q-loop as the secondary low affinity binding site (195) (196). RNA binding stimulates the closing of the hexameric ring (63) and cooperative sequential ATPase activity at the three active sites (174) (173). Within this catalytic cycle the R-loop would function in a similar manner to the F_1-ATPase R-loop by periodically contact the RNA and concomitantly dissociate the P-loop from the active site (122). Thus, the catalytic mechanisms of F_1-ATPase and Rho are likely to be very similar, with the main differences being that the Rho hexamer consists of a single subunit and translocates RNA rather than having a rotating γ-subunit inserted in the central channel of the enzyme hexamer.

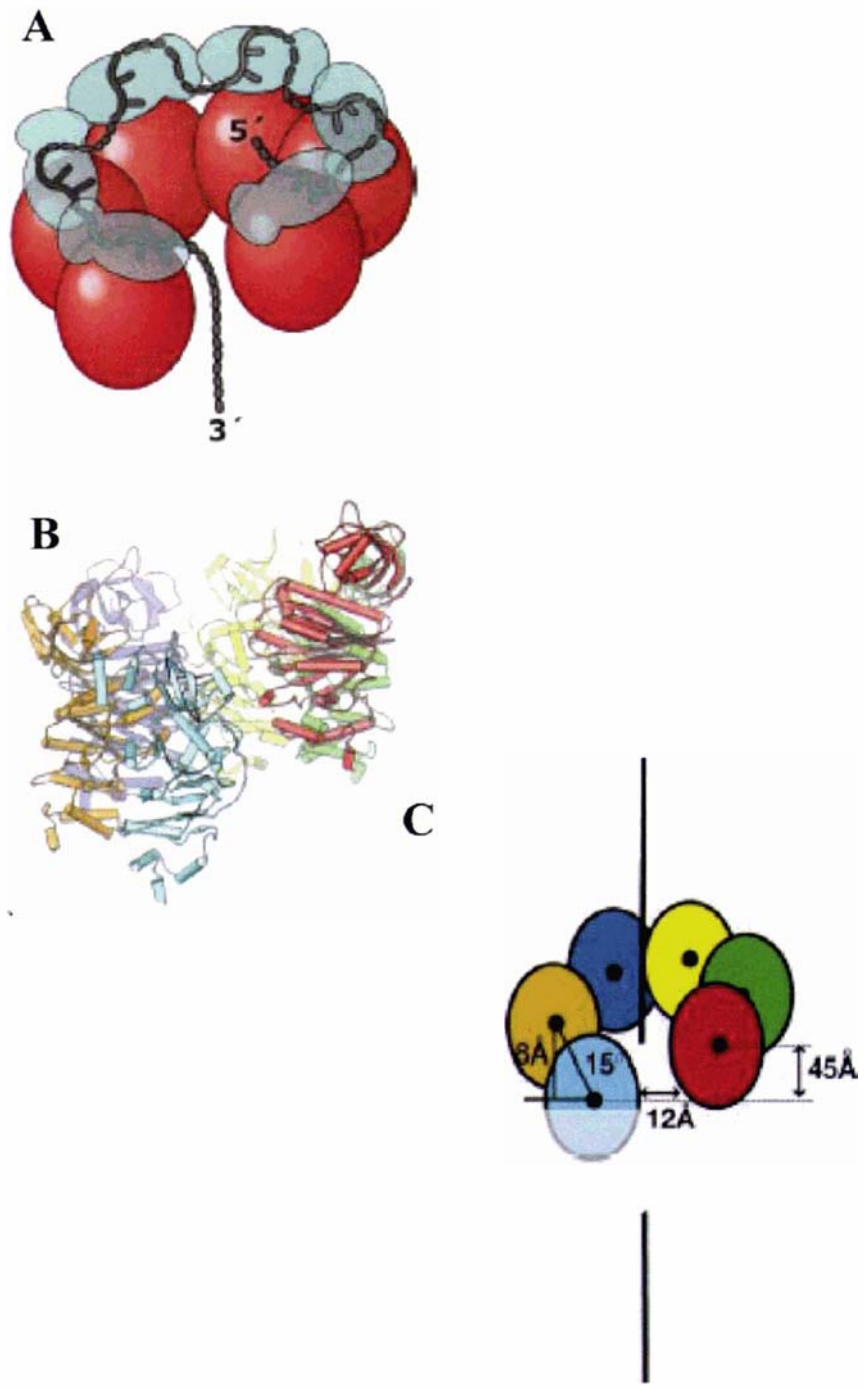


Figure 1.13. The quarternary structure of Rho transcription termination factor. From (164). **A**. Loading of RNA into the Rho hexamer. **B**. Ribbon diagram of the crystal structure of Rho in the "lock washer" conformation. **C**. Schematic showing the relative rise and offset of Rho subunits in the "lock washer" conformation.

The structure of the flagellum export apparatus is unknown. Functionally the export apparatus can be regarded as a complex which recognises and distinguishes exported flagellum substrates, unfolds them and energises their export.

When this project began *S. typhimurium* FliI had been proposed to be an ATPase that energises flagellum protein export. This was supported by catalytic mutations in the Walker A and B boxes that led to aflagellate, non-motile cells in both *S. typhimurium* and *H. pylori* (43) (188) (142) (77) (52). Furthermore, *Salmonella* researchers had also acknowledged significant sequence identity between FliI and the F_1-ATPase β-subunit. But, these researchers also disregarded any structural or functional relationship between the proteins.

The overriding hypothesis of this thesis is that the F_1-ATPase and to a lesser extent the Rho transcriptional terminator are indeed structural homologues of FliI. Furthermore, these proteins are functionally analogous although in different biological contexts.

This hypothesis has several implications for the structure and function of FliI. Firstly, in the absence of a homologue of the α-subunit, FliI must assemble as a homohexamer and during the course of this project this was demonstrated for the *S. typhimurium* FliI (31). Secondly, the FliI hexamer will demonstrate catalytic cooperativity, and such catalytic cooperativity was also demonstrated for *S. typhimurium* during the course of this project (31). Thirdly, in the absence of a homologue of the γ-subunit, a rotary catalysis mechanism for the hydrolysis of ATP will be coupled to the movement, unfolding and export of flagellar proteins through the centre of the FliI hexamer. This coupling of the ATP hydrolysis cycle to the movement and unfolding of flagellar proteins will involve the FliI R-loop contacting the proteins. This third point has yet to be demonstrated.

When FliI and the rotary catalytic mechanism are considered in the broader context of the F_1-ATPase model a second hypothesis can be proposed. That is, another protein of the export apparatus must interact with FliI and act as a stator to prevent rotation of the hexamer with respect to the flagellum substrate being exported. We propose this function is fulfilled by FliH. Indeed, it has been previously recognised that FliH contains amino acid

sequence identity to the b-subunit of the F_1-ATPase stator within the N-terminal half of the protein (212). There is also some sequence identity shared between the C-terminal half of FliH and the δ-subunit of the F_1-ATPase stator.

When this project began FliH was known to be important for flagellum biogenesis as mutants are non-motile and defective in flagellation due to a failure export flagellar proteins (188) (117). A specific function for FliH to explain these observations had not been suggested other than a role in protein export. If FliH is assumed to function as a stator this would imply that it will inhibit ATPase activity to prvent futile uncoupled hydrolysis of ATP. Inhibition of FliI ATPase activity by FliH was demonstrated with the *S. typhimurium* proteins during the course of this project (118). Furthermore, if FliH is analogous to the stator it will form an elongated dimer (213) and exert this inhibitory effect by interacting with FliI. This interaction was demonstrated concurrently in this project and with the *S. typhimurium* FliH and FliI (118).

The interaction of the δ-subunit of the stator with the N-terminus of the α-subunit is quite well characterised (213) (193) (214) (215) (216) in our F_1-ATPase model. Therefore if we extend this analogy it is possible to hypothesize the molecular nature of the FliH-FliI interaction. We would expect this interaction to involve a hydrophobic interface and a helix at the N-terminus of FliI. The major findings of this project support this hypothesis.

2. Materials and Methods

2.1 Bacterial strains and growth conditions

E. coli strains (Table 2.1) were grown in either Luria Bertani (LB) broth at 37°C in an Innova 4330 Refrigerated Incubator Shaker (New Brunswick Scientific) at 220rpm, or on solid LB Agar at 37°C in an incubator. Following growth on LB agar plates, the plates were stored at 4°C for no longer than three weeks. The growth temperature of liquid cultures was varied during protein expression, as described below.

The LB Broth was made in 1 L of deionised H_2O by adding 10 g of Bacto™ Tryptone (Becton, Dickinson and Company, BD), 5 g of Bacto™ Yeast Extract (BD), 10 g of NaCl and stirring to dissolve on a magnetic stirrer. To make LB Agar, 15 g of Difco™ Granulated Agar was additionally added to LB broth. LB broth and LB Agar were sterilised at 121 °C for 40 min in an autoclave. The LB agar was allowed to cool to approximately 50 °C before being poured into petri dishes.

H. pylori strains (Table 2.1) were grown on Chocolate Blood Agar (CBA) at 37 °C in a CO_2 incubator with 5% CO_2. CBA was made with 4.4% (w/v) columbia blood agar base (Difco), after sterilisation the media was allowed to cool to approximately 70 °C, and defibrinated horse blood (Life Technologies) was added to a final concentration of 5%. The media was then placed in a 70 °C and stirred periodically until haemolysis occurred. Cells were also grown in Tryptic Soy Broth (Difco) supplemented with 5% FCS (fetal calf serum, Life Technologies) in bottles placed in anaerobic jars containing CampyGen sachets (Oxoid) to generate the necessary microaerobic atmosphere. The anaerobic jars were then placed in a 37 °C shaker for 24-48 hours.

For long term storage of *E. coli*, 600 µl of a liquid culture were mixed with an equal volume of 70% glycerol and frozen at -80 °C. Similarly, *H. pylori* was mixed with an equal

volume of 70% glycerol and frozen at -80 °C, but the cells were first harvested from a CBA plate with a swab and resuspended in Tryptone Soya Broth (TSB, Difco).

Table 2.1. Bacterial strains used in this study

Bacterial strain	Characteristics	Source / reference
H. pylori CCUG 17874	*H. pylori* type strain, motile	Culture collection University of Gothenburg, Sweden
H. pylori 26695	*H. pylori* genome strain, non-motile	(182)
E. coli DH5α™	F⁻ φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F)U169 *deo*R *rec*A1 *end*A1 *hsd*R17(r_k^-, m_k^+) *pho*A *sup*E44 *thi*-1 *gyr*A96 *rel*A1 λ⁻	(Invitrogen™ Life Technologies)
E. coli Rosetta™	F⁻ *ompT hsdS*$_B$ ($r_B^-m_B^-$) *gal dcm pRARE* (*argU, argW, ileX, glyT, leuW, proL*)	(Novagen)

2.2 DNA Methods

2.2.1 DNA preparation

2.2.1.1 Plasmid purification

Plasmids were purified using the QIAprep® Spin Miniprep Kit, and when a larger scale plasmid preparation was needed for plasmid stocks the plasmid was purified using the QIAGEN® Plasmid Maxi Kit.

Refer to Table 2.3 below for a list of plasmid vectors used in this study and recombinant plasmids created.

1.6.2.2 Preparation of *H. pylori* chromosomal DNA

H. pylori cells were grown on a CBA plate for two days and then harvested with a swab and resuspended in 1.5 ml of 1 x PBS in a sterile cryotube. The cells were then pelleted by centrifugation for 20 seconds in a Beckman Coulter Microfuge® 18 microcentrifuge at top speed (14,000 x g). The supernatant was discarded and the cells were resuspended in 100 µl of 1 x TE (10 mM Tris, pH 8.0, 1 mM EDTA). Then 100 µl of GES (5 M Guanidinium thiocyanate, 100 mM EDTA, 0.5% (w/v) sarkosyl) and 2 µl of 10 mg/ml RNase was added, and mixed thoroughly by inversion of the tube until the mixture became completely clear.

The DNA was then recovered by adding 1 ml of ice cold ethanol (stored at -20 °C) and mixing gently by inversion. The precipitated chromosomal DNA was then removed with a hooked glass capillary, dunked in cold 70% ethanol, and then resuspended in 100-200 µl of 1 x TE in an eppendorf tube. Following the addition of 20 µl of Proteinase K (50 mg/ml, Sigma), the tube was incubated at 50 °C for 8 hours in a waterbath to remove protein associated with the DNA. The solution became clear during this incubation and the DNA was recovered by precipitation as before by the addition of 1 ml of cold 100% ethanol, and the precipitated DNA was washed again in 70% ethanol before being resuspended in 100-200 µl of 1 x TE.

2.2.2 DNA analysis

2.2.2.1 Agarose gel electrophoresis

DNA was sized during cloning by agarose gel electrophoresis. Agarose gels were made by the addition 1.5% agarose (mass/volume) to 30 ml of 1xTAE (40 mM Tris-acetate, 1 mM EDTA) gel running buffer, heated to dissolve the agarose for two minutes in a microwave, allowed to cool and then poured into an Easycast™ Horizontal Electrophoresis (Owl Separation Systems) gel casting tray containing a 10-well comb. Agarose gels were allowed to set for 20 min before the comb was removed, and the gel was placed in the electrophoresis apparatus and submerged in 1xTAE.

DNA samples, 1-5 µl, were mixed with an equal volume of loading buffer (0.25% bromophenol blue, 0.25% xylene cyanol FF, 30% glycerol in water) and loaded on the gel. A DNA molecular weight marker (DirectLoad™ Wide Range DNA Marker, Sigma) was also loaded on the gel. The gel was run at 100 V for 45 min. After electrophoresis the gel was stained in water containing 0.5 µg/ml ethidium bromide. The DNA in the gel was then visualised and photographed on an IS-1000 Digital Imaging System (Alpha Innotech).

2.2.2.2 Restriction Digestion

Restriction endonucleases (from Fermentas, New England Biolabs or Roche) were used according to manufacturers instructions, to prepare PCR products for ligation to plasmid

vectors, to linearise plasmid vectors, and check potentially recombinant plasmids for the presence of an inserted fragment of DNA. A typical reaction mixture consisted of 0.5 µl-1 µl of the restriction endonuclease (such that the glycerol concentration was always less than 10% of the total reaction volume), 1-10 µl target DNA (usually less than 1 µg of total DNA), 2 µl of 10X reaction buffer, H_2O to a final volume of 20 µl.

2.2.2.3 DNA sequencing

Recombinant plasmids (Table 2.3) were sequenced by the MUSEQ, the Massey University Sequencing Facility. The oligonucleotides used for the sequencing reactions either annealed to the DNA sequence of interest or to the vector DNA flanking the multiple cloning site (Table 2.2). Sequences were analysed using the GeneJockey analysis software (Biosoft).

2.3 Cloning procedures

2.3.1 DNA amplification by PCR

PCR was performed in a MyCycler™ Thermal Cycler (Bio-Rad) in a total reaction volume of 20 µl. Each reaction contained 1x reaction buffer (20 mM Tris-HCl (pH 8.8 at 25 °C), 10 mM $(NH_4)_2SO_4$, 10 mM KCl, 0.1% Triton X-100, 0.1 mg/ml BSA, 2 mM $MgSO_4$), 0.2 mM dNTPs, 0.2 µM of the forward primer and 0.2 µM of the reverse primer, 1 ng of template DNA, 0.2 µl (0.5 U) of *Pfu* polymerase (Fermentas), H_2O was added to reach the final volume. A list of primers is presented in Table 2.2.

A typical amplification program consisted of 30 cycles. Each cycle consisted of a 1 min denaturation step at 95 °C, a 30s annealing step at 56 °C, and a 1 min elongation step at 72 °C. In the first cycle the denaturation step is extended to 1 min. In the last cycle the elongation step was extended to 3 min. Adjustments were made to these conditions to accommodate the Tm of the primers and the nature and length of the template. If a genomic DNA template was used the denaturation steps were longer, the annealing temperature was adjusted to keep it at 4 °C below the Tm of the primer, the elongation time was adjusted to allow two minutes for the polymerase to amplify every kb of DNA.

Table 2.2. Oligonucleotides used in this study

Name	Restriction Sites	Sequence[a]	Application
pUC/ M13 FP	na	GTT TTCCCAGTCACGAC	Sequencing the pUC based plasmids used for *HP0256* knock-out mutagenesis
pUC/ M13 RP	na	CAGGAAACAGCTATGAC	Sequencing the pUC based plasmids used for *HP0256* knock-out mutagenesis
ML022	*Bam*HI	CG**GGATCC** CGGGGCGAAAGATTGGAGATTT	Forward primer used for the PCR amplification of arm 1 for the *HP0256* knock-out mutagenesis
ML023	*Eco*RI	CG**GAATTC**CGTTACGCATGCAAGCCCTC	Reverse primer used for the PCR amplification of arm 2 for the *HP0256* knock-out mutagenesis
ML027	*Bgl*II	CCATCGT**AGATCT**GGGCTGCAGCGAATTTTT TCATAGAAAAATCG	Reverse primer used for the PCR amplification of arm 1 for the *HP0256* knock-out mutagenesis
ML028	*Bgl*II	GCAGCCC**AGATCT**ACGATGGGCAATTAAAA AGCGCTCTAAGAAT	Forward primer used for the PCR amplification of arm 2 for the *HP0256* knock-out mutagenesis
ML024	na	GCTAGTA**GCT**GGCGATGATT	Mutagenesis of the *flil* R-loop, E_{311}-A, forward primer

Table 2.2 continued

ML025	na	GCTAGTAGAG**GCT**GATGATT	Mutagenesis of the *fliI* R-loop, G_{312}-A, forward primer
ML026	na	AGTAGAGGGCGCT**GAT**TTGA	Mutagenesis of the *fliI* R-loop, D_{313}-A, forward primer
ML043	na	CAAATCATCGCC**AGC**TACTAGC	Mutagenesis of the *fliI* R-loop, E_{311}-A, reverse primer
ML044	na	CAAATCATC**AGC**CTCTACTAGC	Mutagenesis of the *fliI* R-loop, G_{312}-A, reverse primer
ML045	na	GCTCAAATC**AGC**GCCCTCTACT	Mutagenesis of the *fliI* R-loop, D_{313}-A, reverse primer
ML047	*Bgl*II	GA**AGATCT**TTATCTTAAGATTTCTTCTAATT GCTGAAAGC	Reverse primer for the PCR amplification of full-length *fliI* for R-loop mutagenesis
ML048	na	ATTCAATAAGGTAATTTAATGCCCCTAAAAT CCTTAAAAAACCG	Forward primer for the PCR amplification of full-length *fliI* for R-loop mutagenesis, it also anneals to the 3' end of the *recA* promoter
ML049	*Sal*I	ACGC**GTCGAC**AAGCTGATTGCGCTCACATC ATAA	Forward primer for the amplification of the *recA* promoter used for the expression of R-loop mutants of *fliI*

Table 2.2 continued

ML050	na	AGGATTTTAGGGGCATTAAATTACCTTATTG AATTTGATTTGTTTATTCTATCAAA	Reverse primer for the amplification of the *recA* promoter used for the expression of R-loop mutants of *fliI*, it also anneals to the 5' end of *fliI*
ML051	na	CGATTAGCGCCTGTCATTACA	*fliI* sequencing primer
ML052	na	GAATTTATAGAGAAAAACCTGAAAGGG	*fliI* sequencing primer
ML053	na	CTTTCCTTATTGCCTCAATTAATGGAG	*fliI* sequencing primer
SAM003	*Bam*HI	GACCTG**GGATCC**CGCTATGGGAGCGTGAAA AAAATC	Forward primer used for the PCR amplification of *fliI* DNA encoding the N-terminal domain of FliI without the first 18 amino acids
SAM005	*Eco*RI	GCGCCG**GAATTC**TCAATTCAACCCCTCTTTT AAGAACAG	Reverse primer used for the PCR amplification of *fliI* DNA encoding the N-terminal domain of FliI
ML057	*Bam*HI	CG**GGATTC**CCC**GCG**AAATCCTTAAAAAACC GCTTGAAT	Forward primer for the PCR amplification and mutagenesis of the DNA encoding the *fliI* N-terminal domain, L_3-A
ML058	*Bam*HI	CG**GGATCC**CCCCTAAAATCC**GCG**AAAAACC GCTTGAAT	Forward primer for the PCR amplification and mutagenesis of the DNA encoding the *fliI* N-terminal domain, L_6-A

Table 2.2 continued

ML059	*Bam*HI	CG**GGATCC**CCCCTAAAATCCTTAAAAAAC**G CG**TTGAAT	Forward primer for the PCR amplification and mutagenesis of the DNA encoding the *fliI* N-terminal domain, R_9-A
ML060	*Bam*HI	CG**GGATCC**CCCCTAAAATCCTTAAAAAACC GC**GCG**AATCAG	Forward primer for the PCR amplification and mutagenesis of the DNA encoding the *fliI* N-terminal domain, L_{10}-A
ML063	*Bam*HI	CG**GGATCC**CCCCTAAAATCC**GAA**AAAAACC GCTTGAAT	Forward primer for the PCR amplification and mutagenesis of the DNA encoding the *fliI* N-terminal domain, L_6-E
ML064	*Bam*HI	CG**GGATCC**CCCCTAAAATCCTTA**GCG**AACC GCTTGAAT	Forward primer for the PCR amplification and mutagenesis of the DNA encoding the *fliI* N-terminal domain, K_7-A
ML065	*Bam*HI	CG**GGATCC**CCCCTAAAATCCTTAAAAAAC**G AA**TTGAAT	Forward primer for the PCR amplification and mutagenesis of the DNA encoding the *fliI* N-terminal domain,R_9-E
ML066	*Bam*HI	CGG**GGATCC**CCCCTA**GCG**TCCTTAAAAAAC CGCTTGAAT	Forward primer for the PCR amplification and mutagenesis of the DNA encoding the *fliI* N-terminal domain, K_4-A
ML020	*Bam*HI	GGTTT**GGATCC** TCATTGAATAGCCGTAAGAACTTG	Forward primer for the PCR amplification of *fliH*
ML061	*Bam*HI	GACCTG**GGATCC**GAAAACGCTAAAAACGAC GGCTATAAA	Forward primer for the PCR amplification of *fliH* DNA encoding FliH 94-258

Table 2.2 continued

ML062	*Bam*HI	GACCTG**GGATCC**ACTCACAGCGTGAATGAA GAAAAAAAC	Forward primer for the PCR amplification of *fliH* DNA encoding FliH 117-258
ML021	*Eco*RI	GGTTT**GAATTC**TCACACCTTAAAATTTTCCA ACAC	Reverse primer for the PCR amplification of *fliH*

[a]Restriction sites and codon changes are in bold type, the sequences are written in the 5'-3' direction

2.3.2 Ligation

PCR products were ligated to linearised plasmid vectors using T4 DNA ligase (Invitrogen). A typical reaction mixture consists of 4 µl of 5X reaction buffer (250 mM Tris HCl (pH 7.6), 50 mM $MgCl_2$, 5 mM ATP, 5 mM DTT, 25% (w/v) PEG-8000), 1 µl DNA ligase, insert and vector DNA at a molar ratio of 3:1, H_2O to a final volume of 20 µl. The ligation mixture was incubated at 14 °C for 12 hours in the MyCycler™ Thermal Cycler (Bio-Rad), then used to transform *E. coli* DH5α.

2.3.3 Preparation of competent *E. coli*

The *E. coli* strain to be used as the recipient for transformation was streaked for single colonies on LB agar from a -80 °C glycerol stock, and grown overnight in a 37 °C incubator. The next day a single colony was used to inoculate a 5 ml LB broth culture that was then grown overnight in a shaker at 37 °C. The next day this 500 µl of the overnight culture was diluted in 50 ml of LB broth and grown at 37 °C in a shaker until an $OD_{600\ nm}$ of 0.4-0.6. The cells were then harvested by centrifuging at 6000 x g for 5 min at 4 °C. The supernantant was decanted and the cell pellet was kept on ice at all times during the remainder of the protocol. The cells were resuspended in 10 ml of CM1 (10 mM NaOAc, pH 5.6, 50 mM $MnCl_2$, 5 mM NaCl). The cells were then centrifuged again at 6000 x g for 5 min at 4 °C, the supernatant decanted and the pellet resuspended in 1 ml of CM2 (10 mM NaOAc, pH 5.6, 5 mM $MnCl_2$, 5% glycerol, 70 mM $CaCl_2$). The cells were then transferred to eppendorf tubes in 50 µl aliquots, and frozen at -80 °C until needed.

Throughout this protocol the cells were kept cold following harvesting, all centrifuge tubes were sterile and cold, and all pipette tips were sterile.

2.3.4 Transformation of *E. coli*

Competent *E. coli* cells were transformed using a heat shock method. Briefly, the competent cells were removed from storage at -80 °C, allowed to thaw on ice and then 50 µl of cells were added to the plasmid DNA or ligation in an eppendorf tube. The cells and

the DNA were incubated on ice together for 30 min. The cells were then heat shocked for 1 min at 42 °C in a water bath, then placed back on ice for 2 min. Then 450 µl of LB broth was added to the cells and they were allowed to recover for 40 min at 37 °C in a shaker, before the cells were spread on LB agar containing the appropriate antibiotic to select for transformants. The LB agar plate was placed at 37 °C overnight to allow transformant cells to grow.

2.3.5 Preparation of electrocompetent *H. pylori*

H. pylori cells grown on a CBA plate for 2 days were harvested by swabbing the plate into 1.3 ml of TSB. These cells were then spread on 6 fresh CBA plates and allowed to grow for 24 hours. The cells were then harvested once more with a swab and resuspended in 30 ml of cold (4 °C), sterile deionised water in a centrifuge tube. The cells were then centrifuged at 8,000rpm (9,610 xg) for 15 min at 4 °C. The supernatant was discarded and the cell pellet was resuspended in 15 ml of cold (4 °C) electroporation buffer (15% glycerol, 10% sucrose). The cells were then centrifuged again at 8,000 rpm (9,610 x g) for 15 min at 4 °C, the supernatant discarded and the cells then resuspended in 1 ml of cold electroporation buffer. The cells were then centrifuged again at 9,000 rpm (12,200 x g) for 10 min at 4 °C, the supernatant discarded and the cell pellet resuspended in 300 µl of cold electroporation buffer, dispensed in 50 µl aliquots and frozen at -80 °C until needed.

Throughout this protocol the cells were kept cold following harvesting, all centrifuge tubes were sterile and cold, all pipette tips were sterile.

2.3.6 Transformation of *H. pylori* by electroporation

1 µg of recombinant plasmid DNA was added to 100 µl of electrocompetent cells on ice. The mixture was then transferred to an electroporation cuvette on ice, the cuvette had been stored at -20 °C. The electroporation cuvette was then pulsed in the electroporator (Gene Pulser, Bio-Rad) with the settings 200 Ω, 25 µF, and 1.5 kV. The time constants were usually around 4.8. The electroporated cells were then removed from the cuvette, placed in an eppendorf tube with 800 µl of CO_2 equilibrated TSB broth (Difco). The cells were spread the cells on a non-selective CBA plate and allowed to recover in the CO_2 incubator

at 37 °C and 5% CO_2 for 24 hours. After 24 hours the cells were collected and spread on to selective CBA containing the appropriate antibiotic. After 48-78 hours antibiotic resistant colonies were apparent on the selective CBA. A single colony was picked and spread on a fresh selective CBA plate. Once a lawn of cells had grown on this plate, they were used to make a glycerol stock for long tem storage at -80 °C.

Table 2.3. Plasmids used in this study

Plasmid	Characteristics	Reference/source
pUC19	Ap^r; *oriColE1*; blue/white; MCS	(204)
pRY109	Ap^r; Cm^r; *Campylobacter coli cat* gene in *BamHI* site of pUC18	(205)
pHel3	Kan^r by virtue of *aphA-3* from pILL550; *oriColE1*; *oriT*; *repA*.	(68)
pGEX	Ap^r; pBR322 *ori*; *tac* promoter; *lac* I^q inducer; PreScission, thrombin and factor Xa protease recognition site; glutathione-S-transferase; MCS	(166)
pML016	Ap^r; internally deleted *HP0256* PCR product in pUC19 cut with *BamHI* and *EcoRI*	This study
pML017	Ap^r; Cm^r; *cat* gene from pRY109 in *BglII* digested pML016	This study
pML015	Ap^r; full-length *fliH* PCR from *H. pylori* 26695 in *BamHI/EcoRI* digested pGEX	This study
pML023	Ap^r; *recA* promoter fused to the full-length *fliI* in pGEX digested with *Bam*HI/*Eco*RI	This study
pML024	Ap^r; *recA* promoter fused to the full-length *fliI* containing the E_{311}-A mutation, in pGEX digested with *Bam*HI/*Eco*RI	This study
pML025	Ap^r; *recA* promoter fused to the full-length *fliI* containing the G_{312}-A mutation, in pGEX digested with *Bam*HI/*Eco*RI	This study
pML026	Ap^r; *recA* promoter fused to the full-length *fliI* containing the D_{313}-A mutation, in pGEX digested with *Bam*HI/*Eco*RI	This study

Table 2.3 continued

pML040	Ap^r; *fliI* PCR product encoding amino acids 2-91 and the L_3-A in pGEX digested with *Bam*HI/*Eco*RI	This study
pML041	Ap^r; *fliI* PCR product encoding amino acids 2-91 and the K_4-A in pGEX digested with *Bam*HI/*Eco*RI	This study
pML042	Ap^r; *fliI* PCR product encoding amino acids 2-91 and the L_6-A in pGEX digested with *Bam*HI/*Eco*RI	This study
pML043	Ap^r; *fliI* PCR product encoding amino acids 2-91 and the L_6-E in pGEX digested with *Bam*HI/*Eco*RI	This study
pML044	Ap^r; *fliI* PCR product encoding amino acids 2-91 and the K_7-A in pGEX digested with *Bam*HI/*Eco*RI	This study
pML045	Ap^r; *fliI* PCR product encoding amino acids 2-91 and the R_9-A in pGEX digested with *Bam*HI/*Eco*RI	This study
pML046	Ap^r; *fliI* PCR product encoding amino acids 2-91 and the R_9-E in pGEX digested with *Bam*HI/*Eco*RI	This study
pML047	Ap^r; *fliI* PCR product encoding amino acids 2-91 and the L_{10}-A in pGEX digested with *Bam*HI/*Eco*RI	This study
pML048	Ap^r; *fliH* PCR product encoding amino acids 92-258 in pGEX digested with *Bam*HI/*Eco*RI	This study
pML049	Ap^r; *fliH* PCR product encoding amino acids 117-258 in pGEX digested with *Bam*HI/*Eco*RI	This study

Table 2.3 continued

pSP110a	Ap^r; Cm^r; *cat* gene from pRY109 in a primer introduced *Bam*HI site of pSP106 (pUC based plasmid containing a *fliI* PCR fragment)	(142)
pSM001	Ap^r; *fliI* PCR product encoding amino acids 19-91 in pGEX digested with *Bam*HI/*Eco*RI	Moore Lab
pSM002	Ap^r; *fliI* PCR product encoding amino acids 19-434 in pGEX digested with *Bam*HI/*Eco*RI	Moore Lab
pSM004	Ap^r; *fliI* PCR product encoding amino acids 2-91 in pGEX digested with *Bam*HI/*Eco*RI	Moore Lab
pSM005	Ap^r; *fliI* PCR product encoding amino acids 2-434 in pGEX digested with *Bam*HI/*Eco*RI	Moore Lab
pSM006	Ap^r; *fliH* PCR product encoding amino acids 55-258 in pGEX digested with *Bam*HI/*Eco*RI	Moore Lab

2.4 Microscopy

2.4.1 Phase contrast microscopy

To evaluate motility *H. pylori* cells were grown for 24 hours in liquid culture under microaerobic conditions. Then 10 µl of the culture was placed on a glass slide, covered with a cover-slip, and examined at 400 x magnification using a Nikon Diaphot phase contrast microscope.

2.4.2 Electron Microscopy (EM)

Electron microscopy was used to evaluate the flagellation of *H. pylori* cells. Prior to EM cells were grown for 24 hours on a CBA plate in microaerobic conditions. A plate of cells was then harvested using a swab and resuspended in 1 ml of sterile H_2O. Three drops of this preparation was added to three drops 1% phosphotungstic acid (PTA, adjusted to pH 7.0 with potassium hydroxide) in a test tube. A single drop of this mixture was then placed on a piece of parafilm. A 200-mesh grid (Agar Scientific) covered with Formvar film (0.5% formvar in ethylene dichloride) was then floated on this drop for 1 min. The excess sample was then removed by touching the edge of the EM grid with a piece of filter paper. The grid was then observed using a Philips 201C transmission electron microscope operated at an accelerating voltage of 60 kV under conventional transmission electron microscopy conditions. Photographs of the cells were taken using Agfa Copex Positive PET 10 film. These photos were kindly developed by Doug Hopcroft, at The Horticultural and Food Research Institute of New Zealand, Palmerston North, New Zealand.

2.5 Protein Methods

2.5.1 Protein Purification

All liquid chromatography was performed using an ÄKTA FPLC (Fast Protein Liquid Chromatography) instrument from Amersham Pharmacia Biotech. The instrument included a UPC-900 Monitor, a P920 pump and a Frac-950 fraction collector.

2.5.1.1 Expression of GST-fusion proteins

N-terminal Glutathione-S-transferase-fusions to the proteins of interest were created by cloning the gene of interest into the multiple cloning site of the pGEX-6P-3 expression vector (Pharmacia Biotech).

E. coli Rosetta strains containing the recombinant pGEX plasmid of interest were streaked for single colonies on LB Agar supplemented with Ampicillin (100 μg/ml, Amp^{100}) and Chloramphenicol (34 μg/ml, Cm^{34}) and grown overnight at 37 °C. The following day a single colony was picked with a flame sterilised inoculation loop and used to inoculate a 50 ml LB Amp^{100} broth culture which was then grown overnight at 37 °C in the shaker. The following day 10 mL of the overnight culture was used to inoculate 1L of LB Amp^{100} Cm^{34} LB broth. This culture was grown at 37 °C in the shaker to an optical density at 600 nm ($OD_{600\ nm}$) of 0.6, as measured by the Beckman Coulter DU 640 Spectrophotometer. At this OD, expression of the GST-fusion protein was induced by the addition of IPTG (Ultrapure™ IPTG, Invitrogen Life Technologies) to a final concentration of 0.1 mM. The temperature of the shaker was then decreased to 20 °C and the cells grown for a further 16 hours.

For expression of FliI 2-434 and FliI 19-434, an overnight culture was used to inoculate 1L of Hyper Broth™ (Athena Enzyme Systems™). The cells were subsequently grown and protein expression induced under the same conditions as cells grown in LB broth.

2.5.1.2 Harvesting cells

To recover the cells and the expressed protein they contain, liquid cultures were spun in a Beckman Coulter J2-HS Centrifuge using the JLA-16.250 rotor for 30 min at 6000 x g at 4 °C to pellet the cells. The supernatant was then decanted and the pellet resuspended in 3 ml of lysis buffer (50 mM Tris-HCl (pH 7.0), 150 mM NaCl, 1 mM EDTA) per gram of cell pellet. The cells were then frozen at -80 °C until lysis and purification at a later date.

2.5.1.3 Lysis of cells

Cells were lysed by two passes in a French Press at 6,000 psi. The lysate was centrifuged at 10000 x g for 30 min at 4 °C in the JA-200 rotor to separate the soluble fraction of the lysate from unlysed cells and cell debris. The supernatant from this centrifugation was decanted and centrifugation was repeated under the same conditions. The supernatant from this second spin was decanted and then passed through glass wool in a syringe to remove any flocculant material that may have been resuspended from the pellet during decanting. This clarified lysate supernatant was then subjected to glutathione affinity purification.

2.5.1.4 Affinity purification of GST-fusion proteins

Over-expressed GST-fusion proteins were purified from *E. coli* cell lysates by affinity chromatography using the ÄKTA FPLC and an HR 16/5 column (Amersham Biosciences) packed with 10 ml of Glutathione Sepharose™ 4B (Amersham Pharmacia Biotech). The column was stored in 20% ethanol. Prior to use, it was washed with three column volumes of filtered H_2O to remove the ethanol, and then equilibrated in three column volumes of 1 x PBS (Phosphate buffered saline: 140 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.3).

Clarified lysate supernatant (equivalent to 500 ml of culture) was applied to the column using the 50 ml Superloop™ (Amersham Pharmacia Biotech) at a flow rate of 1 ml/min. The column was then washed with three column volumes of 1 x PBS at 3 ml/min. The

bound GST-fusion protein was then eluted using glutathione elution buffer (50 mM Tris-HCl, 10 mM reduced glutathione, pH 8.0). The GST-fusion protein was detected as it was being eluted by the change in UV absorbance at 280 nm as measured by the UV monitor attached to the FPLC. The eluted fusion protein was collected in 1 ml fractions by the fraction collector.

The Glutathione Sepharose column was cleaned with the following cycle of washes: three column volumes of water, followed by two column volumes of 6M guanidine hydrochloride, five column volumes of water, three column volumes of 70% ethanol, and three column volumes of 20% ethanol for storage.

2.5.1.5 Buffer exchange by dialysis

The buffer change necessary to cleave the GST tag from eluted GST fusion proteins and for other chromatography steps were achieved by dialysis using BioDesign Dialysis Tubing™ with either a 3,500 or 8,000 Molecular Weight Cut Off (MWCO). An appropriate length of the tubing was cut, pre-soaked in the dialysis buffer for an hour, then one end of the tubing was tied and sealed with a clip. The protein solution was added and the tube closed by tying. The protein solution was then dialysed in 3L of the appropriate buffer, with slow stirring, for five hours at 4 °C. The buffer was then changed and the protein solution dialysed for a further 15 hours against the 3L of fresh buffer.

2.5.1.6 Removal of the GST tag from eluted GST-fusion proteins

Following dialysis of an eluted fusion protein into PreScission Cleavage Buffer (50 mM Tris-HCl (pH 7.0), 150 mM NaCl, 1 mM EDTA, 1 mM dithiothreitol (DTT)), the GST tag is proteolytically removed by transferring the fusion protein to a 50 ml tube, adding 40 µl (80 units) of PreScission Protease and DTT to a final concentration of 1 mM, and allowing the solution to mix gently at 4 °C on a Speci-Mix Aliquot Mixer (Barnstead International) for 15 hours.

The free GST was then removed from the protein of interest by re-applying the solution to the Glutathione Sepharose 4B column in the same manner as described previously. However, in this case the free protein of interest was expected to be in the column flow-through fractions and therefore these were collected, pooled and placed at 4 °C for further purification and analysis. The free GST binds to the glutathione sepharose, it was eluted with elution buffer and discarded. The column was then cleaned as described previously.

2.5.1.7 Hydrophobic Interaction Chromatography (HIC)

HIC was performed using a HiPrep 16/10 Phenyl FF (high sub) column, with 20 ml of bed volume. The column was pre-equilibrated in 50 mM Sodium phosphate, pH 8.0, 1 M $(NH_4)_2SO_4$, and the protein sample to be purified was dialysed into this buffer before HIC began (check). The sample was loaded onto the column using the superloop at a flow rate of 1 ml/min. After loading the protein was eluted using an ammonium sulfate gradient from 800 mM to 0 mM over five column volumes at a flow rate of 3 ml/min.

2.5.1.8 Ion Exchange Chromatography (IEC)

Anion exchange chromatography was performed using SOURCE™ 15 Q ion exchange media (Amersham Pharmacia Biotech) packed in an HR 16/5 column (Amersham Biosciences) with a bed volume of 7.5 ml. The column was pre-equilibrated in 50 mM Sodium phosphate, pH 8.5, 100 mM NaCl, and the protein to be purified was dialysed into this buffer before chromatography began. The sample was injected onto the column from a sample loop at a flow rate of 1 ml/min. The proteins were eluted using an NaCl gradient from 100 mM to 500 mM over five column volumes at a flow rate of 1 ml/min. At the end of the gradient the column was regenerated with 50 mM Sodium phosphate, pH 8.5, 1 M NaCl for three column volumes, before being washed with three column volumes of water followed by three column volumes of 20% ethanol for storage.

2.5.1.9 Size Exclusion Chromatography (SEC)

SEC was peformed on either a Superdex™ 75 10/300 GL (Amersham Biosciences) or a Superdex 200 10/300 (Amersham Biosciences) depending on the size of the protein or complex being examined. The SEC columns were equilibrated in 50 mM Sodium phosphate, pH 7.0, 150 mM NaCl. The sample was loaded using a sample loop, the sample was injected and eluted at a flow rate of 0.5 ml/min.

2.5.1.10 Concentration of proteins

Small volumes of protein were concentrated using regenerated cellulose Centricon® YM-10 (10 kDa molecular weight cut off) or YM-5 (5 kDa molecular weight cut off) centrifugal filter devices (Amicon Bioseparations, Millipore) for volumes up to 2 ml. Larger volumes of protein were concentrated using Amicon® Ultra-15 centrifugal filter devices with either a 10 kDa or 5 kDa molecular weight cut off (Millipore).

2.6 Protein Analysis

2.6.1 SDS-PAGE

Proteins were separated by Sodium-dodecyl-sulfate Polyacrylamide gel electrophoresis (SDS-PAGE) by the method of Laemli (95). Gels were cast with either a 12% or 15% separating gel, and a 4% stacking gel using the Hoefer SE245 Mighty Small Dual Gel Caster (Amersham Biosciences), 0.75 mm spacers and 10-well combs.

The capacity of the wells is around 15 µl, but in most cases the samples of 5 µl or less were loaded. Samples were mixed with 5 µl of sample loading buffer (62.5 mM Tris-HCl pH 6.8, 10% glycerol, 2% SDS (w/v), 5% β-mercaptoethanol, 1% (w/v) Bromophenol blue), boiled for 1 min, allowed to cool and then loaded on the gel.

Gels were run at 20 mA constant current for approximately 50 min or until the dye front ran off the bottom of the gel. The running buffer was 25 mM Tris-HCl, 200 mM glycine, 0.1% (w/v) SDS. To visualise the protein bands, the gels were stained with Coomassie blue

(0.1% Coomassie blue, 40% methanol, 10% acetic acid) for 45 min, the stain was then discarded and the gel destained with several changes of a solution of 40% methanol and 10% acetic acid over three hours at room temperature. During staining and destaining the gels were agitated gently on an IKA® VIBRAX VXR Basic Orbital Shaker.

Coomassie blue-stained gels were either photographed on an IS-1000 Digital Imaging System (Alpha Innotech) or dryed on a Bio-Rad Gel Dryer and scanned.

2.6.2 Determination of protein concentration

Protein concentration was estimated using a Beckman Coulter DU 640 Spectrophotometer. The spectrophotometer was blanked using the protein buffer. Then an aliquot of the protein of interest was diluted in a quartz cuvette (Bio-Rad, quartz spectrophotometer cell semi micro 9-Q-10 mm), and an absorbance spectrum measured between the wavelengths of 200 and 300 nm. The concentration of the protein was then calculated using the theoretical absorbance of 1 g/L of protein at 276 nm (ProtParam at the Expasy web site), and the absorbance of the diluted sample at 276 nm.

2.6.3 Far UV CD spectroscopy

Far UV CD spectroscopy was performed on a PiStar-180 Spectrometer (Applied Photophysics Limited, UK), and the data was collected on an attached Acorn computer. Prior to CD spectroscopy the cell was purged with Nitrogen at a flow rate of 8-10 L/min, and Nitrogen was pumped through the cell while CD spectra were collected at a flow rate of 5-6 L/min. Prior to gathering CD spectra the lamp intensity was also allowed to stabilise for 30 min.

The experimental parameters were as follows: elipticity was measured between the wavelengths of 180 and 260 nm, at 0.5 nm intervals over this range. The entry and exit slits were 4 nm wide. Proteins were in a 20 mM Bis-Tris, pH 6.5, 50 mM NaCl buffer. The CD spectrum of each sample was measured five times.

Prior to determination of the CD spectrum, the protein sample was degassed for 30 min in a jar under a vacuum. Water used to dilute protein samples was also degassed for 30 min in a Falcon tube attached to a vaccum in a sonicating water bath. Protein samples (30 µl each) were analysed in a quartz cuvette, with a 0.1 mm pathlength (Hellma, QS-106P).

CD spectra were smoothed using the Savitsky Golay algorithm and 10 reference points in the Origin 7 software package (OriginLab Corporation). The CDNN (23) software was used for the deconvolution of CD spectra.

2.6.4 Dynamic Light Scattering (DLS)

Dynamic light scattering was performed in a Dyna Pro MS-800 (Protein Solutions). Prior to analysis, protein samples were filtered using a 100 µl syringe (SGE Australia) and 13 mm Whatman Anodisc filters with a 0.02µm pore size. The samples (12 µl each) were analysed in the appropriate quartz cuvette. Data was collected and visualised using Dynamics V5.26.60 software.

2.6.5 Calibration of the size exclusion chromatography (SEC) columns

The Superdex 200 and Superdex 75 size exclusion columns were calibrated using High Molecular Weight and Low Molecular Weight Gel Calibration kits (Amersham Pharmacia Biotech). The high molecular weight kit contains Blue Dextran (elutes in the void volume), Thyroglobulin (669 kDa), Ferritin (440 kDa), Catalase (232 kDa), and Aldolase (158 kDa). The low molecular weight kit contains Albumin (67 kDa), Ovalbumin (43 kDa), Chymotrypsinogen A (25 kDa), and Ribonuclease A (13.7 kDa). These samples were eluted from the size exclusion columns and their elution volume was used to calculate a K_{av} for each protein, K_{av} = Ve-Vo/Vt-Vo. This was then used to generate a calibration curve by plotting the K_{av} vs the log molecular weight. To estimate the molecular weight of a protein of interest, the K_{av} was calculated based on the elution volume of the protein. Then by using the equation describing the calibration curve it was possible to the K_{av} calculate the unknown molecular weight of a protein of interest.

2.6.6 Limited Proteolysis

To perform a limited proteolysis experiment 100 ng of Trypsin (1 µl of the stock solution) was added to 100 µg of the protein of interest in an eppendorf tube. Following the addition of the protease the tube was placed at 4 °C and at timed intervals, 5, 10, 30, 60, 120, 240 min 10 µg of the protein was removed from the reaction. This aliquot was immediately mixed with an equal volume of SDS-PAGE loading buffer, boiled for 1 min and then frozen at -20 °C until the end of the time course. At the end of the time course the proteolysis samples were loaded on an SDS-PAGE gel with a molecular weight marker and undigested protein. Trypsin protease was stored at -80 °C, in 1 x PBS containing 10% glycerol, at a concentration of 0.1 mg/ml.

2.6.7 GST-pull-downs

In this assay the interaction between purified proteins (the prey) and a GST-fusion protein (the bait) bound to glutathione sepharose is tested. Glutathione Sepharose 4B was washed with 1 x PBS to remove the 20% ethanol and aliquoted into four eppendorf tubes. The assay consists of two 30 min room temperature incubations, a GST only control, a sepharose only control, and a third control where a GST-fusion protein is bound to sepharose but no prey protein is added.

60 µg of eluted GST-fusion protein was added to two of the tubes, 60 µg of GST was added another tube, no protein was added to the fourth tube. The volume in tubes was made up to 200 µl with 1 x PBS, and they were incubated with gentle rocking for 30 min at room temperature on a Speci-Mix Aliquot Mixer (Barnstead International), to allow binding of the protein to the sepharose. The sepharose was then pelleted and the supernatant removed. The sepharose in each tube was then washed three times by with 500 µl of 1 x PBS, pelleting the sepharose in a microfuge (Beckman Coulter Microfuge® 18 Centrifuge), and removing the supernatant.

