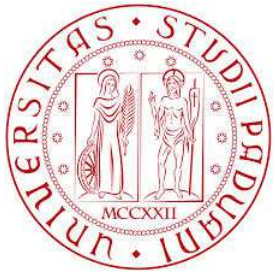

Università degli Studi di Padova

Dipartimento di Biologia

Corso di Dottorato di Ricerca in Bioscienze e Biotecnologie

Curricolo: Biotecnologie

Ciclo XXIX

Structural characterization of proteins involved in *Helicobacter pylori* motility and adhesion

Tesi redatta con il contributo finanziario del MIUR, fondo PRIN
"Unraveling structural and functional determinants behind *Helicobacter pylori* pathogeneses and persistence"

Coordinatore: Ch.mo Prof. Paolo Bernardi

Supervisore: Ch.mo Prof. Giuseppe Zanotti

Dottoranda: Valentina Loconte

Contents

Summary

Helicobacter pylori is a well-known pathogen able to colonize the human stomach of more than half of the world population. Luckely, only the 10% of the affected people manifest severe incidences, like chronic gastritis, peptic ulcers and, among those, only in 1% of cases MALT (Mucosal associated lymphoid tissue) lymphoma and gastric adenocarcinomes are recognized [1].
The project described in this thesis focus mainly on the structural determination and functional characterization of the proteins involved in the assembly of the flagellar architecture. The presence of a flagellate phenotype is paramount for the bacterial colonization processes. Its role is strictly related to the chance of avoiding the acidic milieu in the gastric environment and its periodical mechanical clearances, allowing *H. pylori* to swim through the mucus layer and adhere to gastric epithelial cells [2]. Once the epithelium is reached, the regular colonization process takes place, starting with the adhesion to the cells *via* a set of adhesines, which cover both the bacterial membrane and flagella [3]. Its propulsion in the motion is guaranteed by a tuft of 5-7 polar flagella, long from 2 to 5 μm, and characterized by a S shape [4].
The flagellum is composed by about thirty proteins that can increase up to hundred considering the components involved in chemotatic and regulatory mechanism [5]. The flagellum is a complex rotary nano-machine, which can be divided in two main components: i) the hook-basal bady portion, where the engine is located; ii) the filament, representing the machinery propeller [6].
During the thesis, the proteins from the hook region of the flagellum and the adhesine recovering the filament were analyzed. The strategy adopted included the amplification of the selected genes from the purified *Helicobacter pylori* chromosomal DNA (in particular the strains G27 and P12), the cloning in one or two expression vectors, followed by the expression of the protein in *E. coli* cells and the specific purification from the bacterial lysis supernatant. Unfortunately, the soluble proteins from the hook do not share a significance sequence similarity with any other proteins already known in literature, so their behaviour in solution was not easily predictable. The flagellar proteins analyzed along the work of thesis were: FlgE, FlgE2, FlgK and HpaA (as a flagellar sheath protein). Although all of them were successfully cloned, expressed and purified, only the protein FlgE2 gave sufficiently results to lead to some preliminary structural studies.
In chapter 3 the analysis of the flagellar proteins is reported. FlgK is a *hook-associated protein* involved in the junction between the filament and the hook [7] and FlgE is directly involved in the assembly of the hook [8]. The proteins were characterized in solution and although several crystallization screens were performed, none of them produced crystals suitable for X-ray diffraction analysis. The most relevant limit was in the production of a large amount of protein, that was eluted as an aggregated form during size-exclution

chromatography. The protein FlgE2 is involved in the assembly of the hook as well. It seems to be peculiar for *Helicobacter* and *Campilobater* species, but it has not been characterized further. During the thesis, the protein was successfully purified as a monomer and a tetramer. In both the cases, good quality crystals for the X-ray diffraction studies were produced. Several attempts were performed in solving the structure by molecular replacement, using the orthologous from *Salmonella typhimurium* as template, but none of them was completely successful. However a preliminary model has been designed and is discussed in the chapter. Since the protein does not share a high percentage of similaty with its orthologous proteins, seleniomethionine-labeled protein was produced in order to solve the phase problem and improve the phases obtained by the molecular replacement. The last part of the job is still ongoing. Besides, the functional role of the protein was explored, checking the interaction between the hook structural components and its cognate chaperon, FlgD. The crystal structure of the protein FlgD has been recently solved [9] and demonstrated the presence of a binding site located at the C-terminus. To confirm the binding, a microscale thermophoresis analysis was performed.
To achieve the attachment on the epithelial cell, a set of bacterial adhesines are needed and the role of HpaA protein is presented in chapter 4. *H. pylori* adhesin A is a surface-located lipoprotein that was initially described as a sialic acid binding adhesin, expressed on the surface of gastrointestinal cells [10]. It is supposed to act as a flagellar sheath protein that facilitates the attachment to the cell through flagellar filaments [11]. Owing to its exposed position on the filament, the protein is supposed to be a good candidate as a vaccine target [12]. Due to its high tendency to aggregate, the protein was expressed and purified in a fused form with a super-folder GFP [13]. It produced good quality crystals for X-ray diffraction analysis purpose. However, after the analysis, the crystals resulted to be composed only by sfGFP.
The protein HP1457 is related to a different project and it is reported in chapter 5. The protein belongs to a class of proteins that has been recognized, only recently, to act as a outer-membrane proteins that contributes to the stimulation of the peptidoglycan synthase PBP1B [14]. It is enclosed in a translational unit characterized by the presence of genes coding for high immunogenic, secreted proteins, like HP1454 [15], HP1455 (lipoprotein with unknown function) and the protein HP1456, recognized to be a Lpp20 [16]. The protein was fully characterized in solution and several crystallization trials were performed. Unfortunately, none of those led to obtain protein crystals.

Riassunto

Helicobacter pylori é un batterio Gram-negativo presente in piú del 50% della popolazione mondiale. Tuttavia, i ceppi piú aggressivi del batterio colpiscono solo il 10% delle persone affette, causando gastriti croniche e ulcera peptica. Tra queste, solo l'1% manifesta gravi patologie come MALT linfoma ed adenocarcinoma gastrico [1].
Il progetto descritto in questo lavoro di tesi é finalizzato alla determinazione strutturale ed alla caratterizzazione di proteine coinvolte nell'architettura del flagello di *H. pylori*. Difatti, la presenza del flagello é fondamentale per la colonizzazione da parte del batterio dell'intero ambiente gastrico. Il suo ruole é direttamente collegato all'abilitá del batterio di evitare la permanenza nell'ambiente gastrico, permettengli di attraversare lo strato di muco che aderire all'epitelio gastrico [2]. Grazie ad un set di adesine localizzate sia sulla membrana batterica, sia sui flagelli, *H. pylori* é in grado di aderire alle cellule epiteliali, cominciando il regolare processo di colonizzazione [3]. Inoltre, la presenza di un seti di 5-7 flagelli polarizzati (lunghi tra 2 e 5 μm) e la sua forma ad S garantiscono al batterio la propulsione e la velocitá necessaria per superare la barriera di muco e raggiungere l'epitelio [4].
In generale, la struttura del flagello é costituita da circa trenta diverse proteine, ma il numero puó crescere fino a cento considerando anche le proteine coinvolte nei meccanismi di chemotassia e di regolazione della crescita del flagello [5]. L'architettura generale del flagello é quella di una una complessa nanomacchina, caratterizzata dalla presenza di due principali componenti: i) la porzione dell'uncino e della parte basale, dove é collocato il motore della stuttura; ii) il filamento, che rappresento il vero propulsore della macchina [6].
Nel corso del progetto di dottorato, sono state analizzate le proteine che costituiscono la regione dell'uncino e ricoprono il filamento. La metodica utilizzata prevede l'amplificazione iniziale dei geni target utilizzando il DNA genomico dei ceppi G27 e P12 come templato. I geni amplificati sono stati inseriti in uno o due vettori di espressione, e le proteine ricombinanti sono state prodotte in colture di *E. coli* e purificate, tramite diversi metodi cromatografici, dalla parte solubile della sospensione di lisi batterica. Le proteine dell'uncino di *H. pylori* non presentano un'elevata similaritá di sequenza con le altre proteine giá riportate in letteratura, pertanto la loro reattivitá in soluzione non é facilmente prevedibili a priori. FlgE, FlgE2, FlgK and HpaA (come proteina di rivestimento del flagello) sono le proteine del flagello analizzate durante questo lavoro di dottorato. Sebbene, tutte siano state clonate, espresse e purificate, solo FlgE2 ha fornito dei risultati sufficienti a condurre degli studi strutturali preliminari.
I vari studi condotti sulle proteine del flagello sono riportati nel capitolo 3. FlgK é defina una *hook-associated protein*, che funge da giuntura tra il filamento e l'uncino [7], mentre FlgE é direttamente coinvolta nella polimerizzazione dell'uncino [8]. Entrambe le pro-

teine sono state ampiamente caratterizzate in soluzione ma, sebbene siano stati preparati numerosi screen di cristallizzazione, nessuna delle due ha portato ad ottenere cristalli analizzabili tramite diffrazione a raggi-X. Il limite maggiore osservato nella produzione delle due proteine é l'eluizione di una grande quantitá di forma aggregata, durante il processo di purificazione via cromatografia ad esclusione dimensionale. Anche la proteina FlgE2 é coinvolta nel processo di polimerizzazione dell'uncino. Tuttavia questa sembra essere peculiare solo per le specie batteriche di *Helicobacter* e *Campilobacter* e, pertanto, non si hanno numerosi informazioni a riguardo. Durante il lavoro di tesi, la proteina é stata purificata sia in forma monomerica che in forma tetramerica e, in entrambi i casi, sono stati ottenuti cristalli analizzabili tramite diffrazione da raggi-X. Tuttavia, molti tentavi sono stati svolti per risolvera la struttura tramite *molecular replacement*, utilizzando come templato la struttura della proteina ortologa di *Salmonella typhimurium*, ma nessuno di questo ha avuto interamente successo. Tuttavia, é stato possibile descrivere un modello preliminario, che verrá successivamente discusso all'interno del capitolo. Poiché *Hp*FlgE2 non presenta un'elevata similaritá con *St*FlgE2, é stata prodotta una variante in cui sono state sostituite le metionine con seleniometionine; in questo modo si é provato a migliorare calcolando le fasi relative agli atomi pesanti (Se) e successivamente espandendole tramite i dati raccolti per la proteina nativa. Quest'ultima parte del lavora é ancora in via di definizione. Inoltre, date le scarse conoscenze riportate in letteratura riguardanti la proteina, si é cercato di definirne la reale funzione a partire dall'interazione con lo specifico chaperone FlgD. La struttura cristallografica della proteina FlgD é stata risolta recentement [9], rivelando la presenza di un motivo lineare collocato nella parte C-terminale, rappresentativo di una zona di legame con la proteina coniugata. Il legame é stato confermato tramite termoforesi in microscala.