To test the interaction with the fusion protein, a purified (untagged) protein was added to a tube containing Glutathione sepharose-GST-fusion. The same purified protein was also added to a tube containing only washed Glutathione Sepharose, as a control to test background binding to the sepharose, and to the Glutathione Sepharose with GST bound as a control to test for binding of the purified protein to GST. The tubes were made to an equal final volume of 200 µl with 1 x PBS, and incubated for 30 min at room temperature on the Aliquot Mixer. The sepharose was then washed again as above, the supernatant removed and 30 µl of SDS-PAGE loading buffer added to the sepharose. The sepharose loading buffer mixture was resuspended and 20 µl run on an SDS-PAGE gel.

2.6.8 Protein sequencing

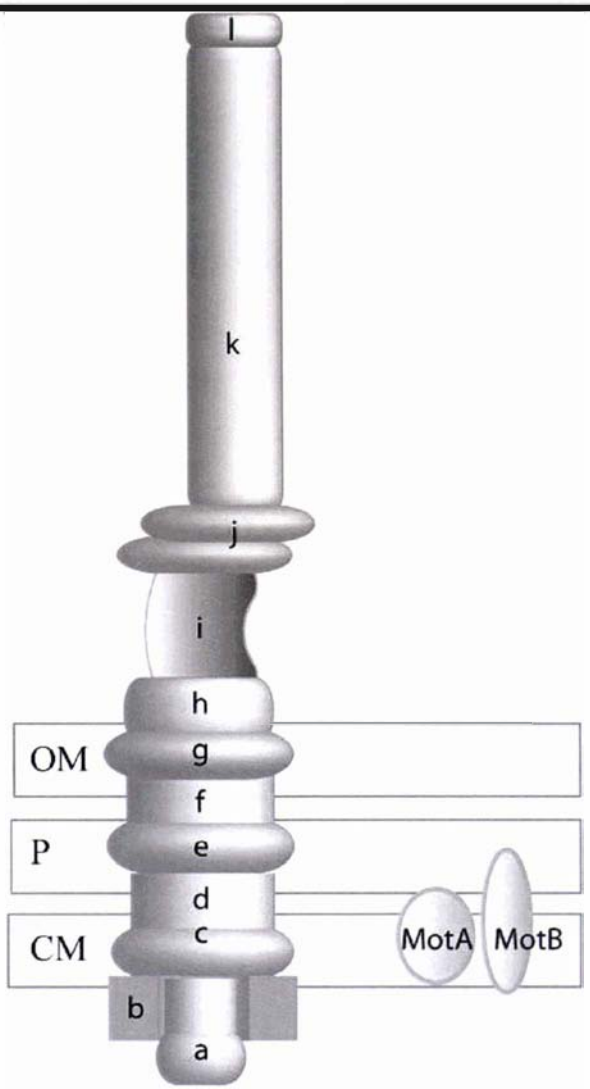
Following a GST pull-down experiment, an SDS-PAGE gel was run. The proteins were transferred from this SDS-PAGE gel to a PVDF membrane (Bio-Rad) using a Bio-Rad transblot apparatus. The transfer buffer was 10mM CAPS, pH 11.0, 10% methanol, and the proteins were transferred for 45 min at 250 mA constant current with an ice pack for cooling and constant stirring of the buffer. The membranes were then stained with Coomassie blue and destained. The dry membranes were then sent for N-terminal sequencing at the Protein Microchemistry Facility, Department of Biochemistry, Universtiy of Otago, Dunedin, New Zealand.

3. Results

3.1 Bioinformatic Analyses of the *H. pylori* Genome

The sequencing of the *H. pylori* genome in 1996 (182) has facilitated the identification of homologues of the majority of the *S. enterica* serovar *typhimurium* genes encoding flagellar proteins, on the basis of their sequence identity (Figure 3.1) (132). These include all of the proteins required for the assembly of the flagellum substructures- the motor/switch, basal body, the hook and the filament. The only variation upon the established flagellum structure of *S. typhimurium* revealed by sequencing and protein studies is the presence of three flagellins FlaA, FlaB and FlaG, and the lipid sheath which envelopes *H. pylori* flagellar the hook and filament (61) (175) (182).

Based on structural similarity to the *S. typhimurium* flagellum, the assembly of the *H. pylori* flagellum is assumed to proceed in the established linear sequential manner from the proximal membrane proteins to the distal extracellular proteins of the filament. Apart from FliF of the MS ring, the P-ring protein FlgI and the L-ring lipoprotein FlgH, the flagellar proteins are exported via the flagellar Type III export pathway facilitated by a somewhat loosely defined export apparatus. Both integral membrane proteins and soluble proteins are involved as part of an export apparatus complex, the stoichiometry and structure of which are still to be determined. There are six integral membrane proteins, FlhA, FlhB, FliO, FliP, FliQ and FliR. The soluble proteins include the specific chaperones FliS, FlgN and FliT, and the more general components FliI, FliH and FliJ which are involved in the export of all substrates.



	S. typhimurium protein	*H. pylori* homologue
l).	FliD	[HP0752]
k).	FliC	[HP0601, HP0115]
	Not annotated	[HP0327, HP0751]
j).	FlgL	[HP0295]
i).	FlgK	[HP1119]
h).	FlgE	[HP0870, HP0908]
g).	FlgG	[HP1092][HP1585]
f).	FlgH	[HP0325]
e).	FlgF	Not annotated
d).	FlgI	[HP0246]
	FlgB	[HP1559]
	FlgC	[HP1558]
	FliE	[HP1557]
c).	FliF	[HP0351]
b).	MotB	[HP0816]
	MotA	[HP0815]
	FliG,	[HP0352]
	FliM, FliN	[HP0584][HP1031]
a).	FliO	Not annotated
	FliP	[HP0685]
	FliQ	[HP1419]
	FliR	[HP0173]
	FlhA	[HP1041]
	FlhB	[HP0770]
	FliH	[HP0353]
	FliI	[HP1420]

Figure 3.1. Schematic diagram of the major structural components of the *H. pylori* flagellum compared to that of Salmonella. *Gene numbers from the TIGR sequence are used to indicate *H. pylori* gene homologues encoding flagellar proteins, except where previously published studies had already named flagellar genes. Relative positions of outer membrane (OM), periplasmic space (P) and cytoplasmic membrane (CM) are indicated.

S. typhimurium FliJ had been proposed as a flagellum specific chaperone based on the physical characteristics of size and predicted α-helicity which had been previously noted as common features of Type III export system substrate-specific chaperones (192). Subsequent experiments have demonstrated a role for *S. typhimurium* FliJ that is consistent with this proposition. FliJ interacts with, and is necessary for the export of rod/hook and filament type substrates in addition to interacting with export apparatus components FlhA, FliI, and FliH (119) (114) (59). Thus, the absence of an annotated FliJ gene in the *H. pylori* genome was striking in the context of flagellar protein export and motility.

A group of 19 FliJ homologues and their derived sequence characteristics, from a diverse set of organisms, are listed in Table 3.1. FliJ is usually encoded by a gene immediately downstream of *fliI*. Like other Type III substrate specific chaperones the FliJ homologues are small but in contrast to Type III chaperones most have a basic pI. It has also been implied previously that FliJ homologues contain a coiled-coil domain within their N-terminus (172) (114). Table 3.1 shows that only 11 of the homologues have a high probability of forming a coiled coil domain as predicted by the Multicoil algorithm (199) and some of the predicted coiled-coils are not N-terminal. Furthermore, in studies of *S. typhimurium* FliJ the N-terminal coiled-coil domain is not essential for the function of FliJ or its interactions with itself or FliH (58). Therefore the physical characteristics of FliJ homologues are of limited predictive value and a bioinformatics approach was taken to try and identify an *H. pylori* FliJ homologue.

Table 3.1 FliJ Homologues

Organism[a]	Size (kDa)	pI	Coiled-coil	5' gene	3' gene
S. typhimurium [1253494]	17.3	8.5	80% N-terminus	*fliI*	*fliK*
S. flexneri [1025160]	13.7	9.9	-	*fliI*	*fliK*
Y. pestis [1174662]	17.5	6.5	-	*fliI*	*fliK*
E. carotovora [2882111]	17.5	10	60% N-terminus	*fliI*	*fliK*
B. pertussis [2665502]	17.1	8.5	80% N-terminus	*fliK*	*fliI*
P. aeruginosa [879381]	17.4	9.7	80% N-term., >70% C-terminus	*fliI*	hypothetical
V. cholerae [2613385]	17.9	8.7	80% N-terminus	*fliI*	*fliK*
P. profundum [3125099]	17.6	7.6	-	*fliI*	hypothetical
S. oneidensis [1170916]	17.6	10.1	-	*fliI*	*fliK*
N. europaea [1083048]	17.4	9.4	-	*fliI*	*fliK*
L. pneumophila	17.9	10.1	-	*fliI*	conserved hypothetical
X. campestris [998649]	17.3	9.6	80% N-terminus	*fliI*	*fliK*
C. acetobutylicum [1118341]	17.2	9.6	-	*fliI*	*fliK*
T. denticola [2740113]	18.3	9.1	>70% Central	*fliI*	conserved hypothetical
T. tengcongensis [997886]	17.6	8.6	100% N-term., 80% C-terminus	*folC*	hypothetical
G. sulfurreducens [2686520]	17.4	9.4	70% C-terminus	*fliI*	hypothetical
L. interrogans [1151936]	22.3	9.1	-	*fliI*	hypothetical
B. subtilis [2633997]	17.9	8.7	-	*fliI*	*ylxF*
C. crescentus [942501]	15.7	5.9	80% N-terminus	*fliI*	*xynA*
HP0256 [899183]	16.7	6.2	>80% C-terminus	*purA*	conserved hypothetical secreted protein

[a]The NCBI-GeneID numbers for each *fliJ* homologue are indicated in square brackets.

H. pylori fliJ has not been annotated. The lack of clustering of the flagellar genes in *H. pylori* complicates gene annotation efforts and furthermore, the FliJ amino acid sequence is generally poorly conserved as can be seen in the following multiple alignment and table of percentage identity (Figure 3.2 and Table 3.2). The web-based BLAST search engine at NCBI was used to search for a FliJ homologue. While PSI-BLAST is designed to identify protein sequences that are distantly related to the query sequence, it begins with the same BLAST algorithm that is used by BLAST-P to search databases for significant similarities. Consequently, it was decided to explore the effects of search parameters using BLAST-P and the *S. typhimurium* FliJ as a query sequence to identify the set of annotated FliJs in the multiple alignment, before using PSI-BLAST. There are two main parameters in BLAST searches, the search matrix and the Expect value. The first BLAST-P search utilised the default parameters, Expect 10 and BLOSUM 62, and the user-defined parameters of a non-redundant database, low complexity filter, and bacterial genomes only. This search resulted in a total of 204 hits but only identified nine of the FliJ sequences above the threshold and a further three below the threshold of the nineteen that were used in the multiple alignment. The search was repeated using the BLOSUM 45 search matrix as this is designed to give higher E-values to more distantly related sequences, but this only changed the E-values of the FliJs slightly and did not identify any more FliJ homologues. The search was repeated with Expect values of 1, 100 and 1000 and this resulted in 93, 508 and 534 hits respectively. While the expect value changed the number of hits it did not identify any further FliJ homologues either above or below the significance threshold. Subsequently, the non-redundant database was searched for FliJ homologues using PSI-BLAST.

Table 3.2 Percentage amino acid sequence identity of FliJ homologues in the multiple alignment, Figure 3.2

Query/Hit	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1. *S. typhimurium*	-	68	54	33	29	18	14	14	16	56	25	23	10	17	21	14	15	28	21	16
2. *S. flexneri*	68	-	45	25	24	15	10	13	16	17	21	20	23	19	9	14	19	11	21	13
3. *Y. pestis*	54	45	-	36	32	19	14	16	15	65	27	19	24	18	10	14	17	22	26	14
4. *B. pertussis*	33	25	36	-	30	22	16	14	11	39	28	25	25	17	14	16	15	31	26	15
5. *P. aeruginosa*	29	24	32	30	-	22	10	14	13	31	27	28	29	16	14	16	17	25	23	14
6. *V. cholerae*	18	15	19	22	22	-	13	12	9	24	31	23	21	17	15	15	17	24	60	18
7. *B. subtilis*	14	10	14	16	10	13	-	15	12	18	18	15	16	20	18	15	19	17	19	18
8. *H. pylori*	14	13	16	14	14	12	15	-	16	16	9	13	16	16	15	14	12	13	13	11
9. *C. crescentus*	16	16	15	11	13	9	12	16	-	16	30	17	14	12	12	14	14	12	24	10
10. *E. carotovora*	56	17	65	39	31	24	18	19	16	-	30	19	25	19	16	12	16	26	24	14
11. *S. oneidensis*	25	21	27	28	27	31	18	18	9	30	-	17	22	20	17	18	18	29	38	17
12. *X. campestris*	23	20	19	25	28	23	15	13	17	19	17	-	20	17	15	18	18	23	15	15
13. *L. pneumophila*	10	23	24	25	29	21	16	16	14	25	22	20	-	16	13	16	13	22	24	19
14. *C. acetobutylicum*	17	19	18	17	16	17	20	16	12	19	20	17	16	-	24	25	22	17	16	26
15. *L. interrogans*	21	9	10	14	14	15	18	15	12	16	17	15	13	24	-	25	25	14	13	21
16. *T. tengcongensis*	14	14	14	16	16	15	15	14	14	12	18	18	16	25	25	-	26	16	17	21
17. *G. sulfurreducens*	15	19	17	15	17	17	19	12	14	16	18	18	13	22	25	26	-	13	19	18
18. *N. europaea*	28	11	22	31	25	24	17	13	12	26	29	23	22	17	14	16	13	-	23	23
19. *P. profundum*	21	21	26	26	23	60	19	12	13	24	38	15	24	16	13	17	19	23	-	22
20. *T. denticola*	16	13	14	15	14	18	18	11	10	14	17	15	19	26	21	21	18	23	22	-

Salmonella enterica serovar *Typhimurium* LT2	[NT01ST2463]	---
Shigella flexneri 2a str. 301	[NT01SF2408]	---
Yersinia pestis CO92	[NT01YP2106]	---
Bordetella pertussis Tohama I	[NT03BP1444]	---
Pseudomonas aeruginosa PAO1	[NT03PA1215]	---
Vibrio cholerae O1 biovar eltor str. N16961	[VC2129]	MRQ
Bacillis subtilis subsp. *subtilis* str. 168	[NT01BS2069]	---
Helicobacter pylori 26695 AAD07334	[HP0256]	---
Caulobacter crescentus CB15	[CC3041]	---
Erwinia carotovora subsp. *atroseptica*	[NT06EC1784]	---
Photobacterium profundum SS9	[NT01PP0972]	---
Shewanella oneidensis MR-1	[SO3224]	---
Nitrosomonas europaea ATCC 19718	[NT01NE2308]	---
Legionella pneumophila subsp. *pneumophila* str. Philadelphia 1	[NT04LP1847]	---
Xanthomonas campestris pv. *campestris* str. ATCC 33913	[NT01XC2620]	---
Clostridium acetobutylicum ATCC 824	[NT01CA2380]	---
Treponema denticola ATCC 35405	[TDE1219]	---
Thermo tengcongensis MB4(T)	[NT01TT0852]	---
Geobacter sulfurreducens PCA	[GSU0414]	---
Leptospira interrogans serovar Copenhageni Fiocruz L1-130	[NT03LI1681]	---

```
                                                679999876889**999999**9998*9*9*8**999989999888
S. typhimurium      ----------MAQHGALETLKDLAEKEVDDAARLLGEMRRGCQQAEEQLKMLIDYQNEYRS 51
S. flexneri         ------------------------------------------MAEEQLKMLIDYQNEYRN 18
Y. pestis           ----------MKSQSPLVTLCDLAQKAVEQASTQLGHVRQSYQNAEQQLTMLLTYQDEYRE 51
B. pertussis        ----------MPSQLPLDMLIGLAKDSTDEAARELGRLSAERNNAEQQLNMLQDYRQDYLQ 51
P. aeruginosa       ----------MEKRAARLAPVVDMASKAERDAATQLGRCQQQLLAAQQKLAELERYRNDYQQ 52
V. cholerae         HAAPCTQLITWSIEMDNALDFLLEQAQESEDKAVLALSKARSELDSYYHQLRQIEQYRLEYCQ 66
B. subtilis         ----------MAYQFRFQKLLELKENEKDQSLSEYQQSVSEFENVAEKLYENMSKKELLEQ 51
H. pylori           ----------MKKFASVLVQLKTLALEKIEQKLESKRLELQQNEREVLDKQAQLSAFKN 49
C. crescentus       ----------MTKWAASLIRISNHEVETLQKRLAEITERRMAAEMRVTLLDAEAEAEAK 49
E. carotovora       ----------MKTQSPLVTLRELAQKDVEKAAGQLGQVRQAHQQAEQQLNMLLNYQDDYRQ 51
P. profundum        ----------MSNNALNLILEHAKEEEHQASLALNQARIERQNYLQQLQQIEQYRLDYCK 50
S. oneidensis       ----------MANTDPLLLVLKLALDAEEQAALLLKSAQLECQKRQNQLDALNNYRLDYMK 51
N. europaea         ----------MATPHSLKLLLDHARKQTDDAAINLGKLNLKQQFAEKTLQLLVEYRENYQS 51
L. pneumophila      ----------MSDRLDRLIQLLKIKQEATQQAYIELVKAREQFNQNKARHEQLVGYRQDYLQ 52
X. campestris       ----------MMQSKRIDPLLRRAQEQEDKVARDLAERQRALDTHQSRLEELRRYAEEYAS 51
C. acetobutylicum   ----------MAGFGFRLQKLLDLRIQKEEESKIEFKKSQDVKKEIEKNLLSMTENYHEYSL 52
T. denticola        ----------MKRFEFRLEKLLNLREFYEHQAEIDLAHAIAHKDYIDLQLKQIAKLKVKTGT 52
T. tengcongensis    ----------MKKFEFTLQSVLNLKEQSEKIEKENLAKIMKEIERE-REKLENLKKHLQEVTK 52
G. sulfurreducens   ----------MTHKHGFQLQQVLNYRKEVEKVKKLEFATARHEFEHATDVLSRHEAEADRARV 53
L. interrogans      ----------MKRFQFSLEPVLNLRKKKEDEKLKAFSKVAGEINQIRNSILENEKQIEHLTG 52
```

```
                          8865    5    58888888888999*9**999999*99****9989***999777888888
S. typhimurium            NLNTDMGNGIASNRWINYQQFIQTLEKAIEQHRLQLTQWTQKVDLALKSWREKKQRLQAWQTL 114
S. flexneri               NLNSDMSAGMTSNRWINYQQFIQTLEKAITQHRQQLNQWTQKVDIALNSWREKKQRLQAWQTL 81
Y. pestis                 RLNDTLCNGMASSSWQNYQQFIQTLEQAIDQHRKQLAQWSIKVEQAVKYWQEKQQRLNAFETL 114
B. pertussis              RMQTAMQSGMSAADCHNYQRFIATLDDAIGQQRHVLHRAEAHLNDGRLNWQQQKRKLNSFDTL 114
P. aeruginosa             QWISQGQKGVSGQWLMNYQRFLSQLETAVAQQANSVTWHREAVDKARLNWQERYARLEGLRKL 115
V. cholerae               QLVERGKSGLTASQYGHLNRFLTQLDETLAKQKSAEQHFRLQVENCEQHWMNMRQKRKSYQWL 129
B. subtilis               NKEKKLKSGMSVQEMRHYQQFVSNLDNTIYHYQKLVIMKRNQMNQKQEILTEKNIEVKKFEKM 114
H. pylori                 PELG------GMSLFLQTQQLKSALRMEIEYYQQESENLNKDLKILEKDYLLANQELEKAKII 106
C. crescentus             NAEGDPSAGWYMIGYRE------GSKRRRADMLVQIEQCQQEEAGARDALSEAFENLKKYEHV 106
E. carotovora             KLNSTMSSGMANNSWQNYQQFIRTLDGAIEQHRQQLSQWTSRLDLAMKTWQEKQQRLNAFEKL 114
P. profundum              QLSTRGQEGLTASSYGHLQKFLTQLDETLAKQKEAGHQFEFQIEQCSEHWNEVRKKRRSIEWL 113
S. oneidensis             QMQSQQGQAISASHYHQFHRFIRQIDEAIAQQQRVVADGEKQKNYRQHYWLEKQKKRKAVELL 114
N. europaea               QFMESAGSGISPVEWRNFKAFICKLDTAIQSQQRLVTMTQQHTEAGSTQYHAHRQKLKSYDTL 114
L. pneumophila            QLEVLGQQGSYVGPLRNRINFINHLDTALVQLNSYLSQLAKNRMKADLNYKQAKTSEEGISKL 115
X. campestris             SQMS----GTSAVALSNRRAFLDRLDSAVLQQAQTVESNRAKVEAERTRLLLASREKQVLEQL 110
C. acetobutylicum         KRLSG-----TVIEQKITQNYLNALNVLIDEASDNLEKQKKVVEDKRNVLIKKQVERKTVEVL 110
T. denticola              EFNPET-DKIDITDLHNAQNYIILLDKKKDELLEKLVLAEQIIEEKRKIYIEAASKRKVISKL 114
T. tengcongensis          RAKEEVEEGTLMYKLAETEAYIMKIREMIEKQANYILKLEKEAEKIREGLLKVSKEKKALENL 115
G. sulfurreducens         EYNNKQAAGTTANELQLYADFFARKHMDIQFQRIEVDNLNRKMSEKREDLMDAAKDKKALELL 116
L. interrogans            --ESHTLHGASLRDYQLHQGYIRSLITKNENLESDIENRKSELDSKRADLILAQKDRKILEIL 113
```

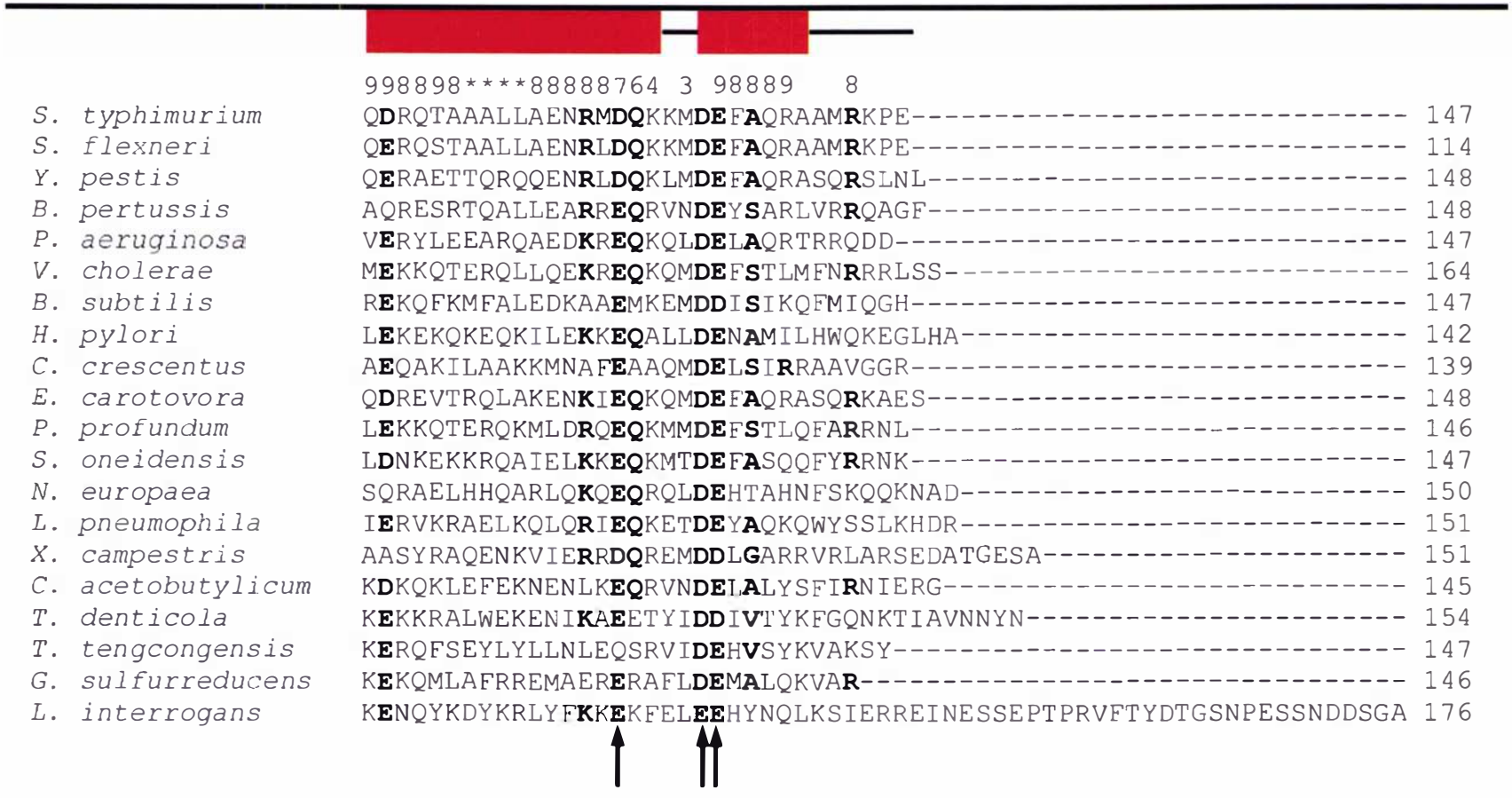


Figure 3.2 CLUSTAL W (1.82) multiple sequence alignment of FliJ homologues

The gene numbers from the TIGR database are listed beside sequence names in square brackets. The secondary structure of FliJ as predicted using the PHD algorithm is indicated above the sequences: the red box represents helix, the black line loop. The probability of the indicated secondary structure at each position is indicated by the number, 9=90%, *=100%.

Following this initial BLAST search, PSI-BLAST generates a Position Specific Scoring Matrix (PSSM) based on a multiple alignment from the first database search. This PSSM can then be iteratively improved by adding new sequences below the significance threshold and removing false positives at or above the threshold. Each of the iterations generates a new multiple alignment and PSSM, which is in turn used for a new database search (6). In the first round of PSI-BLAST searches, each of the 19 aligned and annotated FliJ homologues was used as the query sequence with an Expect value of 10 and the BLOSUM 45 matrix. The results are summarised in Table 3.3 Under these conditions 18 of the 19 FliJ homologues were identified in 4 iterations or less. *C. crescentus* FliJ was not identified by any of the query sequences under these conditions and it was also a poor query sequence as it only identified itself.

The PSI-BLAST searches were then repeated with the same query sequences, but with an Expect value of 100. The results are summarised in Table 3.4 Again 18 of the 19 homologues were consistently identified above the significance threshold in 4 iterations or less, but with an Expect value of 100 the *S. flexneri* and *X. campestris* FliJ query sequences identified *C. crescentus* FliJ above the significance threshold. Furthermore, when *T. denticola*, *L. interrogans*, or *T. tengcongensis* FliJs were used as the query sequence an *H. pylori* protein, HP0256, was identified above the significance threshold.

Table 3.3 Iterations required to identify FliJ homologues above the significance threshold using PSI BLAST searches.
(Expect 10, BLOSUM 45, threshold of 0.005).

Query/Hit	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
1. *S. typhimurium*	1	1	1	1	1	2	2	-	-	1	2	2	2	2	3	3	3	1	2	3
2. *S. flexneri*	1	1	1	1	3	3	3	-	-	1	3	2	3	3	-	4	4	3	3	4
3. *Y. pestis*	1	1	1	1	2	2	3	-	-	1	2	2	2	2	3	3	3	2	2	3
4. *B. pertussis*	1	1	1	1	1	1	2	-	-	1	1	2	2	2	3	3	3	1	1	3
5. *P. aeruginosa*	1	-	-	1	1	-	2	-	-	-	-	-	-	3	3	3	3	-	-	3
6. *V. cholerae*	2	2	2	1	2	1	2	-	-	2	1	2	2	2	3	3	3	2	1	3
7. *B. subtilis*	3	3	3	2	3	3	1	-	-	3	3	3	3	3	3	3	3	3	3	3
8. *H. pylori* 0256	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-
9. *C. crescentus*	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-
10. *E. carotovora*	1	1	1	1	2	2	2	-	-	1	2	2	2	3	3	3	3	2	2	3
11. *S. oneidensis*	2	2	2	1	2	1	2	-	-	2	1	2	2	2	3	3	3	2	1	3
12. *X. campestris*	3	3	3	2	3	3	3	-	-	3	3	1	3	4	4	4	4	3	3	4
13. *L. pneumophila*	2	2	2	2	2	2	3	-	-	2	2	2	1	3	4	3	3	2	2	3
14. *C. acetobutylicum*	3	4	3	3	3	3	4	-	-	3	3	3	4	1	3	3	3	3	3	2
15. *L. interrogans*	4	4	4	3	3	4	3	-	-	4	4	4	4	3	1	3	2	4	4	3
16. *T. tengcongensis*	3	3	3	3	3	3	3	-	-	3	3	3	3	3	2	1	2	3	3	3
17. *G. sulfurreducens*	4	4	4	3	4	3	4	-	-	3	3	4	4	3	3	3	1	4	3	3
18. *N. europaea*	1	2	2	1	2	2	2	-	-	2	2	2	2	3	3	3	3	1	2	3
19. *P. profundum*	2	2	2	2	2	1	2	-	-	2	1	2	2	2	4	3	3	2	1	3
20. *T. denticola*	3	4	3	3	3	3	3	-	-	3	3	3	3	2	3	3	3	3	3	1

Table 3.4 Iterations required to identify FliJ homologues above the significance threshold using PSI BLAST searches (Expect 100, BLOSUM 45, threshold of 0.005).

Query/Hit	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1. *S. typhimurium*	1	1	1	1	1	2	2	-	-	1	2	2	2	2	3	3	3	1	2	4
2. *S. flexneri*	1	1	1	1	2	2	3	-	3	1	2	2	2	3	4	4	4	2	2	4
3. *Y. pestis*	1	1	1	1	2	2	2	-	-	1	2	2	2	2	3	3	3	2	2	3
4. *B. pertussis*	1	1	1	1	1	1	2	-	-	1	1	2	2	2	3	2	3	1	1	3
5. *P. aeruginosa*	1	2	2	1	1	2	2	-	-	2	2	2	2	2	3	2	3	2	2	3
6. *V. cholerae*	2	2	2	1	2	1	2	-	-	2	1	2	2	2	3	2	2	2	1	3
7. *B. subtilis*	2	2	2	2	2	2	1	-	-	2	2	2	2	2	2	2	2	2	2	2
8. *H. pylori* 0256	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-
9. *C. crescentus*	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-
10. *E. carotovora*	1	1	1	1	1	2	2	-	-	1	2	2	2	2	3	2	3	1	2	3
11. *S. oneidensis*	2	2	2	1	2	1	2	-	-	2	1	2	2	2	3	2	3	2	1	3
12. *X. campestris*	2	2	2	2	2	2	3	-	3	2	2	1	2	3	3	3	3	2	2	3
13. *L. pneumophila*	2	2	2	2	2	2	2	-	-	2	2	2	1	3	3	3	3	2	2	3
14. *C. acetobutylicum*	3	3	3	3	3	3	4	-	-	3	3	3	4	1	3	3	3	3	3	3
15. *L. interrogans*	3	3	3	3	3	3	3	4	-	3	3	3	3	2	1	2	2	3	3	2
16. *T. tengcongensis*	3	3	3	3	3	3	3	4	-	3	3	3	3	3	2	1	2	3	3	3
17. *G. sulfurreducens*	3	3	3	3	3	3	3	-	-	3	3	3	3	2	2	2	1	3	3	3
18. *N. europaea*	1	2	2	1	2	2	2	-	-	2	2	2	2	2	3	2	3	1	2	3
19. *P. profundum*	2	2	2	2	2	1	2	-	-	2	1	2	2	2	3	3	3	2	1	3
20. *T. denticola*	3	3	3	3	3	3	3	6	-	3	3	3	3	2	2	2	2	3	3	1

An alternative approach to searching for a potential *H. pylori* FliJ homologue is to use a pattern based search such as PHI-BLAST within BLAST-P or PSI-BLAST. The database is searched for sequences containing the PHI pattern and then BLAST searches any hits for further significant sequence identity around the pattern. Although there is limited sequence conservation to be observed in the multiple alignment of FliJ homologues presented here (Figure 3.2), there are three residues near the FliJ C-terminus that are highly conserved. Within the *S. typhimurium* FliJ sequence these residues are D130, D135 and E136. At the equivalent positions in all of the other FliJ sequences in the multiple alignment the amino acids are either D or E. This conservation can be converted to the following pattern for use with PHI-BLAST: [ED]x(4)[ED][ED], where any four amino acids separate an E or a D from two consecutive residues that are either E or D. This pattern was used with BLAST-P or PSI-BLAST and either *S. typhimurium* FliJ or *T. denticola* FliJ as the query sequence, the results are summarised in Table 3.5.

PHI-BLAST did not improve the effectiveness of either PSI or BLAST-P searches for identifying an *H. pylori* homologue of FliJ. But it did identify JHP0240 and an unknown *H. pylori* protein below the significance threshold. These proteins are both homologues of *H. pylori* 26695 HP0256 in strains J99 and 17874 respectively. An alternative to the PHI-BLAST pattern search is the use of a saved PSSM from a PSI-BLAST search. To test this approach the position specific scoring matrices from iteration 4 of an *S. typhimurium* PSI-BLAST search and a PSSM from a *T. denticola* PSI-BLAST search were used to search the *H. pylori* database. Using the *S. typhimurium* PSSM, HP0256, JHP0240 and the 17874 unknown were identified with highly significant scores of $1x10^{-8}$, $1x10^{-7}$, and $1x10^{-7}$ respectively. A similar search with a *T. denticola* PSSM failed to identify any hits of significance. When HP0256 was previously identified above the significance threshold it had a score between 0.001 and 0.005, barely above the threshold.

Table 3.5 BLAST searches incorporating sequence patterns

Search algorithm	Database	Expect value	Query sequence	Significant hits	E-value
PSI/PHI-BLAST	NR-Bacteria	100	*S. typhimurium*	*S. typhimurium*	$4x10^{-70}$
				S. flexneri	$4x10^{-45}$
				E. carotovora	$9x10^{-39}$
				Y. pestis	$1x10^{-34}$
				B. pertussis	$3x10^{-16}$
				P. aeruginosa	$1x10^{-12}$
				N. europaea	$4x10^{-12}$
				S. oneidensis	$2x10^{-9}$
				V. cholerae	$3x10^{-6}$
				P. profundum	$6x10^{-21}$ (2^{nd} it)*
				L. pneumophila	$1x10^{-15}$ (2^{nd} it)
				X. campestris	$5x10^{-10}$ (2^{nd} it)

Table 3.5 continued

				C. acetobutylicum	$5x10^{-7}$ (2nd it)
				T. tengcongensis	$1x10^{-5}$ (2nd it)
				B. subtilis	$2x10^{-5}$ (2nd it)
				T. denticola	$9x10^{-14}$ (3rd it)
				L. interrogans	$1x10^{-13}$ (3rd it)
				G. sulfurreducens	$3x10^{-13}$ (3rd it)
	NR-Bacteria	100	*T. denticola*	No significant sequence similarity	-
BLAST-P/PHI	*H. pylori*	100	*S. typhimurium*	JHP240	0.38
	H. pylori	100	*T. denticola*	No significant sequence similarity	-

Table 3.5 continued

BLAST-P/PSSM	*H. pylori*	100	*S. typhimurium* PSSM, iteration 4	HP0256 unannotated 17874 protein JHP0240	$1x10^{-8}$ $1x10^{-7}$ $1x10^{-7}$
	H. pylori	100	*T. denticola* PSSM, iteration 7.	36 hits below the significance threshold, no HP0256	-

*it is used in this table as an abbreviation of the word iteration

As HP0256 was the only *H. pylori* sequence that occurs above the significance threshold it was decided to investigate it further. HP0256 encodes a 16.7 kDa protein with a pI of 7.8 that was found to be 15% identical to *B. subtilis* FliJ. It is also predicted to be α-helical (150) and has a high probability of forming a C-terminal coiled-coil between amino acid residues 100 and 130 (199). Further investigation of *B. subtilis* FliJ reveals the characteristic flagellar gene organisation; FliJ is clustered with other flagellar genes, and the gene immediately upstream of *fliJ* is *fliI*. Although *H. pylori* flagellar genes are not clustered, when the genes surrounding HP0256 are examined it is found that HP0257 shares 24% identity with *frzCD*, the gene immediately downstream of *B. subtilis fliJ* (Figure 3.3). FrzCD is an atypical methyl-accepting chemotaxis protein that is involved in vegetative swarming and fruiting body formation in *M. xanthus* (28).

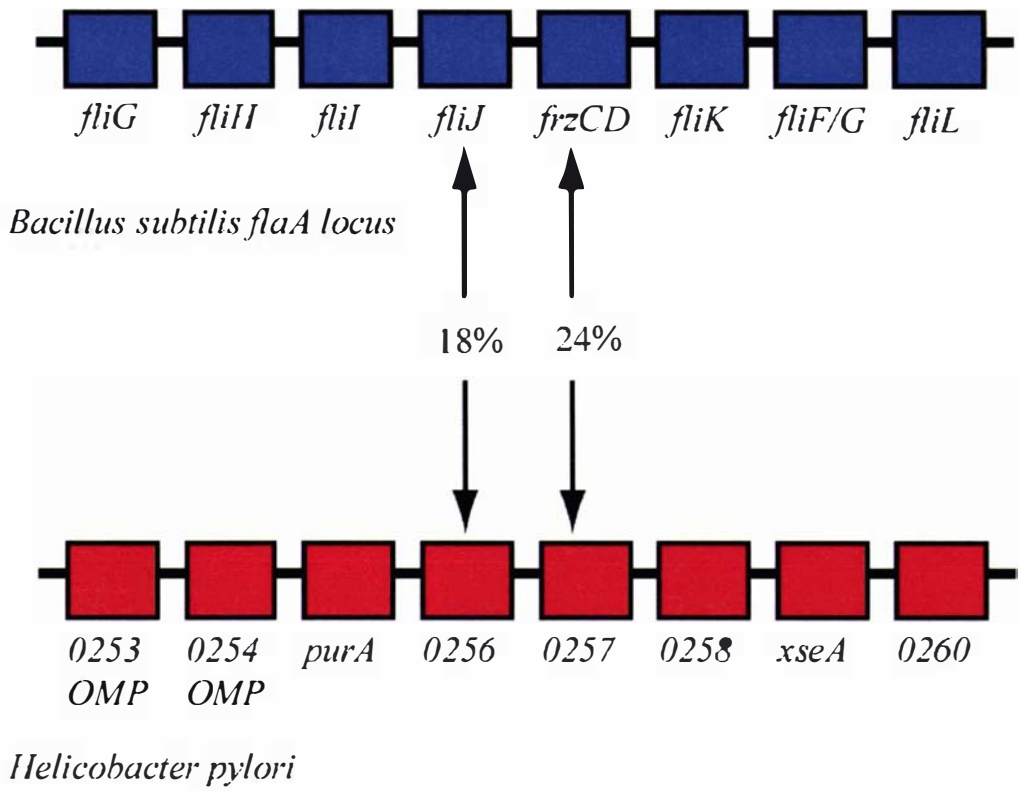


Figure 3.3 Synteny of flagellar genes in *B. subtilis* and *H. pylori*.

To investigate the role of HP0256 in flagellar assembly and motility, a deletion strain has been created by allelic exchange. Briefly, HP0256 was cloned into pUC19, internally deleted by PCR, and a chloramphenicol resistance cassette was cloned into the centre of the gene (Figure 3.4). Following transformation of *H. pylori* with the resulting plasmid, integration of the deleted gene and replacement of the wild type gene was confirmed by PCR. Examination of the mutant cells using a light microscope revealed they were still motile, and electron microscopy revealed that the mutant bacteria are flagellated (Figure 3.5).

If HP0256 had a significant role in flagellar export, the HP0256 knock-out mutants would be expected to be non-motile and aflagellate, However, this was not the case. The observed results do not exclude the possibility that HP0256 is an *H. pylori* FliJ homologue, as experiments in *S. typhimurium* have shown that all FliJ mutations are "leaky" and flagellar export still occurs (114). Therefore it is possible that flagellar protein export in the HP0256 mutant is reduced but still occurs to a significant degree such that flagellation and motility can still be observed but to a reduced degree. To investigate the possibility of such aberrant motility, Dr P.W. O'Toole and colleagues recently subjected the HP0256 knock-out mutant created in this study to more detailed analysis using BacTracker, a system allowing direct microscopic observation and characterisation of the degree and type of motility. This investigation revealed that the HP0256 mutant of *H. pylori* does exhibit an aberrant form of motility, providing further evidence that HP0256 could indeed be a FliJ homologue, or at the very least is involved in motility.

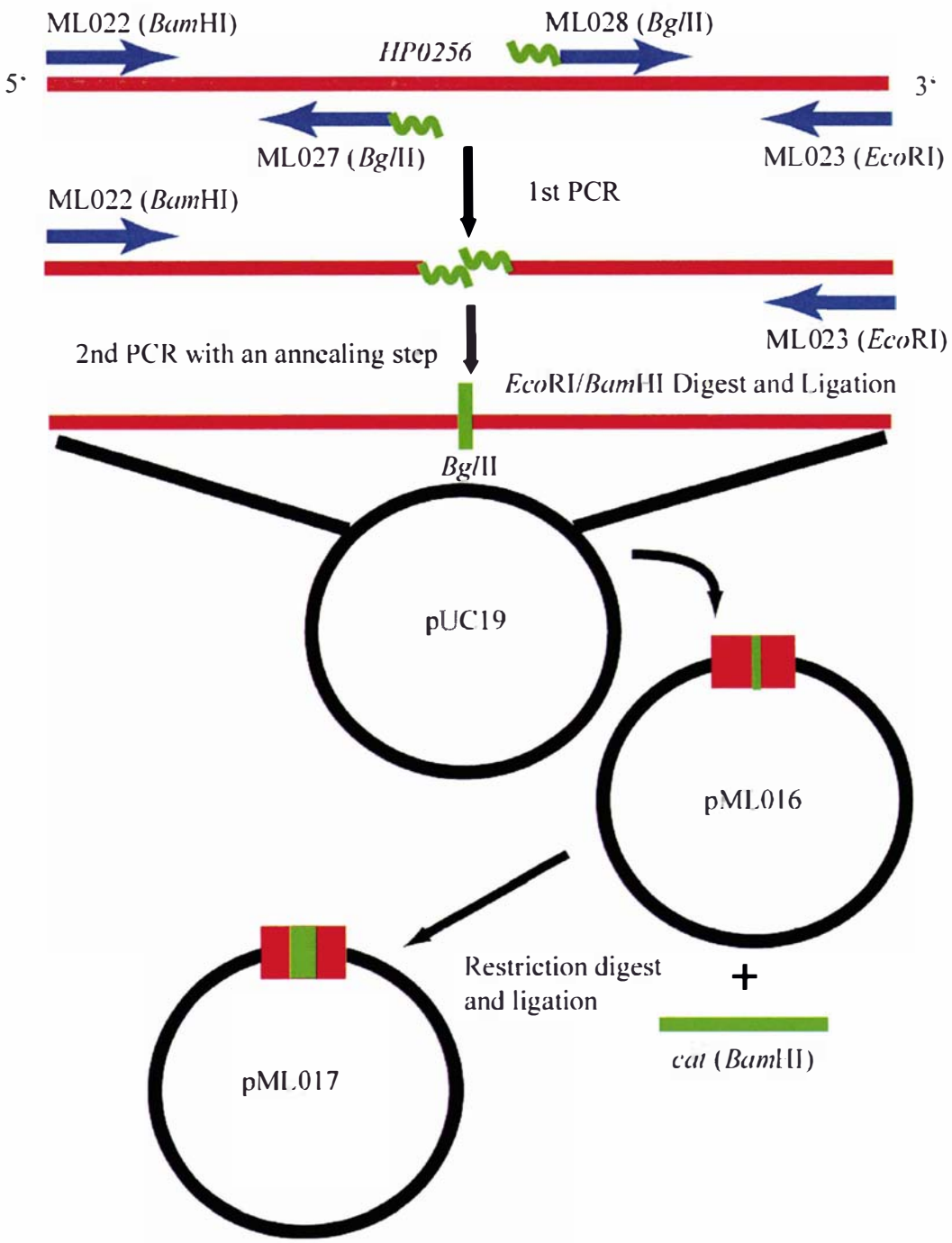


Figure 3.4 Internal deletion and cloning of *HP0256* from *H. pylori* 26695.

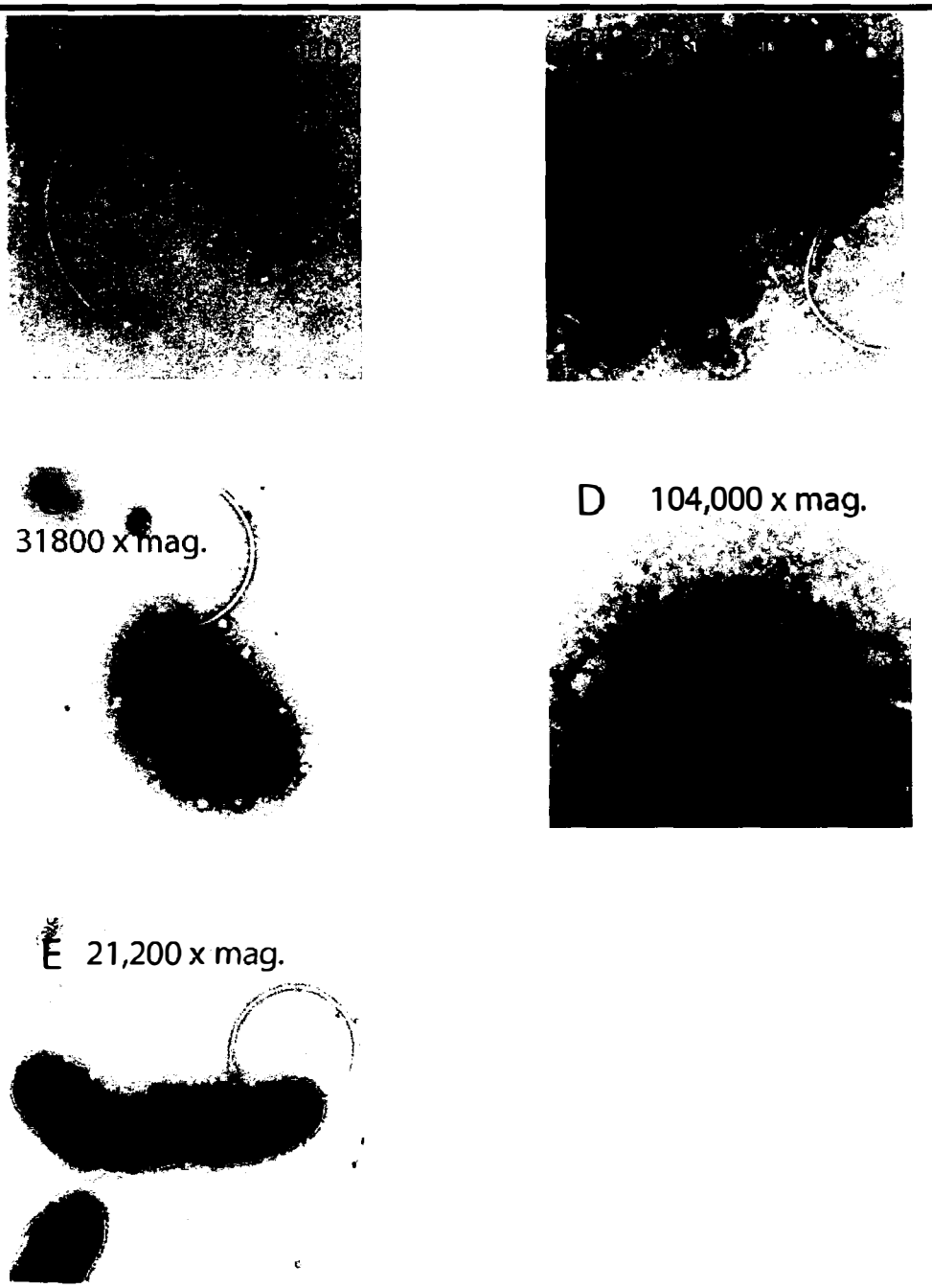


Figure 3.5 Electron micrographs of the *HP0256* deletion mutant and wild type *H. pylori* strain 17874. Panel A, *HP0256* flagellum filament; B, *HP0256* mutant with attached flagella; **C and D,** *HP0256* mutant with the flagellum attached and the L-ring and hook structure visible; **E,** wild type *H. pylori* 17874 with the flagellum attached.

3.2 Characterisation of *H. pylori* FliI

As described in detail in the introduction (Sections 1.5 and 1.6), FliI of *H. pylori* is a cytoplasmic membrane-associated protein required for flagellar assembly (142) (77).

Given the predicted tertiary structure of FliI (Figure 1.7 and 1.8) and our understanding of the function of the homologues Rho transcriptional terminator and F_1-ATPase, it is possible to suggest a hypothesis for the structure and function of *H. pylori* FliI. This hypothesis has three tenets. Firstly FliI will oligomerise to form a hexamer analogous to the pseudo-hexamers of Rho and F_1-ATPase, and secondly hexameric FliI will demonstrate cooperative ATPase activity. Third, FliI will couple ATP hydrolysis to protein export in a similar manner to the Rho transcriptional terminator, which couples ATP hydrolysis to translocation of RNA.

To examine this set of tenets, structure-function analysis of FliI was undertaken using genetic and biochemical approaches.

3.2.1 The role of the FliI R-loop in *H. pylori* flagellar assembly and function

The R-loop in Rho and the F_1-ATPase has a role in the coupling of ATPase activity to the movement of RNA through the centre of the pseudo-hexamer and the rotation of the γ-subunit, respectively. To test whether these conserved residues (E_{311}, G_{312}, and D_{313}) are also required for the function of *H. pylori* FliI , they were each mutated to Ala, and the mutant proteins were examined for their ability to complement a *fliI* knock-out mutant strain of *H. pylori* for flagellation.

The mutations were first created by site-specific mutagenesis, then cloned into *E. coli-H. pylori* shuttle vector pHel3, introduced into the *fliI* knock-out mutant (142) by transformation and finally the flagellation of the transformed strains was examined by electron microscopy.

The R-loop of *H. pylori* FliI (residues 308-318, see alignment in Figure 1.7) was mutated using a PCR-based method of directed mutagenesis (Figure 3.6). For this method of

mutagenesis/cloning, four consecutive rounds of PCR were used. The objective of this exercise was to generate and clone three mutant *fliI* genes, with mutations of E_{311}-A, G_{312}-A, and D_{313}-A. In addition the PCR was designed to fuse the *H. pylori recA* promoter to the mutant genes to allow expression *in vivo*.

The first round of PCR involved the amplification of the *recA* promoter from *H. pylori* 26695 genomic DNA (see the sequence for this promoter in Figure 3.7), and *fliI* was amplified as two arms from *H. pylori* 17874 genomic DNA. Arm 1 was amplified with a forward primer containing a 5' non-annealing sequence complementary to the 3' end of the *recA* promoter sequence and a reverse primer that anneals to the DNA sequence encoding the R-loop with a mismatch that introduces one of the three mutations. Arm 2 was amplified with a forward primer that anneals to the R-loop sequence and contains the same mismatch and sequence complimentary to the 3' end of arm 1, and a reverse primer annealing to the 5' end of *fliI*. The *recA* promoter was amplified with a forward primer and a reverse primer containing non-annealing sequence complimentary to the 5' end of Arm 1. The second round of PCR annealed arm 1 and 2 of *fliI* to generate a PCR template for creation of the full-length *fliI* containing R-loop mutations. The third round of PCR annealed the *recA* promoter to the 5' end of mutated *fliI*.

The 1586 bp products of the three rounds of PCR and wild-type *fliI* fused to the *recA* promoter were then digested with *Bgl*II and *Eco*RI, and ligated to gel-purified linear pHel3 digested with the same enzymes, and therefore with compatible ends. This gave plasmids of 7.1 kb designated pML023-26.

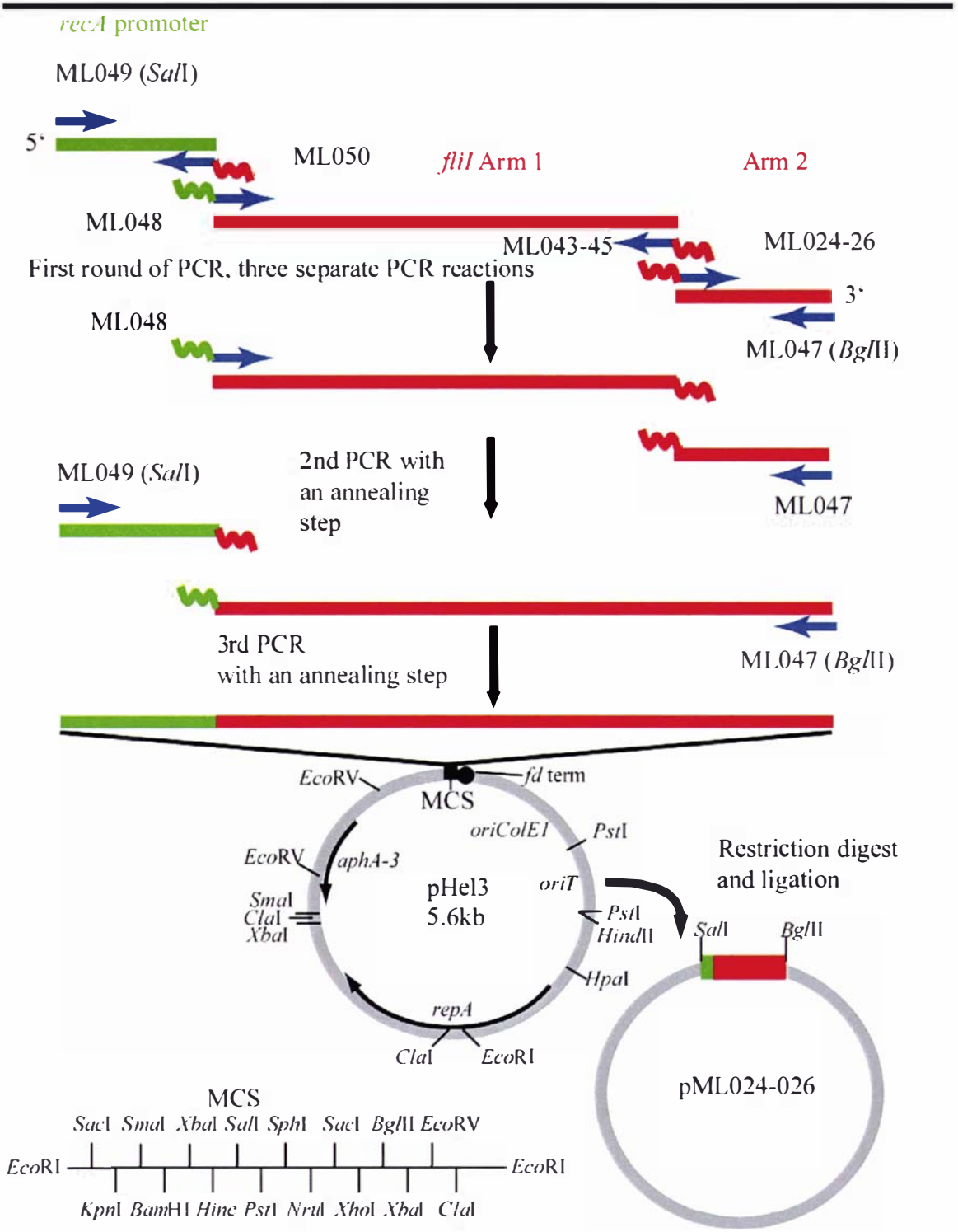


Figure 3.6 PCR based strategy for the mutagenesis of the *fliI* R-loop and cloning of the mutants. The *recA* gene is indicated in green, *fliI* in red and the wavy lines indicate sequence added by the primers that is not present in the target sequence, but required for annealing to PCR products from other reactions. The sequences of the oligonucleotides are shown in Table 3.6.

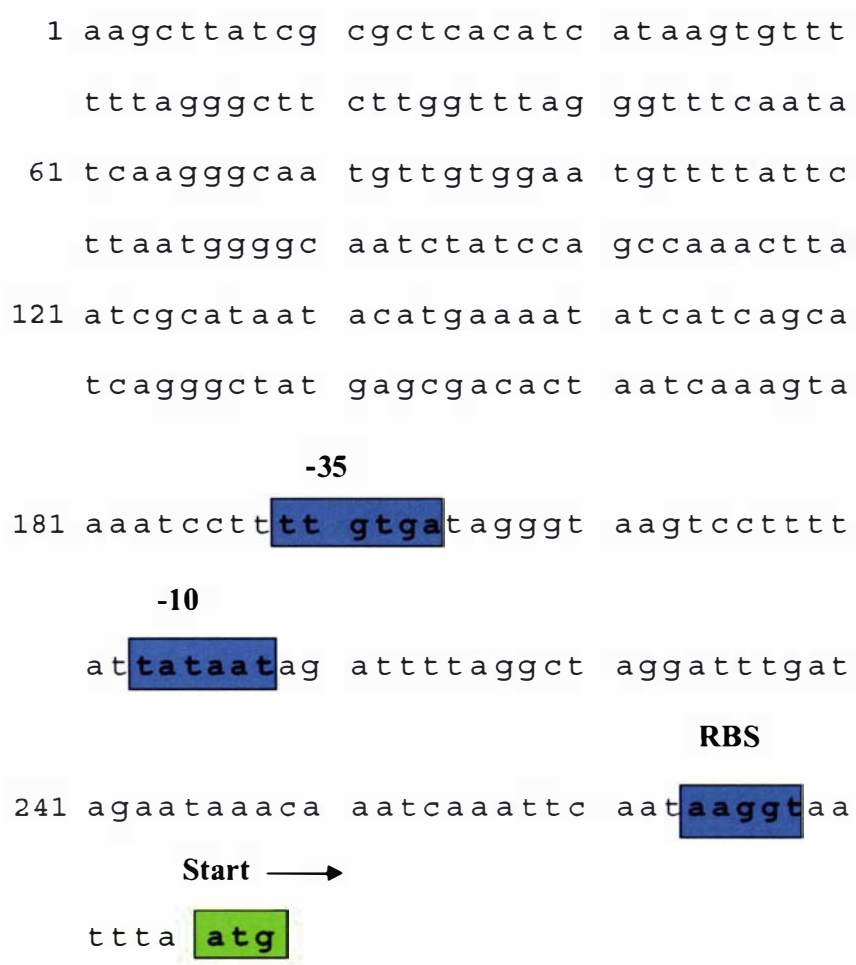


Figure 3.7 *H. pylori recA* Promoter Sequence Used to Express FliI R-loop Mutants (From *recA* sequence from *H. pylori* strain 69A as submitted by Schmitt, W, *et al*, NCBI accession number Z35478 (162)).

Table 3.6. Oligonucleotides used for R-loop mutagenesis in this study

Name	T_m (°C)	Restriction Sites	Annealing Position in Sequence[a]	Sequence[b]	Description
ML024	57.3	na	924 *fliI*	5' GCTAGTA**GCT**GGCGATGATT (forward primer)	E_{311}-A
ML025	55.3	na	924 *fliI*	5' GCTAGTAGAG**GCT**GATGATT (forward primer)	G_{312}-A
ML026	57.3	na	927 *fliI*	5' AGTAGAGGGCGCT**GAT**TTGA (forward primer)	D_{313}-A
ML043	60.3	na	946 *fliI*	5' CAAATCATCGCC**AGC**TACTAGC (reverse primer)	E_{311}-A
ML044	60.3	na	946 *fliI*	5' CAAATCATC**AGC**CTCTACTAGC (reverse primer)	G_{312}-A
ML045	62.3	na	949 *fliI*	5' GCTCAAATC**AGC**GCCCTCTACT (reverse primer)	D_{313}-A
ML047	60.5	*BglII*	1306 *fliI*	5' GA**AGATCT**TTATCTTAAGATTTCTTCTAATTGCTGAAAGC (reverse primer)	na
ML048	61.8	na	256 *recA* promoter, 1 *fliI*	5' ATTCAATAAGGTAATTTAATGCCCCTAAAATCCTTAAAAAACCG (forward primer)	na
ML049	59.3	*SalI*	1 *recA*	5' ACGC**GTCGAC**AAGCTGATTGCGCTCACATCATAA (forward primer)	na
ML050	60.2	na	1 *fliI*, 274 *recA* promoter	5'AGGATTTTAGGGGCATTAAATTACCTTATTGAATTTGATTTGTTTA TTCTATCAAA 3' (reverse primer)	na

[a]Primer annealing positions are relative tom the ATG codon (A=1) See Figure 3.6 for a schematic of the annealing positions and directions of the primers.

[b]Restriction sites and codon changes are in bold type.

To examine the accuracy of constructs, the insert DNA in pML023, pML024, pML025 and pML026 was amplified by PCR using the primers ML049 and ML047 (see Tables 3.6 and 3.8 for the primer details and Table 3.7 for the plasmid details) and the PCR products submitted for sequencing. The primers used for sequencing were at nucleotide sequence positions 256 of *recA*, and positions 1, 358, 598, and 847 of *fliI*, and produced overlapping sequence data. The sequence data demonstrated that the *recA* promoter sequence is correct, and that the *fliI* sequence in each of the recombinant plasmids contains five point mutations relative to the *H. pylori* 26695 *fliI* sequence. Two of these mutations, I_{140}-V and T_{323}-A have already been published by Porwollik and colleagues for the *H. pylori* strain 17874 (142), and are assumed to represent typical strain to strain variation. Three point mutations, S_{52}-T, I_{360}-T and T_{424}-Q, have not been noted before. As they are present in all inserts, each of which was created by an independent PCR reaction, it could be concluded that these mutations are due to colony variation within the strain. The other point mutations are those deliberately introduced by PCR to mutate the R-loop: pML023 *fliI* is wild type, pML024 *fliI* has an E_{311}-A mutation, pML025 *fliI* has a G_{312}-A mutation, and pML026 *fliI* has a D_{313}-A mutation (Table 3.7).

Table 3.7. Plasmids created for FliI R-loop mutagenesis studies

Plasmid name	Vector	Mutagenic primers	Mutation
pML023	pHel3	na	None, wild-type
pML024	pHel3	ML024, ML043	E_{311}-A
pML025	pHel3	ML025, ML044	G_{312}-A
pML026	pHel3	ML026, ML045	D_{313}-A

Table 3.8. Oligonucleotides used for FliI mutant sequencing

Name	T_m (°C)	Annealing Position in Sequence	Sequence
ML048	61.8	256 (*recA*)	5'ATTCAATAAGGTAATTTAATGCCCCTAAAATCCTTAAA AAACCG
ML049	59.3	1 (*fliI*)	5' ACGC**GTCGAC**AAGCTGATTGCGCTCACATCATAA
ML051	57.8	358 (*fliI*)	5' CGATTAGCGCCTGTCATTACA
ML052	58.9	598 (*fliI*)	5' GAATTTATAGAGAAAAACCTGAAAGGG
ML053	60.4	847 (*fliI*)	5' CTTTCCTTATTGCCTCAATTAATGGAG

3.2.2 Testing the effect of the R-loop mutations on motility

Plasmids pML024-26 which express FliI R-loop mutants and pML023 which expresses the wild-type FliI were introduced *in trans* into a *fliI* knock-out strain, *H. pylori* Δ pSP110 (142). The same plasmids were also introduced into wild type *H. pylori* 17874. The transformants were selected for acquisition of the kanamycin resistance marker encoded by the vector portion of the plasmids. The success of transformation was confirmed by purification of plasmids of the correct size from bacteria grown on selective media, by alkaline lysis.

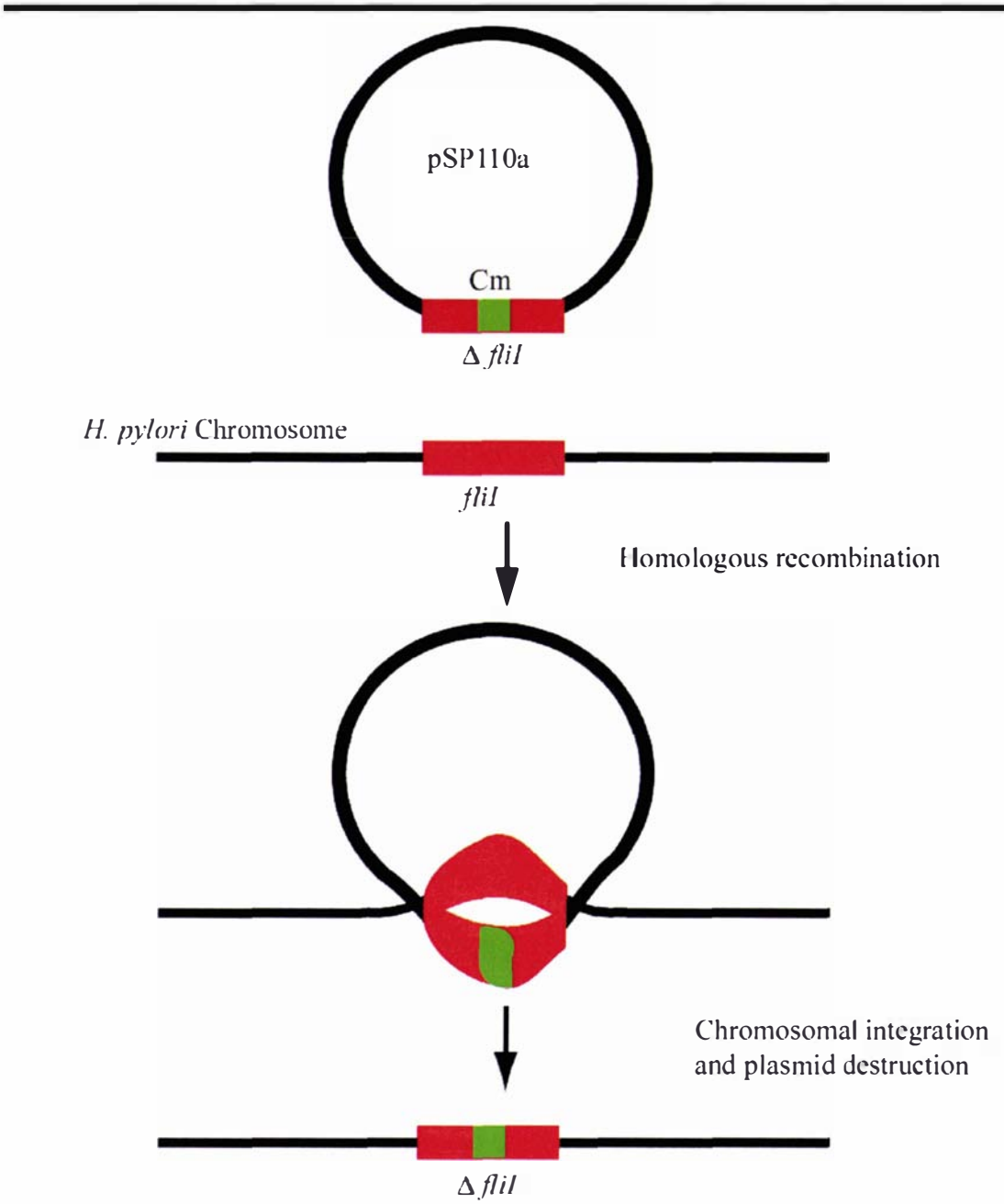


Figure 3.8 Electroporation mediated allelic exchange of *flil*

To examine the influence of the R-loop mutations on motility and flagellation, the *H. pylori fliI* knock-out strain ΔpSP110a, carrying pML023-026 plasmids were observed by confocal light microscopy and electron microscopy (EM). All of the recombinant strains containing *fliI* R-loop mutants were motile and flagellated (Figure 3.9). However, the negative control, *fliI* knock-out strain ΔpSP110a, transformed with the vector pHel3 (no *fliI* insert) was also motile and flagellated. The motility and flagellation of the *fliI* deletion mutant contradicts observations for this very mutant strain created by Porwollik and colleagues (142). This strain was originally generated by electroporation-mediated allelic exchange, utilising a pUC19-based suicide plasmid which carried a *fliI* gene interrupted by a chloramphenicol resistance gene (Figure 3.8). The deletion strain was re-created in this work using the same method and constructs used by Porwollik and colleagues, however this new *fliI* knock-out strain was also motile and flagellated. The reason for this discrepancy is not clear. One possibility is that there is redundancy in *fliI*-like genes in *H. pylori*, and the other possibility is that the *fliI* deletion results in growth disadvantage and consequently intergenic compensatory mutation(s) that also restore flagellation and motility are rapidly selected.

To investigate whether any of the R-loop mutants have a dominant effect over the wild-type *fliI*, the wild-type *H. pylori* 17874 strain was also transformed with pML023-26 and the resulting transformants were observed by confocal light microscopy and EM to monitor motility and flagellation (Figure 3.9). The pML025 and pML026 transformants (carrying mutations G_{312}-A and D_{313}-A, respectively) were motile and flagellated, showing no influence on the phenotype. Surprisingly, the pML024 transformant (carrying E_{311}-A mutation), was hyper-flagellated. In transcription terminator Rho, mutations of the equivalent glutamate produce hyper-terminating mutants. These mutants are characterised by a decrease in their K_M for RNA and a higher RNA (substrate)-dependent ATPase activity. By analogy, FliI mutant could be hyperproductive in flagellar protein export and flagellation in the FliI E_{311}-A R-loop mutant.

Another unexpected observation was the hypo-flagellation of the wild-type strain transformed with the plasmid pML023 which expresses the wild-type FliI. This observation was difficult to explain.

At this stage, due to the unexpected phenotypes of the wild-type and the *fliI* knock-out strains carrying the plasmids pML023 and pML024, the interpretation of the complementation data could only be extremely speculative. Therefore, further research was focused on biochemical characterisation of FliI, with the aim of demonstrating functions such as multimerisation and ATPase activity.

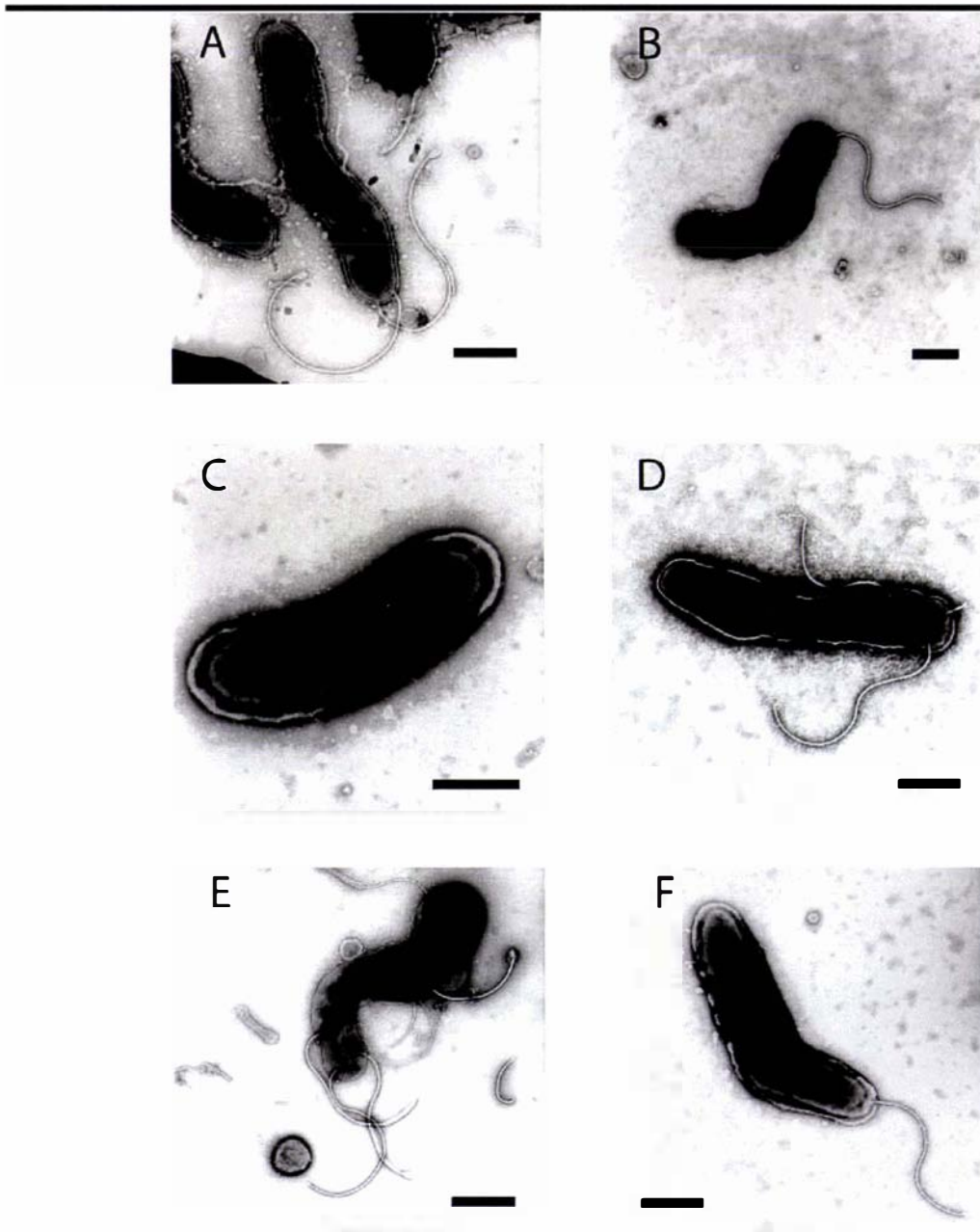


Figure 3.9 Electron Micrographs of wild-type *H. pylori* and *flil* mutants

A. *H. pylori* 17874.

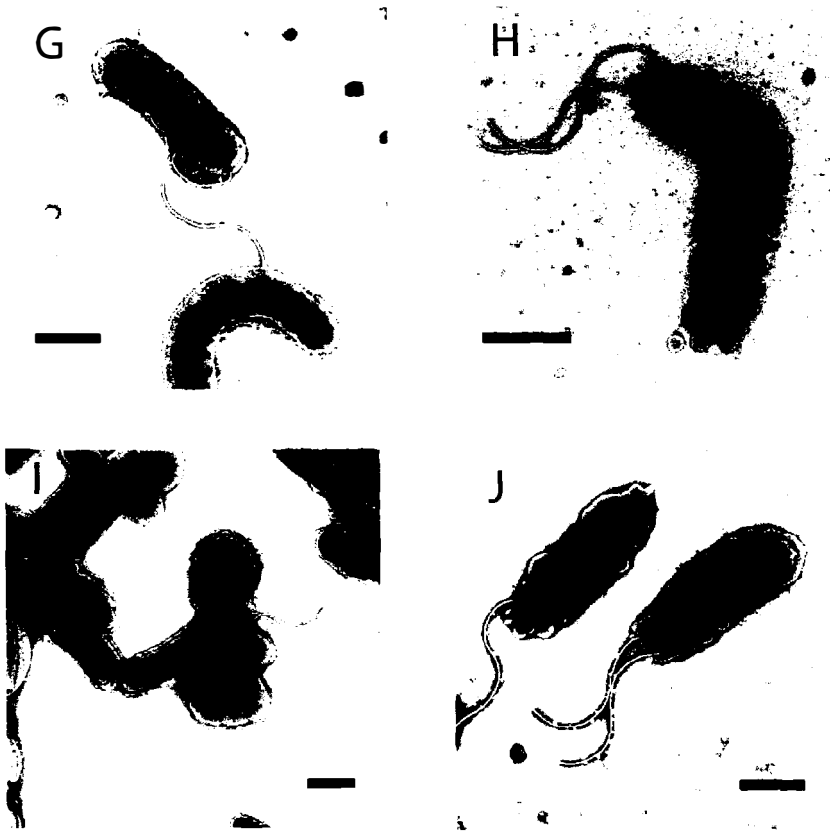
B. *H. pylori* 17874 Δ pSP110a (*flil* knock-out mutant)

C. *H. pylori* 17874(pML023).

D. *H. pylori* 17874 Δ pSP110a(pML023)

E. *H. pylori* 17874(pML024).

F. *H. pylori* 17874 Δ pSP110a(pML024)



G. *H. pylori* 17874(pML025)

H. *H. pylori* 17874 Δ pSP110a(pML025).

I. *H. pylori* 17874(pML026)

J. *H. pylori* 17874 Δ pSP110a(pML025)

The bar in each electron micrograph is equivalent to 1 μm. The magnifications were 15,300, 11,200, 21,200, 15,300, 15,300, 15,300, 15,300, 21,200, 11,200, 15,300X for panels A-J respectively. The cells were negatively stained with 1% PTA and visualised on a Philips 201C transmission electron microscope.

3.2.3 *In vitro* Characterisation of *H. pylori* FliI

3.2.3.1 Affinity Purification of FliI

The gene encoding full-length FliI and truncated versions of the *fliI* gene were cloned as N-terminal fusions to the glutathione-S-transferase (GST) gene of *S. japonicum* in the pGEX-6P3 vector (GM Healthcare). When expressed from this plasmid, the GST-FliI fusion can be purified via the high affinity of GST for glutathione immobilised on sepharose beads. The GST tag can then be removed by PreScission protease that cleaves a specific recognition sequence within the linker between GST and FliI, to release free FliI for further purification. The same GST-affinity purification system was used to purify all proteins examined in this study. This purification is detailed in the Methods chapter and summarised in Figure 3.10.

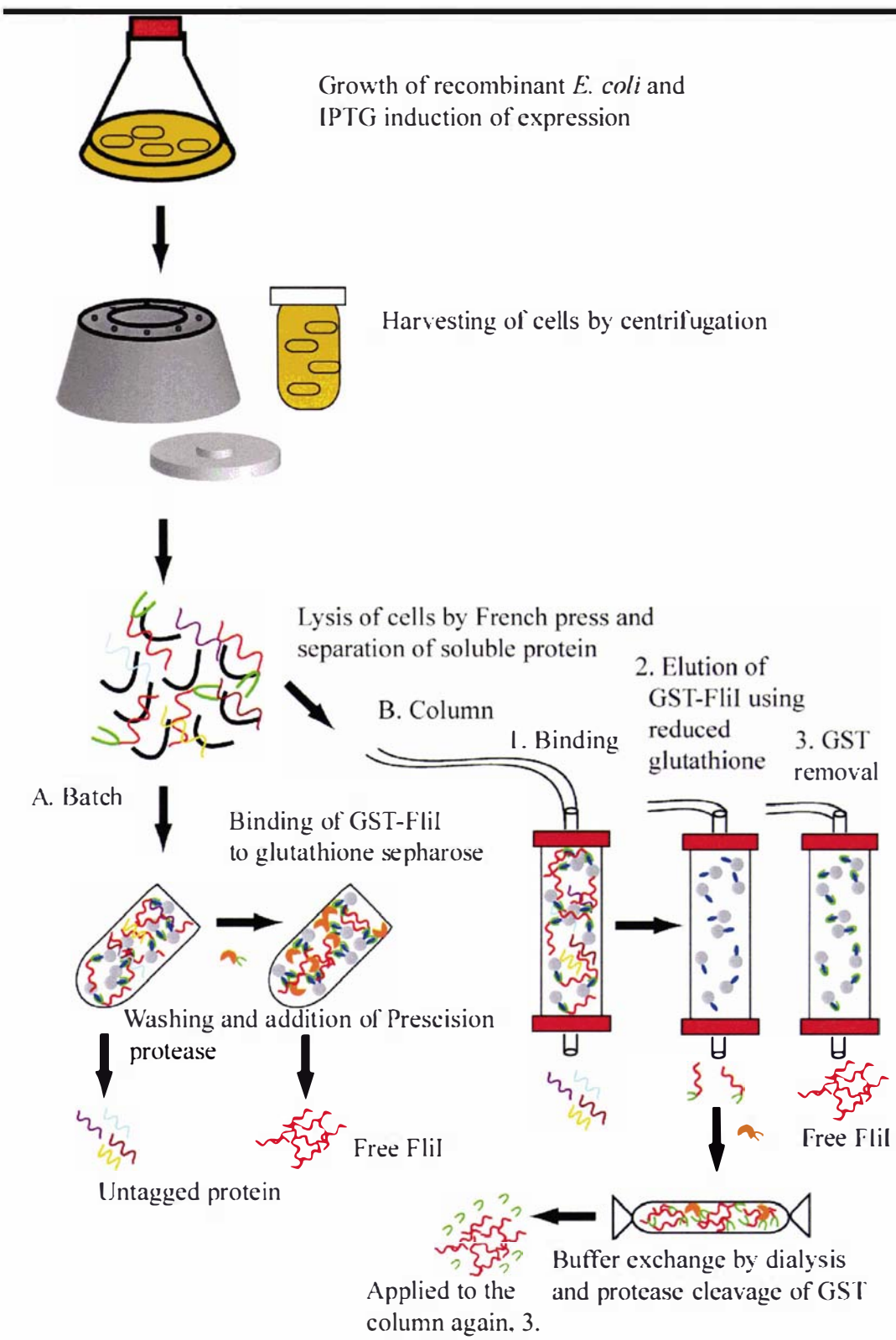


Figure 3.10 GST Affinity purification of FliI (Pharmacia Biotech)

3.2.3.2 Over-expression of FliI 2-434 and FliI 19-434

The F_1-ATPase α- and β- subunits contain an N-terminal extension that is disordered in the crystal structure and protease sensitive *in vitro* (1). Based on the FliI multiple alignment with the F_1-ATPase subunits (Figure 1.7) this disordered region may correspond to the first 18 residues of *H. pylori* FliI. As these residues may affect the behaviour of the protein, *fliI* was also cloned using the same method with a forward primer annealing at position 55 of the DNA sequence (Table 2.3, Methods section). When expressed, this truncated FliI will lack the first 18-residues. FliI expressed from this construct is termed FliI 19-434. Members of the Moore lab created both the FliI 2-434 and FliI 19-434 plasmids before this project began, and the constructs were sequenced to confirm that *fliI* has no PCR-induced mutations.

For expression purposes, *E. coli* Rosetta strain was transformed with the FliI 19-434 or 2-434 plasmids, and the transformant *E. coli* streaked for single colonies on LB Agar Amp^{100} Cm^{10} plates. These plates were incubated at 37 °C overnight. A single colony was then used to inoculate 50 ml of LB Amp^{100} Cm^{10} that was grown overnight at 37 °C in a shaking incubator. The next day 2L of Amp^{100} Cm^{10} Hyper broth was inoculated with 20 mL of this overnight and grown in a 37 °C shaking incubator until an $OD_{600\ nm}$ of 0.6 AU. Expression of FliI was then induced by the addition of IPTG to the flask to a final concentration of 0.1 mM. The incubator temperature was changed to 20 °C and the cells were grown for a further 18 hours. The cells were then harvested by centrifugation and lysed by French press, and separated into soluble and insoluble fractions by centrifugation. Samples of the lysate, lysate supernatant and lysate pellet were kept for analysis and run on an SDS-PAGE gel (Figure 3.11). This gel shows that FliI expression has occurred but most of the protein (possibly 90%) is insoluble and resides in the lysate pellet.

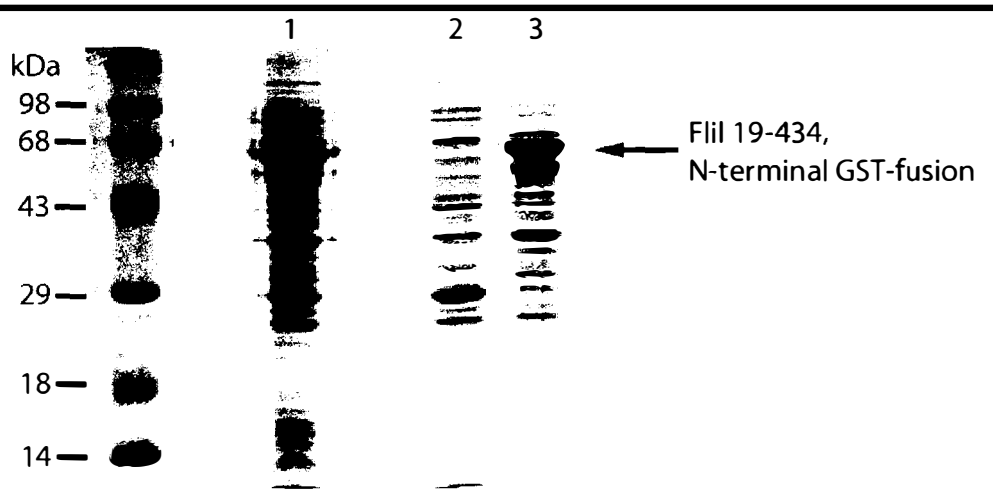


Figure 3.11 Expression and solubility of GST-FliI 19-434. FliI 19-434 was expressed in *E. coli* Rosetta grown at 20 °C for 18 hours after induction with 0.1 mM IPTG at an $OD_{600\ nm}$ of 0.6. Samples were analysed by 12% SDS-PAGE.

Lane:

1, *E. coli* total cell lysate, 10 µl;

2, Lysate supernatant (soluble proteins), 10 µl;

3, Lysate pellet (insoluble proteins), 10 µl

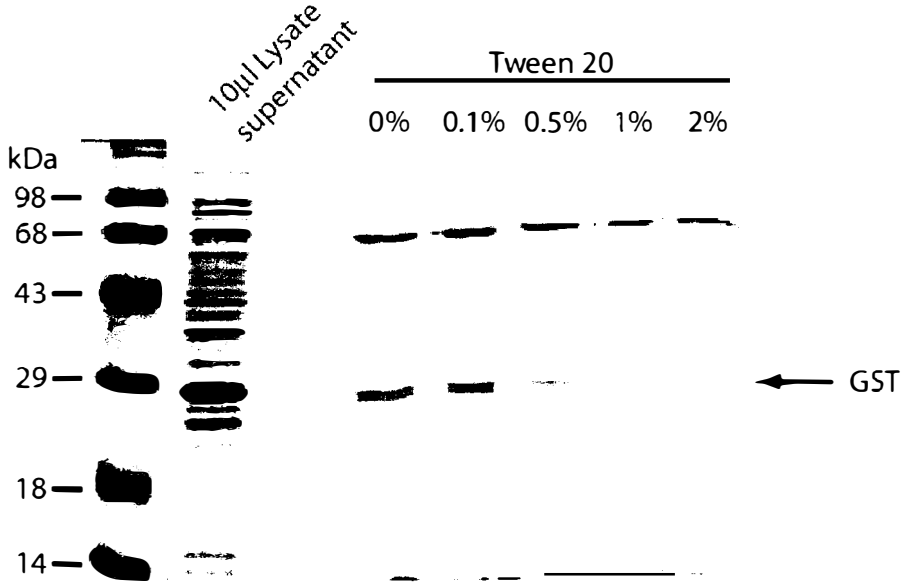


Figure 3.12 The effect of Tween 20 concentration in the wash buffer on the impurities present in FliI purifications. 12% SDS-PAGE gel of GST-FliI washed with buffer containing increasing concentration of Tween-20.

The GST-FliI fusion was then bound to 3 ml of washed Glutathione-sepharose, by mixing the lysate supernatant with the glutathione sepharose in a 50 ml tube on a rocker at 4 °C for 4 hours. The lysate supernatant was then removed and the sepharose washed ten times with 10 volumes of GST cleavage buffer: 50 mM Tris, 150 mM NaCl, 1 mM EDTA, pH 8.0. When a sample of the glutathione sepharose is run on an SDS-PAGE gel after GST-FliI is bound, significant non-specific protein binding can be seen. It is important to reduce this non-specific binding to improve the purity of the protein, and also to remove background proteins to facilitate interaction studies by GST-pulldown. To this end the GST-FliI bound sepharose was either washed with varying concentrations of Tween-20 (Figure 3.12), or varying concentrations of Triton X-100 (not shown) were added while GST-FliI was incubated with the glutathione sepharose.

However, treatment with these weak detergents had little effect on non-specific binding of proteins to the glutathione sepharose. Another attempt to prevent non-specific protein binding was made by using gelatin as a non-specific blocking agent (Figure 3.13). As seen in this gel, this was also ineffective and actually added further impurities. It was decided that the FliI impurities would have to be removed in subsequent purification steps.

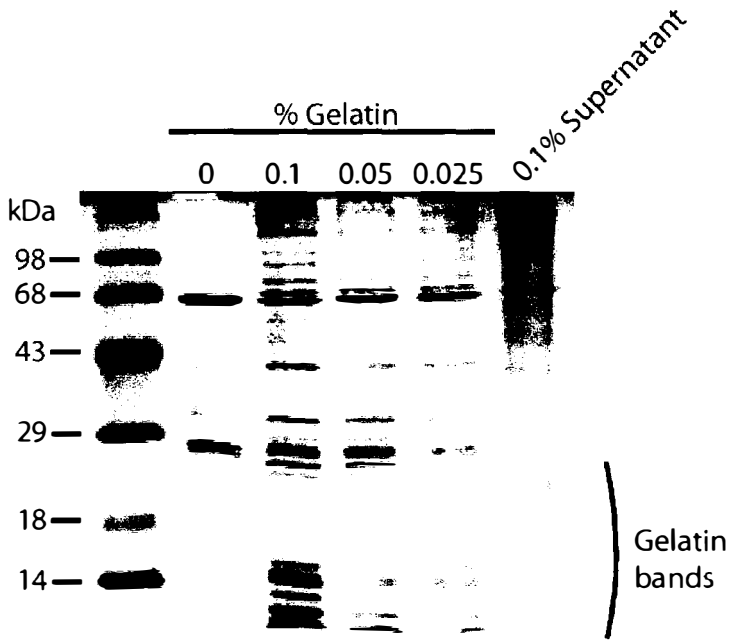


Figure 3.13 12% SDS-PAGE gel of glutathione-sepharose blocked with gelatin during GST-FliI binding. GST-FliI was bound to glutathione-sepharose in the same manner as Figure 3.12, except that a gelatin solution was added to the percentages stated during the 1 hour binding step.

Proteolytic cleavage of the GST-fusion protein was subsequently initiated by the addition of an equal volume of GST cleavage buffer containing 1 mM DTT, and 80 units of PreScission protease / ml of bed volume to the glutathione sepharose. Cleavage was allowed to continue overnight on a rocker at 4 °C. The next day the supernatant containing the free FliI was removed and placed at 4 °C until it was subjected to further purification by chromatography. A small aliquot of the washed cleavage supernatant containing FliI was kept for analysis on by SDS-PAGE (Figure 3.14).

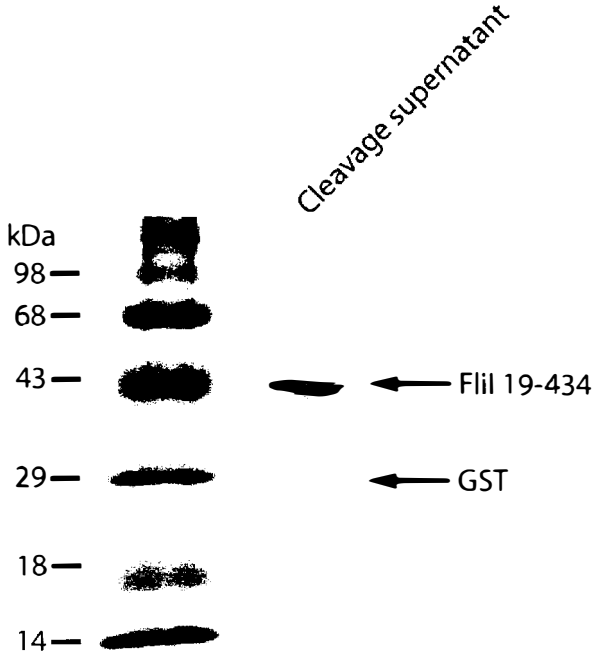


Figure 3.14 Separation of FliI 19-434 from the GST tag. 12% SDS-PAGE gel of the FliI 19-434 after the GST tag has been cleaved away using the PreScission protease and then partially depleted using Glutathione sepharose.

3.2.3.3 Purification of FliI 2-434 and FliI 19-434 by fast protein liquid chromatography (FPLC)

FliI 2-434 and FliI 19-434 have predicted isoelectric points of 6.0 and 5.5 respectively. Consequently, it was decided that anion exchange using a 10 ml Source-Q ion exchange column at pH 8.0 would be an appropriate first purification step. The column was equilibrated in 50 ml of start buffer, 50 mM Tris-HCl, 100 mM NaCl, pH 8.0. Although this buffer differs from the cleavage buffer by 50 mM NaCl, it was found that this made no difference to the purification of FliI 2-434 or 19-434. Consequently, cleaved FliI was injected onto the column without prior equilibration into start buffer. In the example below, 2.25 ml of FliI 2-434 cleaved from 3 ml of glutathione sepharose was injected onto the Source-Q column at a flow rate of 1 ml/min (Figure 3.15). Following injection, the flow-rate was changed to 2 ml/min and another 30 ml of start buffer was pumped down the column to remove any unbound protein and to re-equilibrate the column. Then any bound protein was eluted by applying a linear gradient of 100-500 mM NaCl over 50 ml of elution

buffer: 50 mM Tris-HCl, 1 M NaCl, pH 8.0. The column was then stripped of any remaining protein with 1 M NaCl by pumping elution buffer alone.