Per potersi legare alle cellule epitaliali, *H. pylori* ha bisogno di un set di proteine (adesine) che ne permettano l'adesione. La proteina HpaA (*Helicobater pylori* adhesine A), analizzata nel capitolo 4, é una lipoproteina, localizzata sulla superfice del flagello ed inizialmente é stata ipotizzata affine all'acido sialico, esposto sulla superficie delle cellule gastriche [10]. Si suppone che la proteina possa operare come proteina di rivestimento del flagello e che faciliti il legame sulla superficie delle cellule attraverso l'adesione dei flagelli [11]. Inoltre, data la sua posizione esposta, HpaA é stata considerato un ipotetico target per lo svilluppo di vaccini [12]. Infine, la proteina da sola presenta un'elevata tendenza a creare degli aggreagati in soluzione, pertanto é stata espressa e purificata in fusione con una super-folder GFP [13]. La proteina é stata successivamente cristallizata ed é stata analizzata tramite diffrazioni da raggi-X. Sfortunatamente, l'analisi dei dati di diffrazione ha portato alla conclusione che i cristalli erano formati esclusivamente da sfGFP.

Infine, nel capitolo 5 viene analizzata la proteina HP1457. Questa appartiene ad una classe di proteina recentemente riconusciute operare come *outer-membrane proteine*, coinvolte nella stimolazione della peptidoglicano-sintasi, PBP1B [14]. Inoltre, la proteina HP1457 appartiene ad un operone caratterizzato dalla presenza di geni che codificano per proteine di secrezione altamente immunogeniche. Tra queste: HP1454 [15], HP1455 (una lipoproteina di funzione ancora sconosciuta), HP1456 (identificata come la lipoproteina Lpp20) [16]. HP1457 é stata ampiamente caratterizzata in soluzione e sono stati effettuati numerosi tentativi di cristallizzazione su diversi costrutti, ma nessuno di questi ha portato ad ottenere cristalli.

1 | *Helicobacter pylori*

1.1 Introduction

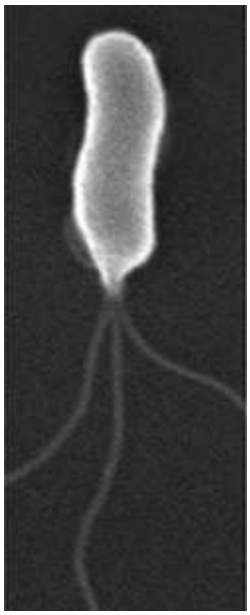

Figure 1.1: Electron microscopy of the Helicobacter pylori. Spiral bacterial bodies with a bundle of sheathed flagella.

Helicobacter pylori was isolated for the first time from the stomach of patients affected by gastric ulcers in 1984 by two scientists, Marshall and Warren [17]. They spotted the cause of gastric pathologies to be a Gram-negative, spiral-shaped, microaerophilic bacterium, from 2 to 5 μm long and with a tuft of 5-7 polar flagella. Tomb *et al.* confirmed in 1997 the presence of *H. pylori* in most of the gastric niche of the primates [18], also supporting its wide spreading diffusion during the Cooper Age sustained by Maixner *et al.* [19], and the hypothesis of co-evolution reported by Falush *et al.* [20].
Since nineties, *H. pylori* was estimated to have infected about half of the human population, with a persistence onset in developing countries. It is usually acquired during childhood by oral-oral, fecal-oral and gastric-oral infection and its colonization is significantly long: it usually lasts life spanning [21]. Most of the infected people are asymptomatic, with mild inflammation detectable only histologically. However, an important minority (about 10%) develops during their life time severe gastric pathologies, like MALT lymphoma, adonocarcinomas and duodenal ulcers [2]. Furthermore, the World Health Organization declared *H. pylori* a Class 1 carcinogenic bacterium in 1994, due to its responsibility in 5.5% of the gastric cancers all over the world [22].
The bacterium shows up a high genetic variability: comparing the sampled strains from South American, African and Asian people, *H. pylori* has shown to have a genetic heterogeneity, likely induced by the need to adapt to a highly changing environment such as the stomach niche [23]. Indeed, to successful adapt to the singular human microenvironment, the bacterium can develop a high number of point mutations, variation in the genetic structure (for instance, inversion, translocation, duplication of the same gene), gene fusion, slit and fragmentation, which help the bacterium to overcome the immunity response of the host [24]. Plus, the gene order in the whole bacterial genome is extremely variable among the bacterial strains [25]. The high diversification in bacterial genome can take place thanks to the reduced dimension of the bacterial plasmid, long only 1.5 up to 1.7 Mbp. Up to now, several bacterial strains have their genome totally sequencend: HP26695, the standard strain, used to named and identify every single gene in the other *H.pylori* strains; J99, P12, G27, Shi470, B38 and 52. All of them were isolated from patients with various diseases from different area in the world. From the comparison of the sequenced species, it seems evident that the strains shares a core genome of about 1200 genes [26].
In generale, the human stomach is well protected from bacterial infections. Nevertheless, *H. pylori* can survive in it thanks to several factors which helps to buffer the harsh acidic conditions and to install a persistent colonization. Among these factors, flagella, urease and adhesins play a relevant role. The colonization stage induces an activation of the innate and adaptive immune system, with the consequential infiltration into the gastric mucosa of leukocytes, lymphocytes and macrophages.

1.2 Bacterial Pathogenic Factors

1.2.1 Urease

H. pylori is a neutrophilic bacterium unable to survive during its life time in the gastric lumen. It can live in such acidic environment only for short periods that allow it to pass over the viscous mucosa and reach the gastric epithelium, where it can find the basic nutrients. To survive longer in the gastric environment, the bacterium has dedicated a set of genes to the development of highly competent enzymes defined urease. The genes *ure*AB and *ure*EFGHI are responsible of the biosynthesis of the corresponding proteins, involved in the maintenance of the bacterial cytoplasmic pH at a neutral level for its survival [27]. This situation is satisfied by the hydrolysis of urease in ammonia and carbonic anydhrase, that generate a nearly neutral microenvironment in the gastric lumen. *H. pylori* urease is associated with the surface of the organism upon spontaneous lysis, followed by adsorption of the enzyme onto the surface of the bacterial membrane. This process is the so-called "altruistic autolysis" and allows the 30% of bacterial enzymes to be exposed [28].

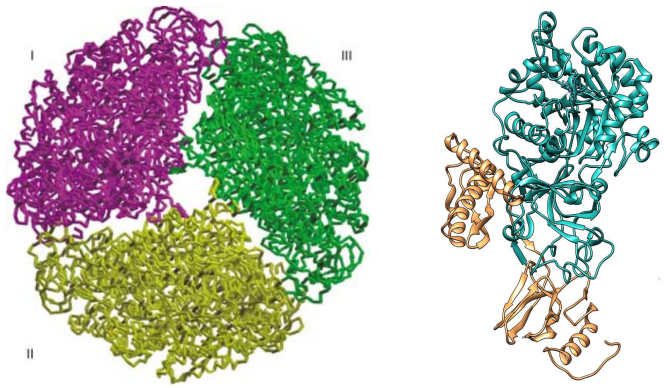

Figure 1.2: On the left: Dodecameric assembly of H. pylori urease; The crystal packing of H. pylori urease shows a spherical assembly composed of four $(\alpha\beta)_3$ units, related by the crystallographic two- and three-fold symmetry axis. The structure is observed perpendiculary to the two-fold axis. On the right: α (gold) and β (cyan) subunits from H. pylori urease (taken from [28]).

The external urease is essentiale for pH maintenance. In general, free or surfaced bacterial urease has an optimum activity at a pH between 7.5 and 8.0, but its inactivation becomes irreversible at pH<4.0. Cytoplasmic ones have a low activity at neutral pH, which increases from 10 to 20 times as the external pH falls to 2.5 [29]. From a structural point of view, the *H. pylori* urease is a dodecamer composed by two distint subunits, α and β, with a molecur weight of 61.5 kDa and 26.5 kDa respectively. They are organized together to form a massive complex of 600 kDa, which has been demonstated by Ha *et al.* to form a double ring dodecamer of 1.2 MDa, with 13 nm in diameter [28]. The urea intake in bacterial cytoplasm is regulated by a supplementary proton-channel protein, called UreI (HP0071). It is triggered to open at acidic pH values, allowing the diffution of urea insiede the channel, but it is closed at neutral pH to avoid over-alkalinization of cytoplasm. Finally, UreE (HP0070), UreF (HP0069), UreG (HP0068) and UreH (HP0067) participate as chaperones in the complex assembly and activation/inactivation [30].

1.2.2 VacA

VacA (Vacuoling Cytotoxin A) is a pathogenic protein present in the active form as a 95 kDa toxin, responsable of the major vacuolization of the epithelial cells [31]. It is initially produced as a 140 kDa protoxin, secreted through a type V pathway. The proteolytic cleavage induces the formation of the mature toxin and its organization in a flower-shaped dodecameric complex with a 30 nm of diameter [32]. Among the autotransporters, VacA is the only so far known to oligomerize in order to create a pore in the host membrane. As most of the proteins belonging to the same family, it is mainly composed of β-motifs [33], and in the mature toxin two main domains can be identified: p33 and p55. Both of them are cleaved after a two-step process that induces the cut of 33 kDa signal peptide at the N-terminus and directs the secretion from the cytoplasm into the periplasm, whilst the C-terminus, composed of a 45 kDa fragment, acts as an autotransporter inducing the export through the membrane [34]. Later, the towing can assemble in the oligomeric state in the lipid bilayer to form an anionic selective channels into both epithelial and immune cells, inducing apoptosis and interfering with the process of antigen presentation and the inhibition of T cells proliferation [35][36][37].