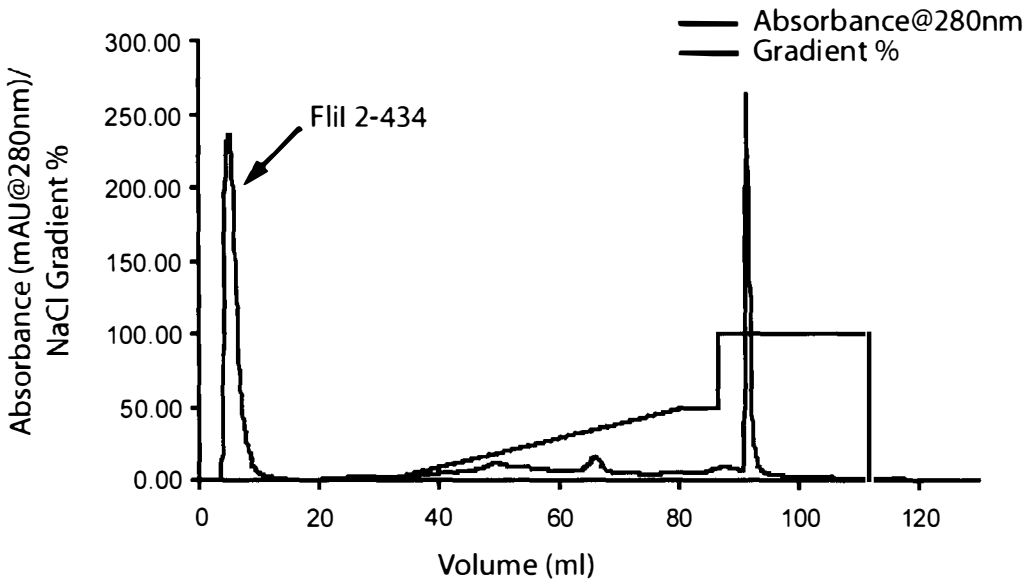


Figure 3.15 Purification of FliI 2-434 using Source-Q Anion Exchange Chromatography.

The start buffer was 50 mM Tris-HCl, 100 mM NaCl, pH 8.0; the elution buffer was 50 mM Tris-HCl, 1 M NaCl, pH 8.0; the flow rate of injection was 1 ml/min all other steps were performed at 2 ml/min. The bed volume of the column was 10 ml.

As can be seen in Figure 3.15 and Figure 3.17, FliI 2-434 (and 19-434) do not bind to Source-Q sepharose under these conditions but rather pass through the column in the void volume, in a volume of approximately 6 ml. But in spite of this poor binding Figure 3.17 shows that a degree of purification is still achieved, presumably because other protein contaminants bind to the column.

Following anion exchange chromatography FliI 2-434 was concentrated using a Millipore Amicon Ultra-4 10,000 Molecular Weight Cut Off (MWCO) centricon to a final volume of 390 µl. To purify FliI further it was decided to apply the protein to a Superdex 200 gel filtration column. This column is also useful for determining the size of a protein or a protein complex, and therefore can indicate if FliI is behaving as a hexamer in solution. To this end, FliI was incubated under conditions thought to be conducive to the formation of hexamers at 4 °C on a rocker overnight prior to size exclusion chromatography. The buffer was 50 mM Tris-HCl, 100 mM NaCl, 5 mM $MgCl_2$, 100 mM Urea, 10 mM ATP, 1

mM DTT, pH 8.0. The column was pre-equilibrated with three column volumes of a 50 mM Sodium phosphate, 150 mM NaCl, 5 mM $MgCl_2$, pH 8.0 buffer. Then FliI 2-434 was injected and eluted from the Superdex 200 column at a flow rate of 0.5 ml/min (Figure 3.16).

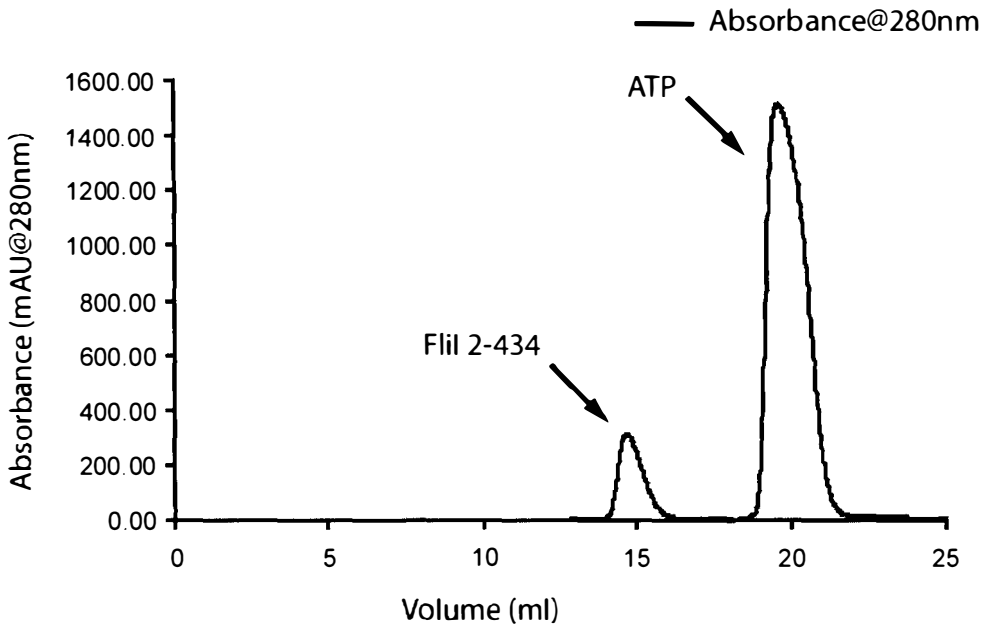


Figure 3.16 Purification of FliI 2-434 using a Superdex 200 Size Exclusion Column.
The column was equilibrated in 50 mM Sodium phosphate, 150 mM NaCl, 5 mM $MgCl_2$, pH 8.0 buffer. The volume of FliI injected was 320 µl, and the flow rate of the injection and elution was 0.5 ml/min.

The SDS-PAGE gel in Figure 3.17 illustrates that Source-Q anion exchange chromatography followed by size exclusion chromatography has resulted in purification of FliI 2-434 almost to homogeneity. The same purification method has also been used successfully to purify FliI 19-434. But this purification demonstrates two important points that are worth noting. Firstly, although the purification produces very pure protein the yield is low and this problem is due to the poor solubility of FliI which seems to be an inherent problem with *H. pylori* FliI and its homologues (eg. *S. typhimurium* FliI (43) (52)). This can be illustrated by reference to the Figure 3.17, which shows the protein present at each stage of the purification as the gel is examined from Lanes 1-8. Lanes 1 and 2 of the gel are 1 µl of the lysate supernatant before and after mixing with the glutathione sepharose. There

is no obvious over-expressed protein in the supernatant before incubation with the glutathione sepharose (Lane 1). Most of the FliI is in the pellet, and after the supernatant is removed from the sepharose there is no obvious decrease in any of the protein bands (Lane 2). Despite the low levels of soluble GST-FliI protein in the lysate supernatant, GST-FliI from the lysate supernatant does bind to the glutathione sepharose and the GST tag can be removed by proteolysis (Lane 3). The cleaved protein visible in Lane 5 is purified from 4L of *E. coli* and thus the yield is low, furthermore the free FliI precipitates during cleavage of the GST tag. This precipitate was pelleted prior to the Source-Q step, resuspended in 2.25 ml of cleavage buffer, and a 5ul aliquot loaded in Lane 4 of the gel in Figure 3.17. After the Source-Q step, FliI 2-434 was concentrated to 370 µl using a centricon and the concentration was estimated to be 19.7 mg/ml by UV absorbance at 276 nm (Absorbance of a 1 mg/ml solution@276 nm=0.312, based on the abundance of Tyrosine and Tryptophan residues in the FliI amino acid sequence). The concentration of FliI in the proteolysis supernatant could not be determined by similar UV spectroscopy because of protein contaminants, but 5 µl was loaded in lane 5 and the intensity of the Coomassie-stained band gives a rough estimate of the relative protein concentration. Assuming that the concentration of FliI is half the concentration of FliI after Source-Q anion exchange, and the volume of the cleavage supernatant is seven times greater, then there was approximately 20 mg of protein in the cleavage supernatant. Therefore, approximately two thirds of the protein have been lost during either during the chromatography steps or as a precipitate when the FliI was stored at 4 °C between purification steps. Indeed after the size exclusion chromatography and subsequent concentration of FliI (Lane 8), the protein continues to precipitate at 4 °C until it reaches a concentration of 2 mg/ml. Precipitation of *S. typhimurium* FliI has also been observed (43).

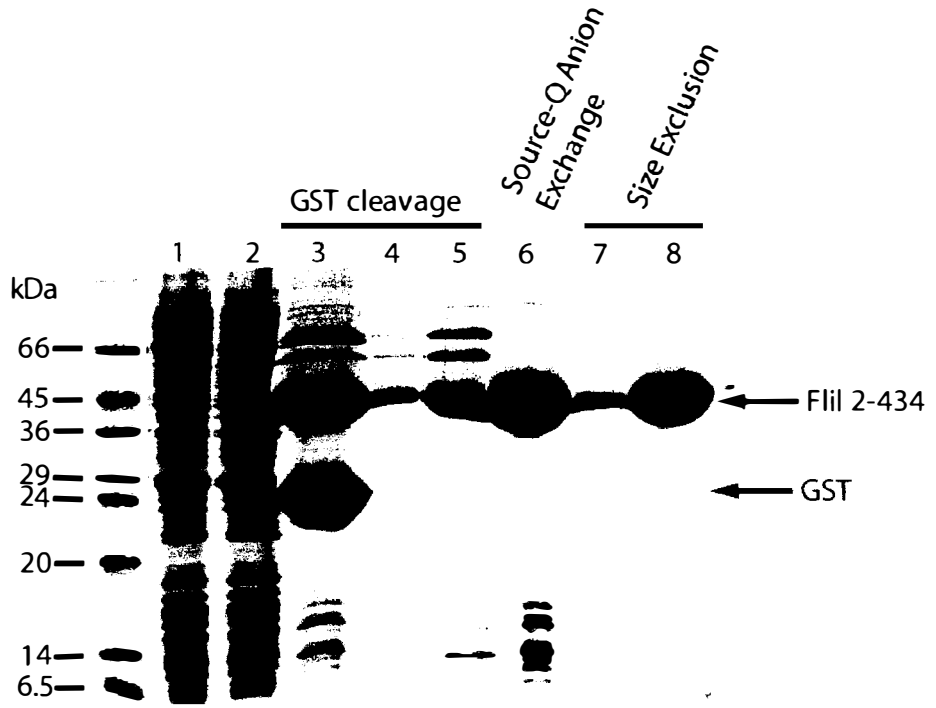


Figure 3.17 Purification of FliI 2-434. Purity of the FliI 2-434 at various stages of purification is monitored by analysis on 15% SDS-PAGE

Lane: 1, *E. coli* lysate supernatant, 1 μl;

2, *E. coli* lysate supernatant after incubation with glutathione sepharose, 1 μl;

3, glutathione sepharose after PreScission protease cleavage of the GST affinity tag, 5 μl;

4, protein precipitate from the GST cleavage supernatant, 5 μl;

5, GST cleavage supernatant, 5 μl;

6, flow-through peak from Source-Q anion exchange of FliI 2-434 (Figure 3.2.8) after concentration, 5 μl;

7, first size exclusion peak after concentrated, 10 μl;

8, second size exclusion peak concentrated, 5 μl.

Attempts were also made to purify FliI using hydrophobic interaction chromatography (Figure 3.18 A and B). This was very effective for purifying FliI almost to homogeneity in a single step, but again the yield of purified protein was low.

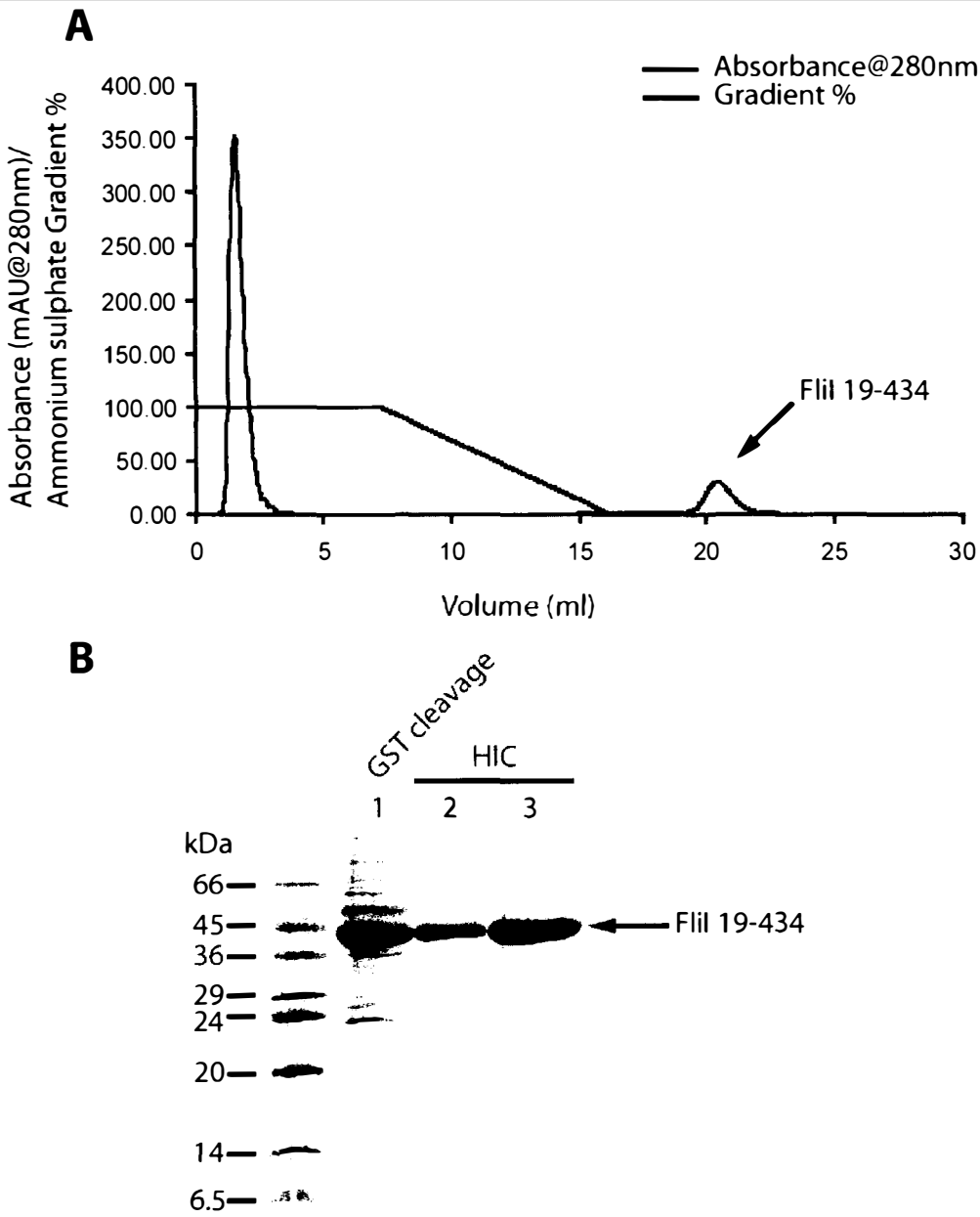


Figure 3.18 Purification of FliI 19-434 by Hydrophobic Interaction Chromatography. **A.** Chromatogram of FliI 19-434 purification using a 1 ml HiTrap Phenyl HP HIC column equilibrated in 50 mM Sodium phosphate, 800 mM Ammonium sulphate, pH 7.0. A volume of 500 µl of partially purified FliI was injected at a flow-rate of 1 ml/min, FliI 19-434 was eluted in 2 ml at the end of an 800 mM-0 mM gradient over five column volumes. **B.** 15% SDS-PAGE gel of FliI 19-434 HIC purication.

Lane: 1, Partially purified FliI after GST affinity chromatography and proteolytic cleavage, 10 µl;

2, HIC peak concentrated to 200 µl, 4 µl; 3, 10ul of the HIC peak.

The second point to note from the purification of FliI 2-434 is the elution volume of FliI during size exclusion chromatography, which as previously mentioned, can be used to estimate the molecular mass of a protein. These predictions are based on the linear relationship between log molecular mass and normalised elution volume in the form of K_{av}. If proteins of known molecular weight are plotted against their elution volumes, a linear equation can be used to determine an unknown molecular mass based on elution volume. Such a calibration curve for the Superdex 200 column used in this study is shown below in Figure 3.19 and the values from which it was derived in Figure 3.20.

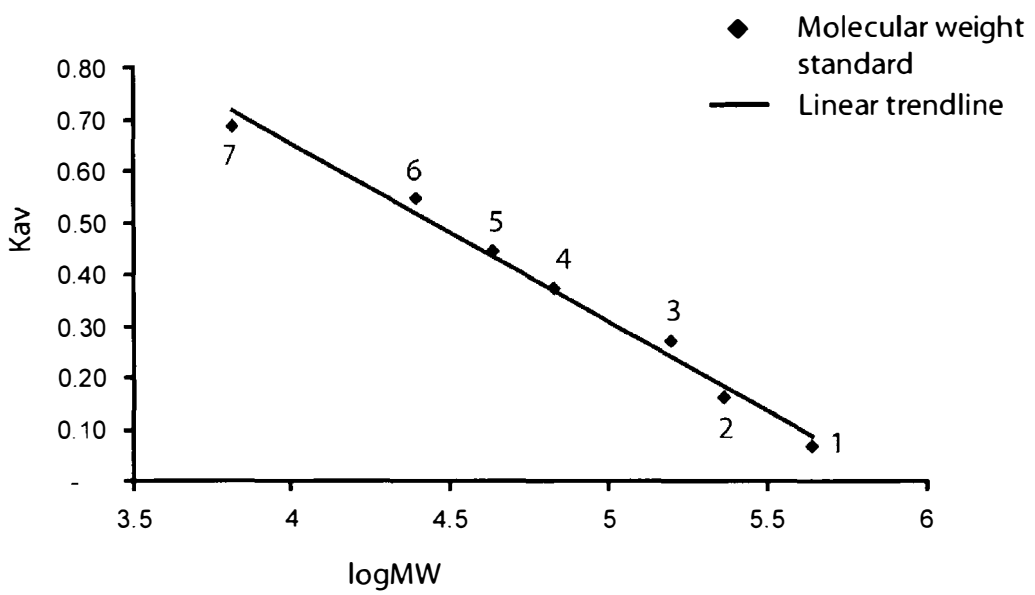


Figure 3.19. Superdex 200 Size Exclusion Column calibration curve

$K_{av}= V_e-V_0/V_t -V_0$

V_e= elution volume, V_0= column void volume, V_t = column volume.

The void volume is determined by the elution volume of Blue Dextran.

1= Ferritin, 2= Catalase 3= Aldolase, 4= Albumin, 5= Ovalbumin, 6= Chymotrypsinogen A, 7= Aprotinin.

The equation of the line: $y=-0.3442x + 2.0313$ $R^2=0.986$.

	V_e	V_0	V_t	K_{av}	MW	logMW
Peak 1 Ferritin	9.12	8.04	24	0.07	440,000	5.64
Peak 2 Catalase	10.64	8.04	24	0.16	232,000	5.37
Peak 3 Aldolase	12.40	8.04	24	0.27	158,000	5.20
Peak 4 Albumin	14.00	8.04	24	0.37	67,000	4.83
Peak 5 Ovalbumin	15.12	8.04	24	0.44	43,000	4.63
Peak 6 Chymotrypsinogen	16.80	8.04	24	0.55	25,000	4.40
Peak 7 Aprotinin	19.01	8.04	24	0.69	6,500	3.81

Figure 3.20. Superdex 200 Size Exclusion Column calibration data used in Figure 3.19

Most of the FliI 2-434 eluted at a volume of 14.7 ml in Figure 3.16 and therefore has a predicted molecular weight of 48 kDa. This suggests that despite being incubated in a buffer that should be conducive to the formation of hexamers, FliI remains primarily monomeric in solution. However there was a second peak of 10mAU that eluted at a volume of 13.3 ml, but this is not visible in Figure 3.16 because of the scale. At this elution volume the size of FliI would be 88 kDa, and therefore a small percentage of FliI may have been present as a dimer. This is unlikely to be physiologically relevant as a FliI dimer has been previously reported in *S. typhimurium* (52) as an experimental artifact in the absence of a reducing agent, and it is not seen every time *H. pylori* FliI 2-434 is eluted from the size exclusion column. To explore the possibility that the buffer conditions were inappropriate for hexamerisation, purified FliI 2-434 was also incubated with AMP-PNP instead of ATP, and also with a different buffer: 50 mM HEPES, 150 mM Potassium acetate, 20 mM Magnesium acetate, pH 8.0 in the presence anionic phospholipids and ATP. Neither of these conditions was conducive to hexamerisation based on elution volumes from the size exclusion column. Yet, despite the lack of a demonstrable *H. pylori* FliI 2-434 hexamer, Hughes and colleagues have recently demonstrated the formation of *S. typhimurium* FliI

hexamers under similar experimental conditions (31). In their study formation of hexamers was enhanced by the presence of AMP-PNP as an alternative to ATP and phospholipids.

The first attempt to purify *S. typhimurium* FliI involved refolding the protein from insoluble inclusion bodies formed during overexpression (43). FliI purified in this manner could bind ATP but it was not a functional ATPase, and may have been misfolded. Given the failure to observe FliI hexamers, the folding of purified FliI 2-434 was investigated by Far UV Circular Dichroism (CD) Spectroscopy. The major feature of the CD spectrum of purified FliI 2-434 is a trough with two minima, the first is -1.36 mdeg at 220.5 nm then the elipticity rises slightly to -1.27 mdeg before dropping to the second minimum of -1.31 mdeg at 210.5 nm. The spectrum then rises to a maximum elipticity of 1.43 mdeg at 190 nm without reaching a discernable peak. The spectrum may have peaked around 190 nm but measurements could not be obtained between 180 and 190 nm. For comparison the characteristic CD spectrum of an α-helical secondary structure consists of a trough with two minima at approximately 225 and 210 nm. An anti-parallel β-sheet has a trough at 215 nm and a peak at 195 nm (78). Therefore, the FliI spectrum suggests FliI is folded with elements of both α-helical and β-sheet secondary structure. This assertion is supported by analysis of the FliI CD spectrum with two deconvolution algorithms, CDNN and K2d (23) (12). These predict FliI has 7-11% helical content, 45% β-structure (parallel, anti-parallel and turn) and 33-48% random coil. Although the model of *H. pylori* FliI based on the structure of F_1-ATPase suggests 31% α-helix, 24% β-sheet and 44% random coil (loop), this discrepancy could be error due to the lack of data between 180 and 195 nm, as CD spectra of secondary structural elements contain a characteristic feature within this range. Furthermore, random coil secondary structure has a characteristic CD spectrum with a small positive peak at 220 nm and a trough with a minimum elipticity at 197.5 nm. Neither of these features are evident in the FliI CD spectrum. The CD spectrum of FliI was also measured following the addition of Urea to 0.6 and 2.4 M. In both cases the trough changed shape such that there was only one minimum, and minimum elipticity decreased to 1.9 and 1.98 mdeg at 210 and 206 nm respectively. This suggests that the protein is unfolding and possibly contains significant random coil in the presence of Urea. In conclusion our

purified recombinant *H. pylori* FliI is probably folded but this data is not robust enough to be conclusive.

3.2.3.4 Over-expression of the FliI N-terminal Domain

Given the poor solubility of FliI and the low yield from the purification, it was decided to clone the N-terminal domain of FliI, as a smaller, more soluble fragment should be easier to study. In addition less is known about this domain. It has previously been suggested that it may have a flagellum-specific function (118). Based on a FliI/F_1-ATPase multiple sequence alignment (Figure 1.7) and the known structures of the F_1-ATPase subunits, the N-terminal domain extends from residue 2 to 91 of *H. pylori* FliI and consists of two sub-structures. Residues 2-18 of *H. pylori* FliI are predicted to correspond to the N-terminal extension that is disordered and protease sensitive in the Bovine $F_1\alpha$ and β structures (1). Residues 19-91 of *H. pylori* FliI are predicted to form a six-stranded β-barrel structure. In *E. coli* rho the strands of this β-barrel have a different connectivity and it resembles an OB-fold but it does contain a function-specific feature, the primary RNA binding site (42), and in the Type III ATPase InvC loss of function mutations have been identified within the sequence corresponding to our predicted β-barrel domain (2). Furthermore a temperature sensitive L12P mutation of FliI has also been identified in *S. typhimurium* that results in reduced filament regrowth after shearing, implicating the N-terminal extension in the function of FliI. However, the significance of this mutation requires further investigation as it will likely have changed the structure of the FliI N-terminus dramatically.

To investigate the structure and function of the *H. pylori* FliI N-terminal domain the DNA sequences encoding amino acid residues 2-91 and 19-91, were amplified and cloned as N-terminal GST fusions in the pGEX-6P3 over-expression vector in the same manner as FliI 2-434 and 19-434. This cloning was completed by our technician Stacey McDonald.

FliI 2-91 and 19 -91 were expressed under the same conditions as FliI 2-434 and 19-434, with the only difference being the use of LB instead of Hyper Broth for growth of cells and expression of protein. A small sample of the lysate supernatant from the second spin and the pellet from the first spin were analysed by SDS-PAGE (Figure 3.21). This gel shows

that, in contrast to FliI 2-434 and 19-434, both of the N-terminal constructs FliI 2-91 and 19-91 are moderately soluble.

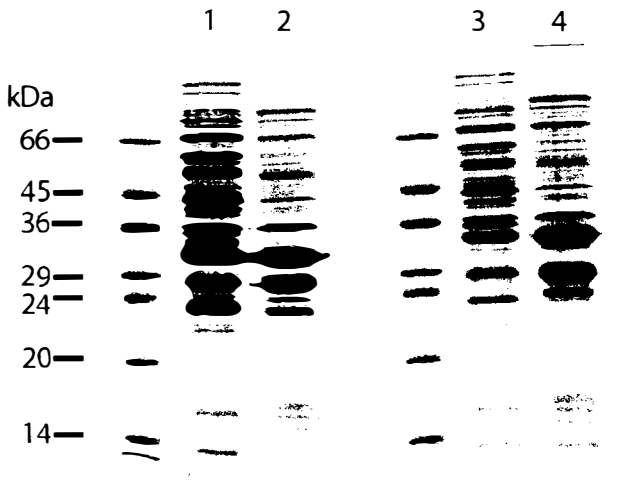


Figure 3.21 GST-FliI 2-91 and GST-FliI 19-91 expression. The proteins in the lysate are analysed by 12% SDS-PAGE.

Lane 1, Lysate supernatant from *E. coli* cells expressing GST-FliI 19-91
Lane 2, Lysate pellet from *E. coli* cells expressing GST-FliI 19-91
Lane 3, Lysate supernatant from *E. coli* cells expressing GST-FliI 2-91
Lane 4, Lysate pellet from *E. coli* cells expressing GST-FliI 2-91

Both FliI 2-91 and 19-91 can be purified in the same manner, firstly GST affinity purification followed by hydrophobic interaction chromatography (HIC) and finally size exclusion chromatography (SEC) with the Superdex 75 column. Affinity purification can be achieved by incubating the cell lysate with glutathione sepharose to bind the fusion proteins to the sepharose and then cleaving the GST tag with PreScission protease, in the same manner already described for FliI 2-434 and 19-434 (Figure 3.10). Due to the much greater solubility of FliI 2-91 and 19-91, the lysate from 500 ml of bacterial culture is sufficient to saturate 10 ml of glutathione sepharose. Recently GST-FliI 2-91 and 19-91 have been purified from the supernatant of cell lysates by using reduced glutathione to elute the fusion proteins from a glutathione sepharose column using FPLC (Figure 3.22). An example of a typical elution profile is shown below in Figure 3.22. Following buffer exchange to remove the glutathione and PreScission protease cleavage of the GST tag, the cleavage mixture is pumped onto the glutathione sepharose column to remove the free GST

tag, and the free FliI 2-91 or 19-91 flows through the column while the GST binds to the glutathione sepharose. The GST is then removed from the column with elution buffer. An example of a typical elution profile following the cleavage of GST from a fusion protein and re-application of the sample to a glutathione sepharose column is shown in Figure 3.22. Small samples of the protein were also kept at the various stages of this purification and analysed by SDS-PAGE to monitor the effectiveness of the procedure (Figure 3.23).

This method of removing the GST tag has been found to be more efficient as it requires less protease and the glutathione sepharose can be reused after it has been washed with 6M Guanidinium-HCl and 70% Ethanol.

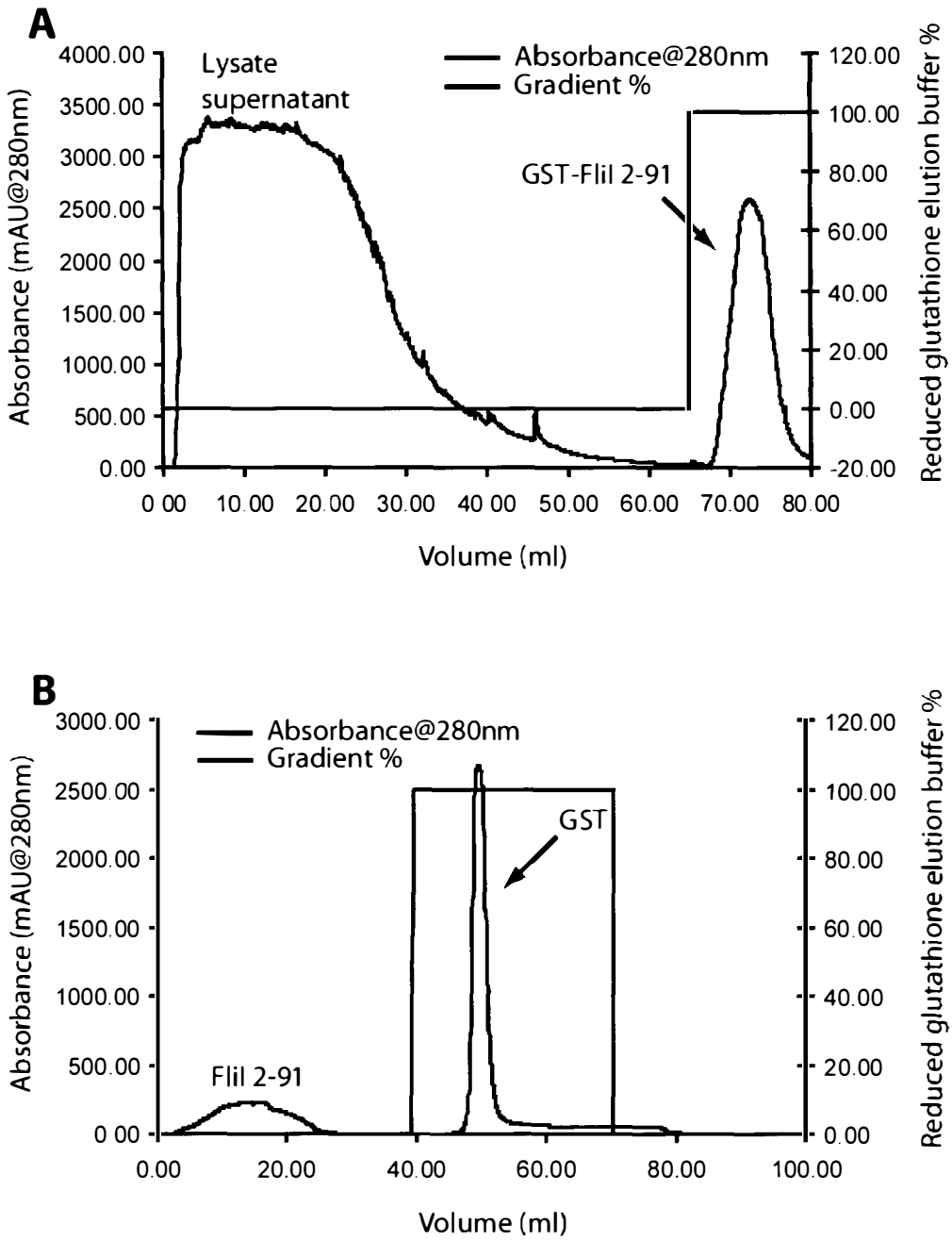


Figure 3.22 Affinity purification of FliI 2-91 using a GST tag The flow-rate was 1 ml/min and the bed volume was approximately 10 ml. **A).** Elution profile of GST-FliI 2-91. Approximately 20 ml of lysate supernatant was loaded on the column. The elution volume of the fusion protein was 73 ml. **B).** Elution profile of FliI 2-91 after cleavage of the GST tag. The elution volume of FliI 2-91 was 16 ml, GST eluted at 50 ml.

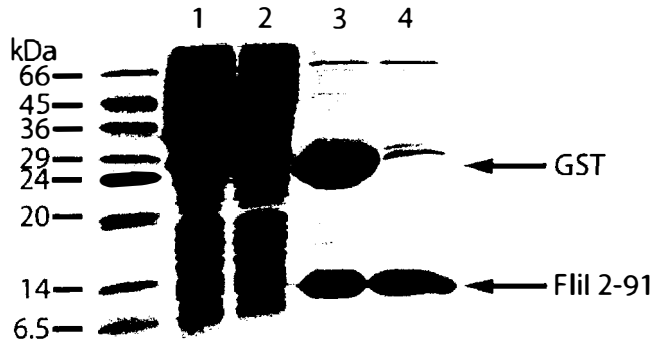


Figure 3.23. Affinity purification of FliI 2-91. Samples at successive steps during affinity purification were analysed by SDS-PAGE on 20% gel.

Lane: 1. Lysate supernatant before binding to the glutathione-sepharose column.

2. The protein that flows through the column during binding, this sample was taken after approximately 10 ml of the sample had been loaded onto the column.

3. A sample of the eluted GST-FliI 2-91 after it had been dialysed and cleaved.

4. FliI 2-91 peak that flowed through the glutathione-sepharose column when the cleaved fusion protein (lane 3) was applied.

Following affinity purification both FliI 2-91 and 19-91 were further purified by HIC using a HiPrep Butyl FF 16/10 column with a bed volume of 24 ml. The column was pre-equilibrated with five column volumes of start buffer: 50 mM Sodium phosphate, 800 mM ammonium sulphate, pH 7.0. Prior to loading the proteins on the column they were adjusted to 500 mM ammonium sulphate by slowly adding solid ammonium sulphate while stirring continuously. The sample was then loaded on the column using the superloop (capacity = 50 ml) at a flow rate of 1 ml/min, and then the flow rate was changed to 3 ml/min and any protein that bound was eluted by applying a salt gradient from 800 mM to 0 mM ammonium sulphate over five column volumes. A chromatogram from one such HIC purification of FliI 2-91 and a gel of the purified protein are shown below (Figure 3.24).

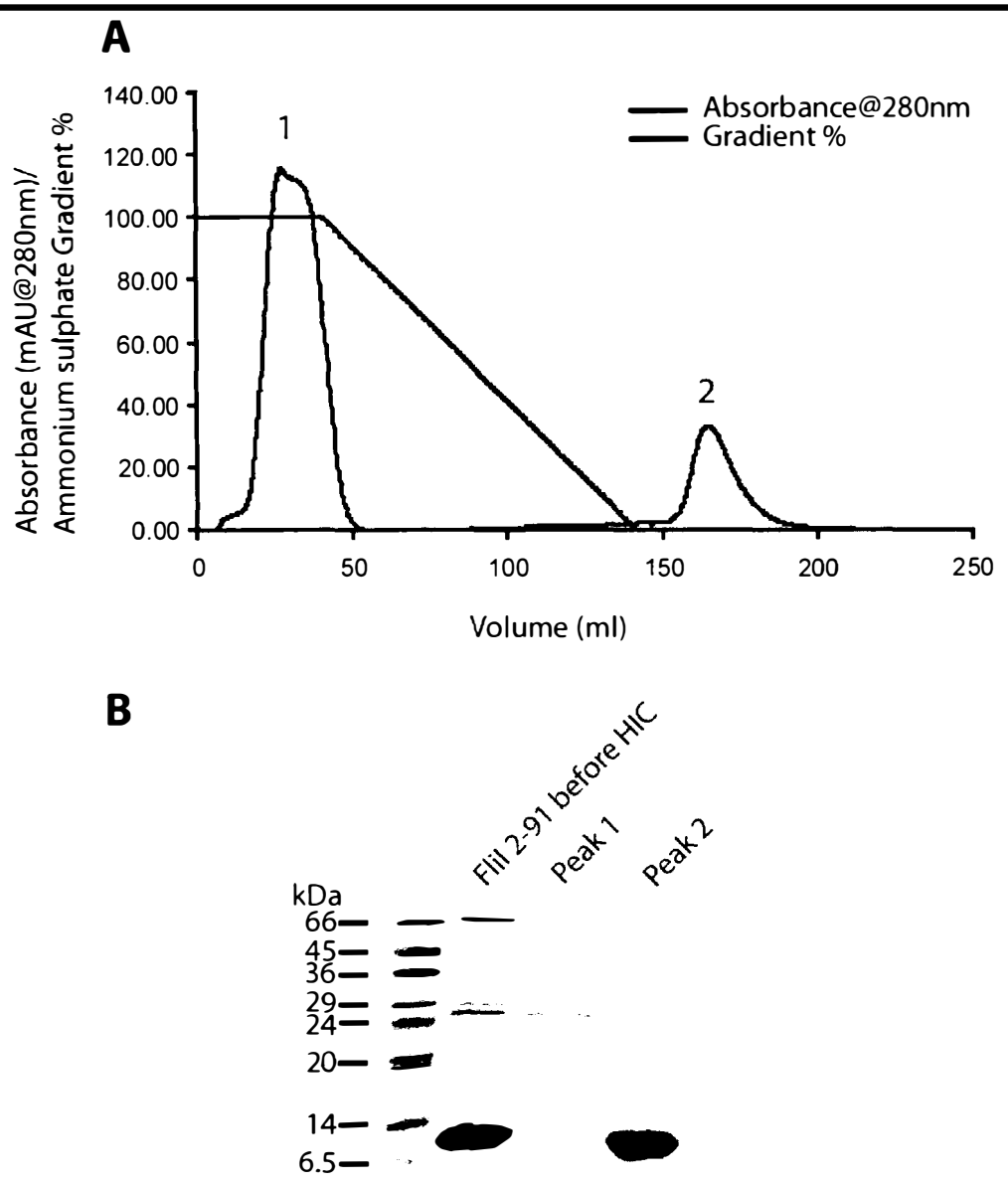


Figure 3.24 HIC Purification of FliI 2-91. A). Elution profile of FliI 2-91 from the Butyl FF 16/10 HIC column, FliI 2-91 eluted at 164 ml. The peaks that were analysed by SDS-PAGE are designated by numbers. **B).** 20% SDS-PAGE analysis of FliI 2-91 before HIC, and the peaks that eluted from the HIC column.

As can be seen from the gel above, HIC purified FliI 2-91 and FliI 19-91 (not shown) almost to homogeneity. Despite this FliI 2-91 and 19-91 were purified further by SEC using

the Superdex 75 HR 10/30 column (Figure 3.25 and 3.26), and this also allowed the size of the purified proteins to be estimated by calibrated size exclusion.

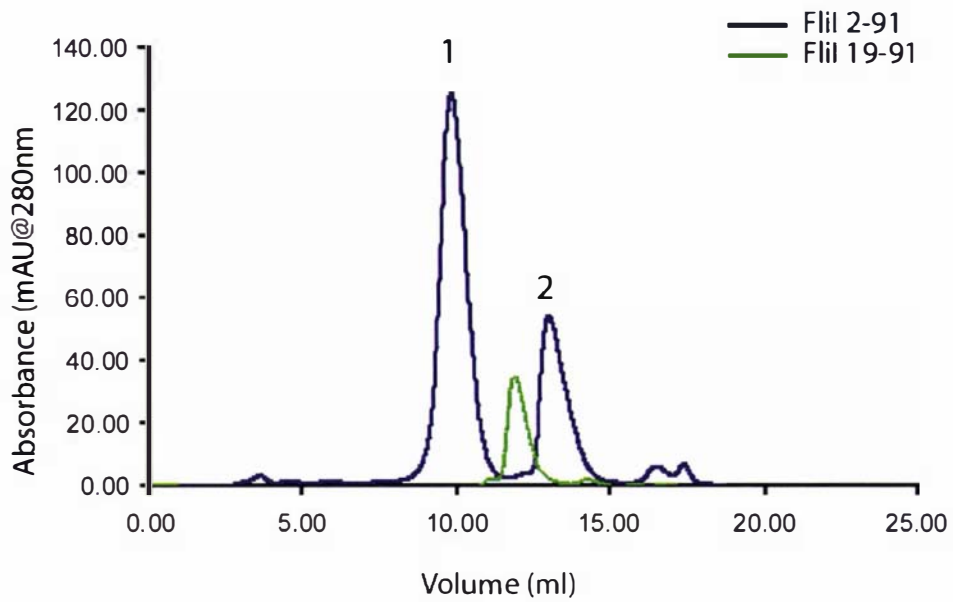
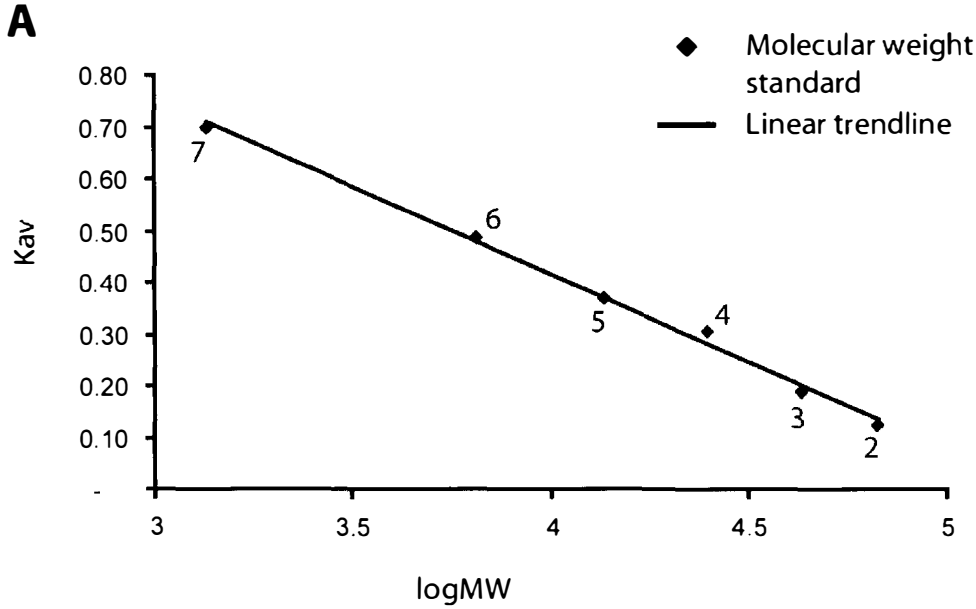


Figure 3.25 Purification of FliI 2-91 and 19-91 by Size Exclusion Chromatography using a Superdex 75 HR 10/30 gel filtration column. FliI 2-91 eluted as two peaks designated 1 and 2 on the elution profile, with elution volumes of 9.92 and 13.1 ml respectively. FliI 19-91 eluted as a single peak with an elution volume of 11.9 ml.

A



B

	V_e	V_0	V_t	K_{av}	MW	logMW
Peak 1 Blue dextran	7.84	7.84	24	-	-	-
Peak 2 Albumin	9.84	7.84	24	0.12	67,000	4.83
Peak 3 Ovalbumin	10.88	7.84	24	0.19	43,000	4.63
Peak 4 Chymotrypsinogen	12.8	7.84	24	0.31	25,000	4.40
Peak 5 Ribonuclease	13.8	7.84	24	0.37	13,700	4.14
Peak 6 Aprotinin	15.73	7.84	24	0.49	6,500	3.81
Peak 7 Vitamin B12	19.15	7.84	24	0.70	1,355	3.13

Figure 3.26 A. Calibration curve for the Superdex 75 HR 10/30 column

B. The measurements from which the calibration curve was derived

$K_{av}= V_e-V_0/V_t-V_0$ V_e= elution volume, V_0= column void volume, V_t = column volume.

The void volume is determined by the elution volume of Blue Dextran.

The equation of the line: $y=-0.3393x + 1.7729$ $R^2=0.9947$

The SDS-PAGE analysis of the protein that eluted from the size exclusion column (Figure 3.27 A and B) shows that both FliI 2-91(3.27 A) and FliI 19-91 (3.27 B) have been purified to homogeneity. FliI 19-91 elutes at a volume of 11.9 ml as a single peak, and although the FliI 19-91 has a mass of approximately is approximately 8 kDa, mass estimates based on the gel filtration elution volume suggest that the protein species is 28 kDa in mass and therefore most likely a trimer. Interestingly FliI 2-91 eluted from the size exclusion column in two peaks. These two peaks were combined before being examined by SDS-PAGE, and FliI 2-91 is the only protein evident suggesting that FliI 2-91 interacts with itself in solution. Furthermore, FliI 2-91 is approximately 10 kDa in mass based on the amino acid sequence, but estimates of the mass of the proteins in these peaks by calibrated size exclusion suggest that peak one and two contain 66 kDa and 18 kDa species respectively. Assuming peak 2 is the FliI 2-91 monomer, then peak 1 represents trimeric FliI 2-91 in solution. For reference, the predicted masses of FliI 2-91 and 19-91 oligomers are tabulated in Figure 3.28.

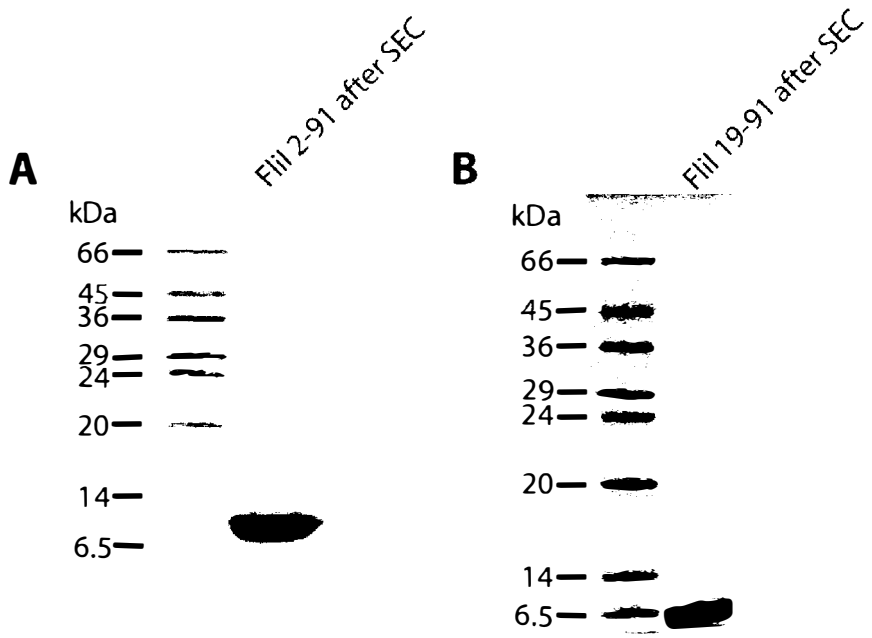


Figure 3.27 A. 20% SDS-PAGE of FliI 2-91 B. 15% SDS-PAGE of FliI 19-91 after SEC.

FliI 2-91 peaks 1 and 2 were pooled before a sample was loaded on this gel.