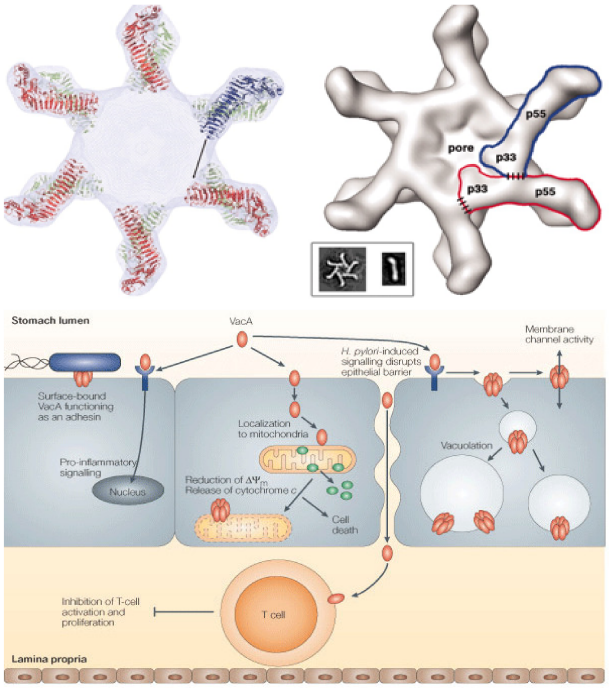

Figure 1.3: On the top: twelve p55 subunits are docked into a 19-Å cryo-EM map of a VacA dodecamer (take from [38]). On the bottom: effects of VacA on gastric epithelial cells include alterations in mitochondrial membrane permeability and apoptosis, stimulation of pro-inflammatory signalling, increased permeability of the plasma membrane and alterations in endocytic compartments (Adapted from [39]).

1.2.3 HP-NAP

The *H. pylori* Neutrophilic Activating Protein is a highly immunogenic protein, responsible for the activation of neutrophils and monocytes by chemotactic attraction of the bacterium itself to the sites of infection [40]. It induces the attachment of the human neutrophils to the inflammation site, causing the production of reactive oxygen radicals (ROS) and the activation of IL-12 and IL-23, which favors T cells differentiation [41]. The inflammation site is also invoved in the induction of leukocytes in the adhesion to the epithelium [42] and the activation of the mast cells [43]. The toxin interacts with epithelium with via toll-like receptors, which induce the activation of the nuclear factor-kB [44].
HP-NAP is arranged in a dodecameric fashion, with a total molecular weight of 180 kDa. Each monomer is characterized by a four-helix-bundled domain [45]. It has a 30 nm diameter highly resembling other DNA binding proteins [46] and is able to store more than five hundred iron ions.

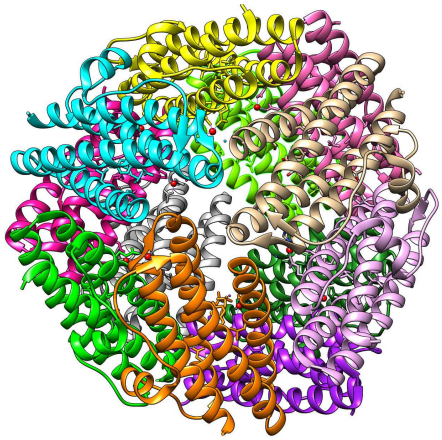

Figure 1.4: Crystal structure of HP-NAP protein (PDB 1JI4). The dodecameric structure is represeted with the 3-fold axis running approximately perpendicular to the plane of the paper in the center of the image, where a putative tunnel for the iron binding should be located.

1.2.4 *cag*PAI

The pathogenic activity of *H. pylori* is directly correlated to the presence of the *cytotoxin-associated gene* Pathogenic Island (*cag*-PAI), a 40-kb DNA sequence involved in the transmigration of the effector protein CagA from the bacteria to the host cells. CagA is a protein with a molecular weight of 140 kDa and its corresponding gene is present in approximately 50 to 70% of *H. pylori* strains [47]. Patients affected by *H. pylori* $CagA^+$ strains usually show high inflammatory response [2].
Secretion of CagA is attributed to the *cag*-PAI secretion system, a type IV secretion apparatus, which forms a syringe-like structure, capable of penetrating gastric epithelial cells and facilitating CagA translocation. Significant homologies have been found between the genetic structure of *H. pylori* cag-PAI and *Agrobacterium tumefaciens* T4SS. The last is a secretion system used by the plant pathogen to deliver DNA and proteins into the cells, causing tumor outbreaks. The transporter machine is composed of twelve proteins, identified as VirB1-11 and VirD4, and categorized in three groups: cytoplasmic or inner membrane proteins, a core complex located into the periplasm, and a pilus (a surface

structure) beyond the outer membrane, involved in the release of toxins [48]. Two AT-Pases (VirB4 and VirB11) and VirD4 (a "coupling protein") are classified as cytoplasmic or inner membrane proteins. VirB11 and VirD4 proteins assemble into hexameric structures, whereas VirB4 is a homodimer. VirB7, VirB8, VirB9 and VirB10 are important components of the core complex. The pilus of the T4SS is composed of VirB2 and VirB5, which are considered major and minor subunits, respectively [49]. The secretion system T4SS is encoded by *cag*-PAI in *H. pylori* and allows the CagA translocation into human epithelial cells [48]. With respect to the system from *A. tumefaciens*, the *H. pylori* system encodes for 27 proteins and, among 32 genes carried by *cag*-PAI, only 18 are involved in CagA translocation [50]. The remaining 14 genes are supposed to stimulate a cag-PAI dependent IL-8 response [51]. The *H. pylori* proteins that form the core complex are CagT, CagX and CagY, respectively homologues of VirB7, VirB9 and VirB10. CagT, together with CagW, is the protein associated with the pilus base. CagC, homologue of VirB2, is associated with the pilin and CagY is the sheet protein, containing variable numbers of repeats that avoid the host immune-response. CagL plays a role in adhesion, attaching the T4SS to the host cells surface [49][52] . Finally, CagF is a chaperon, binding the C-terminus secretion signal of CagA and that helps the protein to be translocated [53].
Once delivered into the cell, CagA is phosphorylated and interferes with the physiological signal transduction and causes pathological cellular responses, such as the increase of cell proliferation, motility, apoptosis and morphological changes in epithelial cells [47]. After tyrosin phosphorylation, CagA binds specifically to SPH-2, a cytoplasmic tyrosine phosphatase. The complex formation induces a conformational change, responsible of SPH-2 phosphatase activity, whose activity is important in cell morphogenesis and motility [54]. Deregulation of CagA induces an abnormal proliferation of gastric cells, represented by cellular elongation (called "*hummeringbird phenotype*"). Deregulation of cells adhesion is associated with the increase of cells mobility. Both the events play an important role in pathogenic activity of cag^+ strains of *H. pylori* [48]. The *cag*-PAI may also affect the immune response, due to its ability to induce apoptosis of T cells [47]. CagA itself can induce ROS production in gastric epithelial cell. Excessive ROS production in gastric epithelial cells can cause DNA damage and thus might be involved in gastric carcinogenesis [55].

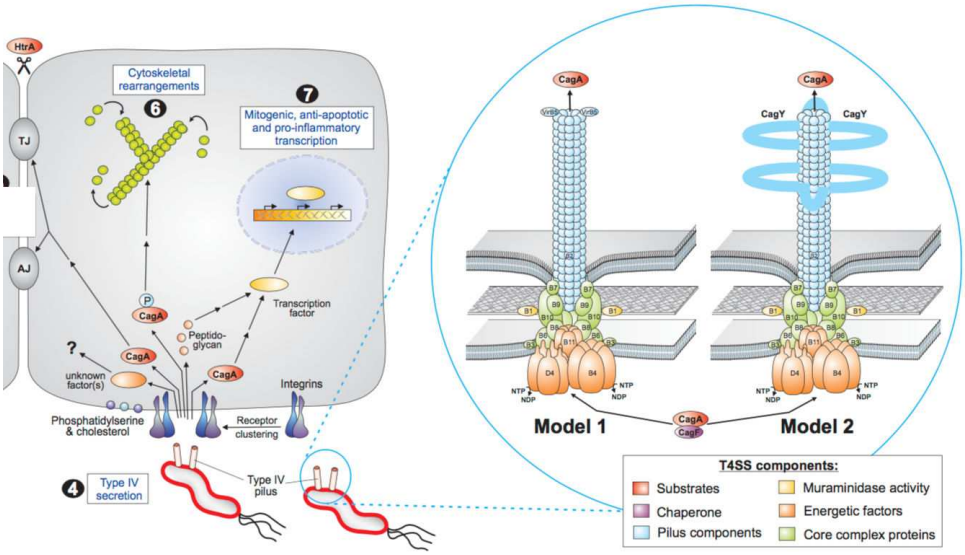

Figure 1.5: Type IV Secretion System and its role in CagA secretion in human cells. Upon cell contact, H. pylori assembles T4SS pili at their surface enabling delivery of molecules CagA and peptidoglycan, from bacterial cytoplasm into host cells. cagPAI proteins (CagA, CagI, CagL and CagY) interact with integrin receptors, inducing numerous cellular effects including disruption of cell-to-cell junctions, cytoskeletal rearrangements and nuclear signalling. Two models for the assembled T4SS machinery in H. pylori were proposed. Model-1 assumes VirB1- 11 proteins, the coupling factor VirD4 and accessory factors such as CagF assemble in a similar fashion to that proposed for A. tumefaciens T4SS. Model-2 assumes that the T4SS pilus surface is covered with CagY (VirB10) molecules (from [56]).

1.2.5 Lipopolysaccharides

Several Gram-negative bacteria use LPS as a powerful tool to develop inflmmatory response in the host organism. Nevertheless, Moutiala *et al.* have noticed the role of LPS in *H. pylori* is 500-fold less endotoxic than in the other Gram-negative species (like *Salmonella typhimurium*) [57]. Indeed, LPS from *H. pylori* are less stimulating macrophages in cytokines, prostaglandines and nitric oxide production, involved in the inflammatory response. This fact is probably due to the composition of Lipid A, characterized by the presence of a long 3-hydroxy fatty acid, a D-glucosammine disaccharide as a backbone and the expressed O-antigenes related with the Lewis blood antigenes group, expressed in human cells [57][36][58]. This can reduce the bacterial pathogenic effect, by inducing a mimicring effect that protects the bacterium from immune recognition [59].