Oligomer	FliI 2-91 predicted molecular mass	FliI 19-91 predicted molecular mass
Monomer	18	8
Dimer	36	16
Trimer	54	24
Tetramer	72	32
Pentamer	90	40
Hexamer	108	48

Figure 3.28. Predicted molecular mass of FliI 2-91 and 19-91 oligomers based of the presumed elution volumes (predicted masses) of the monomers

To further verify the predicted size of FliI 2-91, dynamic light scattering (DLS) was also performed with this protein (Figure 3.29). Like SEC, DLS suggests that FliI 2-91 is larger than its estimated monomeric size in solution with all of the particles in solution estimated to have a molecular weight greater than 22 kDa. The DLS data also provides some evidence for a larger dimeric FliI 2-91 species in solution with approximately 40% of the particles predicted to be larger than 30 kDa. However, it is difficult to be conclusive about the oligomerisation state of FliI 2-91 in solution without further experiments to determine the shape and size of the protein such as inline multi-angle light scattering and analytical ultracentrifugation.

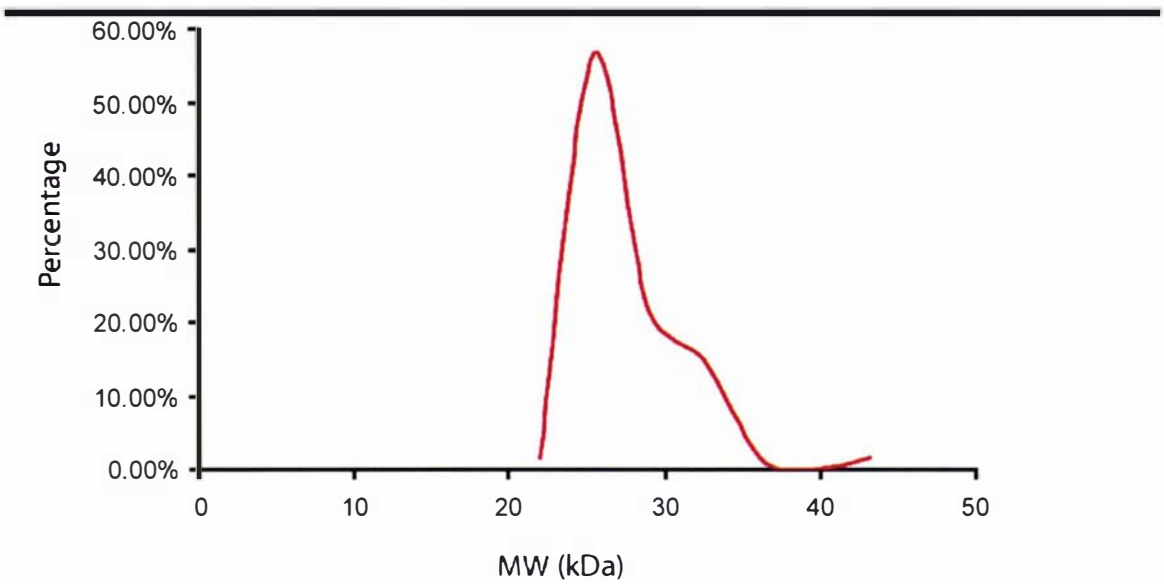
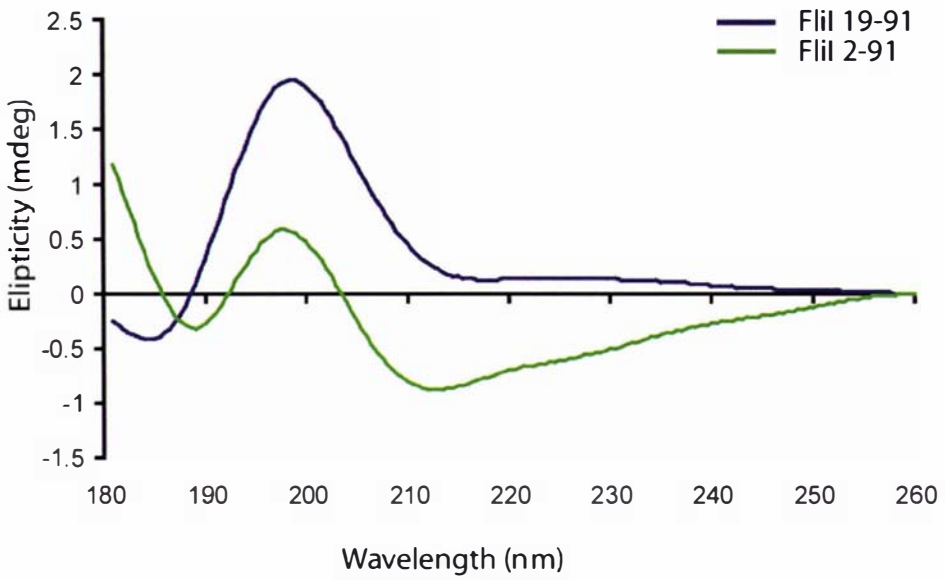


Figure 3.29 FliI 2-91 molecular mass abundance profile as predicted by DLS

To further characterise the structure of the FliI N-terminal domain and verify that FliI 2-91 and 19-91 are folded, the CD spectra of these proteins were measured (Figure 3.30 A). Based on the model of *H. pylori* FliI, if the over-expressed and purified N-terminal domain is folded correctly it should be predominantly β-sheet. The CD spectrum of both FliI 2-91 and 19-91 display a trough and peak with elipticity minimum and maximum at 210 nm and 195 nm respectively, characteristic of β-sheet (Figure 3.30 A). Furthermore, deconvolution of the CD data estimates that FliI 2-91 and 19-91 contain 52 and 58% anti-parallel β-sheet respectively (Figure 3.30 B).

However, there is considerable error in the CD deconvolution data because the model of *H. pylori* FliI predicts the FliI N-terminal domain contains approximately 45% β-sheet, and the total secondary structure predicted is greater than 100%. Furthermore, the deconvolution estimates that the the secondary structure of FliI 2-91 and 19-91 is approximately the same while there is a significant difference between the two spectra. While the elipticity of the FliI 19-91 CD spectrum remains positive between 260 and 205 nm, the elipticity of FliI 2-91 is negative and has a trough with a minimum at approximately 210 nm. Such a trough is characteristic of α-helical secondary structure, and as this trough is absent in the spectrum of FliI 19-91, may indicate that the first 19 amino acids of *H. pylori* FliI form an α-helix.

A



B

	FliI 2-91	FliI 19-91
Anti-parallel β-sheet	52%	58%
Parallel β-sheet	4%	4%
β-turn	18%	16%
Random coil	28%	26%
α-helix	9%	8%
Total %	111%	112%

Figure 3.30 A. Far-UV CD Spectrum of FliI 19-91 and 2-91
B. CDNN Deconvolution of FliI 2-91 and 19-91

To investigate the structure and folding of the N-terminal domain FliI 2-91 was also subjected to Trypsin limited proteolysis and samples of the protein analysed by SDS-PAGE at various time points during the proteolysis (Figure 3.31).

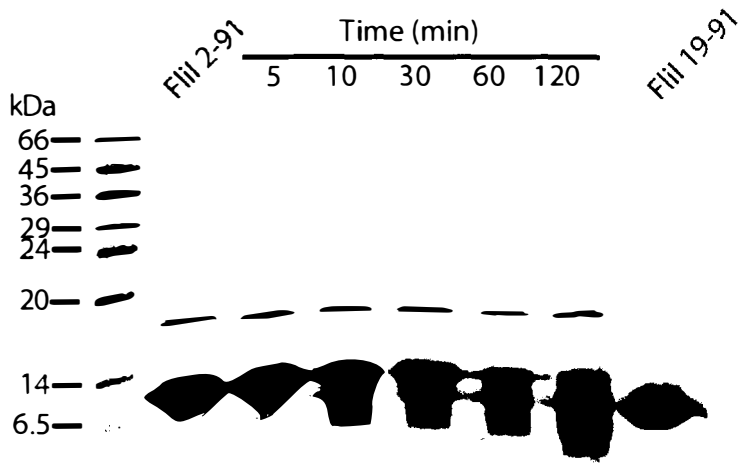


Figure 3.31 SDS-PAGE of Trypsin limited proteolysis of FliI 2-91

FliI 2-91 usually migrates with a molecular weight of approximately 10 kDa on SDS-PAGE. But after 10 minutes of Trypsin proteolysis a second band is visible by SDS-PAGE at approximately 8 kDa and corresponding in size to FliI 19-91. After 120 minutes almost all of the FliI 2-91 is proteolysed to this 8 kDa species. As residues 2-18 are thought to correspond to the disordered N-terminal extension of Bovine $F_1\alpha$ and β, and are therefore likely protease sensitive, this 8 kDa species is assumed to be the product of proteolysis of the first 17 residues of FliI 2-91 by Trypsin, resulting in FliI 19-91.

In summary, the FliI N-terminal domain can be expressed as a GST fusion, is soluble and can be purified to homogeneity. There is evidence that the purified protein is folded, contains β-sheet and may oligomerise to form trimers. However, the stoichiometry of the *H. pylori* FliI 2-91 and 19-91 oligomers await further experimental characterisation because estimates of protein mass based on size exclusion and DLS are not independent of protein shape and therefore do not correspond to the theoretical mass of the protein.

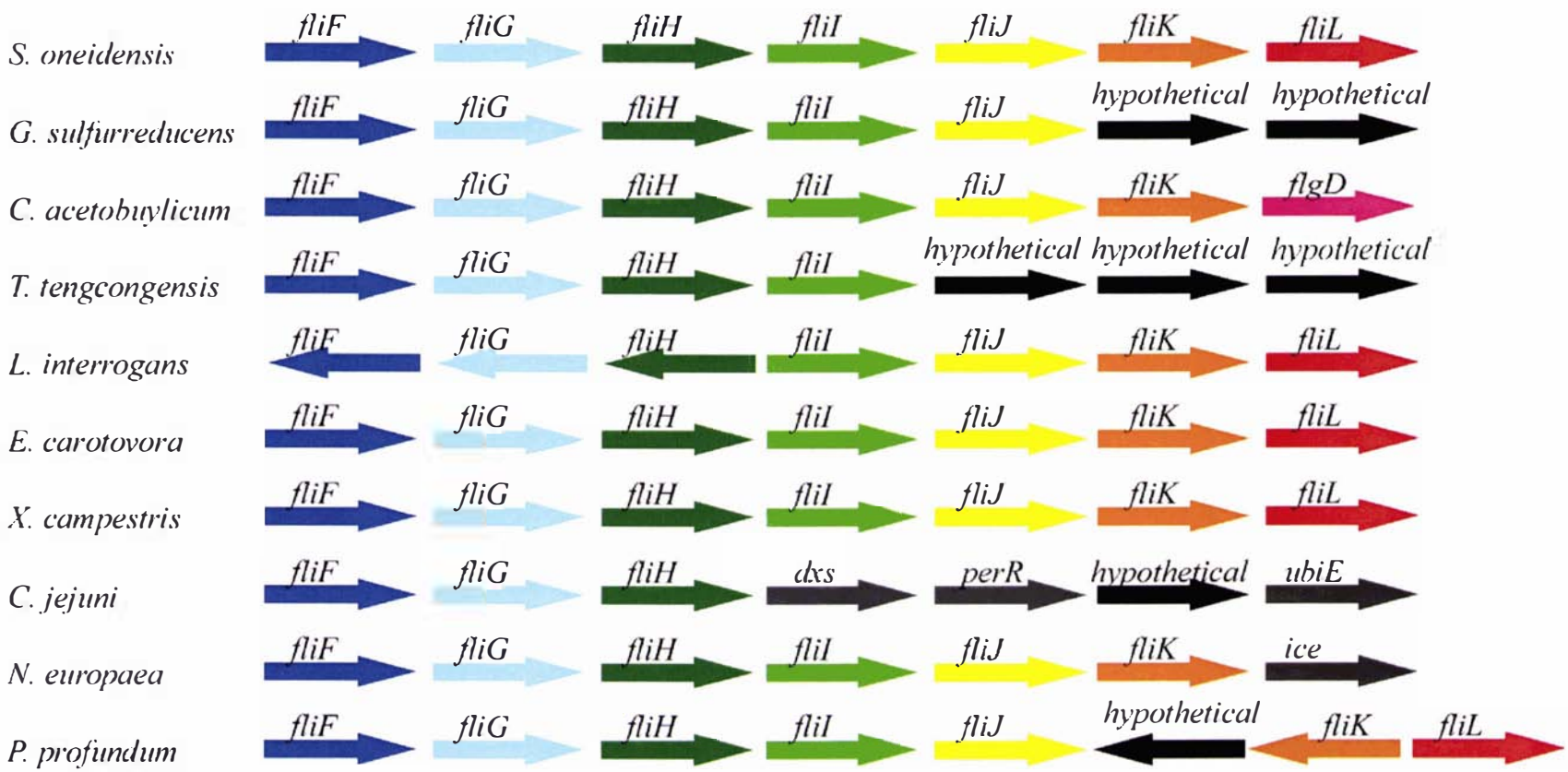
3.3 Characterisation of *H. pylori* FliH

FliH of *Salmonella typhimurium* is located in the same structure with FliI. Furthermore, genes encoding these two proteins are located next to each other in the *S. typhimurium* genome and at the flagellar loci of other bacteria (Figure 3.32).

When this project started very little was known about the structure or function of FliH. However, its importance for flagellation had been implied by non-motile and aflagellate *S. typhimurium* FliH mutants (188) (90). These observations failed to distinguish between a role for FliH in flagellar assembly or flagellar export, and subsequent experiments have shown that *S. typhimurium* FliH is required for the export of FlgE, FlgD and flagellin (117).

It is possible that FliI and FliH cooperate in flagellar protein export and assembly, and therefore it was decided to investigate the structure and function of *H. pylori* FliH and the interactions between FliI and FliH.

A multiple alignment of the *H. pylori* FliH and other FliH homologues is shown in Figure 3.33. The N-terminus of the FliH homologues is divergent in sequence and consequently difficult to align. For example, *H. pylori* has 10% identity with *S. typhimurium* FliH from residue 1 to 150. Secondary structure predictions by the PHD algorithm predict residues 1-119 of *H. pylori* FliH to be a mixture of α-helix and loop but there is considerable variation in the structural predictions for the N-terminal region of the polypeptide amongst the homologues that reflects the sequence divergence in this part of the protein. In contrast to the N-terminus, the C-terminus of FliH homologues is more highly conserved and this conservation is apparent in the multiple alignment, but the sequence identity is still low. For example *H. pylori* and *S. typhimurium* FliH share 20% identity from residue 151 to 258. Furthermore the PHD secondary structure predictions for the twenty homologues are very consistent and allow a position-specific average PHD probability to be calculated (Figure 3.33). PHD predicts residues 120-187 of *H. pylori* are α-helical and residues 188-258 are a mixture of α-helix and β-sheet.


S. oneidensis
fliF fliG fliH fliI fliJ fliK fliL
G. sulfurreducens
fliF fliG fliH fliI fliJ hypothetical hypothetical
C. acetobuylicum
fliF fliG fliH fliI fliJ fliK flgD
T. tengcongensis
fliF fliG fliH fliI hypothetical hypothetical hypothetical
L. interrogans
fliF fliG fliH fliI fliJ fliK fliL
E. carotovora
fliF fliG fliH fliI fliJ fliK fliL
X. campestris
fliF fliG fliH fliI fliJ fliK fliL
C. jejuni
fliF fliG fliH dxs perR hypothetical ubiE
N. europaea
fliF fliG fliH fliI fliJ fliK ice
P. profundum
fliF fliG fliH fliI fliJ hypothetical fliK fliL

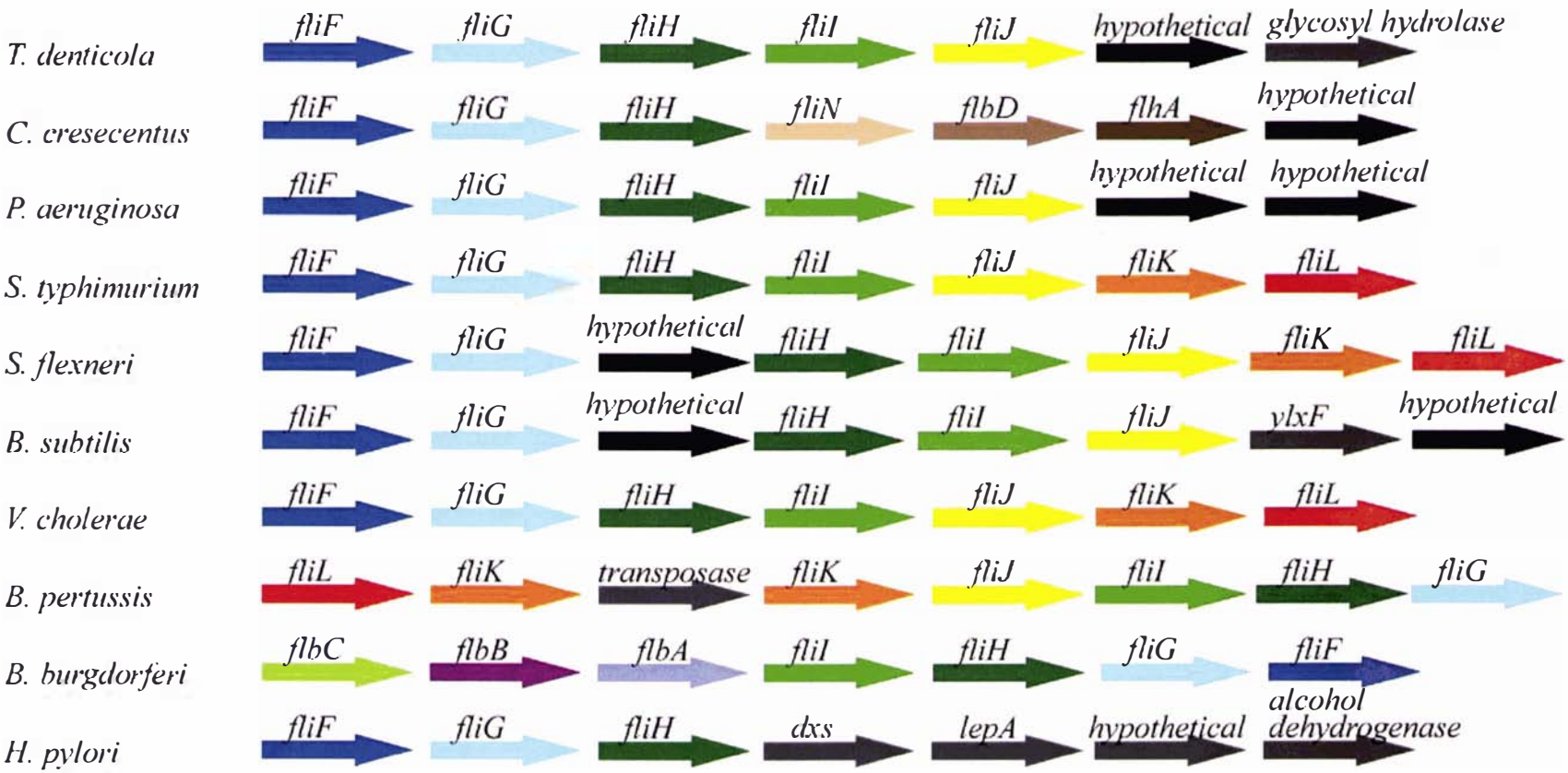
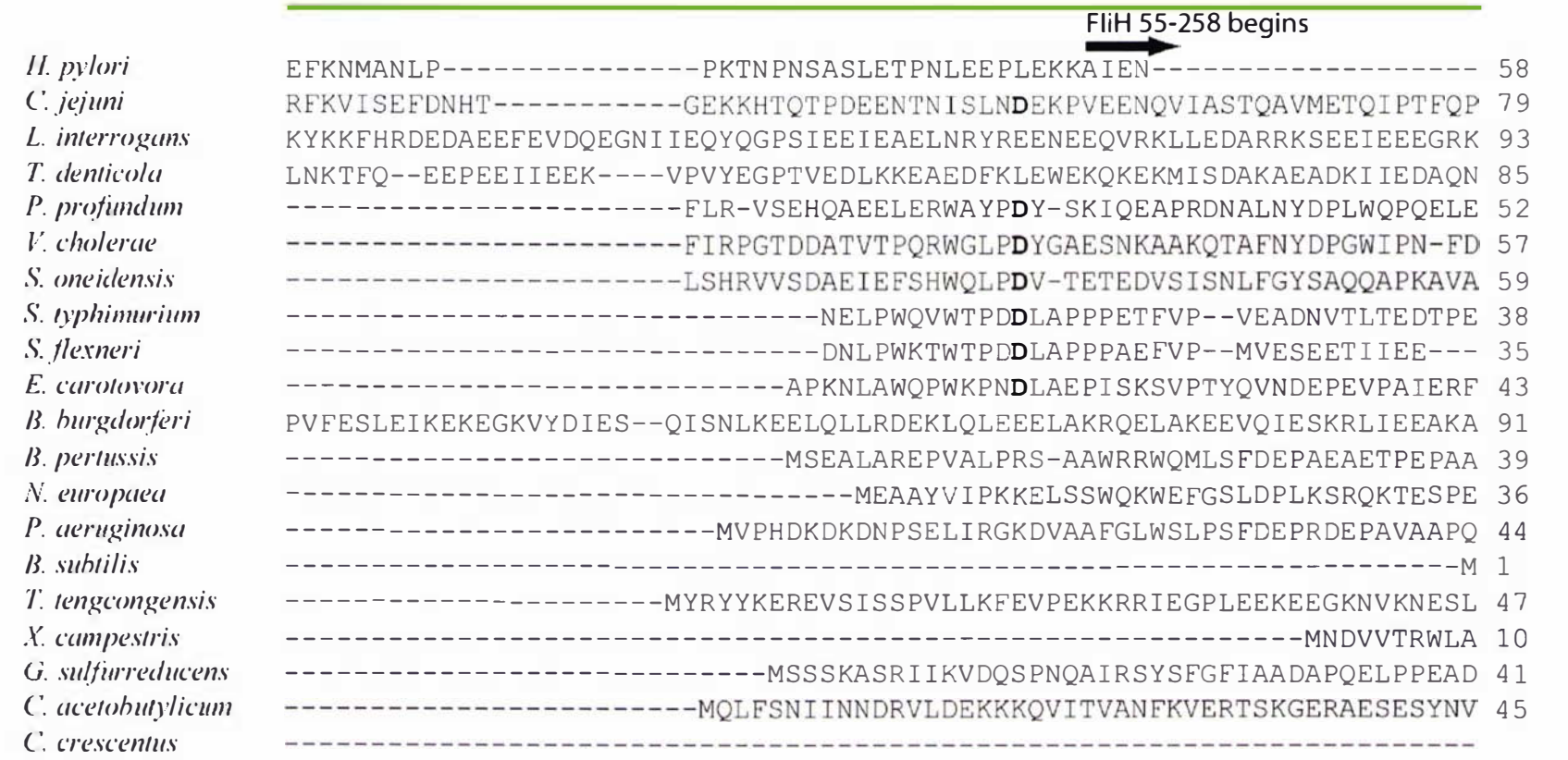


Figure 3.32 Genes surrounding FliH in twenty organisms. The arrows indicate genes in the 5'-3' direction. Black arrows represent hypothetical open reading frames. The gray arrows represent annotated genes that are not part of the flagellum or the flagellum export apparatus, *dxs* = 1-deoxyxylulose-5-phosphate synthase, *perA*= peroxide stress regulator, *ubiE* = ubiquinone/menaquinone biosynthesis methyltransferase, *ice*= nucleation protein, *ylxF*= ?, *lepA*= GTP-binding membrane protein.

Helicobacter pylori 26695	[15644981]	-MSLNSRKNLIQKDHLNKHDIQKY 23
Campylobacter jejuni subsp. *jejuni*	[15791688]	---MVNRSNVISGGASDQHVVEGY 21
Leptospira interrogans serovar Copenhageni Fiocruz L1-130	[24215289]	MAKLVFKPIQISDMKDQVELAIPD 24
Treponema denticola ATCC 35405	[41817041]	MAKTIFRGFEVN--KNNNDVVFLQ 22
Photobacterium profundum SS9	[46912544]	----------MNDDRRRG------ 8
Vibrio cholerae 01 biovar eltor str. 16961	[15642130]	------MRITMSGERKRG------ 12
Shewanella oneidensis MR-1	[24374738]	----MSSSNRSDNQLDDK------ 14
Salmonella typhimurium LT2	[16420506]	----------MS----------- 2
Shigella flexneri 2a str. 301	[56480006]	----------MS----------- 2
Erwinia carotovora subsp. atroseptica	[49611183]	----------MSN---------- 3
Borellia burgdorferi B31	[15594634]	MPKVLYKSSEVDNLVKFEFVEIAK 24
Bordetella pertussis Tohama I	[33572146]	------------------------
Nitrosomonas europaea ATCC 19718	[30139132]	------------------------
Pseudomonas aeruginosa PA01	[15596300]	------------------------
Bacillus subtilis subsp. *subtilis* str. 168	[2633995]	------------------------
Thermoanaerobacter tengcongensis MB4 (T)	[20807887]	------------------------
Xanthomonas campestris pv. *campestris* str. ATCC 35405	[21113035]	------------------------
Geobacter sulfurreducens PCA	[39995520]	------------------------
Clostridium acetobutylicum ATCC 824	[15025153]	------------------------
Caulobacter crescentus CB15	[16125159]	------------------------


FliH 55-258 begins
H. pylori EFKNMANLP----------PKTNPNSASLETPNLEEPLEKKAIEN---------- 58
C. jejuni RFKVISEFDNHT----------GEKKHTQTPDEENTNISLNDEKPVEENQVIASTQAVMETQIPTFQP 79
L. interrogans KYKKFHRDEDAEEFEVDQEGNIIEQYQGPSIEEIEAELNRYREENEEQVRKLLEDARRKSEEIEEEGRK 93
T. denticola LNKTFQ--EEPEEIIEEK----VPVYEGPTVEDLKKEAEDFKLEWEKQKEKMISDAKAEADKIIEDAQN 85
P. profundum ----------FLR-VSEHQAEELERWAYPDY-SKIQEAPRDNALNYDPLWQPQELE 52
V. cholerae ----------FIRPGTDDATVTPQRWGLPDYGAESNKAAKQTAFNYDPGWIPN-FD 57
S. oneidensis ----------LSHRVVSDAEIEFSHWQLPDV-TETEDVSISNLFGYSAQQAPKAVA 59
S. typhimurium ----------NELPWQVWTPDDLAPPPETFVP--VEADNVTLTEDTPE 38
S. flexneri ----------DNLPWKTWTPDDLAPPPAEFVP--MVESEETIIEE--- 35
E. carotovora ----------APKNLAWQPWKPNDLAEPISKSVPTYQVNDEPEVPAIERF 43
B. burgdorferi PVFESLEIKEKEGKVYDIES--QISNLKEELQLLRDEKLQLEEELAKRQELAKEEVQIESKRLIEEAKA 91
B. pertussis ----------MSEALAREPVALPRS-AAWRRWQMLSFDEPAEAETPEPAA 39
N. europaea ----------MEAAYVIPKKELSSWQKWEFGSLDPLKSRQKTESPE 36
P. aeruginosa ----------MVPHDKDKDNPSELIRGKDVAAFGLWSLPSFDEPRDEPAVAAPQ 44
B. subtilis ----------M 1
T. tengcongensis ----------MYRYYKEREVSISSPVLLKFEVPEKKRRIEGPLEEKEEGKNVKNESL 47
X. campestris ----------MNDVVTRWLA 10
G. sulfurreducens ----------MSSSKASRIIKVDQSPNQAIRSYSFGFIAADAPQELPPEAD 41
C. acetobutylicum ----------MQLFSNIINNDRVLDEKKKQVITVANFKVERTSKGERAESESYNV 45
C. crescentus ----------

FliH 94-258 FliH 117-258

```
H. pylori              DLIDCLLKKT-DELSSHLVKLQMQFEKAQEESKVLIENAKNDG-------YKIGFKEGEEKMRNELTHS 119
C. jejuni              SFVEDLLKKT-DEMSSNIIKLQMQIESQENEFNNRLNSELENA-------KEKFTKEGYEKAKEEFQKE 140
L. interrogans         RAFQMIQDSK-EKVKLEEDTGKAKAEQILDRAKMEVERMIKEAEMKVAEIEHEAYLKGYDAGREVGFKK 161
T. denticola           AAFDEVKRQT-DEAQVIAQNAKKDAEDIIAEAEQKARDIIADSEKNKDSVNRDAYKEGFNRGREEGFKE 153
P. profundum           EEEEQGPPPL-TAADLEHIRQSASDEGFNEGKVTGHTEGFDVGKIEG---LEAGHAEGLEQGLAQGLEQ 117
V. cholerae            EPEQVVEHEF-SEEEIALIRTAAQQEGFEAGQAEGYQQGFEQGKAEG---FQAGHQEGQTQGYQDGVAE 122
S. oneidensis          VETVAPP----TMAEIEDIRAQAEEEGFNEGKTQGHIEGLEQGRLEG---LEQGHKEGFTQGHEQGLET 121
S. typhimurium         PELTAEQQL---EQELAQLKIQAHEQGYNAGLAEGRQKGHAQG----------YQEGLAQGLEQGQAQ 93
S. flexneri            ----AEPSL---EQQLAQLQMQAHEQGYQAGIAEGRQQGHEQG----------YQEGLAQGLEQGLAE 86
E. carotovora          NEENAVPTL---EDELASLRESMMQQARETGFSQGHQQGYDAG----------YQEGLAKGQQQGLQN 98
B. burgdorferi         KANEVLEAAK-QEADLLQKEAIYKKESIETESNAEIERLAREYEEKLKTDLEIAIAKGREEGYSKGYES 159
B. pertussis           LEPEPEPAPD-PEELLREWRATAERAGHAAGHQAGQEQGQREGYAAG---HAQGLAAGRAEGHAEGLAQ 104
N. europaea            KTPHTARSADQPENQIAAGAKTTAHEAIVLPTAEQIEQIYQQAR----------EEGKTAGYQEGMQQ 94
P. aeruginosa          VPAVAEPAPA--PPAVEEVELETVKPPTLEEIEAIRQDAYNEG-------FATGERDGFHAGQLKARQE 104
B. subtilis            ARVKEEADRISEQANSHIENIRRQIEQEKNDWAAEKQKLIEEAK----------AEGFEQGVALGKAE 59
T. tengcongensis       TLAKEMIERARQVQREILSKTREDIEKMLKEAEERAHEIEEKYRIKG-------YEEGYRAGYEEGYKK 109
X. campestris          PDLHMAPALP-----EAEFDEPAVYEPVLRPPTLEEIQAIEDAAQ----------QEGFARGHAEGFAQ 64
G. sulfurreducens      GFVPFALGTPVPLPGLQSAFEEPDPDPVVPFNLEGKVVLAEDELQAR--------VDEVFRNGMDEGRRQ 102
C. acetobutylicum      IGEAIIKKARHEAEYIKNKALEESREFYKKAYDEGLEEGRKKG----------YEDAYNETVVKGNME 103
C. crescentus          --MTDIPHRKFAFDTVFDDHGGVAYTAPRVKKSFTPEEVEAAK----------AQAYAEGERSALVR 55
```


99*******************99*99999*******************999 9**9**********9
H. pylori VNEEKNQLLHAITTLD-EKMKKSEDHLMALEKELSAIAIDIAKEVILKEVEDNSQKVALALAEELLKNV 187
C. jejuni LSDFKDKYLKSIAKLD-NACENLENFIEKNEKELADTAIDIAKEVILKELELNSSKIAYALAKDLIGEL 208
L. interrogans GQGEVRRLIDRLGTIVGKAIDIREEIIQASEKQMVEMILIIARKVIKDEII-ERKEIVLNNIREALKRI 229
T. denticola GNLEVQRLTDRLHTIINKTMDRRQEILSETEQQIVDLVLLMTRKVVKVISE-NQRNVVVSNVVHALRKV 221
P. profundum GQAQITEHVSHLTQLVEKLANPLKQLDAEVESQVLTLVTSLTRELIRVEVQTNPQVMLNTIREVIATLP 186
V. cholerae GQALIQEQVKTFMALANQFAQPLDLLNAQVEKQLVDMVLALTKEVVHVEVQTNPQVILDTVKASVEALP 191
S. oneidensis GLAEAKALLSRFEGLLCQFEKPLQLLDGDIEHTLMSLTMALAKSVIGHELKTHPEQILSALRLGVESLP 190
S. typhimurium AQTQQAPIHARMQQLVSEFQNTLDALDSVIASRLMQMALEAARQVIGQTPAVDNSALIKQIQQLLQQEP 162
S. flexneri AKAQQAPIHARMQQLVSEFQTTLDALDSVIASRLMQMALEAARQVIGQTPTVDNSALIKQIQQLLQQEP 155
E. carotovora SLQQQQPIIEQMQQMVTEFQQTLDTLDSVIPARLMQLALTAAKQILGQPPVCDGNALLGQIQQLLQQEP 167
B. burgdorferi GFEDFDKVMRKLHVIIASLIAERKGILESSSGQIVSLVMQIAIKVIKRITD-SQKDIVLENVNEVLKRV 227
B. pertussis GREDARQQAERLHALAQACAASVARLEDNMGQSLLTLALDIAGQVLRTTLAEHPEAMVAAVREVLQINP 173
N. europaea AKHAALVEVKHLQSLTGALEQELKQIDQTMAQDLLTLAIDLARKITSHALEIKPELILPVVEEALRQLP 163
P. aeruginosa AEEALKERLQSLERLMTQLLEPIAEQDALIEQGMVNLVNHVARQVIQRELHMDSSHVRQVLREALKLLP 173
B. subtilis AMKQYAELIGQANTITEMSRKAVEDKLEDANEEIVELAVALAKKVWQ--QKSDDKEAFLLLVQQVINEV 126
T. tengcongensis GEEKAQALIEEARALKEEILEEKKRLYKEAESDMINVILEAVEKIVGKHVE-EDKDLILSLIKKGMENY 177
X. campestris GQSEVRRLTAQIDGILDNFTRPLARLENEVVGALGELAVRIAGSLVGRAYQAEPQLLADLVQEAIDAVG 133
G. sulfurreducens AERGLANVFKSLRDGVAALTGLRSRVMKESEEDLLRLAVMIARKIVQREVAQDPQVLAAIVAAAVGG-C 170
C. acetobutylicum IENLKSIAEQNVANMIKSAKFEVQEYFNDMEAKIKSLSIEIAQHILKERIKEADG--IDNMIYEALEIC 170
C. crescentus TEQEAAQALAEVAHAVQQAFGTLAHVAHEHREGSAMLALACGRKIADAALTHFPEAPVTAALEALAREV 124

```
                 89 9***8    9****999*99 *        898  8 8888  998 88       89*******
H. pylori        LDATDIHLKVNPLDYPYLNERLQNAS-------KIKLESNEAISKGGVMITSSNGSLDGNLMERFKTLK 249
C. jejuni        KGASAIELKVNAEDYEYLKEQFDQNA-------HIKISLDDAISKGSVVIISDAGNIESNLNSRLTKIK 270
L. interrogans   KDRDRVDIRVNFADLELTTAHKDELIKLMESLRKVNIYEDSRVDRGGVIIETDVGAIDARISTQLKEIE 298
T. denticola     KGRGDVVIRVNLADVKMTTEHIQNFISAAENIKNITVVEDSTVDQGGCIIETDFGAVDARIASQLNELE 290
P. profundum     IAERQIRLSLHPDDLENVRQAYGDENLVER---KWTLIGEPSLNRGDLQVQSGDSTVDYYIDDRIRHLL 252
V. cholerae      IAGHAITLKLNPEDVEIIRQAYGEQEIETR---NWTLLSEPALSRGDVQIEAGESSVSYRMEERIRSVL 257
S. oneidensis    IKEQTVNLRMHPDDVTLVETLYSSTQLTRN---QWQLEADPSLTAGDCIISSQRSLVDLTLSSRIDAVF 256
S. typhimurium   LFSGKPQLRVHPDDLQRVEEMLG-ATLSLH---GWRLRGDPTLHHGGCKVSADEGDLDASVATRWQELC 227
S. flexneri      LFSGKPQLRVHPDDLQRVDDMLG-ATLSLH---GWRLRGDPTLHPGGCKVSADEGDLDASVATRWQELC 220
E. carotovora    MFSGKTQLRVHPSDLERVEQYLG-PTLSLH---GWRLLADSQLHPGGCKVSAEEGDLDASLATRWHELC 232
B. burgdorferi   KDKTQITIRVNLDDLDIVRHKKSDFISRFDIIENLEIIEDPNIGKGGCIIETNFGEIDARISSQLDKIE 296
B. pertussis     ATGAAMRLWVHPDDLELVRTHLADELNEAN----WRLQADESITRGGCRTETAYGDVDATLQTRWRRVA 238
N. europaea      AVSQSIRLTLHPDDAARIRDHLENHPAHPK----WHIYEDTQIEPGGCRIESGGCEVDATLATRWQRTL 228
P. aeruginosa    MGAANIRIHVNPQDFERVKALRERHEES------WRILEDDSLLPGGCRIETEHSRIDATIETRLAQAV 236
B. subtilis      KEYDDISIYVDPYYYETIFQQKDEIQQLLYKECRLGIYADEKAQKGTCYIETPFGRVDASVDTQLMQLK 195
T. tengcongensis NAFDKVTVRVSEEDYEHCVKNKDKILKDVEFLDEINIVKDLSLRKGDCIIETNSGVINSGVTTQLKTLK 246
X. campestris    SAGREVEVRLHPDDITALLPHLATSSTT-------RVAPDLTLSRGDLRVHAESVRVDGTLDARLRAAL 195
G. sulfurreducens TERDRVVVRLNPDDYTQVSANRQAFLAGLGEESAITLAPDESIGPGGCLVETATGTVDARIEAQLDEIY 239
C. acetobutylicum KKSKTIIIKCSSNHSEGIRKSLDNWKSTLPYRGDFFVVDDGYMDEGSAIIESDSGKVKVSIDDALEKIK 239
C. crescentus     EAQPRIFVRVSPELEERTQQALENVAAQIGFQGQIVARADGAMAPAAFTFDWGDGRAAFDPDGAAQRVA 193
```

```
          **9
H. pylori         ESVLENFKV---------------------------------------------- 258
C. jejuni         KMVNNE------------------------------------------------- 276
L. interrogans    EAIRNAEPI---------------------------------------------- 307
T. denticola      QKILEISPIKTKIKTGNI-------------------------------------- 308
P. profundum      EQFSGVNSQR-GEAL----------------------------------------- 266
V. cholerae       KSFCGINRHHPGE------------------------------------------- 270
S. oneidensis     ESLRNQQSHLALQQQQRQAALDEENAAKIVPESIKDADELVAQYQGDGEQDAQSPPSTAE 316
S. typhimurium    RLAAPGVL----------------------------------------------- 235
S. flexneri       RLAAPGVV----------------------------------------------- 228
E. carotovora     RLAAPGEL----------------------------------------------- 240
B. burgdorferi    EKFKNFSLLS--------------------------------------------- 306
B. pertussis      ASLGRSAAWEETA------------------------------------------ 251
N. europaea       AVLGQEQPWLV-------------------------------------------- 239
P. aeruginosa     KQLFEQQREQATHPLASDIRIDLDAPGDVDAP----------------------- 268
B. subtilis       DKLLTALEAGAAE------------------------------------------ 208
T. tengcongensis  SLFAGVLNE---------------------------------------------- 255
X. campestris     ETVMRKSGAGL-------------------------------------------- 206
G. sulfurreducens RSLLEERSAPVEPSASPDTDSRADLAFGGEETIAPFKGQGAWVKGSEEKPRDDV------ 293
C. acetobutylicum EVLLKTE------------------------------------------------ 246
C. crescentus     EALEAAIAAEGLHAEPLFT------------------------------------ 212
```

Figure 3.33 Multiple alignment of FliH homologues. The NCBI-GI numbers are listed beside the organism names in the first panel. The predicted secondary structure of *H. pylori* FliH is indicated above the alignment, loop is indicated by the green lines, α-helix by the red boxes and β-strand by the wavy blue lines. The coloured numbers indicate the average PHD probability for that position. The start positions of the expressed truncated proteins are shown.

To begin the characterisation of FliH, the encoding sequence was cloned in pGEX-6P3 and affinity purified as discussed previously (section 3.2.3). Limited proteolysis by another member of the Moore lab revealed the protein was protease sensitive. Consequently, a truncated form of FliH, FliH 55-258 was cloned as a GST fusion in pGEX-6P3 by the Moore lab. Although both full-length FliH and FliH 55-258 were soluble and could be affinity purified and shown to interact with FliI, it was discovered that these recombinant proteins elute in the void volume when purified by size exclusion chromatography. A size exclusion profile for FliH 55-258 is shown in Figure 3.34.

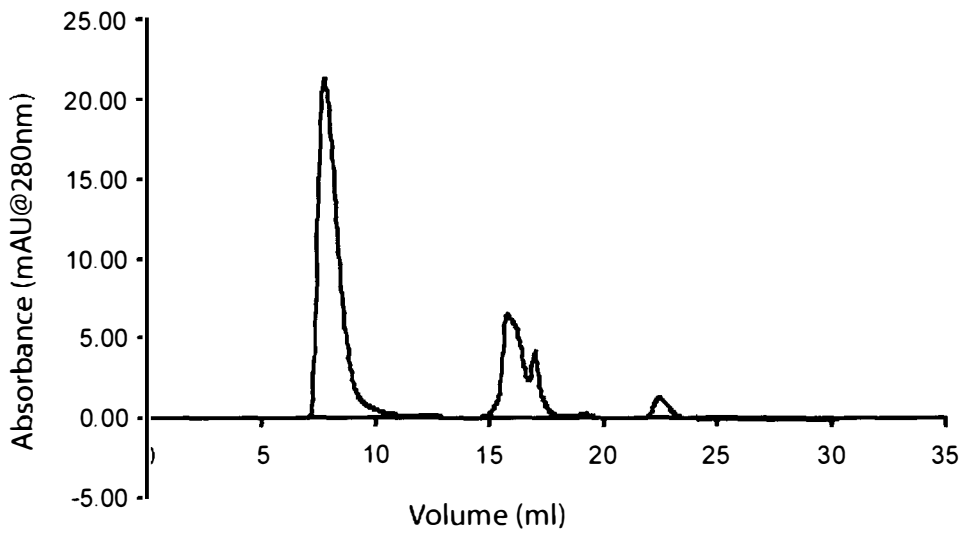


Figure 3.34 Elution profile of FliH 55-258 on Superdex 200

As the exclusion limit of the column is 200 kDa the anomalous elution behaviour of full-length FliH and FliH 55-258 suggests the proteins are at least 200 kDa and probably aggregating as the monomeric proteins have predicted masses of 29 kDa and 23 kDa respectively. The aggregation of FliH 55-258 was further illustrated by DLS (Figure 3.35), which allows a size estimate of 900-1600 kDa. Therefore, FliH and FliH 55-258 expressed and purified in this manner were not useful for further study.

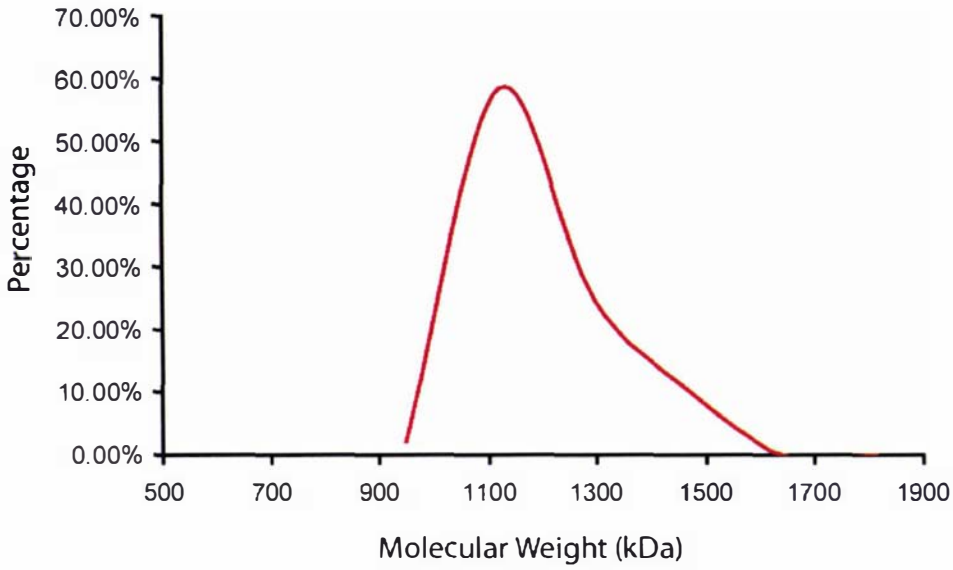


Figure 3.35 DLS profile of FliH 55-258

In an attempt to purify FliH in a form that is functional and physiologically relevant FliH 94-258 and FliH 117-258 truncated mutants were cloned as GST fusions in pGEX-6P3. These fusion proteins were expressed in *E. coli* as previously outlined, and at least 50% of the over-expressed recombinant protein was found to be in the soluble fraction of the cell lysates (Figure 3.36).

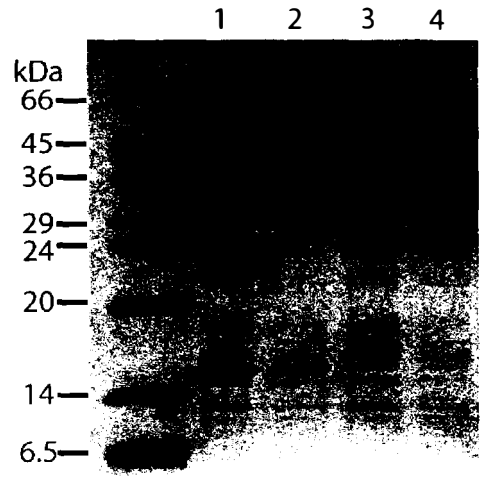


Figure 3.36 Analysis of GST-FliH 94-258 and 117-258 expression by SDS PAGE on a 15% gel.

Lane 1, GST-FliH 94-258 lysate supernatant.

Lane 2, GST-FliH 94-258 lysate pellet.

Lane 3, GST-FliH 117-258 lysate supematant.

Lane 4, GST-FliH 117-258 lysate pellet.

Following expression in *E. coli* the GST-FliH 94-258 and 117-258 were purified by affinity chromatography via the GST tag as outlined previously for GST-FliI fusions (section 3.2.3). The FliH proteins were eluted from glutathione sepharose using reduced glutathione, and the GST tag was cleaved by the addition of PreScission protease and removed by reapplication of the cleaved mixture to a glutathione sepharose column (Figure 3.37).

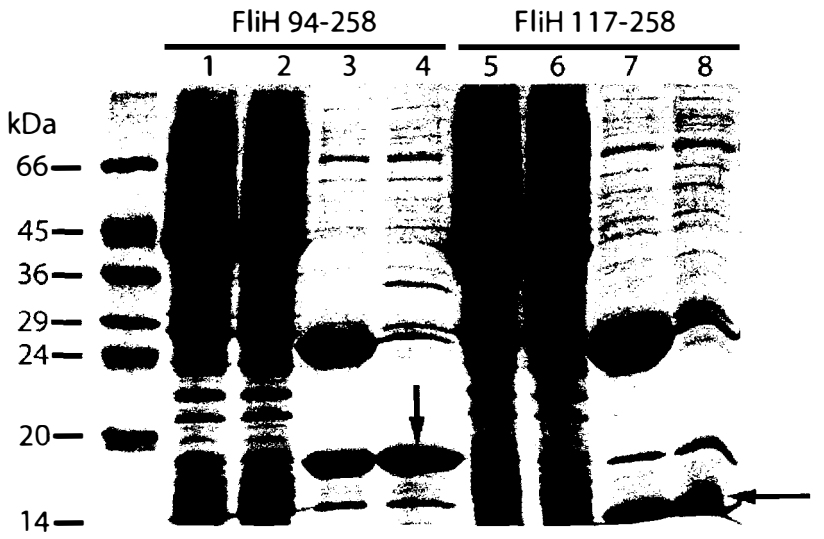


Figure 3.37 Analysis of FliH 94-258 and 117-258 affinity purification by SDS-PAGE on a 15% gel

Lane 1. FliH 94-258, lysate supernatant before purification

Lane 2. FliH 94-258, lysate supernatant flow-through after 10 ml of lysate was applied
Lane 3. FliH 94-258, after glutathione elution and cleavage of the GST tag
Lane 4. FliH 94-258, after the removal of free GST.
Lane 5-8. The same as the first four lanes but with FliH 117-258.