1.2.6 Adhesines

An essential step in *H. pylori* colonization is its adhesion on the epithelial cell surface. Indeed, about 20% of the bacterium in the stomach has been found attached to the mucus layer, protecting epithelial cells. Its ability to colonize the gastric environment strongly depends on the genetic variations of the host's mucins, along with the bacterial strain. The bacterium comes into contact with the gastric epithelium thanks to the propulsion motion of flagella and several chemotactic factors abundant in the mucosa [60]. The adhesion takes place on the human gastric epithelial cells as cellular receptors fucosylated glycoproteins and sialylate glycoproteins [61]. One of the common features shared among *H. pylori* adhesion proteins is their tendency to bind sialic acid in mucus [62]. The best characterized *H. pylori* adhesins are BabA and SabA, two outer membrane proteins, that bind to the fucosylate LeB antigen in MUC5AC [63] and to sialyted glycocompunents in salivary mucin MUC7 [64], respectively. Although SabB has been found out to be involved in the development of duodenal cancer, its specific receptor is still unknown [65]. Other outer membrane proteins identified to be adhesines are AlpA and AlpB, which are able to mediate the adhesion to the cell surface generating macromolecular complexes on the bacterial surface [66].

1.2.7 Flagellum

The *H. pylori* success in host colonization is achieved through the use of tuft of polar flagella, which allows the bacterial motion from the stomach lumen to the epithelium [67]. In fact, they allow the bacterium to avoid the acidic milieu and to survive during host's periodical mechanical clearances, allowing *H. pylori* to swim through the mucus layer and adhere to gastric epithelial cells [68]. Moreover, thanks to its helical shape and the presence of polar flagella, *Helicobacter* and *Campilobacter* species acquire an unsual velocity to pass through the viscous media [6].
H. pylori does not present clusters encoding for flagellar genes. It is known that the expression of flagellar genes is regulated by σ^{70} and σ^{54} factors, respectively involved in the promotion of nine operons including fifteen flagellar genes and five operons containing hook basal body genes [68].
The structural organization of a single flagellum is quite complex. It is composed approximately of thirty different proteins, assisted in the expression and in the assembly processes by about fifteen more ([18]; Table 1.1). The flagellum is a complex rotatory nano-machine that can be divided in two main portions: the hook basal-body and the extracellular filament. The first can be divided in three substructures: (i) the base, localized in the inner membrane and spanning to the cytoplasm; (ii) the rod and ring structures, located in the periplasm; and (iii) the hook, a flexible junction presented on the surface [69][6][70]. Several studies were performed on the bacterial flagellar motor of *Escherichia coli* and *Salmonella typhimurium* and, despite the crystal structure of some of the separated components are known, a complete model of the flagellum has not been proposed yet [71]. For what concerns *H. pylori*, only eight structures are solved. Although some general considerations could be made in homology with the already known structures form others Gram-negative organisms, some specific achievements on the features of the bacterium could be obtained from their study. Indeed, the relevance of studying the molecular architecture of flagellar proteins is crucial for designing new therapeutic targets. On the other

side, the flagellum represents an ideal example of a molecular machine for the conversion of chemical energy into mechanical energy. The bacteria have been able to devise it in billion years of evolution and it can eventually be applied for nano-technological purposes.

The Type 3 Secretion System

Type 3 Secretion System is a specialized machine evolved in many Gram-negative bacteria as a delivery of effector proteins into target eukaryotic cells [72]. It is widespread in nature and has an important role in bacterial infection and symbiotic interactions [73]. Moreover, recently phylogenetic analysis identified the T3SS as the precursor of the export apparatus of bacterial flagella [74]; according to Galan *et al.* the process could have taken place in two steps: 1) the first is the development of a competent structure for protein secretion; 2) the second is the recruitment of secretines [75].
T3SS was widely studied in several bacteria like *Salmonella*, *Shigella*, and *Yersinia spp*, and is known to be at the base of the structural organization the flagellum from *H. pylori*, too. In this case, it modulates the assembly of the whole flagellar structure. In analogy with the system from the other organism, the T3SS is composed by roughly twenty proteins, most of whose located in the inner membrane, and an ATPase associated with the cytoplasmic side of the membrane [76]. T3SS is able to deliver protein from the cytoplasmic side of the membrane to the extracellular environment. Since the activity requires a high amount of energy, two energy source were identified for the system [77] one is the hydrolysis of ATP per each protein translocated, whilst the second is the generation of a proton motive force (PMF) which induces the formation of both a potential gradient and a variation in pH [78]. Translocated proteins undergo to processes after the activation through a secretion signal. The mechanism ensures the proteins can be released to the target cell and not into the extracellular space. Unfortunately, the recruitment mechanism is still unknown [75].
The component of T3SS is a 50-nm-long needle complex spanning through the whole bacterial membrane. The complex is constituted by a base (approximately 25 nm wide and 30 nm long), with two inner rings located in the inner membrane (IR1 and IR2) connected *via* neck with two outer rings in the outer membrane (OR1 and OR2). The base is linked to the needle through an inner rod, enclosing a 20-Å-wide channel for protein translocation [79]. At the end of the needle is located the tip complex, involved in the target cells recognition [80]. Finally, a cluster of inner membrane proteins is located at the center of the inner ring, mediating the passage of the effectors [81].

Localization	**Protein Name**	**Name for HP26695 strain**	**Role and Complex**	**PDB ID**
Export system	FliO	HP0583	Protein associated with FliP	
	FliP	HP0685	Protein associated with FliO and MS-ring	
	FliQ	HP1419	Export component associated with FlaA and FlaB	
	FliR	HP0173	Associated with MS-ring	
	FlhA	HP1041	Target for soluble export proteins	3MYD
	FlhB	HP0770	Target for soluble proteins	
	FliH	HP0353	Negative regulator of FliI	
	FliI	HP1420	ATPase, interacting with FliH	
	FliJ	HP0256		
	FlgJ	HP0245		
MS ring	FliF	HP0351	Transmembrane component of rotor associated with FlhA	
C ring	FliG	HP0352	Rotor switch protein associated with MotA and FliM	3USW; 3USY
	FliM	HP1031	Involved the rotor switch, binds the CheY-P protein, N-terminus interacts FliF	4GC8
	FliN	HP0584	Flagellar export component, role as a switching protein	
	FliY	HP1030	Flagellar export component	
Stator	MotA	HP0815	Forms a proton channel through the membrane	
	MotB	HP0816	Associated with MotA	3SOY
P ring	FlgI	HP0246	Part of the bushing	
L ring	FlgH	HP0325	Part of the bushing	
Proximal rod	FliE	HP1557	Part of the export gate and MS ring rod junction protein	
	FlgB	HP1559		
	FlgC	HP1558		
Distal rod	FlgF	HP1092		
	FlgG	HP1585		
Hook-filament junction proteins	FlgK	HP1119	HAP1	

Localization	**Protein Name**	**Name for HP26695 strain**	**Role and Complex**	**PDB ID**
	FlgL	HP0295	HAP3	
Filament capping	FliD	HP0752		
Hook cap	FlgD	HP0907	Rod-modification protein, required for hook polymerization	4ZZF; 4ZZK
Hook	FliK	HP0906	Hook length regulator	
	FlgE	HP0870	Hook protein	
	FlgE2	HP0908	Hook protein	
Filament	FlaA	HP0601	Flagellin subunit	
	FlaB	HP0115	Flagellin subunit	
		HP1076	Interacts with FliS	3K1H
	FliS	HP0753		3IQC; 3K1I
Regulatory protein	FlgM	HP1122	Encodes for the inhibitor of flagellum σ^{28}-factor	
	FliA	HP1032	Negative regulator of FlgM	
	FlgR	HP0730	Activates transcription with σ^{54}-factor	
Chemotactic factor	CheY	HP1067	Chemotaxis, interacts with FliM	3GWQ; 3HIF; 3HIG; 3HIE
Chaperon	FliT	HP0754	Chaperon FliD	
	FlgA	HP1477	Chaperon involved in the assembly of the P-ring	
Flagellar Biosynthesis	FlhF	HP1035		
Assembly Factor	FliW1	HP1154		
Paralysed Protein	pFlA	HP1274		
Assembly factor	FliW2	HP1377		

Table 1.1: H. pylori proteins involved in the assembly of flagellum.

2 | Materials and Methods

2.1 Materials

2.1.1 Oligonucleotides

Oligonucleotides are designed according the correspondent gene from different *H. pylori* strains. Lyophilized and desalted primers are redissolved in sterile water up to 100 μmol/L and stored at -20°C

Name	Sequence (5' to 3')
pETite-hp1457Δ(1-21)-for	GAAGGAGATATACATATGACTTATCAAAATG TGAATGATG
pETite-hp1457Δ(1-21)-His_6-rev	GTGATGGTGGTGATGATGGCTGCCGCGCGG CACCAGAAACAT
pETite-hp1457Δ(1-49)-for	GAAGGAGATATACATATGGATTTATTGCTA ACCGCTAAC
pETite-hp1457Δ(1-49)-His_6-rev	GTGATGGTGGTGATGATGGCTGCCGCGCGG CACCAGAAACAT
pETite-hp0908-for	GTGATGGTGGTGATGATGGCTGCCGCGCGG CACCAGAAACAT
pETite-hp0908-His_6-rev	GTGATGGTGGTGATGATGTTTTTTCAAGCT AATGGCTTC
pETite-hp0908(85-570)-for	GAAGGAGATATACATATGGAATTTCATATG GCGTATCAA
pETite-hp0908(85-570)-rev	GTGATGGTGGTGATGATGGCTGCCGCGCGGCACC AGAAACATGTTCCCGGCATTCACGTTGCT
pETDuet-hp0908-for	GGGGGGCCATGGGAAACGACACCTTATTAAAC
pETDuet-hp0908-His_6-rev	GGGGGGGCGGCCGCTTATGGCTGTGGT GATGGTGATGTTTTTTCAAGCTAAT
pACYC-hp0907-for	GGGGGGGATATCATGGCTATTGATTTAGCAGAA
pACYC-hp0907-Strep-rev	CCCCCCGGTACCTTATTTTTCAAACTGCGGA TGCTCCATGCTGTCTCTTTAGG
pETite-hp1119-for	GAAGGAGATATACATATGGGCGGAATCTTATCT TCACTC
pETite-hp1119-His_6-rev	GTGATGGTGGTGATGATGTTGTTTAATCCCC AATAAAGT
pETite-hp0870-for	GAAGGAGATATACATATGCTTAGGTCTTTATGG TCTGGT
pETite-hp0870-His_6-rev	GTGATGGTGGTGATGATGACTTTGAAAATAT AAATTTTCTTGCTTAAGATTCAATAGGGT

Table 2.1: Oligonucleotides used as primers to amplify gene sequences.

2.1.2 Vectors

pACYCDuet-1 (*Novagen*) is a 4008 bp long vector. The insertion was designed according to the features of the second MCS, including sites for EcoRV and KpnI. Expression of genes cloned into this vector is under the control of the T7-*lac* promoter. Additionally, the vector carries a Chloramphenicol resistance gene (Cm^R) for selection.

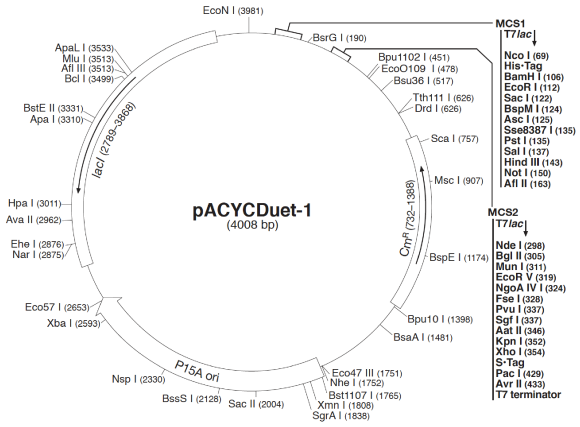

Figure 2.1: pACYC-1 vector map

pETDuet-1 (*Novagen*) is a 5420 bp long bacterial vector for the protein expression under the T7-*lac* promoter. The vector carries an Ampicillin resistance gene (Amp^R) for selection. The insertion was designed according to the features of the first MCS, including sites for NcoI and NotI.

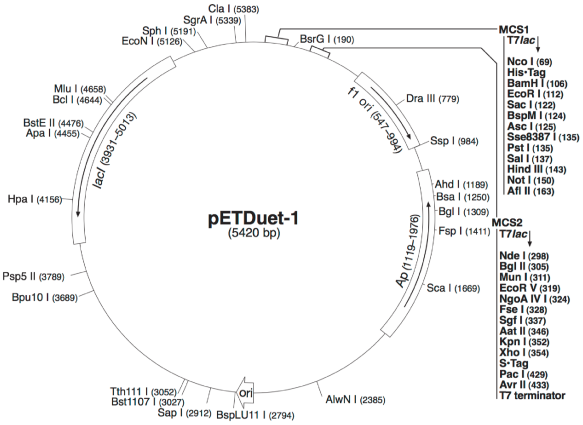

Figure 2.2: pETDuet-1 vector map

pETite C-His (*Lucigen*) is a 2235 bp long pre-processed vector that encodes the C-terminal hexa His-tag. The vector includes a Kanamycin resistance gene (Kan^R) for selection and inducible expression under control of the T7-*lac* promoter.

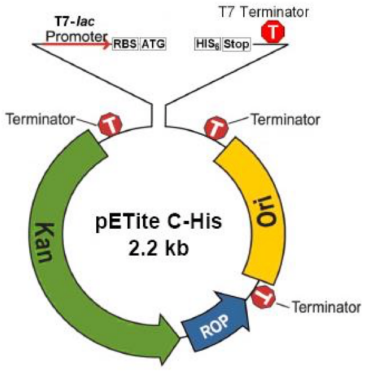

Figure 2.3: pEtite C-His vector map

pRSET-A (*Thermo Scientific*) is a 5700 bp long vector under the induction of a T7-*lac* promoter, including an Ampicillin resistance gene (Amp^R). The vector was modified in order to include the fusion target between two branches of a superfolder GFP1-10 and GFP11 (Prof. Alessandro Negro, Dept. Biomedical Sciences, University of Padova).

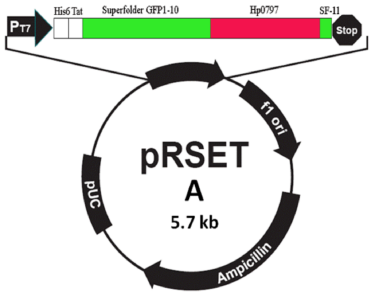

Figure 2.4: pRSET-A vector map

Finally, the designed constructs along with starting and ending position are listed in the following table.

Protein	Template	Vector	Start	End
HP0908 (FlgE2)	*H. pylori* G27	pETite C-His	1	605
		pETDuet-1	1	605
	H. pylori P12	pETite C-His	85	569
HP0908(84-560)	*H. pyloti* P12	pETite C-His	84	560
HP0907 (FlgD)	*H. pylori* G27	pACYCDuet-1	1	316
HP0797 (HpaA)	*H. pylori* 26695	pRSETa	34	203
HP1457Δ(1-31)	*H. pylori* P12	pETite C-His	32	210
HP1457Δ(1-49)	*H. pylori* P12	pETite C-His	50	210
HP0870 (FlgE)	*H. pylori* P12	pETite C-His	1	718
HP1119 (FlgK)	*H. pylori* G27	pETite C-His	1	606

Table 2.2: List of all designed constructs with starting and ending positions of the aminoacid residues.

2.1.3 *E. coli* strains

One Shot TOP10 (F mcrA Δ(*mrr-hsd*RMS-*mcr*BC) ϕ80*lac*ZΔM15 Δ *lac*X74 *rec*A1 *ara*D139 Δ(*araleu*)7697 *gal*U *gal*K *rps*L (StrR) *end*A1 *nup*G) (*Invitrogen*) are chemically competent cells that allow stable replication of high-copy number plasmids.

BL21(DE3) (F *omp*T $hsdS_B$ *(*r_B*-*m_B*-) gal dcm* (DE3)) (*Invitrogen*) are chemically competent cells deficient in both lon and ompT proteases. The designation (DE3) indicates that the host is a lysogen of λDE3, and therefore carries a chromosomal copy of the T7 RNA polymerase gene under control of the lacUV5 promoter. Such strains are suitable for production of protein from target genes cloned in pET vectors using IPTG as an inducer.

BL21pLysS(DE3) (F *omp*T $hsdS_B$ *(*r_B*-*m_B*-) gal dcm* (DE3) pLysS (Cam^R)) (*Invitrogen*) are chemically competent cells used to suppress basal expression of T7 RNA polymerase prior to induction and thus stabilize pET recombinants encoding target proteins that affect cell growth and viability. The pLysS designation is given to hosts carrying a pET-compatible plasmid that encodes T7 lysozyme, which is a natural inhibitor of T7 RNA polymerase.

2.1.4 Equipment

Instrument	textbfManufacture
ÄCTA Purifier	GE Healthcare
Avanti J-25 Centrifuge	Beckman Coutler
Spectrophotometer Cary 50 Bio	Varian
One Shot Cell Breakage System	Constant System Ldt.
Microcentrifuge Z216 M	Hermle
Electrophoresis Mini-PROTEAN Tetra System	BioRad
Nanodrop	Thermo Scientific
Oryx 8 Crystallization Robot	Douglas Instruments
Zetasizer Nano ZS, DLS	Malveren Instrument LDT.
Quadrupole-TOF Spectrometer	Water, Manchester
Spectropolamiter	Jasco Analytical Instruments
Microscope SZX 12	Olympus
UV Fluorescence Microspope MZ16F	Leica
Confocal Microscope SP5	Leica
^{3}Prime Thermal Cycler	Techne
Monolith NT.115 MST	NanoTemper Technologies

Table 2.3: Devices and instruments (basic laboratory equipment is not listed).

2.1.5 Chromatographic Tools

Device	Manufacturer
PD-10 Desalting Columns	GE Healthcare
His-Trap Ni-NTA Agarose	GE Healthcare
Strep-Avidine resin	IBA
Gel-filtration markers	Sigma-Aldrich

Table 2.4: Columns and molecular weight markers used for affinity and size-exclusion chromatography.

2.1.6 Crystallographic Tools

Device	Manufacturer
Pegs II, JCSG Core (I-IV)	Quiagen
Structure Screen 1 & 2, PACT PREMIUM SUITE	Molecular Dimension
Additive Screen	Hampton Research
MRC 96-well sitting drop crystallization plate	Molecular Dimension
MRC MAXI 48-well sitting drop crystallization plate	Molecular Dimension
24-well XRL hanging drop plate	Molecular Dimension
MRC MAXI 48-well sitting drop crystallization plate	Molecular Dimension
Mounted CryoLoops	Molecular Dimension
Mounted Mesh LithoLoop	Molecular Dimension

Table 2.5: Crystallization screens and general equipment.

2.1.7 Kits used for DNA preparation and gel staining

Typology	Manufacturer
GenElute Plasmid Miniprep Kit	Sigma-Aldrich
GenElute Gel Extraction Kit	Sigma-Aldrich
ProteoSilver Stain Kit	Sigma-Aldrich

Table 2.6: Kits used for preparation of DNA and for staining of the gels for visualizing proteins according to the manufacturers' protocols.

2.2 Methods

2.2.1 Transformation of chemically competent cells

Home-made competent cells (100 μL) are thawed out on ice and incubate with 60-300 ng of plasmid DNA for about 30 minutes. The cells are then shocked for 45 seconds at 42°C and further place on ice for 2 minutes. Afterwards the cell suspension is reconstituted shaking at 37°C for 45-60 minutes and finally spread on a LB-agar plate, supplemented with the proper antibiotics. The plate was left overnight at 37°C.

2.2.2 Isolation of the DNA plasmid from *E. coli* cells

The plamism DNA is amplified in One Shot TOP10 cells. After the preparation of a single-colony 5 ml culture, the extraction procedure is performed according to the Sigma-Aldrich's GenElute Plasmid Preparation kit protocol.

2.2.3 Polymerase Chain Reactions

Target-gene amplification is performed using to different type of polymerase:

- The proofreading Q5 high-fidelity polymerase from *Pyrococcus furiosus* is used to amplify the gene from *H. pylori* genome with a low error rate.

- The non-proofreading GoTaq polymerase is used for colony-PCRs.