After the affinity chromatography both FliH 94-258 and 117-258 were purified further using the same methods employed for FliI with essentially the same results. Firstly, both proteins were subjected to Source-Q anion exchange chromatography. The chromatogram for FliH 94-258 is shown in Figure 3.38 A, and the SDS-PAGE analysis of the Source-Q peaks for FliH 94-258 and 117-258 are shown in Figure 3.38 B and C respectively.

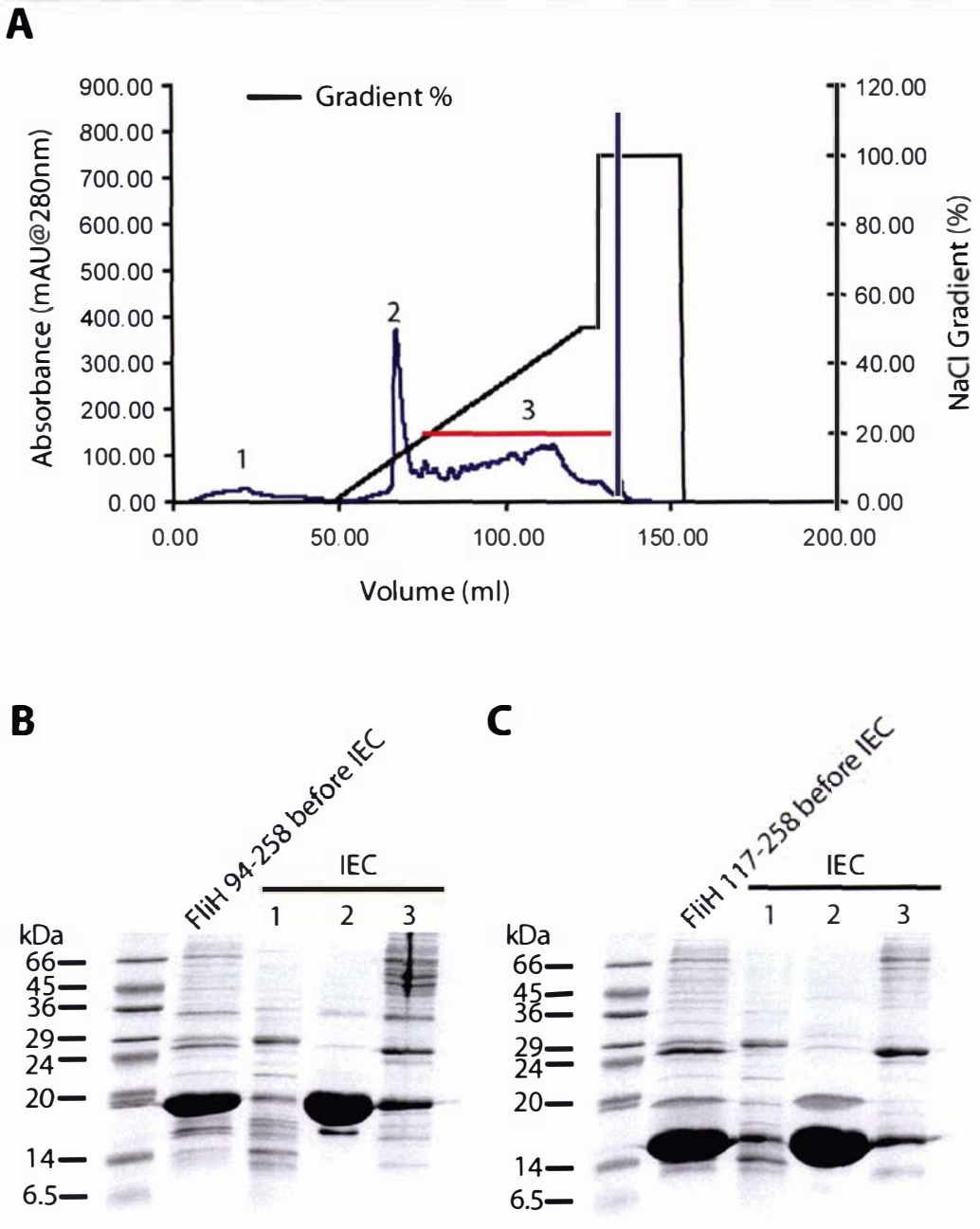


Figure 3.38 Source-Q Ion Exchange Chromatography (IEC) purification of FliH 94-258 and 117-258. A). FliH 94-258 elution profile from Source Q. Protein elution is monitored by Absorbance (mAu) at 280 nm, the peaks examined by SDS-PAGE are indicated with the appropriate numbers. **B and C).** 15% SDS-PAGE of IEC peaks for FliH 94-258 and 117-258 respectively.

As can be seen from the above SDS-PAGE analysis a significant degree of purification was achieved by Source-Q ion exchange for both FliH 94-258 and 117-258, but there were still some protein impurities and it was desirable to remove these. To this end FliH 94-258 and 117-258 were subsequently subjected to hydrophobic interaction chromatography (HIC) (Figure 3.39).

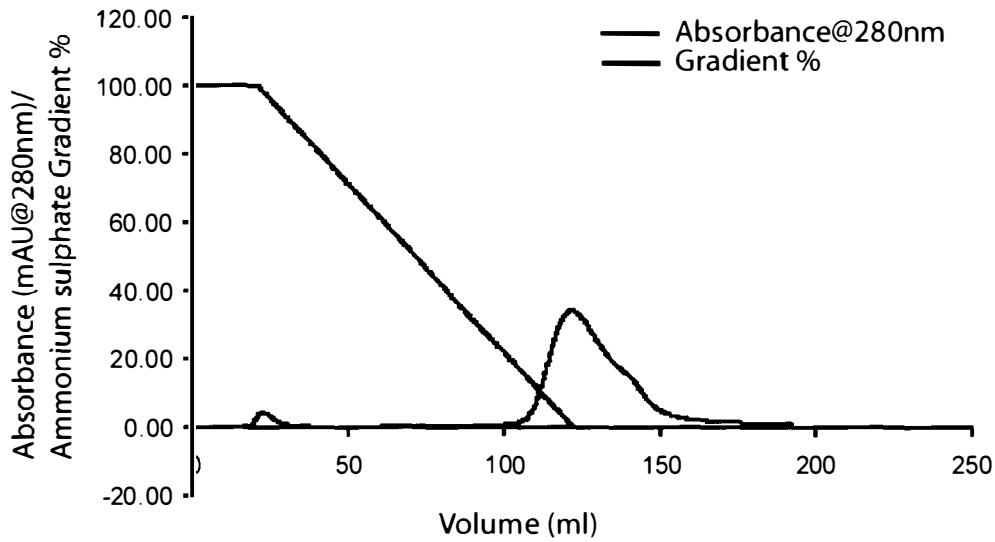


Figure 3.39 Hydrophobic interaction chromatography (HIC) purification of FliH 94-258 and 117-258. Elution profile of FliH 94-258 from the Butyl FF 16/10 HIC column.

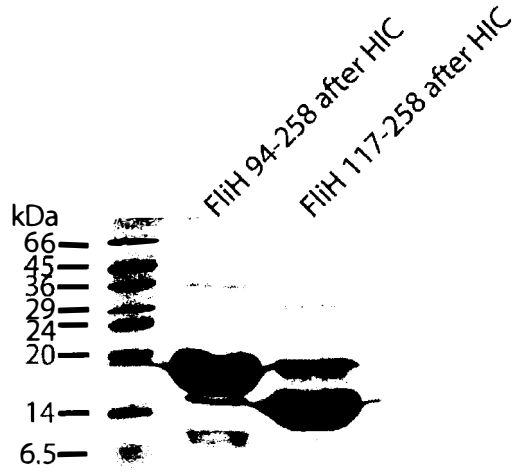


Figure 3.40 20% SDS-PAGE of FliH 94-258 and 117-258 after HIC

Although HIC did remove some of the protein impurities there are still some major contaminants in both FliH 94-258 and 117-258 that are visible by SDS-PAGE (Figure 3.40). Consequently, to attempt further purification of FliH 94-258 and 117-258, both proteins were subjected to size exclusion chromatography using the Superdex 200 column, the combined elution profiles of the two proteins are shown in Figure 3.41 below.

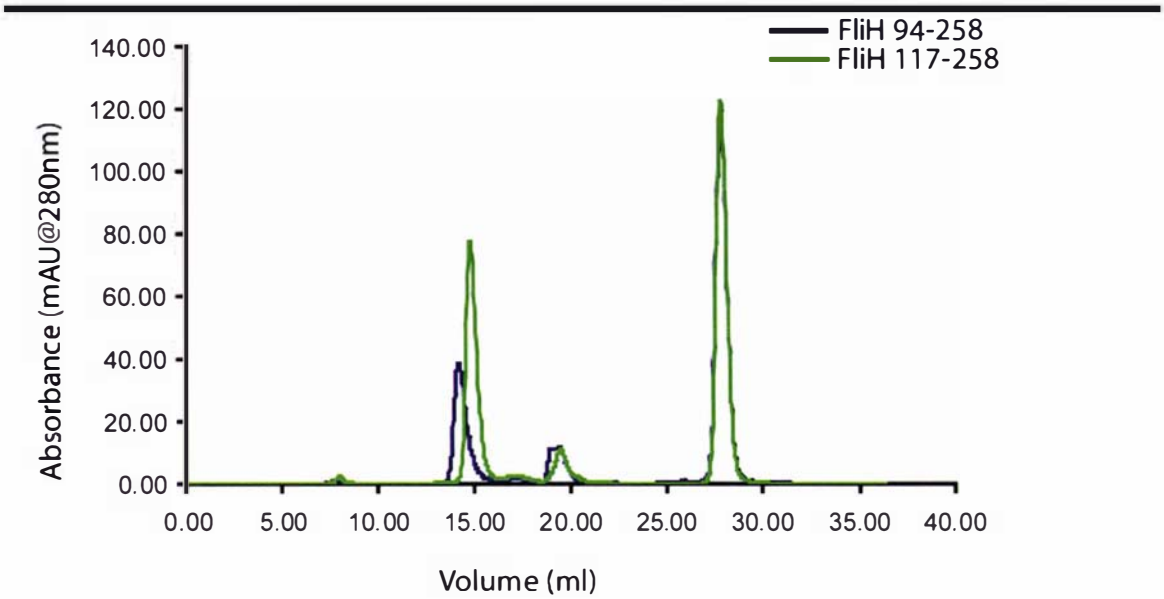


Figure 3.41 Superdex 200 elution profiles of FliH 94-258 and 117-258

The size exclusion chromatography did not improve the purity of FliH 94-258 or 117-258 significantly, but it did allow an estimation of the masses of these proteins by calibrated size exclusion. Based on the calibration curve in Figure 3.19, FliH 94-258 and 117-258 have predicted molecular masses of 60 kDa and 47 kDa respectively. As monomeric FliH 94-258 and 117-258 are predicted to have masses of 17 kDa and 15 kDa respectively, the size exclusion data suggest that FliH is not monomeric in solution and is at least a dimer and possibly forms trimers.

To investigate the structure of FliH and verify that the purified recombinant proteins are folded correctly (i.e. that they are not in a random coil conformation), far UV CD spectra were measured for both FliH 94-258 and 117-258 (Figure 3.42).

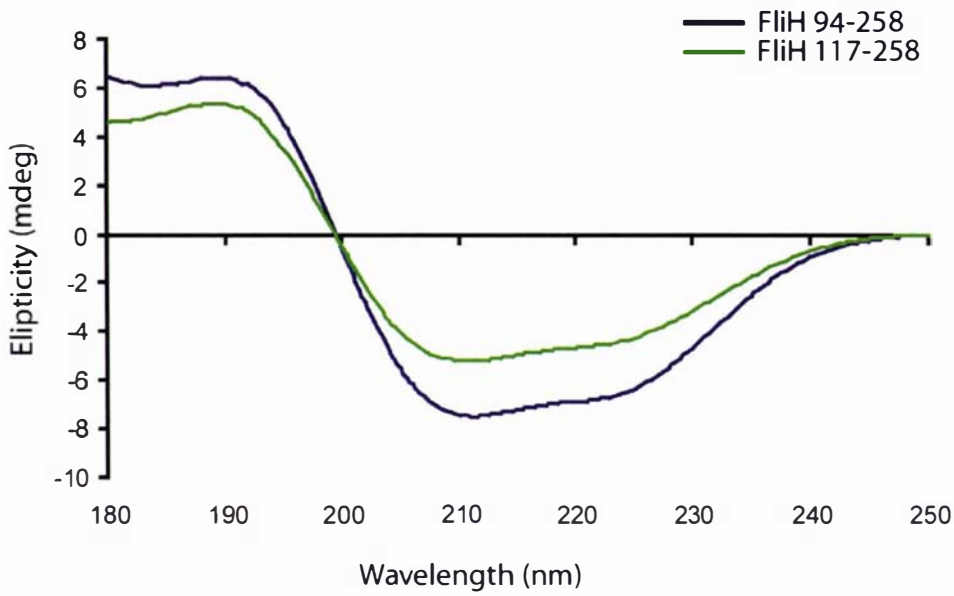


Figure 3.42 Far-UV CD Spectrum of FliH 94-258 and 117-258. The proteins were added to degassed water to a final concentration of 0.5 mg/ml in a quartz cuvette with a pathlength of 0.1 mm. The spectrum was obtained by five scans of the sample at 0.5 nm intervals and then averaged and smoothed using adjacent averaging.

Neither of the CD spectra illustrates the characteristic spectral features of a random coil polypeptide, that is, a peak of positive elipticity at approximately 220 nm and a trough with negative elipticity with a minimum at 195 nm. This indicates that FliH 94-258 and 117-258 are indeed folded in solution. Furthermore, the spectra exhibit the characteristic features of α-helical secondary structure- a trough of negative elipticity with two minima at 222.5 nm and 210 nm, and a peak of positive elipticity at 195 nm. These features of the CD spectra are in agreement with the predominantly α-helical secondary structure predicted by the PHD algorithm, and the CDNN CD-spectral deconvolution which predicts 57% and 42% α-helix for FliH 94-258 and 117-258 respectively (Figure 3.43). The deconvolution also predicts 47% and 37% random coil respectively and some of this could be accommodated within loops, but as for the FliI CD spectral deconvolution there is obviously a large amount of error as the total secondary structure is 125% and 109% for FliH 94-258 and 117-258 respectively.

	FliH 94-258	FliH 117-258
Anti-parallel β-sheet	3%	10%
Parallel β-sheet	4%	3%
β-turn	14%	17%
Random coil	47%	37%
α-helix	57%	42%
Total %	125%	109%

Figure 3.43. CDNN Deconvolution for FliH 94-258 and 117-258.

Anti-parallel β-sheet, minimum at 215 nm, peak at 195 nm.

α -helix trough with two minima, 222.5 nm and 210 nm, peak at 190 nm.

Random coil, peak at 220 nm, trough minimum at 195 nm.

3.4 Interaction of *H. pylori* FliI and FliH

Given that both the FliI N-terminal domain and FliH could be over-expressed and purified it was decided to test for an interaction between the two proteins in order to learn more about the function of both. Initially the interaction of full-length FliH and FliH 55-258 with FliI 2-91 and 19-91 were tested by a GST pull-down method. This involved binding one of the proteins as a GST fusion to glutathione sepharose, then the sepharose was washed with 1 X PBS and subsequently incubated at room temperature after the addition of the second purified protein without the GST tag. As controls the untagged protein was incubated with either glutathione sepharose or GST-bound glutathione sepharose, to check for any background non-specific interactions of the protein. Samples of the glutathione sepharose were then run on SDS-PAGE gels. The results are shown below in Figure 3.44 A-B, and 3.45 A-B.

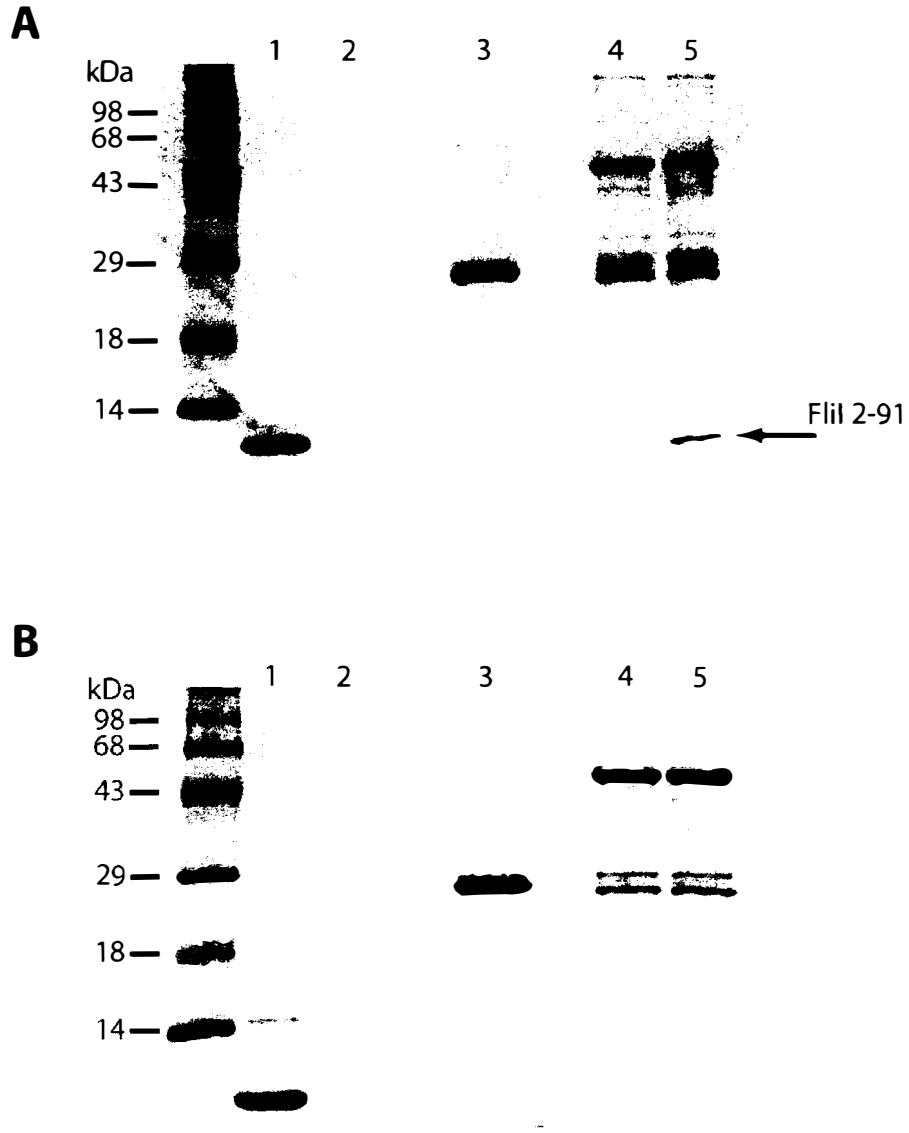


Figure 3.44 Coomassie Blue Stained 15% SDS-PAGE of GST-FliH GST pull-downs.

A. GST-FliH and FliI 2-91.

Lane 1. FliI 2-91, 10 μg.

Lanes 2-5: pull-down experiment with:

Lane 2. FliI 2-91 and glutathione sepharose

Lane 3. GST bound glutathione sepharose and FliI 2-91.

Lane 4 GST-FliH bound sepharose.

Lane 5. GST-FliH bound sepharose and FliI 2-91.

B. The lanes are the same as A, but the sepharose was incubated with FliI 19-91 instead of 2-91.

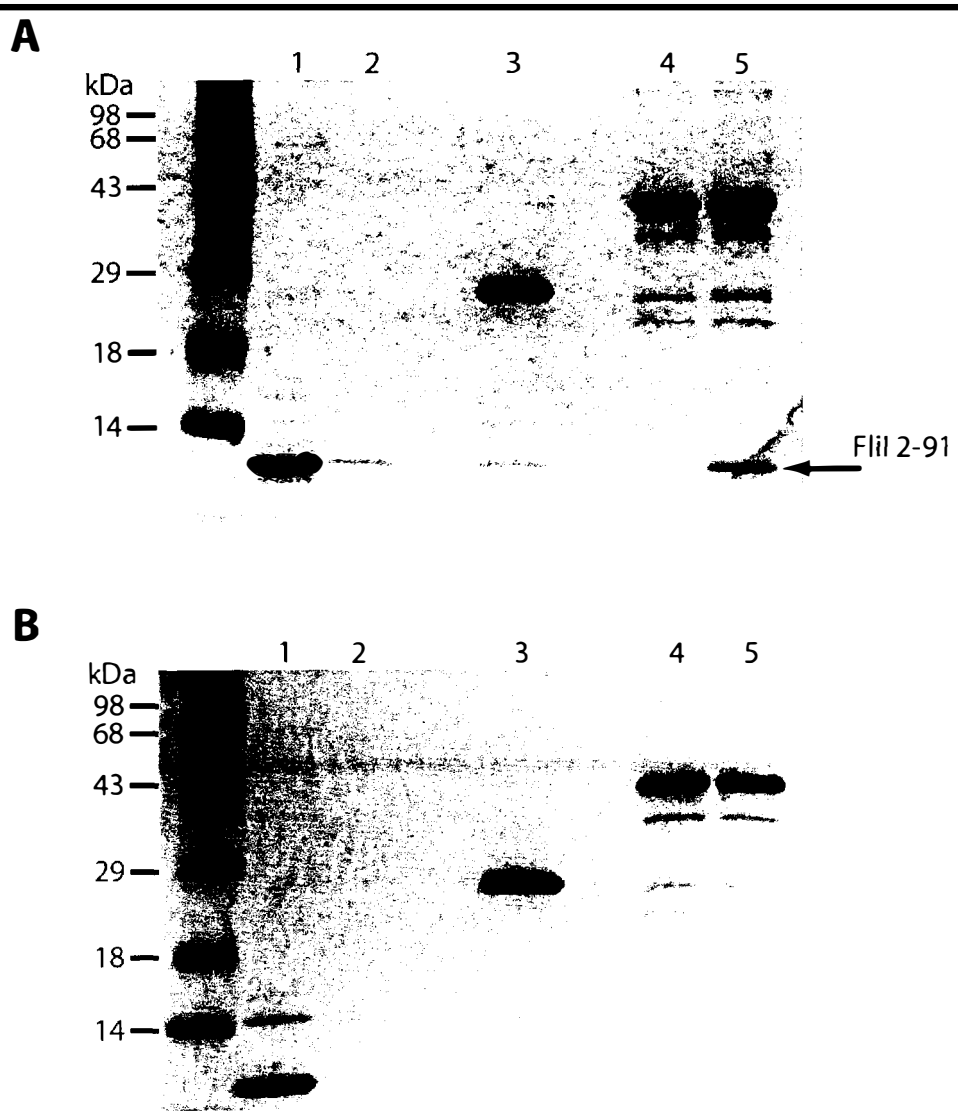


Figure 3.45 Coomassie Blue Stained 15% SDS-PAGE of GST-FliH 55-258 GST pull-downs.

A. GST-FliH 55-258 and FliI 2-91.

Lane 1. FliI 2-91, 10ug.

Lanes 2-5: pull-down experiment with:

Lane 2. FliI 2-91 and glutathione sepharose

Lane 3. GST bound glutathione sepharose and FliI 2-91.

Lane 4 GST-FliH 55-258 bound sepharose.

Lane 5. GST-FliH 55-258 bound sepharose and FliI 2-91.

B. The lanes are the same as A, but the sepharose was incubated with FliI 19-91 instead of 2-91.

As can be seen in Figure 3.44 A and 3.45 A, a band with the same molecular weight as FliI 2-91 remains associated with the FliH/FliH 55-258 bound glutathione sepharose after extensive PBS washing (Lane 5), while being absent from the glutathione sepharose and GST-bound sepharose controls (Lanes 2 and 3) after washing. This proposed FliI 2-91 band in lane 5 of Figure 3.39 A was sequenced to confirm that it is in fact FliI 2-91. This indicates that FliI 2-91 specifically interacts with FliH and FliH 55-258 in spite of the tendency for these proteins to aggregate. This interaction has also been shown to work in reverse using the GST pull-down method, that is a GST-FliI 2-91 fusion can pull down both FliH and FliH 55-258 (not shown). Furthermore, there was no evidence that FliI 19-91 interacts with full-length FliH or FliH 55-258, suggesting that the first 18 residues of FliI are necessary for interaction with FliH.

To verify that the observed interaction was not an artifact of FliH aggregation and to further define the domain of FliH necessary to interact with FliI, the interaction of FliH 94-258 and 117-258 with FliI 2-91 was also investigated using the same GST pull-down method. The results of this experiment are shown below in Figures 3.46 and 3.47.

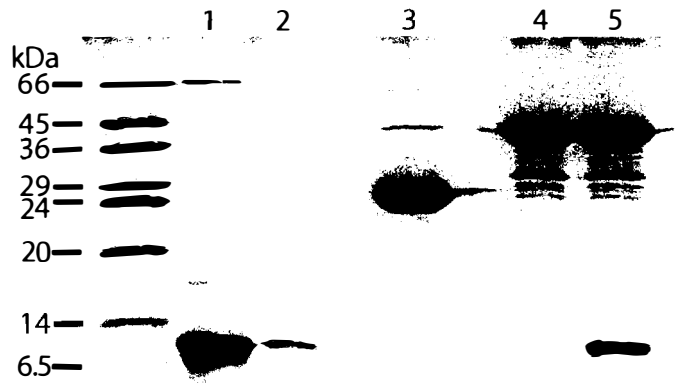


Figure 3.46 Coomassie Blue Stained 15% SDS-PAGE of GST-FliH 94-258/FliI 2-91 pull-down

Lane 1. FliI 2-91, 10 μg.

Lanes 2-5: pull-down experiment with:

Lane 2. FliI 2-91 and glutathione sepharose

Lane 3. GST bound glutathione sepharose and FliI 2-91.

Lane 4 GST-FliH 94-258 bound sepharose.

Lane 5. GST-FliH 94-258 bound sepharose and FliI 2-91.

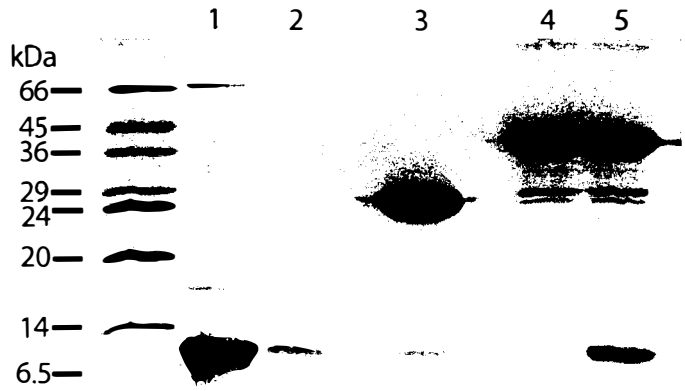


Figure 3.47 Coomassie Blue Stained 15% SDS-PAGE of GST-FliH 117-258/FliI 2-91 pull-down

Lane 1. FliI 2-91, 10 μg.

Lanes 2-5: pull-down experiment with:

Lane 2. FliI 2-91 and glutathione sepharose

Lane 3. GST bound glutathione sepharose and FliI 2-91.

Lane 4 GST-FliH 117-258 bound sepharose.

Lane 5. GST-FliH 117-258 bound sepharose and FliI 2-91.

As can be seen FliI 2-91 interacts with both FliH 94-258 and 117-258, suggesting that the first 117 residues of FliH are not necessary for an interaction with FliI. Furthermore, FliH 94-258 and 117-258 also failed to interact with FliI 19-92 by this method, (these results are not shown) and this again indicates the first 18 residues of FliI are necessary for an interaction with FliH.

To further demonstrate the interaction of the *H. pylori* FliI N-terminal domain and FliH, and to investigate the size and stoichiometry of the FliH 117-258/FliI 2-91 complex, the two proteins were mixed and calibrated SEC was performed (Figure 3.48).

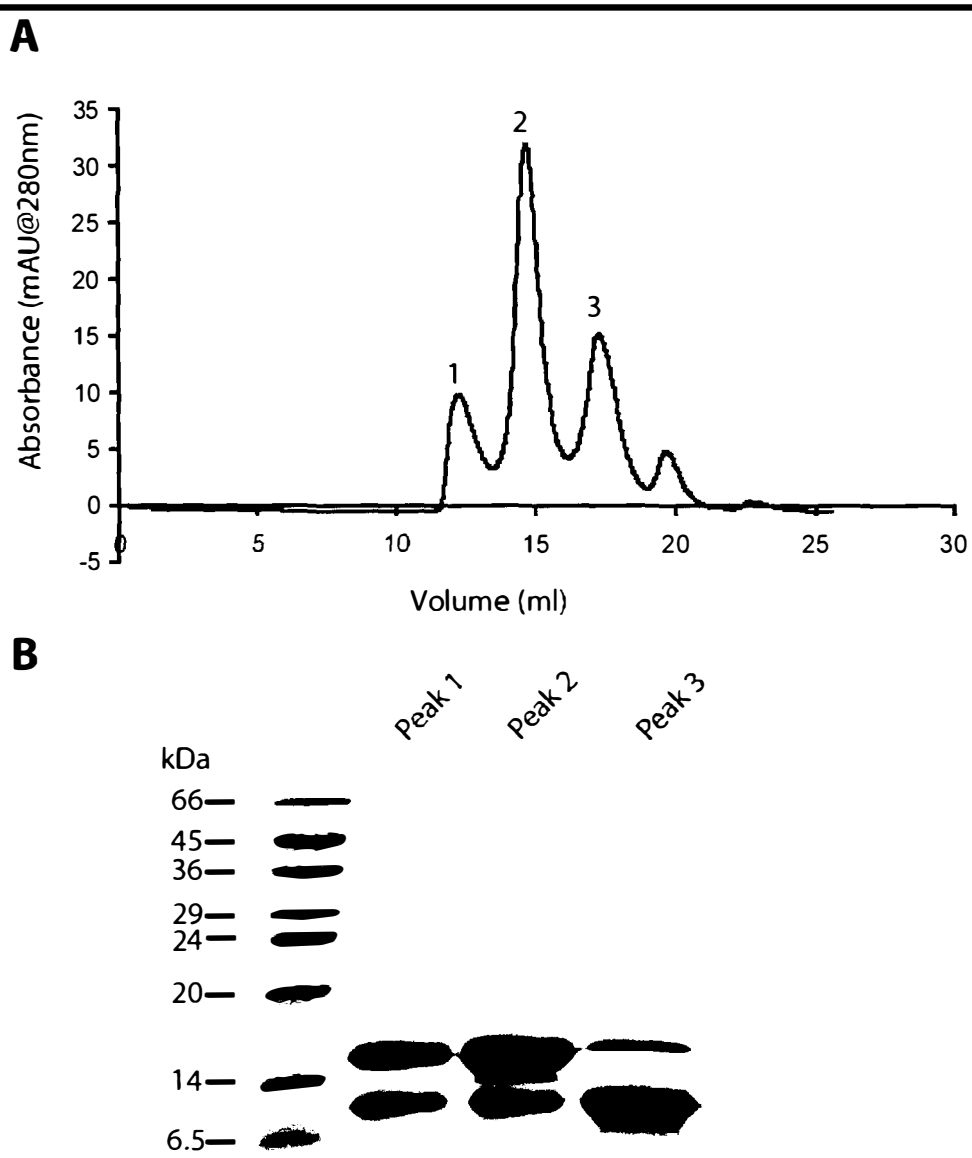


Figure 3.48 Calibrated SEC of the FliH 117-258/FliI 2-91 complex

A. Superdex 200 elution profile of a mixture of FliH 117-258 and FliI 2-91. Both proteins were at a concentration of 2 mg/ml before they were mixed, the mixture contained 500 μg of FliH 117-258 and 370 μg of FliI 2-91, such that the two proteins were present at a 1:1 molar ratio. The mixture was incubated for 30 min on a rocker at room temperature followed by injection into the SEC column.

B. 15% SDS-PAGE of the eluted peaks from A. The peaks were concentrated to the same volume and then 5 μl from each loaded on the gel.

Using the Superdex 200 calibration curve in Figure 3.19 the peaks 1-3 are estimated to be 132, 48 and 16 kDa respectively. Based on these estimated molecular weights and the SDS-PAGE, Peak 3 and 2 probably represent monomeric FliI 2-91 and dimeric FliH 117-258 respectively. The presence of FliI 2-91 in lane 2, and the small quantity of FliH 117-258 in lane 3 of the SDS-PAGE are probably due to incomplete separation of peaks 2 and 3. Peak 1 contains the FliI 2-91/FliH 117-258 complex.

Based on the predicted molecular mass of the complex and the predicted molecular masses of FliI 2-91 and FliH 117-258 of 18 kDa and 23.5 kDa respectively, a number of different complexes are possible. Some of the possible complexes are listed in Figure 3.49.

FliI 2-91:FliH 117-258 Complex	Predicted molecular mass (kDa)
2:2	83
2:3	106.5
2:4	130
3:2	101
3:3	124.5
3:4	148
4:2	119
4:3	142.5
4:4	166
5:2	137
6:2	155
6:3	178.5

Figure 3.49 Predicted molecular mass of theoretical FliI 2-91/FliH 117-258 complexes

It is impossible to determine the exact composition of the complex without further experiments such as multiangle light scattering to determine the exact molecular weight of the complex independently of its shape, and ultimately the determination of the crystal structure of the complex. To this end Peak 1 (Figure 3.48) was isolated, concentrated and then eluted from the Superdex 200 once more (Figure 3.50).

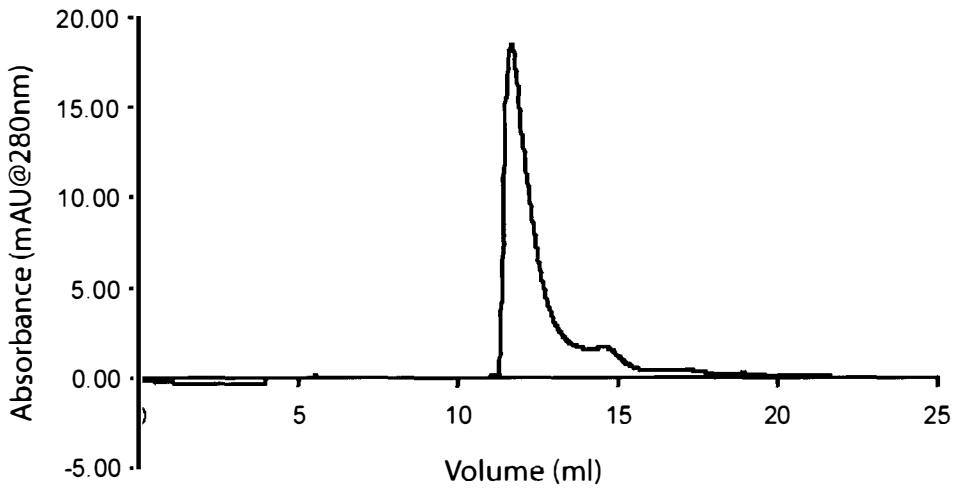


Figure 3.50 FliI 2-91/FliH 117-258 complex peak reapplied and eluted from the Superdex 200 column

This elution profile (Figure 3.50) shows that the complex is stable enough to be purified for crystal screens. A significant proportion of the complex was lost between the first and second elution from the Superdex 200 size exclusion column, suggesting that the complex adheres to the centricon membrane.

FliH 117-258 and FliI 19-91 were also mixed in a 1:1 molar ratio and applied to the Superdex 200 column (Figure 3.51).

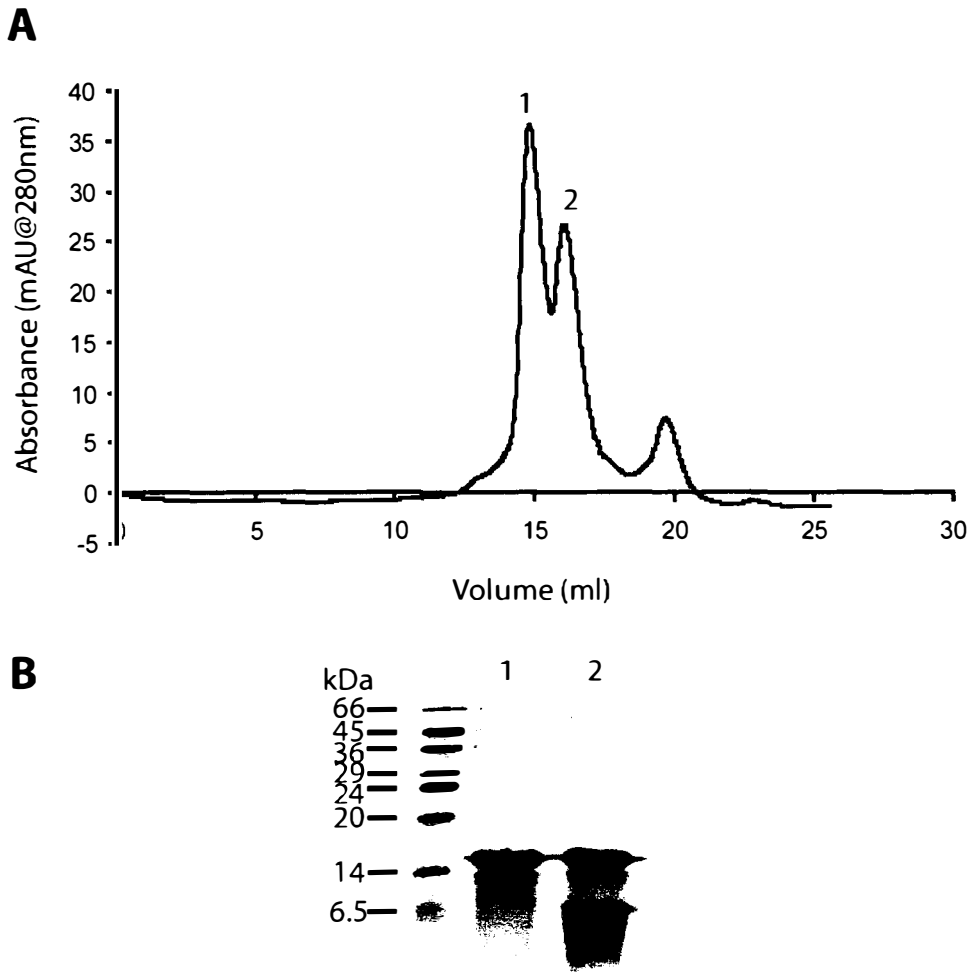


Figure 3.51 Calibrated SEC of a mixture of FliI 19-91 and FliH 117-258

A. Superdex 200 elution profile of a mixture of FliH 117-258 and FliI 19-91. Both proteins were at a concentration of 2 mg/ml before they were mixed; the mixture contained 500 μg of FliH 117-258 and 373 μg of FliI 2-91, such that the two proteins were present at a 1:1 molar ratio. The mixture was incubated for 30 min on a rocker at room temperature and injected into the SEC column. **B.** 15% SDS-PAGE of the eluted peaks from A. The peaks were concentrated to the same volume and then 5 μl from each loaded on the gel.

In contrast to FliI 2-92, when FliI 19-91 is mixed with FliH 117-258 only 2 peaks are seen during size exclusion (Compare Figures 3.48 A and 3.51 A). These correspond to the presumed dimeric FliH 117-258 observed previously (Figure 3.48 A) and trimeric FliI 19-91. A significant proportion of the FliH 117-258 is found in the same peak as trimeric FliI 19-91, this is not because they interact but rather because the two peaks overlap significantly. There is no higher molecular weight peak that would correspond to a complex formed by an interaction between FliI 19-91 and FliH 117-258. This again illustrates the importance of the first 18 residues of FliI 2-91 for the FliI 2-91/FliH 117-258 interaction.

In agreement with these findings, similar but less comprehensive evidence for the FliH-FliI interaction in *S. typhimurium* has been published (119). These authors did not show that the first twenty residues of FliI were critical for the interaction with FliH, but it had previously been shown that an L12P temperature sensitive mutation of FliI affected flagellum assembly (79) (188) (52). Subsequently, the purification of a His-tagged FliI-FliH complex from *S. typhimurium* and the failure to purify a His-FliI L12P/R7C-FliH complex demonstrated the formation of a complex in vivo and the importance of the N-terminus of FliI for this interaction (118).

3.4.1 Chracterisation of the FliI N-terminal Extension

Previous experiments in this project have demonstrated that the first 18 residues of FliI are important for the interaction of the FliI N-terminal domain with FliH. A multiple alignment of the N-terminal extension of FliI homologues is presented in Figure 3.52. Although the average identity to the first 18 residues of *H. pylori* FliI to other homologues is only 19%, alignment of the sequences is facilitated by the periodicity with which hydrophobic residues such as Leucine occur, and to a lesser extent the occurrence of positively charged residues such as Lysine or Arginine.

To determine the significance of this limited sequence conservation to the FliI 2-91-FliH interaction, a PCR based method was used to create seven point mutations of the N-terminal extension of *H. pylori* FliI 2-91. These mutations were L3A, K4A, L6A, L6E, K7A, R9A, R9E, and L10A. The PCR amplified DNA encoding FliI residues 2-91and containing these point mutations was cloned into pGEX-6P3 as N-terminal GST fusions and they were over-expressed and purified in the same manner as was wild-type FliI 2-91. The purification protocol included calibrated size exclusion using the Superdex 75 column. This demonstrated that the point mutants were the same size as wild type FliI 2-91, and that they exhibit the same tendency to form trimers (Figure 3.53).

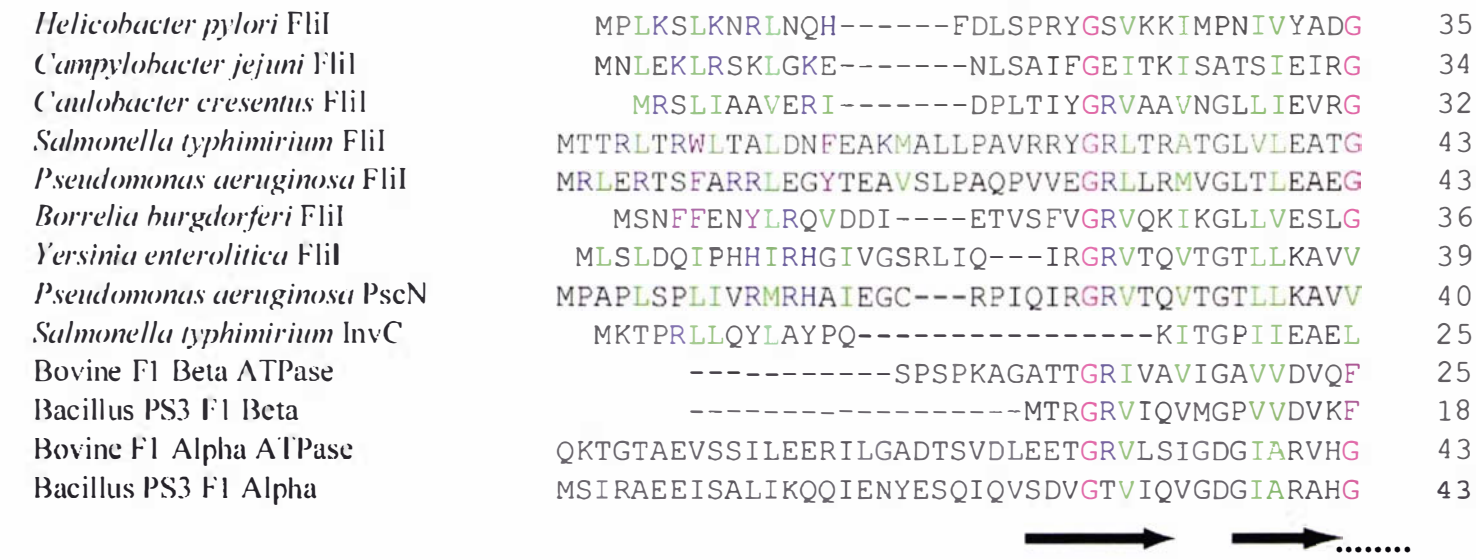


Figure 3.52 N-terminal alignment of FliI homologues, Type III export ATPases and F_1-ATPase subunits

Hydrophobic residues are green, aromatic hyphobic residues are purple, positively charged residues are blue.

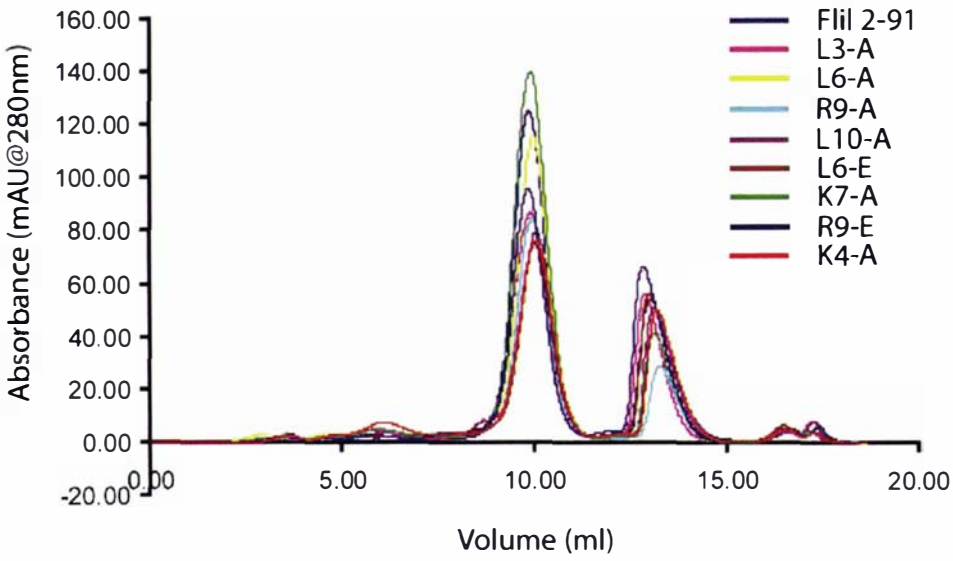


Figure 3.53 Superdex 75 elution profile of FliI 2-91 point mutants

The buffer was 1 x PBS containing 150 mM NaCl, 0.5 ml of each protein was injected, the flow rate was 0.5 ml/min (FliI 2-91= 2.2 mg/ml, L3-A= 2 mg/ml, L6-A= 2 mg/ml, R9-A= 1.3 mg/ml, L10-A= 1.5 mg/ml, L6-E= 1.48 mg/ml, K7-A= 2.2 mg/ml, R9-E= 1.9mg/ml, K4-A= 1.5 mg/ml).