The general PCR protocols used for the gene amplification are reported in the following tables:

Components	**Q5 Protocol**	**GoTaq Protocol**
Template DNA	1 μL	5 μL
Forward Primer (10 μmol\L)	1.25 μL	1 μL
Reverse Primer (10 μmol\L)	1.25 μL	1 μL
Master Mix Reaction Buffer (2X)	12.5 μL	12.5 μL
Nuclease Free Water	10 μL	5.5 μL
Final Volume	25 μL	25 μL

Table 2.7: PCR reactions composition.

Components	Q5	GoTaq
Initial Denaturation	30”/98°C	2’/95°C
Denaturation	5-10”/98°C	1’/95°C
Annealing	10-30”/ primer specific temperature 35 cycles	30-60”/ primer specific temperature 35 cycles
Extension	15-30” per kb/72°C	1’ per kb/72°
Final Extension	2’/72°C	5’/72°C

Table 2.8: Thermal cycling profile for cloning and colony-PCRs.

2.2.4 Cloning and Sequencing

Stardard cloning pACYCDuet-1 and pETDuet-1 vectors are linearized using restriction endonucleases NcoI/NotI and EcoRV-HF/KpnI-HF enzymes, respectivetely. The genes of interest (*hp0907* and *hp0908*) are amplified using primers in Table 2.1.
During the ligation reaction, each vector is mixed with the matching gene and incubated together with T4-DNA ligase at 4°C overnight. After the transformation of the cloned gene into TOP10 cells, the colonies are tested for the successful integration of the insert by colony-PCR. DNA from positive clones is amplified and isolated according to section 2.2.3 and sequenced at the DNA-service center GATC-Biotechnology using the universal proper primers according to the MCS choosen. (pETDuet-1: forward pET-Upstream and reverse Down1. pACYCDuet-1: forward Duet Up2 and reverse T7).

Components	Volume (μL)
Digested Vector (100 ng)	4
Target gene	16
T4 Ligase (5 U^i/μL)	1
T4 Ligase Buffer (10X)	2.5
Nuclease Free Water	1.5
Final Volume	25

Table 2.9: Ligation reaction mixture.

Cloning of *hpaA*-sfGFP in vector pSERTa was performed in collaboration with Prof. Alessandro Negro from the Department of Biomedical Science, University of Padova.

Enzyme-free cloning 1 μL of pre-processed linearized pETite C-His vector (25 ng) is mixed with 5 μL of the target PCR product (50 ng), at room temperature, and transformed directly into 80 μL of chemically competent One Shot TOP10 cells. The target gene has been previously amplified with primers that contained short homology to the ends of the pETite C-His_6 vector allowing recombination within the host cell to fuse the blunt PCR product to the vector.

2.2.5 Expression of recombinant proteins in *E. coli*

Expression of native proteins *E. coli* BL21(DE3) cells are transformed using vectors (pETite, pETDuet-1, pACYCDuet-1) containing T7 promoter, inducible by isopropyl-β-D-1-thiogalactopyranoside (IPTG) and is located upstream with respect to the gene of interest. BL21(DE3) is the T7 host strain used for expression of cloned genes from the bacteriophage T7 promoter. The genome of bacteria harbors the T7 bacteriophage RNA polymerase gene under the control of the IPTG inducible *lacUV5* promoter. The *lac* repressor protein is responsible for maintaining inducible control over the *lacUV5* promoter as well as the T7-*lac* promoter on the vector. In case of protein toxicity (like in the case of HpaA) BL21pLysS(DE3) cells are used for expression of toxic proteins. BL21pLysS(DE3) cells contain pLysS that constitutively expresses low levels of T7 lysozyme, which reduces basal expression of recombinant genes by inhibiting basal levels of T7 RNA polymerase. LB medium is inoculated with a single colony from a fresh plate and treated with the appropriate antibiotic and shaken (180 rpm) at 37°C overnight. Fresh LB medium containing the appropriate antibiotic is inoculated with the overnight culture and incubated with shaking at 37°C. The expression is induced by adding from 0.5 up to 1 mM IPTG, at OD_{600} of 0.6. After induction of 4 h at 28°C (FlgE2, FlgK and the complex FlgD-FlgE2), 37°C (HP1457) or room temperature (HpaA and HP0870), cells are harvested by centrifugation (5000 rpm, 15 min, 4°C) and stored at −20°C. Afterwards, cells are thawed on ice, re-suspended in a protein specific buffer (Table 2.10), and disrupted with the One Shot Cell breakage system (Constant System Ltd.) at 1.35 kbar. To prevent serine protease cleavage, the lysis buffer contains phenylmethylsulfonyl fluoride (PMSF) as an inhibitor. The lysate is later clarified by centrifugation (18000 rpm, 30 min, 4°C) and the soluble fraction is used in the further of purification.

FlgE2 **FlgK** **HP0870**	**HP1457**	**HpaA**	**FlgE2-FlgD**
20 mM Tris-base pH 7.5 150 mM NaCl 20 mM Imidazole 1 mM PMSF	20 mM Tris-base pH 7.4 150 mM NaCl 20 mM Imidazole 1 mM PMSF	30 mM Tris-base pH 8.0 300 mM NaCl 40 mM Imidazole 1 mM PMSF	20 mM Tris-base pH 7.5 150 mM 20 mM Imidazole 1mM PMFS

Table 2.10: Composition of buffers used in cell disruption for each specific protein.

Expression of selenomethionine-labeled FlgE2 (HP0908) Expression of recombinant selenomethionine derivative proteins is carried out using the non-methionine auxotrophic *E. coli* strain BL21(DE3). The bacteria are grown up in the Molecular Dimension kit SelenoMet *Dream* Medium, prepared according to the kit protocol. The expression is performed overnight at 30°C. The expression method takes advantages of the diautoxic growth of *E. coli* in order to auto-induce the target protein expression in higher yields. As in the case of the native protein, lysis is performed using One Shot Cell breakage system at 1.35 kbar and adding 1mM PMSF as inhibitor. The lysate is later clarified by centrifugation and the soluble fraction is used in the further of purification.

2.2.6 Protein Purification

Affinity Chromatography

Affinity chromatography on Ni-NTA agarose resin The purification of polyhistidine affinity-tagged proteins can be facilitated by the use of the nickel-nitrilotriacetic acid resin (Ni-NTA). The nitrilotriacetic acid behaves as a chelator coordinated by the nickel ions, which occupies two out of six coordination sites. The four more available sites are able to strongly coordinate the imidazole ring from the histidine-tag, selecting the fused protein rather than the bacterial lysate.
In this thesis, Ni-NTA affinity purification of proteins (HP1457, FlgE2, HP0870, HP1119 and HpaA) containing the His_6-tag is performed using the 1 mL prepacked columns or loose resin. The resin was firstly equilibrated with protein specific buffer A (Table 2.11) and then the soluble part of lysate is added. In order to obtain high purity samples, the column is connected to an ÄCTA Purifier system, and a washing step of 50 mL of buffer A is performed to avoid unspecifically bound proteins. Finally, the elution is performed in buffer B with a stepwise gradient of imidazole from 20 up to 600 mM.

	FlgE2 FlgK HP0870	**HP1457**	**HpaA**
Buffer A	20 mM Tris, pH 7.5 150 mM NaCl 20 mM Imidazole	20 mM Tris, pH 7.4 150 mM NaCl 20 mM Imidazole	30 mM Tris, pH 8.0 300 mM NaCl 40 mM Imidazole
Buffer B	20 mM Tris, pH 7.0 150 mM NaCl 400 mM Imidazole	20 mM Tris, pH 7.0 150 mM NaCl 400 mM Imidazole	30 mM Tris, pH 8.0 300 mM NaCl 600 mM Imidazole

Table 2.11: Composition of buffers used in Ni-NTA affinity chromatography.

Affinity Chromatography on Strep-tactin Sepharose resin The purification of Strep-tagged protein is based on the principle of the binding of biotin to streptavidin. A eight amino acids peptide (WSHPQFEK) is engineered to bind the biotin binding pocket of streptavidin and to attach to the Step-Tactin resin. In this thesis, Strep-Tactin resin has been used to purified the complex FlgD-FlgE2. FlgD was designed to locate the Strep-tag at the C-terminus. The purification is carried out on 0.5 ml of slurry resin. To proceed with the purification, the resin is first equilibrated with the binding buffer (buffer A in Table 2.12). The supernatant from the cell lysate is added at the resin and incubated at 4°C for one hour. The resin is later washed with a 40 ml of buffer A and eluted after one more hour of incubation in buffer B. The resin is regenerated after the elution with a solution of 1 mM of HABA (2-[4'-hydroxy-benzeneazo]-benzoic acid) in binding buffer.

	FlgE2-FlgD
Buffer A	20 mM Tris, pH 7.5 150 mM NaCl
Buffer B	20 mM Tris, pH 7.5 150 mM NaCl 2.5 mM D-desthiobiotin

Table 2.12: Composition of buffers used in Strep-tactin affinity chromatography

Size-Exclusion Chromatography

Analytical size-exclusion chromatography is used to separate proteins according to the different molecular weight and to determine homogeneity in a protein sample. The columns used are Superdex 200 10/300−16/60 and Superose 12 16/60 for the separation of proteins in the range MW = 10,000-600,000 Da and MW = 1,000−300,000 Da, respectively (GE Healthcare). Prior to use, the column is equilibrated at a flow rate of 0.5−1 mL/min with 1.5−2 CV of the elution buffer. Fractions containing the pure protein are pooled and concentrated for structural characterization assays.
To estimate the oligomeric state of protein, a calibration is carried out with protein standards: vitamin B12 (1.35 kDa), myoglobin (17 kDa), ovalbumin (44 kDa) and γ-globulin (158 kDa), whilst the void volume is determined by using Blue Dextran (GE Healthcare).

FlgE2 **FlgK** **HP0870**	**HP1457**	**HpaA**
20 mM Tris-base, pH 7.5 150 mM NaCl	20 mM Tris-base, pH 7.5 150 mM NaCl	30 mM Tris-base, pH 8.0 300 mM NaCl

Table 2.13: Composition of buffers used for SEC.

2.2.7 Protein biochemical methods

Removal of the His_6-tag from HP1457Δ(1-49) and HP1457Δ(1-21) using Thrombin

Purified His-tagged HP1457 protein is incubated with thrombin (0.1 U/μl) in a protein-protease ratio 1:100, and buffer containing 20 mmol/l Tris-HCl pH 7.4, 150 mmo/l NaCl. The reaction mixture is incubated overnight at 4°C and gently shaken. The reaction solution is passed on a benzamidine sepharose resin (GE Healthcare) to eliminate thrombin and thus prevente secondary enzymatic activity. Finally, the cleaved species are separated from the uncleaved ones performing a size-exclution chromatography.

2.2.8 Protein Characterization

Circular Dichroism Circular dichroism (CD) refers to the differential absorption of the left and right components of circularly polarized light. The CD instrument displays the dichroism at a given wavelength of radiation expressed as either the difference in absorbance of the two components or as the ellipticity in degrees. Chiral arrangement of peptide bonds in secondary structures exhibits distinct far-UV CD spectra absorption.
During the thesis, CD in the far-UV region (195−260 nm) was used for secondary structure analysis of protein. Far-UV CD spectra were recorded by a Jasco spectro-polarimeter in a quartz cuvette with a path length of 0.5 mm and a volume of 70 μL. All measurements were performed with water diluted protein samples (c = 0.1−0.5 mg/mL) at room temperature. All spectra were corrected for buffer absorption prior to analysis. Spectra were recorded with a step resolution of 1 nm and a bandwidth of 2 nm. The scan speed was set to 50 nm/min and 3 spectra were accumulated.

Light Scattering Dynamic and static light scattering techniques (DLS, SLS) are used to investigate the size distribution, stability and aggregation state of protein molecules in solution. The main difference between DLS and SLS is that the DLS instrument measures at a fixed angle (90° to the incident laser light) and can only determine the mean particle size in a limited size range, while the multi-angle SLS instruments can determine the full particle size distribution.

Dynamic Light Scattering (DLS) The basic principle in a dynamic light scattering experiment (also known as photon correlation spectroscopy or quasi-elastic light scattering) is that the sample is illuminated by a laser beam and the fluctuations of the scattered light are detected at a known scattering angle Θ by a fast photon detector. Molecules in solution obey Brownian motion, which means that small molecules move faster than large ones.
The obtained information during the DLS analysis is about the time scale of the movement of the scatterers. The translational diffusion coefficient D_t is derived from these data using an auto-correlation function. Since the hydrodynamic radius R_h is related to the D_T of the particles, R_h can be calculated using the Stokes-Einstein relation: $D_t = \frac{k_B T}{6\pi\eta R_h}$, where k_B is Boltzmann's constant (1.381×10^{-23} J/K) and η the dynamic viscosity of the solvent.
In this study, DLS measurements of proteins (Zetasizer Nano ZS, Malvern Instruments Ltd) were performed to examine the dispersion of samples at concentration of 5−10 mg/mL at 298 K. Prior to use, samples were centrifuged for 10 min at 10,000 rpm, and 40 μL of supernatant was transferred into the quartz ZEN 2112 cuvette.

2.2.9 Thermophoresis Assay

Microscale thermophoresis measures the difference in the molecular motion along with a localized microscopic temperature gradient, which induces a variation in charge, hydration shell and global size of a complex. An infrared laser induces the temperature variation, whilst the movement of the molecules is detected by the conjugation of the target protein with a fluorescence dye.
The samples are inserted in capillaries and are free to move in solution. Each capillary

contains a different concentration of the ligand protein (in this case FlgE2), whilst the concentration of the dye is kept constant. The IR-laser heats each capillary inducing a variation in the fluorescence intensity due to the molecular motion. After the laser is turned off, the molecules can move back and re-establish the fluorescence signal. The signal is recorded in all capillaries with varying concentration of the non-fluorescent ligand. Any change of thermophoretic properties is observed as a change in fluorescence intensity and can be related to the dissociation constant of the complex. In mathematical terms, the steady-state concentration ratio of the fluorescent species is related to the normalized fluorescence, which is defined as $F_{normalized}=\frac{F_{hot}}{F_{cold}}$. In a linear approximation the same equation can be expressed as: $F_{normalized}=1+(\frac{\partial F}{\partial T}-S_T)\Delta T$, where S_T is the Soret coefficient that gives information about the depletion of the number of molecules located at high temperature zones; $\frac{\partial F}{\partial T}$ is the variation in fluorescence induced by a change in temperature and ΔT represents the temperature gradient. Bound and unbound species contribute to the fluorescence variation, so the same equation could be rewritten in a specie-dependent way: $F_{norm}=(1-x)F_{norm}(A)+xF_{norm}(AT)$, with x the ratio of bound and bound molecules, $F_{norm}(A)$ the normalize value for unbound molecules and $F_{norm}(AT)$ the value for bound ones. Using the action mass law, the normalized value of fluorescence can be directly related to the K_D, in order to determine the affinity between one protein and its ligand. The MST analysis was performed in collaboration with prof. Jasmine Rokov Plavec, Dept. of Biochemistry, university of Zagreb.

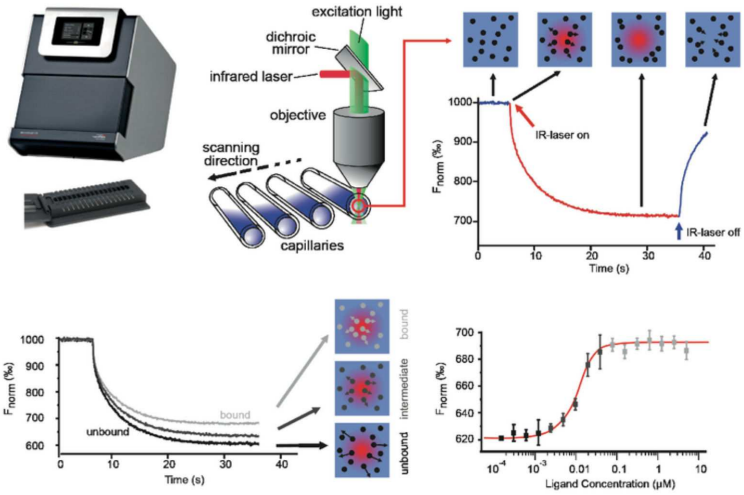

Figure 2.5: Microscale Thermophoresis Equipment. The measurement is performed according to the following steps: 1) one of the binding partner is labeled using standard fluorescent labeling protocols; 2) a titration series of up to 15 dilutions is prepared, where the concentration of the fluorescent binding partner is kept constant and the concentration of the unlabeled molecule is varied; 3) when the reaction is at the equilibrium, it is transferred into a glass capillary. The capillary is placed on the sample tray, which is moved into the instrument; 4) the machinery automatically measures the thermophoresis signal of each sample and calculates the dissociation coefficient.

2.2.10 High-performance liquid chromatography and mass spectrometry

Mass spectrometry is used to monitor the level of degradation of HP1457Δ(1-21). The fractions of the gel-filtered protein are dried out and dissolved back in deionized water. The mass of the soluble form of HP1457 is determined by mass analysis of the peaks isolated by reverse phase chromatography (C4-column, RP−HPLC).The samples are charged on a reverse-phase HPLC column (Synergi MAX−RP 80A, 150 × 4.6 mm, 4 μm (Phenomenex)), applying the following method:

Time (s)	%Buffer B	Flow (ml/min)	Pressure (MPa)
0	0	0.8	400
10	4	0.8	400
40	60	0.8	400
50	100	0.8	400
55	100	0.8	400
56	4	0.8	400
80	4	0.8	400

Table 2.14: Method used to perform the elution of HP147Δ(1-21) on a HPLC C-4 column.

and using the following buffers:

- Buffer A: Pure MilliQ Water, 0.1% of TFA
- Buffer B: 95% CH_3CN, 0.1% of TFA

The collected fractions are dried up and re-suspended in water supplemented with 0.1% formic acid and directly injected in the nano−ESI source. Mass measurements are performed with a quadrupole−TOF spectrometer (Waters, Manchester, UK) (capillary voltage: 2800−3000 V; cone voltage: 45 V; scan time: 1 s; interscan: 0.1 s). Analyses of the spectra were performed by using the MASSLYNX software (Micromass, Wynthenshow, UK).
Both mass and HPLC experiments were performed in collaboration with Prof. P. Polverino de Laureto from the Department of Pharmaceutical Sciences, University of Padova.

2.2.11 Crystallization Method: Vapour Diffusion

Crystallization trials are prepared using the sitting and hanging drop vapour diffusion method. Every crystallization drop is prepared by mixing equal volumes of protein and precipitant solutions (500 nL), and equilibrated against 75 μL of the mother liquid solution. Water evaporates from the protein containing droplet since the precipitant concentration is lower than in the reservoir. The system equilibrates creating a supersaturated condition in the drop, therefore increasing protein and precipitant concentrations and the chance of crystal to grow.
The initial crystallization trials are carried out exploring 96 different conditions from

Molecular Dimension and Quiagen's kits. The drops are prepared by Oryx 8 crystallization robot. Afterwards, the crystallization conditions are improved by optimizing several parameters, like: drop size, volume ratio between precipitant and protein, and ratio of components of the precipitant solution. Additionally, some additives and/or crystals seeds can be included in the precipitant solution, in order to stabilize and facilitate protein crystallization.

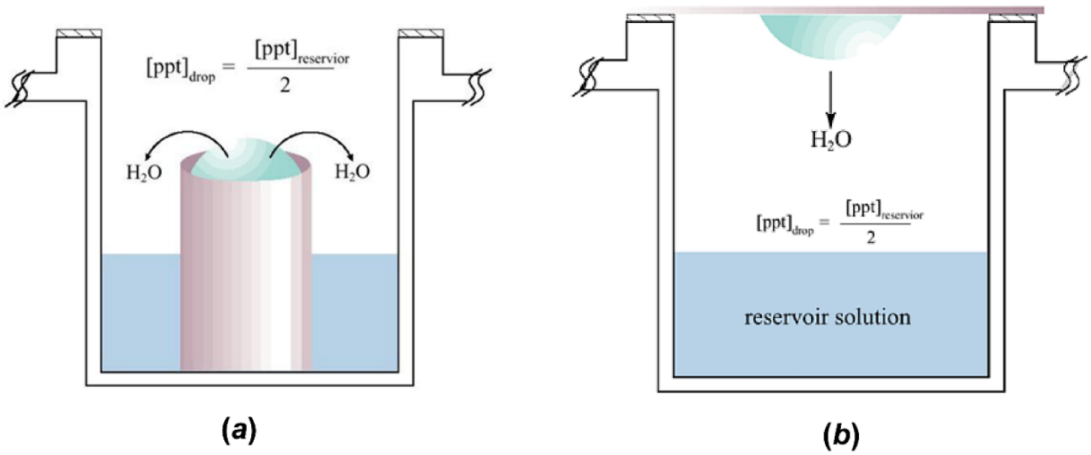

Figure 2.6: Crystallization experiment using the Vapour Diffusion Method in: a) sitting drop and b) hanging drop fashion.