The concentration of the purified FliI 2-91 point mutants was determined by their UV absorbance at 280 nm and then 5 μg of each protein was analysed by SDS-PAGE to demonstrate that their concentration can be estimated accurately. The interaction of these point mutants with FliH 117-258 was then tested using the same GST pull-down method used for wild-type FliI 2-91 and the truncated FliH proteins (Figure 3.54).

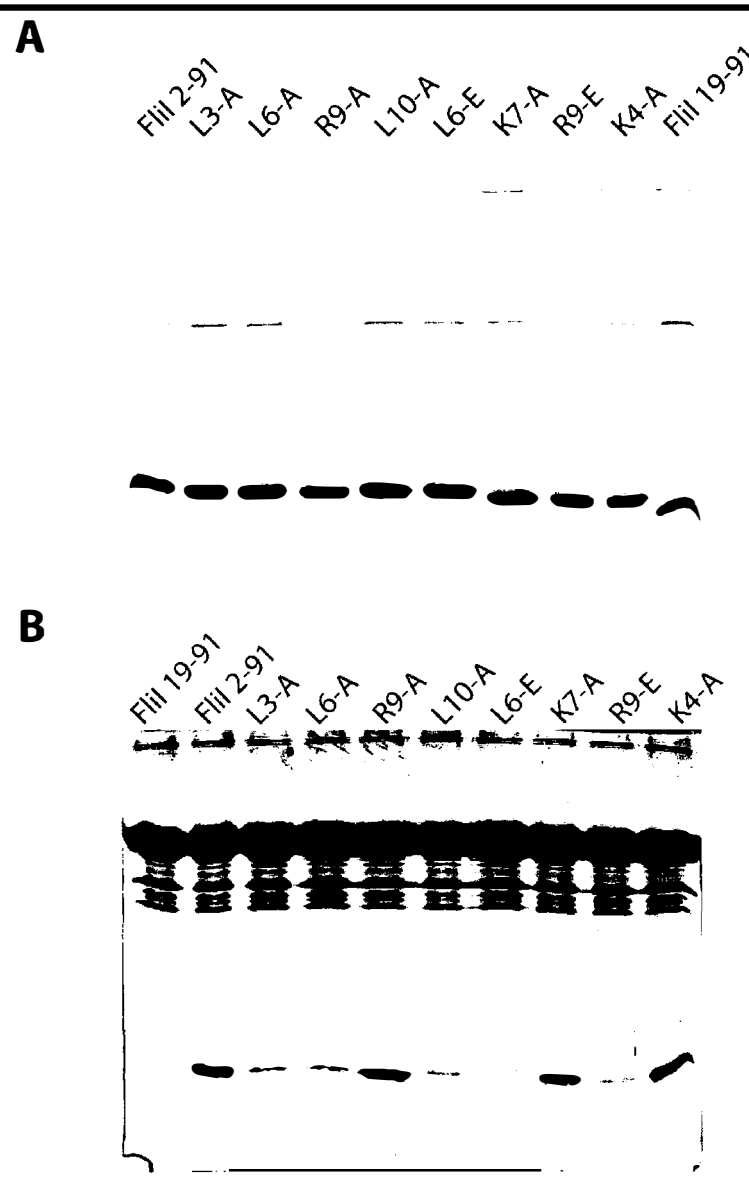


Figure 3.54 Interaction of GST-FliH 94-258 and FliI 2-91 point mutants

A. 15% SDS-PAGE gel showing that the concentration FliI 2-91 point mutants can be estimated accurately and equivalent quantities loaded on a gel or added to a GST pull-down reaction.

B. 15% SDS-PAGE gel of FliH 94-258/mutant FliI 2-91 GST pull-downs

The GST-FliH 94-258 pull-downs with the mutant FliI 2-91 proteins (Figure 3.54 B) show that within the N-terminal extension of *H. pylori* FliI the hydrophobic residues L3, L6 and L10 are important for the interaction with FliH. If any of these residues are changed to alanine, a much smaller hydrophobic residue, the interaction with FliH 94-258 does not occur above the background. Furthermore, FliI 2-92 and FliH 94-258 also fail to interact if L6 is changed to glutamate, a negatively charged residue. In contrast, when the positively charged residues K4, K7 and R9 are changed to alanine the pull-downs suggest the interactions of these mutant proteins with FliH 94-258 are as strong with wild-type FliI 2-91. But when R9 is changed to a glutamate, the mutant FliI 2-91 and FliH 94-258 do not interact. Together these data suggest that the interaction between FliI 2-92 and FliH 94-258 involves the N-terminal FliI extension and the interaction is primarily hydrophobic in nature. There is may be a small contribution from the charged residues of the N-terminal extension but a change to these residues can be tolerated as long as the charge is not reversed such that the proposed ionic interaction becomes repulsive.

To further demonstrate the importance of the hydrophobic residues in the *H. pylori* FliI N-terminal extension for the interaction with FliH, the L3A FliI 2-91 mutant was mixed in a 1:1 molar ratio with FliH 117-258 and calibrated SEC was performed (Figure 3.55).

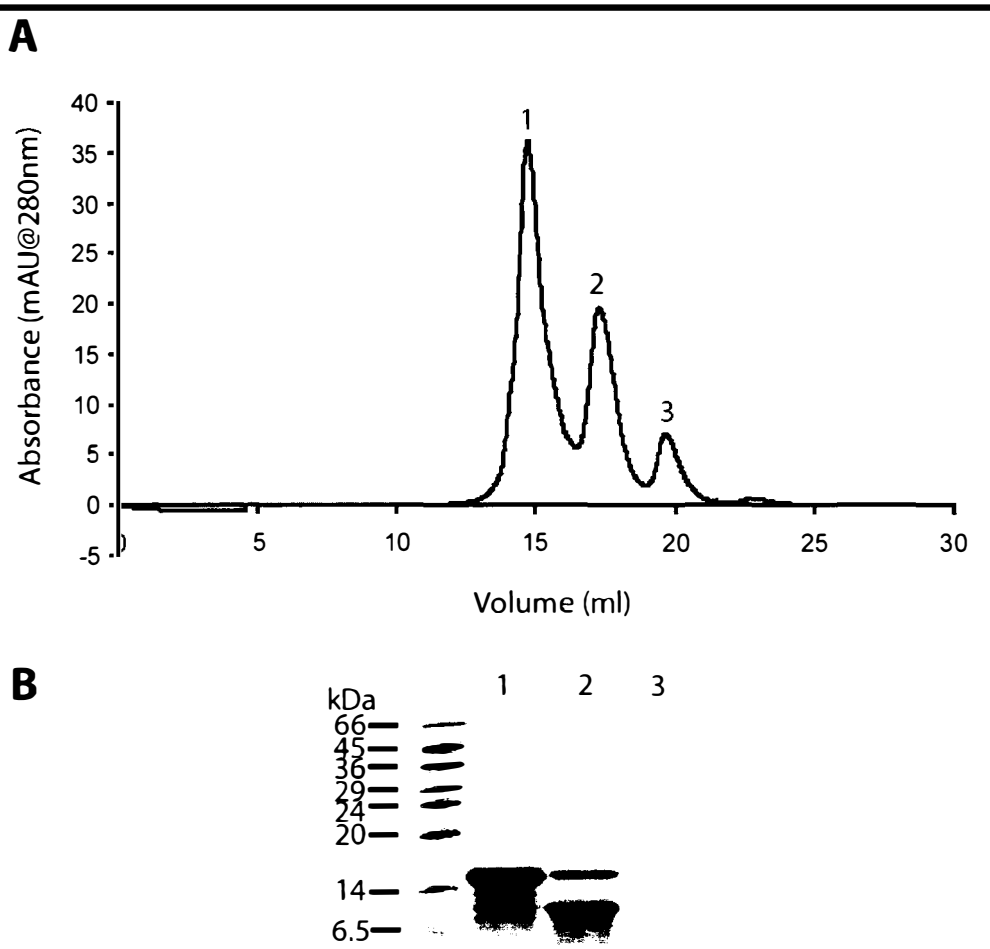


Figure 3.55 Calibrated SEC of a mixture of FliI 2-91 L3A and FliH 117-258

A. Superdex 200 size exclusion elution profile of FliI 2-91 L3A and FliH 117-258 mixed at a 1:1 molar ratio. Both proteins were at a concentration of 2 mg/ml before they were mixed, the mixture contained 500 μg of FliH 117-258 and 370 μg of FliI 2-92, such that the two proteins were present at a 1:1 molar ratio. The mixture was incubated for 30 min on a rocker at room temperature and injected into a SEC column.

B. 20% SDS-PAGE gel of the peaks in the elution profile (A), the peaks were concentrated to the same final volume and 5 μl samples of each peak was loaded on the gel.

When wild-type FliI 2-91 and FliH 117-258 were mixed (Figure 3.48) there were three peaks in the elution profile. The first peak, thought to represent the FliI 2-91/FliH 117-258 complex had an elution volume of 12.21 ml and a predicted molecular weight of 132 kDa. This peak is absent when FliI 2-91 L3A is mixed with FliH 117-258 in Figure 3.55 A, and this implies that the two proteins are not interacting following the mutation of L3 to alanine. Instead, the elution profile in Figure 3.55 more closely resembles that produced when FliI 19-91 was mixed with FliH 117-258 in Figure 3.51. In that experiment Peak 1 and 2 had elution volumes of 14.8 and 17.4 ml respectively, and predicted molecular masses of 44 and 15 kDa respectively. When FliI 2-91 L3A and FliH 117-258 were mixed (Figure 3.55) the peaks 1 and 2 eluted from the Superdex 200 column with volumes of 14.8 and 17.3 respectively, and predicted molecular masses of 45 and 16 kDa respectively.

3.4.2 Investigation of the structure of the FliI N-terminal extension

Although the N-terminal extensions of the α- and β-subunit are disordered in the crystal structure of the Bovine F_1-ATPase (1), recent studies with the *E. coli* enzyme have suggested that the N-terminal 22 residues of the α-subunit form an α-helix that is involved in binding the δ-subunit (193). Theoretically the first 18 residues of *H. pylori* FliI could form an amphipathic α-helix as suggested when these residues are used to create a helical wheel diagram (Figure 3.56). But limited proteolysis of FliI 2-91 previously discussed in this project (Figure 3.31), suggests that the N-terminal 18 residues of *H. pylori* FliI are loosely folded enough to be accessible to proteases and they may be disordered.

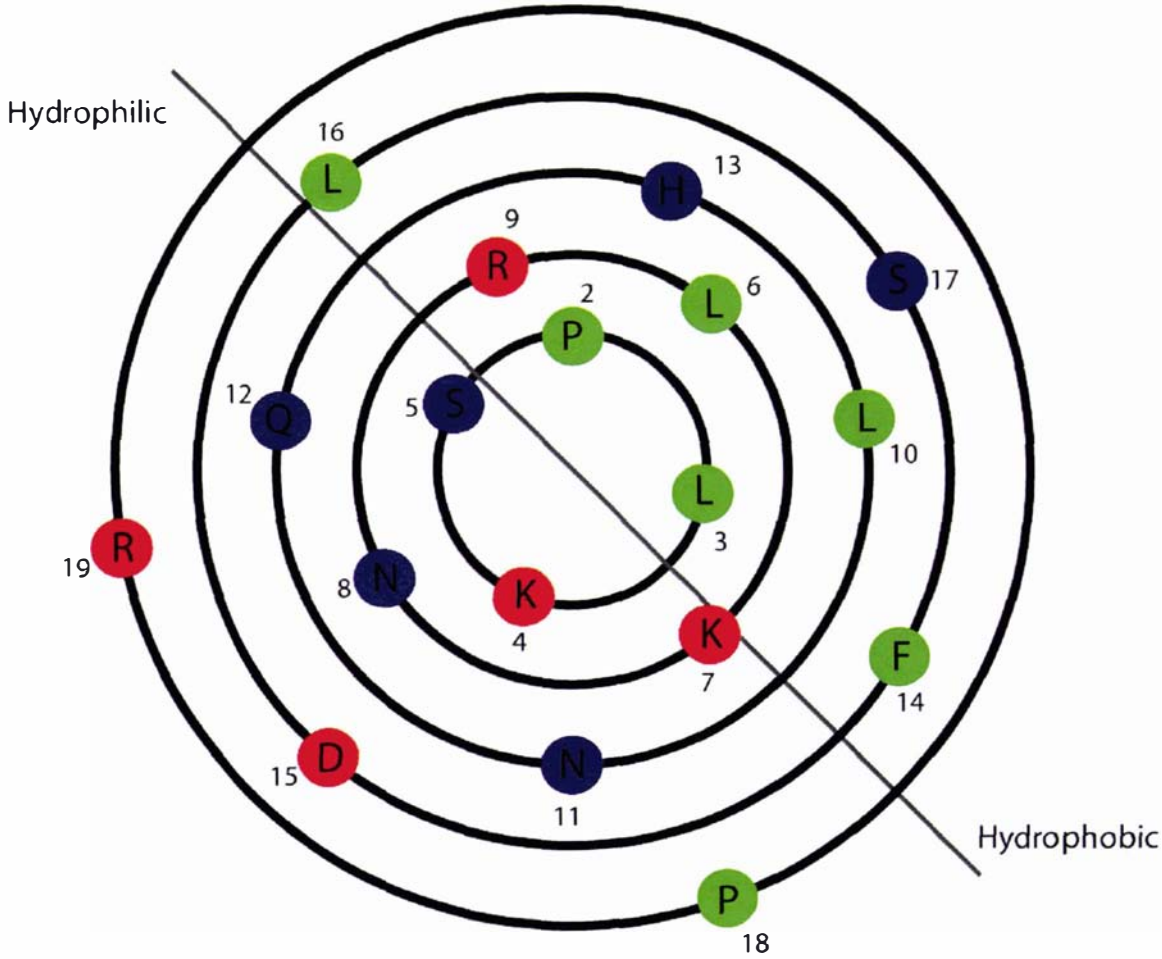


Figure 3.56 Helical wheel for the residues 2-18 of *H. pylori* FliI

Although FliI residues 2-18 are likely to be disordered, it is also possible that they become helical upon interaction with FliH. To investigate this possibility FliI 2-91 and FliH 94-258 were combined in a 2:1 molar ratio and the CD spectra of the mixture was measured (Figure 3.57 A and B).

If two proteins are mixed and the CD spectrum of the mixture is measured the spectrum should be the sum of the spectra of the individual proteins if there is no interaction occurring. If there is an interaction between the two proteins a change in the spectrum would be expected that reflected the structural transitions associated with the interaction, and therefore the spectrum would deviate from the sum of the individual spectra. In the case of the FliI 2-91/FliH 94-258 interaction an increase in the helical content might be reflected in the CD spectrum. The CD spectrum of a helix consists of two characteristic features. The first is a trough with two minima, the first is at approximately 225 nm, then the spectrum rises slightly before falling to a second minimum at approximately 210 nm. The second feature is a peak with a maximum elipticity at approximately 190 nm.

Interestingly, when FliI 2-91 and FliH 94-258 are combined (Figure 3.57) there is a small but significant increase in the magnitude of the negative elipticity around 220 nm. This makes the first minimum more pronounced and may be indicative of an increase in helicity in the FliI 2-91/FliH 94-258 complex. At the same time there is also a decrease in the elipticity of the peak at 190 nm which could be indicative of a decrease in helicity, but changes in the solvent can also change the elipticity below 200 nm.

In contrast when FliI 19-91 and FliH 94-258 are combined the spectra are almost identical at 220 nm and the trough has the same shape. Like the FliI2-91/FliH 94-258 spectrum there is also a significant decrease in the elipticity at 190 nm that is difficult to interpret.

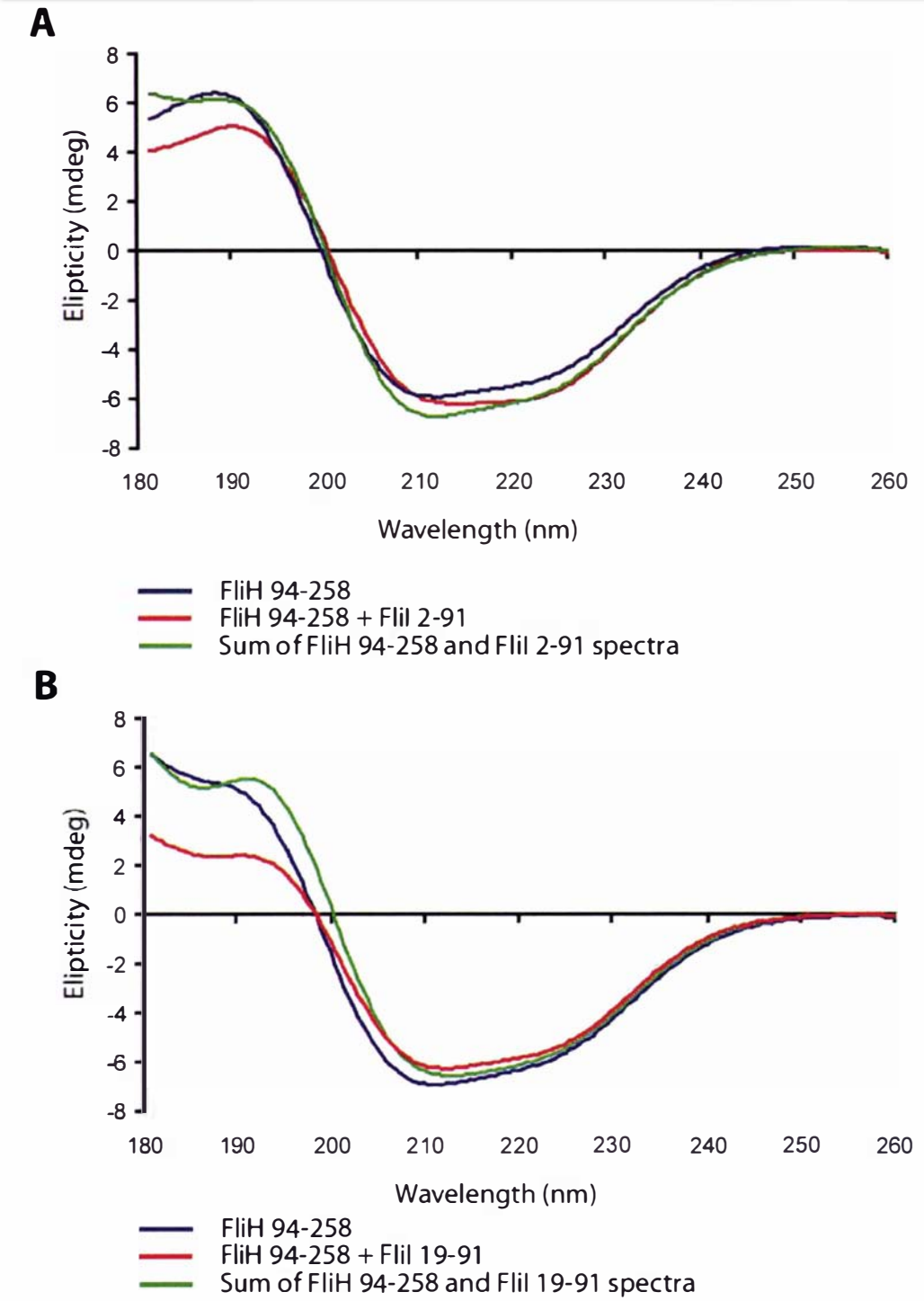


Figure 3.57 Investigation of the FliI/FliH interaction by CD spectroscopy

A. CD spectrum of FliH 94-258 alone, mixed with FliI 2-91 or summed with FliI 2-91.

B. CD spectrum of FliH 94-258 alone, mixed with FliI 19-91 or summed with FliI 19-91.

	FliH 94-258	FliI 2-91 and FliH 94-258 mixture	FliI 2-91 and FliH 94-258 summed
α-helix	29	33	34
Anti-parallel β-sheet	11	10	9
Parallel β-sheet	5	5	5
β-turn	19	17	18
Random coil	31	30	29
Total %	95%	95%	95%

Figure 3.58 CDNN Deconvolution of FliH 94-258/FliI 2-91 mixed CD spectra

	FliH 94-258	FliI 19-91 and FliH 94-258 mixture	FliI 19-91 and FliH 94-258 summed
α-helix	32	30	33
Anti-parallel β-sheet	9	10	10
Parallel β-sheet	5	5	5
β-turn	20	19	18
Random coil	30	31	30
Total %	96%	95%	96%

Figure 3.59 CDNN Deconvolution of FliH 94-258/FliI 19-91 mixed CD spectra

The deconvolution of the CD spectrum suggests that the helical content of the proteins increases by 4% when FliI 2-91 and FliH 94-258 are added (Figure 3.58). Furthermore, the deconvolution of the FliI 19-91/FliH 94-258 data does not show an increase in helical content when these two preoteins are mixed (Figure 3.59).

But two points are of note with regard to this deconvolution, firstly the helical content of the summed spectra is also higher than the FliH 94-258 alone. Furthermore, this CD deconvolution only used the data from 200 to 260nm, when the data from 185 to 260 nm is used the difference between the FliH 94-258 spectrum and the FliH 94-258/FliI 2-91 combined spectrum drops to less than 1%.

In summary the CD spectra provide some evidence for an increase in helicity when FliI 2-91 and FliH 94-258 are incubated together, but the data are not conclusive and further experiments are required. To this end a synthetic peptide comprising FliI residues 2-14 was purchased and the CD spectrum of this peptide was measured (Figure 3.60).

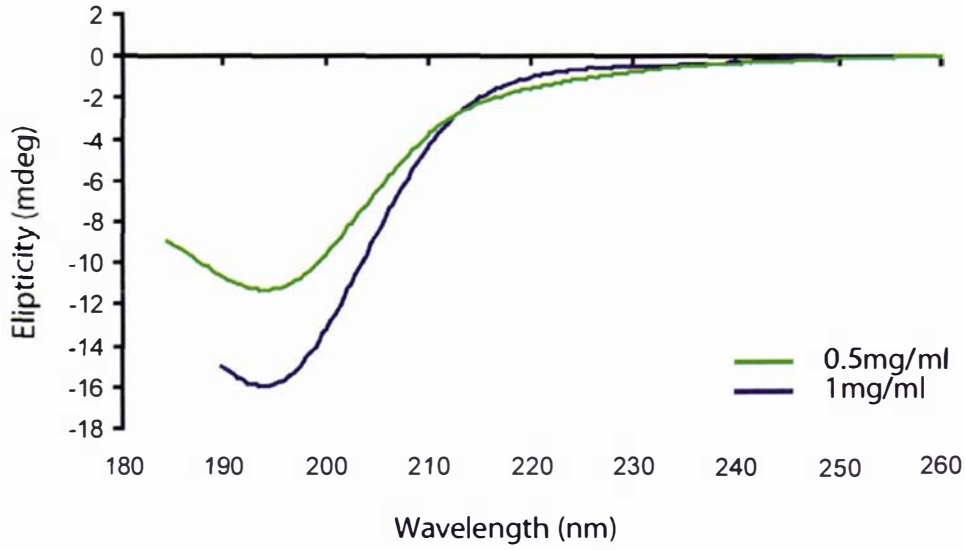


Figure 3.60 CD spectra of a FliI 2-14 synthetic peptide

Contrary to our expectations, the CD spectrum of the FliI 2-14 synthetic peptide was characteristic of a random coil not an α-helix. But it is common for short synthetic peptides to be unstructured and it was decided to test the interaction of FliH 94-258 and FliI 2-14 by calibrated size exclusion (Figure 3.61).

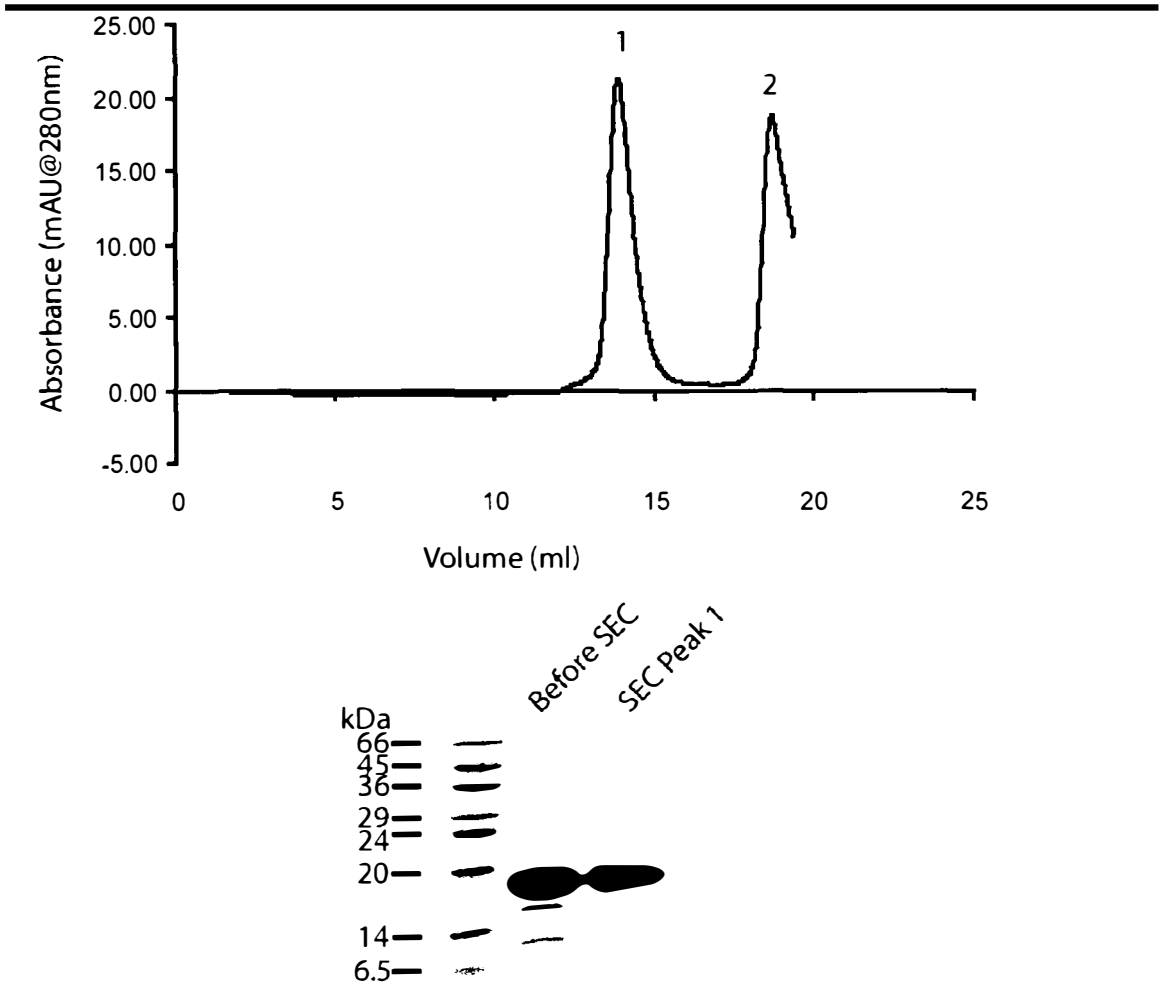


Figure 3.61 Interaction of FliH 94-258 and FliI 2-14 synthetic peptide

A. Superdex 200 elution profile of a mixture of FliH 94-258 and FliI 2-14

B. 20% SDS-PAGE gel of the FliH 94-258/FliI 2-14 mixture before SEC and the first peak to elute during SEC.

Although peak 2 in the SEC elution profile is probably the synthetic peptide and there is no evidence of the peptide being present in peak 1 by SDS-PAGE, peak 1 has a predicted molecular weight of approximately 68 kDa. FliH 94-258 alone has a predicted molecular weight of 60 kDa. This suggests that FliH 94-258 and the synthetic FliI 2-14 do interact, but the structure of FliI 2-14 still remains to be determined.

4. Discussion

This project began with a search for an *H. pylori* homologue of *S. typhimurium* FliJ. A potential homologue, HP0256, was identified using PSI BLAST, but it only has an E-value ≤ the 0.005 significance threshold with the appropriate combination of query sequence, Expect value and number of iterations. That is, *T. denticola, T. tengcongensis,* and *L. interrogans* FliJ homologues, are identified with an Expect value of 100 and 6, 4 and 4 iterations respectively. This bioinformatics search could probably be refined further by considering the phylogenetic relationships of the annotated FliJ homologues prior to searching. A phylogenetic analysis may demonstrate that *T. denticola, T. tengcongensis, L. interrogans* FliJ form a distinct and more closely related phylogenetic cluster within the FliJ homologues. *H. pylori* FliJ may be more closely related to this group than *S. typhimurium* FliJ for example. Examination of the motility and flagellation of HP0256 null mutants of *H. pylori* strain 17874, suggested that HP0256 is not involved in the motility or flagellar assembly. However, more detailed examination of this mutant strain in the lab of P. W. O'Toole shows that the HP0256 null mutant does demonstrate defective motility (O'Toole, personal communication). Therefore this protein remains of future interest, and requires further studies of its interaction with other flagellar proteins, and its effect on the aggregation and export competence of flagellar export substrates.

The major findings of this project relate to the molecular characteristation of *H. pylori* FliI and FliH, and the interactions of these two proteins. Data presented here suggests this interaction is analogous to the interaction between the F_1-ATPase α-subunit and the δ-subunit of the F_1-ATPase stator.

Due to the apparent structural homology of *H. pylori* FliI model to both F_1-ATPase and Rho transcription terminator, an attempt was made to extend this relationship further to the catalytic mechanism of the enzyme by mutating the FliI R-loop and observing the effect on *H. pylori* motility. In dominance experiments one interesting result was obtained, the E_{311}-

A mutant appeared hyper-flagellated by electron microscopy when expressed *in trans*. The interpretation of this result was complicated by the finding that the *fliI* knock-out strain was still flagellated and motile and therefore the effect of the R-loop mutagenesis could not be evaluated by complementation. Thus, the function of the R-loop in the FliI catalytic mechanism remains of future interest but the experiments in this study were inconclusive.

When appropriate controls are found such that the effect of mutating the R-loop of *H. pylori* FliI can be evaluated, *in vitro* experiments will be necessary to explain the observed phenotype in the context of aberrant enzyme activity. To this end *H. pylori* FliI was overexpressed and a purification strategy involving Source-Q anion exchange chromatography was developed. Like *S. typhimurium* FliI (43), the *H. pylori* FliI is poorly soluble but purification of this protein did allow limited in vitro characterisation by calibrated SEC and far UV CD spectroscopy. The spectrum of FliI by far UV CD spectroscopy suggests that the protein has a mixture of α-helical and β-sheet secondary structure. Furthermore, the addition of Urea caused changes in the spectrum that suggested the protein was unfolding. Therefore, *H. pylori* FliI as purified in this study is assumed to be folded. However, calibrated size exclusion in the presence of ATP indicated the size of FliI was 48 kDa. This suggests that although the protein is folded it is primarily a monomer in solution and does not oligomerise under the conditions used. Claret and colleagues (31) have shown that *S. typhimurium* FliI forms a hexamer in the presence of ATP and this hexamerisation was promoted in the presence of the ATP analog AMP-PNP and *E. coli* phospholipids. Therefore, to test further for *H. pylori* FliI oligomerisation, size exclusion was performed in the presence of AMP-PNP instead of ATP and in the presence of ATP and phospholipids. Conditions were not identified that were conducive to hexamer formation. There are at least three possible explanations for the failure of FliI to oligomerise. One of these explanations involves the experimental conditions. Although Claret and colleagues observed oligomerisation by SEC and multi-angle light scattering, the monomeric peak represented 60% of the protein, and the higher molecular mass peaks of 120 kDa and 169 kDa represented 22.5 and 12% of the protein respectively. Furthermore, these higher molecular mass peaks do not represent a hexamer; heaxamerisation was observed by electron microscopy. Therefore it is possible that only a very small proportion

of *H. pylori* FliI is present as a hexamer in solution and SEC is not sensitive enough to detect this minor species. Another technique that fixes FliI like EM may be necessary to visualise the hexamer. Another possible explanation for the failure of FliI oligomerise could be partial misfolding of the protein. Previous experiments with *S. typhimurium* FliI illustrated that although it could be purified, the enzyme could bind but not hydrolyse ATP (43). This *S. typhimurium* FliI was purified by refolding inclusion bodies, subsequent purifications of soluble FliI could hydrolyse ATP and the protein was folded as evaluated by far UV CD spectroscopy (52) (118). A third possible explanation for the failure to observe FliI oligomers concerns the oligomerisation characteristics of the protein. *H. pylori* FliI may require the presence of another protein, such as an exported flagellar protein or a component of the export apparatus like FliH, to oligomerise.

When this study began FliH was a poorly characterised protein. Homologues are now annotated in flagellar gene clusters in most motile organisms despite poor sequence conservation, and it was shown to be involved in motility as *S. typhimurium fliH* mutants are non-motile and aflagellate (188) (90). In this thesis *H. pylori* FliH was purified and characterised *in vitro*.

Initially full-length FliH (2-258) was cloned and over-expressed but it was found by DLS and calibrated SEC that this protein although soluble was aggregated in solution. Some degree of aggregation has also been noted for *S. typhimurium* FliH (118), but this small (though undisclosed) fraction of the expressed protein that was removed by centrifugation at 130,000 xg. The purification protocol used in this study also included a high speed centrifugation step but this was insufficient to remove the aggregates. Subsequently, three more truncated forms of FliH were cloned and over-expressed, FliH 55-258, 94-258 and 117-258. FliH 55-258 was also found to aggregate, but FliH 94-258 and 117-258 were found to be soluble and consequently these two proteins were used in further study.

Far UV CD spectroscopy was performed with both FliH 94-258 and 117-258 and based on the spectra both proteins are predicted to be predominantly α-helical, with CDNN deconvolution estimating 57% and 42% α-helix respectively. Intriguingly, when these

proteins were subjected to calibrated SEC FliH 94-258 and 117-258 had estimated masses of 60 kDa and 47 kDa. The theoretical masses of FliH 94-258 and 117-258 are 17 kDa and 15 kDa respectively. This suggests that either FliH is oligomerising in solution, is elongated, or a combination of the two. *S. typhimurium* FliH has also demonstrated an anomalously rapid elution behavior in SEC, but in this case the SEC was combined with in line multi-angle light scattering to determine the mass of the complex independently of shape. This illustrated that the complex was a dimer, and FliH to be an elongated molecule (118). However, another *S. typhimurium* study has shown that the elongated shape of FliH is primarily due to the N-terminal 102 residues (116). In this thesis, however, the N-terminal residues of *H. pylori* FliI have been removed, therefore it is possible that the shape of the resulting truncated FliH was not elongated. Taking the molecule shape bias into account, the SEC size estimation from this thesis suggest that either FliH forms a trimer or an elongated dimer. These two possibilities cannot be resolved without further experiments.

H. pylori FliI was modeled on the F_1-ATPase ATPase atomic coordinates and this model structure is consistent with the formation of a hexamer. Each FliI monomer was predicted to consist of three domains, a C-terminal bundle of five helices, a central catalytic domain formed by a β-sheet sandwiched by α-helices, and an N-terminal β-barrel with a disordered N-terminal extension. As full-length FliI was difficult to work with it was decided to clone and express the N-terminal β-barrel (residues 2-91) and characterise this domain independently of the remainder of the protein. FliI 2-91 was cloned and expressed and found to be soluble and stable, but limited proteolysis suggested that an N-terminal extension of 19 residues was protease sensitive and therefore possibly disordered. Therefore FliI 19-91 was also cloned and expressed such that the function of the N-terminal extension could also be examined by comparison between the two truncated proteins, FliI 2-91 and 19-91. In this study the far UV CD spectra of FliI 2-91 and 19-91 have been interpreted to suggest that both proteins are folded and predominantly β-sheet, with a CDNN deconvolution of 52 and 58% anti-parallel β-sheet respectively. Despite similar deconvolution the spectra of FliI 2-91 and 19-91 are significantly different. Between 260 and 205 nm the elipticity of the CD spectrum of FliI 19-91 remains positive, while the spectrum of FliI 2-91 forms a trough of negative elipticity over the same wavelengths. This

trough is indicative of α-helical structure and may imply that the first 18 residues of FliI form an α-helix, but this is difficult to confirm without further experiments. However, by analogy to the FliI homologues, F_1-ATPase α-subunit and Rho, it might be expected that the N-terminus of FliI will form an α-helix. N-terminal 22 residues of the F_1-ATPase α-subunit are predicted to form an α-helix by far UV CD spectroscopy (193), and the N-terminal residues of Rho form a three helix bundle (164).

FliI 2-91 and 19-91 were also examined by calibrated SEC. FliI 2-91 has a theoretical molecular mass of 10 kDa, but it elutes from the Superdex 200 column as two peaks with estimated molecular masses of 66 and 18 kDa. This indicates that the N-terminal domain of FliI oligomerises, the 18 kDa peak is assumed to represent the monomeric form and therefore the 66 kDa peak is predicted to be a trimer. Furthermore, the FliI N-terminal domain must be elongated to explain the anomalously fast elution that leads to a larger than expected size prediction. FliI 19-91 eluted as a single peak with a predicted molecular mass of 28 kDa, and therefore it is also assumed to form a trimer with an elongated shape as the theoretical molecular mass of this protein is 8 kDa. To our knowledge this is the first evidence of the FliI N-terminal domain oligomerising, and suggests that the full-length protein may also oligomerise. However, interaction between the N-terminal domains is not unheard of in full-length FliI homologues. In Rho there are interactions between the N-terminal domains that involve a loop connecting α2 and α3 of the three-helix bundle and the linker that joins the N- and C-terminal domains of the adjacent subunit (164).

A GST pull-down assay was used to investigate the possibility of an interaction between FliH and FliI 2-91 N-terminal domain. This assay demonstrated that FliH does interact with FliI 2-91. Furthermore FliI 2-91 also interacts with FliH 55-258, 94-258 and 117-258 in GST pull-downs. In contrast FliI 19-91 did not interact with FliH or the truncated variants of this protein. This demonstrated that the 18 N-terminal residues of the FliI N-terminal domain are required for the interaction with FliH, whereas the N-terminal 117 residues of FliH appear dispensable for this interaction. To further demonstrate the interaction and investigate the stoichiometry of the complex, a mixture of FliH 117-258 and FliI 2-91 were

examined by calibrated SEC. The resulting elution profile consisted of three peaks with estimated molecular masses of 132, 48 and 16 kDa.

Concurrently with this study in *H. pylori* the homologous interaction of *S. typhimurium* full-length FliI and FliH was demonstrated (118). The *S. typhimurium* studies illustrated that like *H. pylori* FliH, the C-terminal half of *S. typhimurium* FliH is important for the interaction with FliI. Unlike this *H. pylori* study the *S. typhimurium* study isolated a $(FliH)_2FliI$ complex. Based of SEC-estimated molecular masses of *H. pylori* FliH 117-258 and FliI 2-91 of 23.5 kDa and 18 kDa respectively, the equivalent complex in *H. pylori* would have a molecular mass of 65 kDa. Therefore, unless the *H. pylori* complex is considerably more elongated, it has a different composition than the *S. typhimurium* complex. Given the estimated molecular mass of the FliI 2-91/FliH 117-258 complex of 132 kDa, and the error in this estimation due to the elongated shape of the complex, the possible stoichiometries include 2:3 (106.5 kDa), 2:4 (130 kDa), 3:2 (101 kDa), 3:3 (124.5 kDa) and 4:2 (119 kDa). Further experiments are required to determine the exact stoichiometry of the complex.

The *S. typhimurium* studies identified L12P and R7C/L12P mutants of FliI that displayed a temperature sensitive loss of motility (188). It was subsequently shown that these mutants fail to interact with FliH (118). Thus, the involvement of the N-terminus of FliI in the interaction with FliH was implied, but not demonstrated directly as it is here with the *H. pylori* proteins.

The N-terminal extension of FliI is not highly conserved but it can be aligned by virtue of conserved hydrophobic amino acids such as Leucine and positively charged residues such as Lysine and Arginine. It was decided to investigate the role of these residues in the *H. pylori* FliI 2-91-FliH interaction. To this end the FliI 2-91 DNA sequence was mutated such that the encoded protein contained one of the following mutations: L3A, K4A, L6A, L6E, K7A, R9A, R9E or L10A. These mutants had the same molecular mass as wild-type FliI 2-91 when examined by calibrated SEC, and they exhibited the same tendency to form trimers. GST pull-down assays were subsequently performed with these mutants. The

interaction of FliH 94-258 with the FliI 2-91 L3A, L6A, L6E, and L10A mutants was reduced to the level of background. This indicates that large hydrophobic sidechains at these positions are important for the FliH-FliI 2-91 interaction. However, mutating the K4, K7 and R9 to A had no effect on the interaction of FliI 2-91 with FliH 94-258, but the R9-E mutation abolished the interaction. This suggests that the interaction is primarily hydrophobic, and although charged residues may not be directly involved in the interaction they are present in the vicinity and the nature of the charge is important. This is analogous to the interaction between the F_1-ATPase α-subunit and the δ-subunit, which involves the hydrophobic residues of the δ-subunit and the N-terminal residues of the α-subunit (215) (216). The significance of these mutations was further demonstrated by mixing FliI 2-91 L3-A and FliH 117-258 and performing calibrated SEC. The 132 kDa peak formed by wild-type FliI 2-91 and FliH 117-258 was absent, illustrating that the proteins do not interact.

As previously mentioned it is possible that the N-terminal 18 residues of *H. pylori* FliI form an α-helix. Again this would be analogous to the N-terminus of the F_1-ATPase α-subunit, which is predicted to form an α-helix upon interaction with the δ-subunit (216) (193). To investigate this further a synthetic peptide was purchased that consisted of residues 2-14 of *H. pylori* FliI, and the secondary structure examined by far UV CD spectroscopy. The spectrum indicated the synthetic peptide was predominantly random coil in aqueous solution. This result is not unexpected, it is consistent with previous observations of the structure of synthetic peptides comprising the N-terminal residues of either the α- or β-subunit of the F_1-ATPase (216). Despite the synthetic peptide being random coil, when FliH 94-258 and FliI 2-14 were mixed and examined by calibrated SEC, the FliH peak had an estimated molecular mass of 68 kDa, 8 kDa larger than FliH 94-258 alone. This suggests that the synthetic peptide can interact with FliH and likely acquires structure upon interaction. To investigate this possibility further FliH 94-258 was mixed with FliI 2-91 and examined by far UV CD spectroscopy. Relative to the spectrum of FliH 94-258 the spectrum of the mixed proteins has difference in shape, which cannot be accounted for by summing the spectra of the two proteins alone. By comparison when FliI 19-91 and FliH 94-258 are mixed there is no change in the shape of the spectrum relative to

FliH 94-258 alone, just slight differences in the magnitude of the signal. Therefore, the addition of FliI 2-91 to FliH 94-258 causes a change in elipticity, and this may be indicative of an increase in FliI 2-91 helicity upon interaction with FliH. However, the change is small and the error in the far UV CD spectroscopy deconvolution appears large so further experiments are required to confirm this.

4.1 Future directions

Further experiments are required to determine the composition of the FliH 117-258-FliI 2-91 complex. In line multi-angle light scattering will be effective to this end as it determines molecular mass independently of shape.

There is limited evidence for the N-terminal extension of FliI forming an α-helix. The best way of being certain of this will be to determine the crystal structure of FliI 2-91. Crystal screens of FliI 2-91 and FliH 94-258/117-258 have thus far yielded no crystals. It is interesting to note that the optimal crystallisation of Rho required co-crystallisation with mild detergents and nucleic acid substrate (164). Co-crystallisation of FliI 2-91 with FliH 94-258 or 117-258 may prove to be a productive line of investigation. Furthermore, the N-terminal extension of FliI 2-91 is more likely to be folded in such a complex.

Bibliography

1. **Abrahams, J. P., Leslie, A. G. W., Lutter, R., Walker, J. E.** 1994. Structure at 2.8 A resolution of F_1-ATPase from bovine heart mitochondria. Nature **370:**621-628.
2. **Akeda, Y., Galan, J. E.** 2004. Genetic analysis of the *Salmonella enterica* type III secretion-associated ATPase InvC defines discrete functional domains. Journal of Bacteriology **186:**2402-2412.
3. **Alami, M., Luke, I., Deitermann, S., Eisner, G., Koch, H-G., Brunner, J., Muller, M.** 2003. Differential interactions between a twin-arginine signal peptide and its translocase in *Escherichia coli.* Molecular Cell **12:**937-946.
4. **Aldridge, P., Karlinsey, J., Hughes, K.T.** 2003. The type III secretion chaperone FlgN regulates flagellar assembly via a negative feedback loop containing its chaperone substrates FlgK and FlgL. Molecular Microbiology **49:**1333-1345.
5. **Allison, T. J., Wood, T.C., Briercheck, D.M., Rastinejad, F., Richardson, J.P., Rule, G.S.** 1998. Crystal structure of the RNA-binding domain from transcription termination factor rho. Nature Structural Biology **5:**352-356.
6. **Altschul, S. F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Research **25:**3389-3402.
7. **Amieva, M. R., Vogelmann, R., Covacci, A., Tompkins, L.S., Nelson, W.J., Falkow, S.** 2003. Disruption of the epithelial apical-junctional complex by *Helicobacter pylori* CagA. Science **300:**1430-1434.
8. **Andersen-Nissen, E., Smith, K.D., Strobe, K.L., Barrett, S.L., Cookson, B.T., Logan, S.M., Aderem, A.** 2005. Evasion of toll-like receptor 5 by flagellated bacteria. Proceedings of the National Academy of Sciences, USA **102:**9247-9252.
9. **Andersen, C., Koronakis, E., Bokma, E., Eswaran, J., Humphreys, D., Hughes, C., Koronakis, V.** 2002. Transition to the open state of the TolC periplasmic tunnel entrance. Proceedings of the National Academy of Sciences, USA **99:**11103-11108.
10. **Andersen, C., Koronakis, E., Hughes, C., Koronakis, V.** 2002. An aspartate ring at the TolC tunnel entrance determines ion selectivity and presents a target for blocking by large cations. Molecular Microbiology **44:**1131-1139.
11. **Anderson, D. M., Schneewind, O.** 1997. A mRNA signal for the type III secretion of Yop proteins by *Yersinia enterocolitica.* Science **278:**1140-1143.
12. **Andrade, M. A., Chacon, P., Merelo, J. J., Moran, F.** 1993. Evaluation of secondary structure of proteins from UV circular dichroism spectra using an unsupervised learning neural network. Protein Engineering **6:**383-390.
13. **Aspholm-Hurtig, M., Dailide, G., Lahmann, M., Kalia, A., Ilver, D., Roche, N., Vikstrom, S., Sjostrom, R., Linden, S., Backstrom, A., Lundberg, C., Arnqvist, A., Mahdavi, J., Nilsson, U.J., Velapatino, B., Gilman, R.H., Gerhard, M., Alarcon, T., Lopez-Brea, M., Nakazawa, T., Fox, J.G., Correa, P., Dominguez-Bello, M.G., Perez-Perez, G.I., Blaser, M.J., Normark, S., Carlstedt, I., Oscarson, S., Teneberg, S., Berg, D.E., Boren, T.** 2004. Functional adaptation of

BabA, the *H. pylori* ABO blood group antigen binding adhesin. Science **305:**519-522.

14. **Auvray, F., Ozin, A.J., Claret, L., Hughes, C.** 2002. Intrinsic membrane targeting of the flagellar export ATPase FliI: interaction with acidic phospholipids and FliH. Journal of Molecular Biology **318:**941-950.
15. **Auvray, F., Thomas, J., Fraser, G. M., Hughes, C.** 2001. Flagellin polymerisation control by a cytosolic export chaperone. Journal of Molecular Biology **308:**221-229.
16. **Barnhart, M. M., Pinkner, J.S., Soto, G.E., Sauer, F.G., Langermann, S., Waksman, G., Frieden, C., Hultgren, S.J.** 2000. PapD-like chaperones provide the missing information for folding of pilin proteins. Proceedings of the National Academy of Sciences, USA **97:**7709-7714.
17. **Beier, D., Frank, R.** 2000. Molecular characterization of two-component systems of *Helicobacter pylori*. Journal of Bacteriology **182:**2068-2076.
18. **Bennett, J. C. Q., Thomas, J., Fraser, G. M., Hughes, C.** 2001. Substrate complexes and domain organization of the *Salmonella* flagellar export chaperones FlgN and FliT. Molecular Microbiology **39:**781-791.
19. **Berger, B. R., Christie, P.J.** 1994. Genetic complementation analysis of the *Agrobacterium tumefaciens virB* operon: *virB2* through *virB11* are essential virulence genes. Journal of Bacteriology **176:**3646-3660.
20. **Berks, B. C.** 1996. A common export pathway for proteins binding complex redox factors? Molecular Microbiology **22:**393-404.
21. **Binet, R., Wandersman, C. .** 1995. Protein secretion by hybrid bacterial ABC-transporters: specific functions of the membrane ATPase and the membrane fusion protein. EMBO Journal **14:**2298-2306.
22. **Bogden, C. E., Fass, D., Bergman, N., Nichols, M. D., Berger, J. M.** 1999. The structural basis for terminator recognition by the rho transcription termination factor. Molecular Cell **3:**487-493.
23. **Bohm, G., Muhr, R., Jaenicke, R.** 1992. Quantitative analysis of protein far UV circular dichroism, spectra by neural networks. Protein Engineering **5:**191-195.
24. **Bourzac, K. M., Guillemin, K.** 2005. *Helicobacter pylori*- host cell interactions mediated by type IV secretion. Cellular Microbiology **7:**911-919.
25. **Boyer, P. D.** 1997. The ATP Synthase-A Splendid Molecular Machine. Annual Review of Biochemistry **66:**717-749.
26. **Brahmachary, P., Dashti, M.G., Olson, J.W., Hoover, T.R.** 2004. *Helicobacter pylori* FlgR is an enhancer-independent activator of σ^{54}-RNA polymerase holoenzyme. Journal of Bacteriology **186:**4535-4542.
27. **Briercheck, D. M., Wood, T.C., Allison, T.J., Richardson, J.P., Rule, G.S.** 1998. The NMR structure of the RNA binding domain of *E. coli* Rho factor suggests possible RNA-protein interactions. Nature Structural Biology **5:**393-399.
28. **Bustamante, V. H., Martinez-Flores, I., Vlamakis, H. C., Zusman, D. R.** 2004. Analysis of the Frz signal transduction system of *Myxococcus xanthus* shows the importance of the conserved C-terminal region of the cytoplasmic chemoreceptor FrzCD in sensing signals. Molecular Microbiology **53:**1501-1513.
29. **Cheng, L. W., Anderson, D.M., Schneewind, O.** 1997. Two independent type III secretion mechanisms for YopE in *Yersinia enterocolitica*. Molecular Microbiology **24:**757-765.

30. **Christie, P. J.** 2004. Type IV secretion: the *Agrobacterium* VirB/D4 and related conjugation systems. Biochimica et Biophysica Acta **1694:**219-234.
31. **Claret, L., Calder, S.R., Higgins, M., Hughes, C.** 2003. Oligomerization and activation of the FliI ATPase central to bacterial flagellum assembly. Molecular Microbiology **48:**1349-1355.
32. **Claret, L., Hughes, C.** 2002. Interaction of the atypical prokaryotic transcription activator $FlhD_2C_2$ with early promoters of the flagellar gene hierarchy. Journal of Molecular Biology **321:**185-199.
33. **Claros, M. G., Brunak, S., von Heijne, G.** 1997. Prediction of N-terminal protein sorting signals. Current Opinion in Structural Biology **7:**394-398.
34. **Cornelius, G. R., Van Gijsegem, F.** 2000. Assembly and function of type III secretory systems. Annual Review of Microbiology **54:**735-774.
35. **Cristobal, S., de Gier, J-W., Nielsen, H., von Heijne, G.** 1999. Competition between Sec- and TAT-dependent protein translocation in *Escherichia coli*. EMBO Journal **18:**2982-2990.
36. **Daughdrill, G. W., Chadsey, M.S., Karlinsey, J.E., Hughes, K.T., Dahlquist, F.W.** 1997. The C-terminal half of the anti-sigma factor, FlgM, becomes structured when bound to its target, sigma 28. Nature Structural Biology **4:**285-291.
37. **De Reuse, H., Mendz, G., Ball, G., Labigne, A., Thiberg, J.M., Bury-Mone, S.** 2005. The β carbonic anhydrase is required for efficient colonisation of the mouse stomach by *Helicobacter pylori*. ASM 105th General Meeting **B-142**.
38. **Debarbieux, L., Wandersman, C.** 2001. Folded HasA inhibits its own secretion through its ABC exporter. EMBO Journal **20:**4657-4663.
39. **Delepelaire, P.** 2004. Type I secretion in gram-negative bacteria. Biochimica et Biophysica Acta **1694:**149-161.
40. **Ding, Z., Atmakuri, K., Christie, P. J.** 2003. The outs and ins of bacterial type IV secretion substrates. Trends in Microbiology **11:**527-535.
41. **Doig, P., De Jonge, B. L., Alm, R. A., Brown, E. D., Uria-Nickelsen, M., Noonan, B., Mills, S. D., Tummino, P., Carmel, G., Guild, B. C., Moir, D. T., Vovis, G. F., Trust, T. J.** 1999. *Helicobacter pylori* physiology predicted from genomic comparison of two strains. Microbiology and Molecular Biology Reviews **63:**675-707.
42. **Dombroski, A. J., Brennan, C.A., Spear, P., Platt, T.** 1988. Site-directed alterations of rho protein affect its activities as a termination factor. The Journal of Biological Chemistry **263:**18802-18809.
43. **Dreyfus, G., A. W. Williams, I. Kawagishi, and R. M. Macnab.** 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. Journal of Bacteriology. **175:**3131-3138.
44. **Driessen, A. J.** 1993. SecA, the peripheral subunit of the *Escherichia coli* precursor protein translocase, is functional as a dimer. Biochemistry **32:**13190-13197.
45. **Duong, F.** 2003. Binding, activation and dissociation of the dimeric SecA ATPase at the dimeric SecYEG translocase. EMBO Journal **22:**4375-4384.
46. **Durand, E., Bernadac, A., Ball, G., Lazdunski, A., Sturgis, J.N., Filloux, A.** 2003. Type II protein secretion in *Pseudomonas aeruginosa*: the pseudopilus is a multifibrillar and adhesive structure. Journal of Bacteriology **185:**2749-2758.

47. **Eaton, E., Brooks,C.l., Morgan, D.R., and S. Krakowka.** 1991. Essential role of urease in pathogenisis of gastitis Induced by *Helicobacter pylori* in gnotobiotic piglets. Infection and Immunity **59:**2470 - 475.
48. **Eaton, K. A., Morgan, D.R., Krakowka, S.** 1989. *Campylobacter pylori* virulence factors in gnotobiotic piglets. Infection and Immunity **57:**1119-25.
49. **Eaton, K. A., Morgan, D.R., Krakowka, S.** 1992. Motility as a factor in the colonisation of gnotobiotic piglets by *Helicobacter pylori.* Journal of Medical Microbiology **37:**123-127.
50. **Eaton, K. A., Suerbaum, S., Josenhans, C., Krakowka, S.** 1996. Colonization of gnotobiotic piglets by *Helicobacter pylori* deficient in two flagellin genes. Infection and Immunity **64:**2445-48.
51. **Fan, F., K. . Ohnishi, N. R. Francis, and R. M. Macnab.** 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. Molecular Biology **26:**1035 - 1046.
52. **Fan, F., Macnab, R. M.** 1996. Enzymatic characterization of FliI. The Journal of Biological Chemistry **271:**31981-31988.
53. **Fernandez, L. A., Berenguer, J.** 2000. Secretion and assembly of regular surface structures in Gram-negative bacteria. FEMS Microbiology Reviews **24:**21-44.
54. **Ferris, H. U., Furukawa, Y., Minamino, T., Kroetz, M.B., Kihara, M., Namba, K., Macnab, R.M.** 2005. FlhB regulates ordered export of flagellar components via autocleavage mechanism. The Journal of Biological Chemistry **280:**41236-41242.
55. **Fillingame, R. H., Dmitriev, O.Y.** 2002. Structural model of the transmembrane F_0 rotary sector of H^+-transporting ATP synthase derived by solution NMR and intersubunit cross-linking in situ. Biochimica et Biophysica Acta **1565:**232-245.
56. **Filloux, A.** 2004. The underlying mechanisms of type II protein secretion. Biochimica et Biophysica Acta **1694:**163-179.
57. **Fischer, W., Puls, J., Buhrdorf, R., Gebert, B., Odenbreit, S., Haas, R.** 2001. Systematic mutagenesis of the *Helicobacter pylori* cag pathogenicity island: essential genes for CagA translocation in host cells and induction of interleukin-8. Molecular Microbiology **42:**1337-1348.
58. **Fraser, G. M., Gonzalez-Pedrajo, B., Tame, J. R. H., Macnab, R. M.** 2003. Interactions of FliJ with the *Salmonella* type III flagellar export apparatus. Journal of Bacteriology **185:**5546-5554.
59. **Fraser, G. M., Hirano, T., Ferris, H.U., Devgan, L.L., Kihara, M., Macnab, R.M.** 2003. Substrate specificity of type III flagellar protein export in *Salmonella* is controlled by subdomain interactions in FlhB. Molecular Microbiology **48:**1043-1057.
60. **Galan, J. E.** 2001. *Salmonella* interactions with host cells: type III secretion at work. Annual Review of Cell and Developmental Biology **17:**53-86.
61. **Geis, G., Leying, H., Suerbaum, S., Mai, U., Opferkuch, W.** 1989. Ultrastructure and chemical analysis of *Campylobacter pylori* flagella. Journal of Clinical Microbiology **27:**436-441.
62. **Gierasch, L. M.** 1989. Signal sequences. Biochemistry **28:**923-930.
63. **Gogol, E. P., Seifried, S.E., von Hippel, P.H.** 1991. Structure and assembly of the *Escherichia coli* transcription termination factor Rho and its interaction with RNA. I. Cryoelectron microscopic studies. Journal of Molecular Biology **221:**1127-1138.

64. **Gonzalez-Pedrajo, B., Fraser, G. M., Minamino, T., Macnab, R. M.** 2002. Molecular dissection of *Salmonella* FliH, a regulator of the ATPase FliI and the type III flagellar protein export pathway. Molecular Microbiology **45:**967-982.

65. **Gophna, U., Ron, E.Z., Graur, D.** 2003. Bacterial type III secretion systems are ancient and evolved by multiple horizontal transfer events. Gene **312:**151-163.

66. **He, S. Y., Nomura, K., Whittam, T.S.** 2004. Type III protein secretion mechanism in mammalian and plant pathogens. Biochimica et Biophysica Acta **1694:**181-206.

67. **Henderson, I. R., Navarro-Garcia, F., Nataro, J.P.** 1998. The great esape: structure and function of the autotransporter proteins. Trends in Microbiology **6:**370-378.

68. **Heuermann, D., and R. Haas.** 1998. A stable shuttle vector system for efficient genetic complementation of *Helicobacter pylori* strains by transformation and conjugation. Molecualr and General Genetics **257:**519-528.

69. **Holland, I. B., Blight, M.A.** 1999. ABC-ATPases, adaptable energy generators fuelling transmembrane movement of a variety of molecules in organisms from bacteria to humans. Journal of Molecular Biology **293:**381-399.

70. **Homma, M., DeRosier, D. J., Macnab, R. M.** 1990. Flagellar hook and hook-associated proteins of *Salmonella typhimurium* and their relationship to other axial components of the flagellum. Journal of Molecular Biology **213:**819-832.

71. **Homma, M., Kutsukake, K., Hasebe, M., Iino, T., Macnab, R. M.** 1990. FlgB, FlgC, FlgF and FlgG. A family of structurally related proteins in the flagellar basal body of *Salmonella typhimurium.* Journal of Molecular Biology **211:**465-477.

72. **Hong, W., Sano, K., Morimatsu, S., Scott, D.R., Weeks, D.L., Sachs, G., Goto, T., Mohan, S., Harada, F., Nakajima, N., Nakano, T.** 2003. Medium pH dependent redistribution of the urease of *Helicobacter pylori.* Journal of Medical Microbiology **52:**211-216.

73. **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:**1277-1280.

74. **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* adehesin binding fucosylated histo-blood group antigens revealed by retagging. Science **279:**373-377.

75. **Irikura, V. M., Kihara, M., Yamaguchi, S., Sockett, H., Macnab, R.M.** 1993. *Salmonella typhimurium fliG* and *fliN* mutations causing defects in assembly, rotation, and switching of the flagellar motor. Journal of Bacteriology **175:**802-810.

76. **Jacob-Dubuisson, F., Fernandez, R., Coutte, L.** 2004. Protein secretion through autotransporter and two-partner pathways. Biochimica et Biophysica Acta **1694:**235-257.

77. **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 Microbiology Letters **152:**205-211.

78. **Johnson, C. W. J.** 1989. Protein secondary structure and Circular Dichroism: A practical guide. Proteins: Structure, Function, and Genetics **7:**205-214.

79. **Jones, C. J., and R. M. Macnab.** 1990. Flagellar assembly in *Salmonella typhimurium*: analysis with temperature sensitive mutants. Journal of bacteriology. **172:**1327-1339.

80. **Karamanou, S., Vrontou, E., Sianidis, G., Baud, C., Roos, T., Kuhn, A., Politou, A.S., Economou, A.** 1999. A molecular switch in SecA protein couples ATP hydrolysis to protein translocation. Molecular Microbiology **34:**1133-1145.

81. **Karlinsey, J. E., Lommer, J., Brown, K. L., Hughes, K. T.** 2000. Translation/secretion coupling by type III secretion systems. Cell **102:**487-497.

82. **Katayama, E., Shiraishi, T., Oosawa, K., Baba, N., Aizawa, S-I.** 1996. Geometry of the flagellar motor in the cytoplasmic membrane of *Salmonella typhimurium* as determined by stereophotogrammetry of quick-freeze deep-etch replica images. Journal of Molecular Biology **255:**458-475.

83. **Kihara, M., Minamino, T., Yamaguchi, S., Macnab, R.M.** 2001. Intergenic suppression between the flagellar MS ring protein FliF of *Salmonella* and FlhA, a membrane component of its export apparatus. Journal of Bacteriology **183:**1655-1662.

84. **Kim, D.-E., Patel, S. S.** 2001. The kinetic pathway of RNA binding to the *Escherichia coli* transcription termination factor Rho. The Journal of Biological Chemistry **276:**13902-13910.

85. **Konninger, U. W., Hobbie, S., Benz, R., Braun, V.** 1999. The haemolysin-secreting ShlB protein of the outer membrane of *Serratia marcescens*: determination of surface-exposed residues and formation of ion-permeable pores by ShlB mutants in artificial lipid bilayer membranes. Molecular Microbiology **32:**1212-1225.

86. **Koronakis, V., Hughes, C., Koronakis, E.** 1993. ATPase activity and ATP/ADP-induced conformational change in the soluble domain of the bacterial protein translocator HlyB Molecular Microbiology **8:**1163-1175.

87. **Koronakis, V., Sharff, A., Koronakis, E., Luisi, B., Hughes, C.** 2000. Crystal structure of the bacterial membrane protein TolC central to multidrug efflux and protein export. Nature **405:**914-919.

88. **Krause, S., Barcena, M., Pansegrau, W., Lurz, R., Carazo, J. M., Lanka, E.** 2000. Sequence-related protein export NTPases encoded by the conjugative transfer region of RP4 and by the *cag* pathogenicity island of *Helicobacter pylori* share similar hexameric ring structures. Proceedings of the National Academy of Sciences, USA **97:**3067-3072.

89. **Kubori, T., Matsushima, Y., Nakamura, D., Uralil, J., Lara-Tejero, M., Sukhan, A., Galan, J.E., Aizawa, S-I.** 1998. Supramolecular structure of the *Salmonella typhimurium* type III protein secretion system. Science **280:**1998.

90. **Kubori, T., Shimamoto, N., Yamaguchi, S., Namba, K., Aizawa, S.I.** 1992. Morphological pathway of flagellar assembly in *Salmonella typhimurium*. Journal of Molecular Biology **226:**433-446.

91. **Kubori, T., Sukhan, A., Aizawa, S.-I., Galan, J. E.** 2000. Molecular characterization and assembly of the needle complex of the *Salmonella typhimurium* type III protein secretion system. Proceedings of the National Academy of Sciences, USA **97:**10225-10230.

92. **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. Journal of Bacteriology **179:**813-817.

93. **Kutsukake, K., Ikebe, T., Yamamoto, S.** 1999. Two novel regulatory genes, *fliT* and *fliZ*, in the flagellar regulon of *Salmonella*. Genes and Genetic Systems **74:**287-292.

94. **Kutsukake, K., Iyoda, S., Ohnishi, K., Iino, T.** 1994. Genetic and molecular analyses of the interaction between the flagellum-specific sigma and anti-sigma factors in *Salmonella typhimurium*. The EMBO Journal **13:**4568-4576.

95. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227:**680-685.

96. **Lane, M. C., O'Toole, P.W., Moore, S.A.** 2006. Molecular basis of the interaction between the flagellar export apparatus proteins FliI and FliH from *Helicobacter pylori*. Journal of Biological Chemistry **281:**508-517.

97. **Lee, S. H., Galan, J. E.** 2004. *Salmonella* type III secretion-associated chaperones confer secretion-pathway specificity. Molecular Microbiology **51:**483-495.

98. **Li, H., Qian, L., Chen, Z., Thibault, D., Liu, G., Liu, T., Thanassi, D.G.** 2004. The outer membrane usher forms a twin-pore secretion complex. Journal of Molecular Biology **344:**1397-1407.

99. **Liu, X., and P. Matsumura.** 1994. The FlhD/FlhC complex, a transcriptional activator of the *Escherichia coli* flagellar class II operons. Journal of Bacteriology **176:**7345-7351.

100. **Lloyd, S. A., Norman, Rosqvist, R., Wolf-Watz, H.** 2001. *Yersinia* YopE is targeted for type III secretion by N-terminal, not mRNA, signals. Molecular Microbiology **39:**520-531.

101. **Lloyd, S. A., Tang, H., Wang, X., Billings, S., Blair, D.F.** 1996. Torque generation in the flagellar motor of *Escherichia coli*: evidence of a direct role for FliG but not for FliM or FliN. Journal of Bacteriology **178:**223-231.

102. **Luirink, J., Sinning, I.** 2004. SRP-mediated protein targeting: structure and function revisited. Biochimica et Biophysica Acta **1694:**17-35.

103. **Lux, R., Kar, N., Khan, S.** 2000. Overproduced *Salmonella typhimurium* flagellar motor switch complexes. Journal of Molecular Biology **298:**577-583.

104. **Macnab, R. M.** 1996. Flagella and Motility, p. 123-145. *In* F. C. Neidhardt (ed.), *Escherichia coli* and *Salmonella typhimurium*: Cellular and Molecular Biology. ASM, Washington, DC.

105. **Macnab, R. M.** 1992. Genetics and biogenesis of bacterial flagella. Annual Review of Genetics **26:**131-58.

106. **Mahdavi, J., Sonden, B., Hurtig, M., Olfat, F.O., Forsberg, L., Roche, N., Angstrom, J., Larsson, T., Teneberg, S., Karlsson, K-A., Altraja, S., Wadstrom, T., Kersulyte, D., Berg, D. E., Dubois, A., Petersson, C., Magnusson, K-E., Norberg, T., Lindh, F., Lundskog, B. B., Arnqvist, A., Hammarstrom, L., Boren, T.** 2002. *Helicobacter pylori* SabA adhesin in persistent infection and chronic inflammation. Science **297:**573-578.

107. **Makishima, S., Komoriya, K., Yamaguchi, S., Aizawa, S-I.** 2001. Length of the flagellar hook and the capacity of the type III export apparatus. Science **291:**2411-2413.

108. **Marcus, E. A., Moshfegh, A.P., Sachs, G., Scott, D.R.** 2005. The periplasmic α-carbonic anhydrase activity of *Helicobacter pylori* is essential for acid acclimation. Journal of Bacteriology **187:**729-738.

109. **Marlovits, T. C., Kubori, T., Sukhan, A., Thomas, D.R., Galan, J.E., Unger, V.M.** 2004. Structural insights into the assembly of the type III secretion needle complex. Science **306:**1040-1042.

110. **Marshall, B.** 2002. *Helicobacter pylori*: 20 years on. Clinical Medicine **2:**147-152.

111. **Marshall, B. J., and J. R. Warren.** 1984. Unidentified curved bacilli in the stomach of patients with gastritis and peptic ulceration. The Lancet.:1311-14.

112. **Marykwas, D. L., Schmidt, S.A., Berg, H.C.** 1996. Interacting components of the flagellar motor of *Escherichia coli* revealed by the two-hybrid system in yeast. Journal of Molecular Biology **256:**564-576.

113. **Menard, R., Dehio, C., Sansonetti, P.J.** 1996. Bacterial entry into epithelial cells: the paradigm of *Shigella*. Trends in Microbiology **4:**220-226.

114. **Minamino, T., Chu, R., Yamaguchi, S., Macnab, R. M.** 2000. Role of FliJ in flagellar protein export in *Salmonella*. Journal of Bacteriology **182:**4207-4215.

115. **Minamino, T., Gonzalez-Pedrajo, B., Kihara, M., Namba, K., Macnab, R.M.** 2003. The ATPase FliI can interact with the type III flagellar protein export apparatus in the absence of its regulator, FliH. Journal of Bacteriology **185:**3983-3988.

116. **Minamino, T., Gonzalez-Pedrajo, B., Oosawa, K., Namba, K., Macnab, R. M.** 2002. Structural properties of FliH, an ATPase regulatory component of the *Salmonella* type III flagellar export apparatus. Journal of Molecular Biology **322:**281-290.

117. **Minamino, T., Macnab, R. M.** 1999. Components of the *Salmonella* flagellar export apparatus and classification of export substrates. Journal of Bacteriology **181:**1388-1394.

118. **Minamino, T., Macnab, R. M.** 2000. FliH, a soluble component of the type III flagellar export apparatus of *Salmonella*, forms a complex with FliI and inhibits its ATPase activity. Molecular Microbiology **37:**1494-1503.

119. **Minamino, T., Macnab, R. M.** 2000. Interactions among components of the *Salmonella* flagellar export apparatus and its substrates. Molecular Microbiology **35:**1052-1064.

120. **Minamino, T., Saijo-Hamano, Y., Furukawa, Y., Gonzalez-Pedrajo, B., Macnab, R.M., Namba, K.** 2004. Domain organisation and function of *Salmonella* FliK, a flagellar hook-length control protein. Journal of Molecular Biology **341:**491-502.

121. **Minamino, T., Tame, J. R. H., Namba, K., Macnab, R. M.** 2001. Proteolytic analysis of the FliH/FliI complex, the ATPase component of the Type III flagellar export apparatus of *Salmonella*. Journal of Molecular Biology **312:**1027-1036.

122. **Miwa, Y., Horiguchi, T., Shigesada, K.** 1995. Structural and functional dissections of transcription termination factor rho by random mutagenesis. Journal of Molecular Biology **254:**815-837.

123. **Mobley, H. L., Island, M.D., Hausinger, R.P.** 1995. Molecular biology of microbial ureases. Microbiological Reviews **59:**451-480.

124. **Molinari, M., Galli, C., Norais, N., Telford, J.L., Rappuoli, R., Luzio, J.P., Montecucco, C.** 1997. Vacuoles induced by *Helicobacter pylori* toxin contain both late endosomal and lysosomal markers. Journal of Biological Chemistry **272:**25339-25344.

125. **Molinari, M., Salio, M., Galli, C., Norais, N., Rappuoli, R., Lanzavecchia, A., Montecucco, C.** 1998. Selective inhibition of Ii-dependent antigen presentation by *Helicobacter pylori* toxin VacA. Journal of Experimental Medicine **187:**135-140.

126. **Montecucco, C., de Bernard, M.** 2003. Molecular and cellular mechanisms of action of the vacuolating cytotoxin (VacA) and neutrophil-activating protein (HP-NAP) virulence factors of *Helicobacter pylori*. Microbes and Infection **5:**715-721.

127. **Morgan, W. D., Bear, D.G., Litchman, B.L., von Hippel, P.H.** 1985. RNA sequence and secondary structure requirements for rho-dependent transcription termination. Nucleic acids research **13:**3739-3754.

128. **Muramoto, K., Makishima, S., Aizawa, S-I., Macnab, R.M.** 1999. Effect of hook subunit concentration on assembly and control of length of the flagellar hook of *Salmonella*. Journal of Bacteriology **181:**5808-5813.

129. **Ng, T. W., Akman, L., Osisami, M., Thanassi, D.G.** 2004. The usher N-terminus is the initial targeting site for chaperone-subunit complexes and participates in subsequent pilus biogenesis events. Journal of Bacteriology **186:**5321-5331.

130. **Niehus, E., Gressmann, H., Ye, F., Schlapbach, R., Dehio, M., Dehio, C., Stack, A., Meyer, T.F., Suerbaum, S., Josenhans, C.** 2004. Genome-wide analysis of transcriptional hierarchy and feedback regulation in the flagellar system of *Helicobacter pylori*. Molecular Microbiology **52:**947-961.

131. **Noji, H., Yasuda, R., Yoshida, M., Kinoshita, K, Jr.** 1997. Direct observation of the rotation of F_1-ATPase. Nature **386:**299-302.

132. **O'Toole, P. W., Lane, M. C., Porwollik, S.** 2000. *Helicobacter pylori* motility. Microbes and Infection **2:**1-8.

133. **Ohnishi, I., Kutsukake, K., Suzuki, H., Iino, T.** 1992. A novel transcriptional regulatory mechanism in the flagellar regulon of *Salmonella typhimurium*: an anti-sigma factor inhibits the activity of the flagellum-specific sigma factor σ F. Molecular Microbiology **6:**3149-3157.

134. **Ohnishi, K., F. Fan, G. J. Schoenhals, M. Kihaea, and R. M. Macnab.** 1997. The FliO, FliP, FliQ, and FliR Proteins of *Salmonella typhimurium:* Putative Components for Flagellar Assembly. Journal of Bacteriology **179:**6092 - 6099.

135. **Ohnishi, K., K. Kutsukake, H. Suzuki, and T. Iino.** 1990. Gene *fliA* encodes an alternative sigma factor specific for flagellar operons in *Salmonella typhimurium*. Molecular and General Genetics **221:**139-147.

136. **Olfat, F. O., Zheng, Q., Oleastro, M., Voland, P., Boren, T., Karttunen, R., Engstrand, L., Rad, R., Prinz, C., Gerhard, M.** 2005. Correlation of the *Helicobacter pylori* adherence factor BabA with duodenal ulcer disease in four European countries. FEMS Immunology and Medical Microbiology **44:**151-156.

137. **Oomen, C., van Ulsen, P., van Gelder, P., Feijen, M., Tommassen, J., Gros, P.** 2004. Structure of the translocator domain of a bacterial autotransporter. EMBO Journal **23:**1257-1266.

138. **Owen, R. J., Xerry, J.** 2003. Tracing clonality of *Helicobacter pylori* infecting family members from analysis of DNA sequences of three housekeeping genes (*ureI*, *atpA* and *ahpC*), deduced amino acid sequences, and pathogenicity associated markers (*cagA* and *vacA*). Journal of Medical Microbiology **52:**515-524.

139. **Ozin, A. J., Claret, L., Auvray, F., Hughes, C.** 2003. The FliS chaperone selectively binds the disordered flagellin C-terminal D0 domain central to polymerisation. FEMS Microbiology Letters **219:**219-224.

140. **Patterson-Delafield, J., Martinez, R.J., Stocker, B.A., Yamaguchi, S.** 1973. A new *fla* gene in *Salmonella typhimurium -flaR-* and its mutant phenotype-superhooks. Archiv fur Mikrobiologie **90:**107-120.

141. **Porcelli, I., de Leeuw, E., Wallis, R., van den Brink-van der Laan, E., de Kruijff, B., Wallace, B.A., Palmer, T., Berks, B.C.** 2002. Characterization and membrane assembly of the TatA component of the *Escherichia coli* twin-arginine protein transport system. Biochemistry **41:**13690-13697.

142. **Porwollik, S., B. Noonan, and P. W. O'Toole.** 1999. Molecular Characterization of a Flagellar Export Locus of *Helicobacter pylori*. Infection and Immuity **67:**2060 -70.

143. **Ramamurthi, K. S., Schneewind, O.** 2003. *Yersinia yopQ* mRNA encodes a bipartite type III secretion signal in the first 15 codons. Molecular Microbiology **50:**1189-1198.

144. **Rashkova, S., Spudich, G.M., Christie, P.J.** 1997. Mutational analysis of the *Agrobacterium tumefaciens* VirB11 ATPase: identification of functional domains and evidence for multimerisation. Journal of Bacteriology **179:**583-589.

145. **Richardson, L. V., Richardson, J. P.** 1996. Rho dependent termination of transcription is governed primarily by the upstream Rho utilization (*rut*) sequences of a terminator. The Journal of Biological Chemistry **271:**21597-21603.

146. **Rieder, G., Fischer, W., Haas, R.** 2005. Interaction of *Helicobacter pylori* with host cells: function of secreted and translocated molecules. Current Opinion in Microbiology **8:**67-73.

147. **Robien, M. A., Krumm, B.E., Sandkvist, M., Hol, W.G.J.** 2003. Crystal structure of the extracellular protein secretion NTPase EpsE of *Vibrio cholerae.* Journal of Molecular Biology **333:**657-674.

148. **Robinson, C., Bolhuis, A.** 2001. Protein targeting by the twin-arginine translocation pathway. Nature Reviews **2:**350-356.

149. **Rohde, M., Puls, J., Buhrdorf, R., Fischer, W., Haas, R.** 2003. A novel sheathed surface organelle of the *Helicobacter pylori cag* type IV secretion system. Molecular Microbiology **49:**219-234.

150. **Rost, B., Sander, C.** 1993. Journal of Molecular Biology **232:**584-599.

151. **Ryan, K. A., Karim, N., Worku, M., Penn, C.W., O'Toole, P.W.** 2005. *Helicobacter pylori* flagellar hook-filament transition is controlled by a FliK functional homolog encoded by the gene HP0906. Journal of Bacteriology **187:**5742-5750.

152. **Sagi, Y., Khan, S., Eisenbach, M.** 2003. Binding of the chemotaxis response regulator CheY to the isolated, intact switch complex of the bacterial flagellar motor. The Journal of Biological Chemistry **278:**25867-25871.

153. **Sagulenko, Y., Sagulenko, V., Chen, J., Christie, P.J., .** 2001. Role of *Agrobacterium* VirB11 ATPase in T-pilus assembly and substrate selection. Journal of Bacteriology **183:**5813-5825.

154. **Samatey, F. A., Matsunami, H., Imada, K., Nagashima, S., Shaikh, T.R., Thomas, D.R., Chen, J.Z., DeRosier, D.J., Kitao, A., Namba, K.** 2004. Structure of the bacterial flagellar hook and implication for the molecular universal joint mechanism. Nature **431:**1062-1068.

155. **Sargent, F., Gohlke, U., de Leeuw, E., Stanley, N.R., Palmer, T., Saibil, H.R., Berks, B.C.** 2001. Purified components of the *Escherichia coli* Tat protein transport

system form a double-layered ring structure. European Journal of Biochemistry **268:**3361-3367.

156. **Sauer, F. G., Barnhart, M., Choudhury, D., Knight, S.D., Waksman, G., Hultgren, S.J.** 2000. Chaperone-assisted pilus assembly and bacterial attachment. Current Opinion in Structural Biology **10:**548-556.
157. **Sauer, F. G., Pinkner, J.S., Waksman, G., Hultgren, S.J.** 2002. Chaperone priming of pilus subunits facilitates a topological transition that drives fiber formation. Cell **111:**543-551.
158. **Saulino, E. T., Bullitt, E., Hultgren, S.J.** 2000. Snapshots of usher-mediated protein secretion and ordered pilus assembly. Proceedings of the National Academy of Sciences, USA **97:**9240-9245.
159. **Savvides, S. N., Yeo, H-J., Beck, M.R., Blaesing, F., Lurz, R., Lanka, E., Buhrdorf, R., Fischer, W., Haas, R., Waksman, G.** 2003. VirB11 ATPases are dynamic hexameric assemblies: new insights into bacterial type IV secretion. The EMBO Journal **22:**1969-1980.
160. **Schesser, K., Frithz-Lindsten, E., Wolf-Watz, H.** 1996. Delineation and mutational analysis of the *Yersinia pseudotuberculosis* YopE domains which mediate translocation across bacterial and eukaryotic cellular membranes. Journal of Bacteriology **178:**7227-7233.
161. **Schmitt, L., Benabdelhak, H., Blight, M.A., Holland, I.B., Stubbs, M.T.** 2003. Crystal structure of the nucleotide-binding domain of the ABC-transporter haemolysin B: identification of a variable region within ABC helical domains. Journal of Molecular Biology **330:**333-342.
162. **Schmitt, W., Odenbreit, S., Heuermann, D., Haas, R.** 1995. Cloning of the *Helicobacter pylori recA* gene and functional characterization of its product. Molecular and General Genetics **248:**563-572.
163. **Scott, D. R., Marcus, E.A., Weeks, D.L., Lee, A., Melchers, K., Sachs, G.** 2000. Expression of the *Helicobacter pylori ureI* gene is required for acidic pH activation of cytoplasmic urease. Infection and Immunity **68:**470-477.
164. **Skordalakes, E., Berger, J. M.** 2003. Structure of the Rho transcription terminator: mechanism of mRNA recognition and helicase loading. Cell **114:**135-146.
165. **Skouloubris, S., Thiberge, J.-M., Labigne, A., De Reuse, H.** 1998. The *Helicobacter pylori* UreI protein is not involved in urease activity but is essential for bacterial survival in vivo. Infection and Immunity **66:**4517-4521.
166. **Smith, D. B., Johnson, K.S.** 1988. Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. Gene **67:**31-40.
167. **Sockett, H., Yamaguchi, S., Kihara, M., Irikura, V.M., Macnab, R.M.** 1992. Molecular analysis of the flagellar switch protein FliM of *Salmonella typhimurium.* Journal of Bacteriology **174:**793-806.
168. **Sory, M. P., Boland, A., Lambermont, I., Cornelius, G. R.** 1995. Identification of the YopE and YopH domains required for secretion and internalization into the cytosol of macrophages, using the *cyaA* gene fusion approach. Proceedings of the National Academy of Sciences, USA **92:**11998-12002.
169. **Spohn, G., Scarlato, V.** 1999. Motility of *Helicobacter pylori* is coordinately regulated by the transcriptional activator FlgR, an NtrC homolog. Journal of Bacteriology **181:**593-599.

170. **Stanley, N. R., Palmer, T., Berks, B.C.** 2000. The twin arginine consensus motif of Tat signal peptide is involved in Sec-independent protein targeting in *Escherichia coli*. The Journal of Biological Chemistry **275:**11591-11596.

171. **Stein, M., Bagnoli, F., Halenbeck, R., Rappuoli, R., Fantl, W.J., Covacci, A.** 2002. c-Src/Lyn kinases activate *Helicobacter pylori* CagA through tyrosine phosphorylation of the EPIYA motifs. Molecular Microbiology **43:**971-980.

172. **Stephens, C., C. Mohr, C. Boyd, J. Maddock, J. Gober, and L. Shapiro.** 1997. Identification of the *fliI* and *fliJ* Componenents of the *Caulobacter* Flagellar Type III Protein Secretion System. Journal of Bacteriology **179:**5355-5365.

173. **Stitt, B. L.** 2001. *Escherichia coli* transcription termination factor Rho binds and hydrolyzes ATP using a single class of three sites. Biochemistry **40:**2276-2281.

174. **Stitt, B. L., Xu, Y.** 1998. Sequential hydrolysis of ATP molecules bound in interacting catalytic sites of *Escherichia coli* transcription termination protein rho. The Journal of Biological Chemistry **273:**26477-26486.

175. **Suerbaum, S., C. Josenhans, and A. Labigne.** 1993. Cloning and Genetic Characterization of the *Helicobacter pylori* and *Helicobacter mustalae flaB* Flagellin Genes and Construction of *H. pylori flaA-* and *flaB*-Negative Mutants by Electroporation-Mediated Allelic Exchange. Journal of Bacteriology. **175:**3278-3288.

176. **Sukhan, A., Kubori, T., Wilson, J., Galan, J. E.** 2001. Genetic analysis of assembly of the *Salmonella enterica* serovar typhimurium type III secretion-associated needle complex. Journal of Bacteriology **183:**1159-1167.

177. **Suzuki, H., Yonekura, K., Murata, K., Hirai, T., Oosawa, K., Namba, K.** 1998. A structural feature in the central channel of the bacterial flagellar FliF ring complex is implicated in type III protein export. Journal of Structural Biology **124:**104-114.

178. **Suzuki, H., Yonekura, K., Namba, K.** 2004. Structure of the rotor of the bacterial flagellar motor revealed by electron cryomicroscopy and single-particle image analysis. Journal of Molecular Biology **337:**105-113.

179. **Szabo, I., Brutsche, S., Tombola, F., Moschioni, M., Satin, B., Telford, J.L., Rappuoli, R., Montecucco, C., Papini, E., Zoratti, M.** 1999. Formation of anion-selective channels in the cell plasma membrane by the toxin VacA of *Helicobacter pylori* is required for its biological activity. EMBO Journal **18:**5517-5527.

180. **Terry, K., Williams, S.M., Connolly, L., Ottemann, K.M.** 2005. Chemotaxis plays multiple roles during *Helicobacter pylori* animal infection. Infection and Immunity **73:**803-811.

181. **Thanabalu, T., Koronakis, E., Hughes, C., Koronakis, V.** 1998. Substrate-induced assembly of a contiguous channel for protein export from *E. coli*: reversible bridging of an inner-membrane translocase to an outer membrane exit pore EMBO Journal **17:**6487-6496.

182. **Tomb, J. F., O. White, A. R. Kerlavage, R. A. Clayton, G. G. Sutton, R. D. Fleischmann, K. A. Ketchum, H. P. Klenk, S. Gill, B. A. Dougherty, K. Nelson, J. Quakenbush, L. Zhou, E. F. Kirkness, S. Peterson, B. Loftus, D. Richardson, R. Dodson, H. G. Khalak, A. Glodek, K. McKenney, L. M. Fitzgerald, N. Lee, M. D. Adams, E. K. Hickey, D. E. Berg, J. D. Gocayne, T. R. Utterback, J. D. Peterson, J. M. Kelley, M. D. Cotton, J. M. Weidman, C. Fujii, C. Bowman, L. Watthey, E. Wallin, W. S. Hayes, M. Borodovsky, P. D. Karp, H. O. Smith, C.**

M. Fraser, and J. C. Venter. 1997. The complete genome sequence of the gastric pathogen *Helicobacter pylori*. Nature **388:**539-547.

183. **Tombola, F., Carlesso, C., Szabo, I., de Bernard, M., Reyrat, H.M., Telford, J.L., Rappuoli, R., Montecucco, C., Papini, E., Zoratti, M.** 1999. *Helicobacter pylori* vacuolating toxin forms anion-selective channels in planar lipid bilayers: possible implications for the mechanism of cellular vacuolation. Biophysical Journal **96:**1401-1409.
184. **Umehara, S., Higashi, H., Ohnishi, N., Asaka, M., Hatakeyama, M.** 2003. Effects of *Helicobacter pylori* CagA protein on the growth and survival of B lymphocytes, the origin of MALT lymphoma. Oncogene **22:**8337-8342.
185. **Vakharia, H., German, G.J., Misra, R.** 2001. Isolation and characterization of *Escherichia coli tolC* mutants defective in secreting enzymatically active alpha-hemolysin. Journal of Bacteriology **183:**6908-6916.
186. **Veenendaal, A. K. J., van der Does, C., Driessen, A. J. M.** 2004. The protein-conducting channel SecYEG. Biochimica et Biophysica Acta **1694:**81-95.
187. **Vignon, G., Kohler, R., Larquet, E., Giroux, S., Prevost, M.C., Roux, P., Pugsley, A.P.** 2003. Type IV-like pili formed by the type II secreton: specificity, composition bundling, polar localization, and surface presentation of peptides. Journal of Bacteriology **185:**3416-3428.
188. **Vogler, A., M. Homma, V. M. Irikura, and R. M. Macnab.** 1991. *Salmonella typhimurium* Mutants Defective in Flagellar Filament Regrowth and Sequence Similarity of FliI to F_0 F_1 Vacuolar and Archaebacterial ATPase Subunits. Journal of Bacteriology **173:**3564 - 72.
189. **Voland, P., Weeks, D.L., Marcus, E.A., Prinz, C., Sachs, G., Scott, D.R.** 2002. Interactions among the seven *Helicobacter pylori* proteins encoded by the urease gene cluster. American Journal of Physiology- Gastrointestinal and Liver Physiology **284:**G96-G106.
190. **Vrontou, E., Economou, A.** 2004. Structure and function of SecA, the perprotein translocase nanomotor. Biochimica et Biophysica Acta **1694:**67-80.
191. **Wang, H. W., Chen, Y., Yang, H., Chen, X., Duan, M.X., Tai, P.C., Sui, S.F.** 2003. Ring-like pore structures of SecA: implication for bacterial protein-conducting channels. Proceedings of the National Academy of Sciences, U.S.A **100:**4221-4226.
192. **Wattiau, P., B. Bernier, P. Deslee, T. Michiels, and G. R. Cornelius.** 1994. Individual chaperones required for Yop secretion by *Yersinia*. Proceedings of the National Academy of Sciences USA **91:**10493-10497.
193. **Weber, J., Muharemagic, A., Wilke-Mounts, S., Senior, A. E.** 2003. F_1F_0-ATP Synthase. Binding of δ subunit to a 22-residue peptide mimicking the N-terminal region of α subunit. The Journal of Biological Chemistry **278:**13623-13626.
194. **Weeks, D. L., Eskandari, S., Scott, D.R., Sachs, G.** 2000. A H^+-gated urea channel: the link between *Helicobacter pylori* urease and gastric colonisation. Science **287:**482-485.
195. **Wei, R. R., Richardson, J. P.** 2001. Mutational changes of conserved residues in the Q-loop region of the transcription factor Rho greatly reduce secondary site RNA binding. Journal of Molecular Biology **314:**1007-1015.

196. **Wei, R. R., Richardson, J.P.** 2001. Identification of an RNA-binding site in the ATP binding domain of *Escherichia coli* Rho by H_2O_2/Fe-EDTA cleavage protection studies Journal of Biological Chemistry **276:**28380-28387.

197. **Weiner, J. H., Bilous, P.T., Shaw, G.M., Lubitz, S.P., Frost, L., Thomas, G.H., Cole, J.A., Turner, R.J.** 1998. A novel and ubiquitous system for membrane targeting and secretion of cofactor-containing proteins. Cell **93:**93-101.

198. **Wickner, W., Schekman, R.** 2005. Protein translocation across biological membranes. Science **310:**1452-1456.

199. **Wolf, E., Kim, P. S., Berger, B.** 1997. Multicoil: A program for predicting two- and three- stranded coiled coils. Protein Science **6:**1179-1189.

200. **Woodbury, R. L., Hardy, S.J., Randall, L.L.** 2000. Complex behaviour in solution of homodimeric SecA. Protein Science **11:**875-882.

201. **Wu, J., A. K. Benson, and A. Newton.** 1995. Global regulation of a sigma 54-dependent flagellar gene family in *Caulobacter crescentus* by the transcriptional activator FlbD. Journal of Bacteriology **177:**3241-3250.

202. **Wu, J., and A. Newton.** 1997. Regulation of the *Caulobacter* flagellar gene hierarchy; not just for motility. Molecular Biology **24:**233 -239.

203. **Yang, F. L., Braun, V.** 2000. ShlB mutants of *Serratia marcescens* allow uncoupling of activation and secretion of the ShlA hemolysin. International Journal of Medical Microbiology **290:**529-538.

204. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33:**103-119.

205. **Yao, R., R. A. Alm, T. J. Trust, and P. Guerry.** 1993. Construction of new *Campylobacter* cloning vectors and a new mutational *cat* cassette. Gene **130:**127-130.

206. **Yeo, H.-J., Savvides, S. N., Herr, A. B., Lanka, E., Waksman, G.** 2000. Crystal structure of the hexameric traffic ATPase of the *Helicobacter pylori* type IV secretion system. Molecular Cell **6:**1461-1472.

207. **Yeo, H.-J., Waksman, G.** 2004. Unveiling molecular scaffolds of the type IV secretion system. Journal of Bacteriology **186:**1919-1926.

208. **Yokoseki, T., Kutsukake, K., Ohnishi, K., Iino, T.** 1995. Functional analysis of the flagellar genes in the *fliD* operon of *Salmonella typhimurium.* Microbiology **141:**1715-1722.

209. **Yonekura, K., Maki-Yonekura, S., Namba, K.** 2003. Complete atomic model of the bacterial flagellar filament by electron cryomicroscopy. Nature **424:**643-650.

210. **Zhou, J. D., Lloyd, S.A., Blair, D.F.** 1998. Electrostatic interactions between rotor and stator in the bacterial flagellar motor. Proceedings of the National Academy of Sciences, USA **95:**6436-6441.

211. **Zhu, K., Gonzalez-Pedrajo, B., Macnab, R.M.** 2002. Interactions among membrane and soluble components of the flagellar export apparatus of *Salmonella.* Biochemistry **41:**9516-9524.

212. **Ge, Y., Old, I., Saint Girons, 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:**73-75.

213. **Del Rizzo, P. A., Bi, Y., Dunn, S. D., Shilton, B. H.** 2002. The second stalk of *Escherichia coli* ATP synthase: structure of the isolated dimerization domain. Biochemistry **41:**6875-6884.

214. **Weber, J., Wilke-Mounts, S., Senior, A. E.** 2002. Quantitative determination of binding affinity of δ-subunit in *Escherichia coli* F_1-ATPase. The Journal of Biological Chemistry **277:**18390-18396.

215. **Weber, J., Wilke-Mounts, S., Senior, A. E.** 2003. Identification of the F_1-binding surface on the δ-subunit of ATP synthase. The Journal of Biological Chemistry **278:**13409-13416.

216. **Carbajo, R. J., Kellas, F. A., Runswick, M. J., Montgomery, M. G., Walker, J. E., Neuhaus, D.** 2005. Structure of the F_1-binding domain of the stator of bovine F_1F_0-ATPase and how it binds an α-subunit. Journal of Molecular Biology **351:**824-838.