2.2.12 Data processing and structure determination

Diffraction data of the native orthorombic form of FlgE2 were measured at the ID23-1 beamline at the European Synchrotron Radiation Facility (ESRF, Grenoble, France). Preliminary diffraction tests were performed at the PXIII beamline of the Swiss Synchrotron Light Source (SLS, Villigen, Switzerland). In case of SeMet-FlgE2 a fluorescence emission scan at the selenium edge was performed to optimize the energy for anomalous measurements. Diffraction data of FlgE2 were indexed and integrated by the softwares MOSFLM [82] and XDS [83]. All datasets were merged and scaled with Scala, contained in the CCP4 [84] crystallographic package. Initial phases were obtained using single-wavelength anomalous dispertion (SAD) method, through SHARP/autoSHARP pipeline [85], conducting density modification and model building. The model was then used for refinement against the native data set at higher resolution by the software Phenix [87], and checked and manually adjusted with the graphic software Coot [88]. For the same purpose, another model was achieved undergoing the native data set to the software Phaser [86] and solving the phase problem by molecular replacement using the structure of FlgE from *Salmonella typhimurium* as a template.

2.2.13 Cell culture preparation and confocal microscopic analysis

In order to check the adhesion of sfGFP-HpaA to cell surface, a AGS culture was prepared. Cells are grown in Ham's F12 medium (Thermo Scientific), supplied with 10% of FBS (Euroclone), 100 units/ml of penicillin, 100 μ/ml of streptomycin and 4 mM of Hepes.

Cells are seeded on glass coverslips and allowed to grow up to 60% of confluence, roughly. After 24 hours, the cells are treated with two different concentrations of both sfGFP-HpaA and sfGFP (as a negative control), 120 μg/ml and 80 μg/ml. The cells are firstly incubated for 30 min at 4°C, and secondly fixed using 3.7% of formaldehyde in phosphate-buffered saline (PBS; 140 mM NaCl, 2 mM KCl, 1.5 mM KH_2PO_4, 8 mM Na_2HPO_4, pH 7.4). They are washed twice by PBS. GFP fluorescence is checked out by Leica Confocal SP5 microscope and images are acquired by using Leica AS software. Finally, raw images are smoothed by applying the Gaussian blur filter of the Image J 1.50e software.

3 | Structural studies of proteins from *H. pylori* flagellum

3.1 Introduction

3.1.1 Flagellar architecture

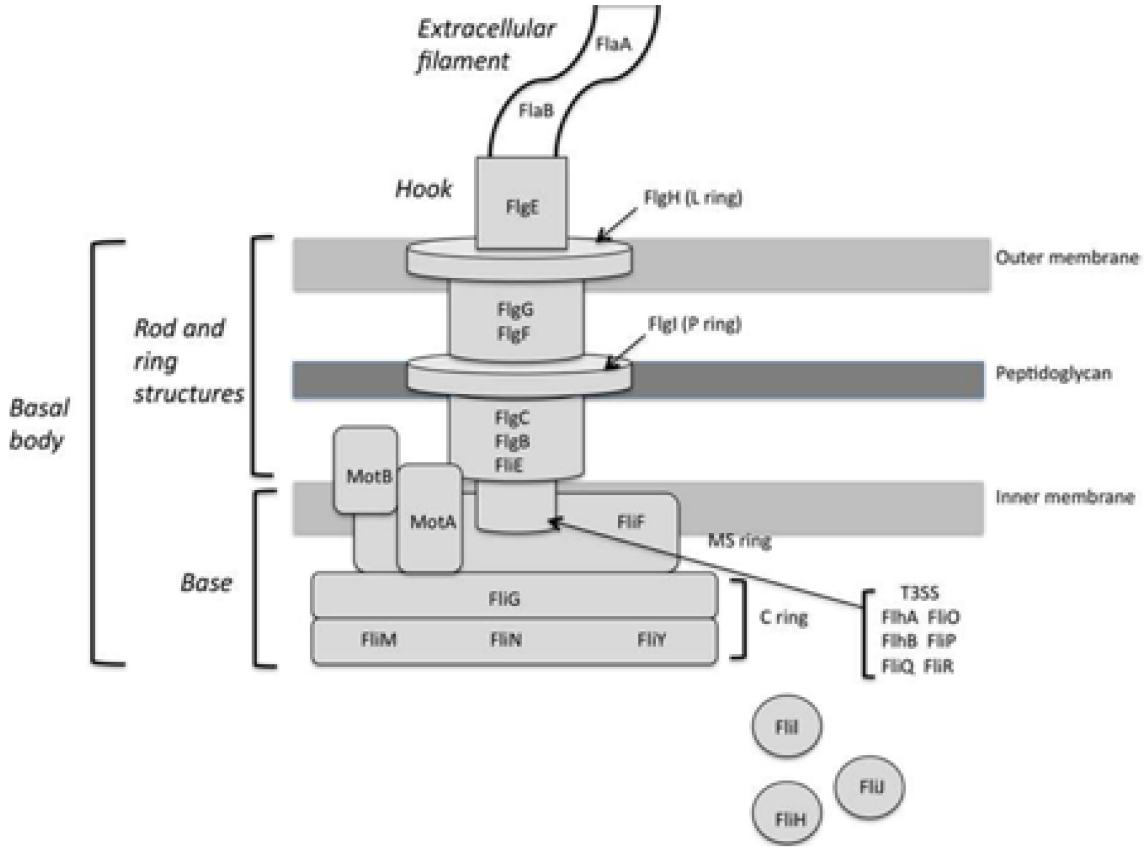

Figure 3.1: General flagellar basal body (Scheme based on [6]).

H. pylori infection usually takes place at the level of the mucus protecting the gastric epithelium. During the colonization process, the bacterial flagellum is an essential determinant, which enables the motion from within the mucus to a microenvironment necessary for the bacterial growth. Its mutants were demonstrated to attenuate their colonization in human hosts [89]. Moreover, *H. pylori* flagella guarantee an unsual high velocity to the bacterium, which corresponds to roughly twice the velocity of those of rod-shaped bacteria (ranging from 22 to 40 μm/s) [90]. Indeed, flagella provide a torque propulsion, that together with the helical shape of the bacterium, creates a corkscrew-like rotation and promotes motily across the high viscosity mucus [91]. Approximately thirty proteins actively participate to the structural composition of flagellum, but from 40 to 100 are required to build up a chemoresponsive machinery [92]. Unfortunately, none of them is clustered in a specific operons [6].
Several studies were conducted on the bacterial flagellar motor of *Escherichia coli* and *Salmonella typhimurium* [71] and, despite the crystal structure of some of the separated components is known, still a complete model of the flagellum has not been proposed yet. Since a high number of flagellar proteins are exposed outside the cell surface, they can

become new targets for the vaccine treatment of bacterial infections. From a technological point of view, the understanding of the working principles of flagellum could open new path in the design and assembly of new nano-machines bio-inspired and bio-energetics.
In this chapter, several crystallization attempts of some soluble proteins from *H. pylori* flagellum will be presented and a preliminary structure for the protein HP0908, a flagellar component, will be proposed.
The general organization of *H. pylori* is conformed with the one already proposed for *E. coli* [91] and *Salmonella typhimurium* [93][94]. It can be divide in two main components: 1) the hook-basal body (HBB) and the extracellular filament, masked by a sheath. Further, the HBB is composed by: 1) a base located in the cytoplasm; 2) a periplasmic rod with the associated rings and 3) the hook, localised on the cell surface. The base is composed by the MS ring, the Type 3 Secretion System (T3SS), the C ring and the motor.

The basal body

In Gram-negative bacteria the basal body is embedded in the cell envelope and consists of a set of membrane rings, called MS ring and a P ring, that passes through the periplasmic space and reaches the outer part of the membrane until the L ring (Figure 3.1). It has the role of transmitting the torque force from the motor to the filament [95]. The bacterial motor plays a paramount role for the conversion of the chemical energy, due to the translocation of cation through the cell membrane, in torque motion. The flagellar motor is basically composed by two parts: the stator, a complex non-convalentely attached to the peptidglycan layer, and a rotor, in contact with the MS ring and responsible for the torque generation [91].
The stator is composed by two transmembrane proteins, MotA and MotB, which associate together forming the complex $MotA_4MotB_2$ [96] (Figure 3.2 and Figure 3.3). Its function is to create a proton flux across the cell membrane, which induces a potential difference used by the system to generate the rotational motion. Indeed, the complex interacts with FliG in the MS ring inducing the torque [95]. Zhou *et al.* observed that in *Salmonella typhimurium* one transmembrane helix from MotA together with two from MotB contribute to the formation of the proton channel [97][98]. The complex MotA-MotB is supposed to remain inactive in the bacterial membrane untill its incorporation into the motor. At this stage, the linkers connecting the transmembrane helices to the peptidoglycan-binding domain are extedend [99]. The periplasmic P ring is proposed to be a binding site for the stator $MotB_2$ complex. Moreover, exploring the structure of MotB (3SOY) from *H. pylori*, the following mechanism was proposed for the conformational change of the mature protein in the complex: 1) unfolding of the linkers; 2) interaction with the peptidoglycan layer; 3) opening of the channel [100]. The lack of the anchoring mechanism will compromise the proper mechanism of the motor [101][98].
The rotor is responsible of the torque mechanism and it is usually oriented in a counterclockwise direction. Nevertheless, a switch complex is responsable of the torque reversal [102]. The motion is controlled by a chemotactic path, consisting in the proteins CheW/CheA and CheY. In biref, the chemoreceptor CheW breaks its bound to the histidine kinase CheA, which phosphorilate and becomes active, activating CheY which interacts with the N-termina of FliN and switches the motion from counterclockwise to clockwise [103].

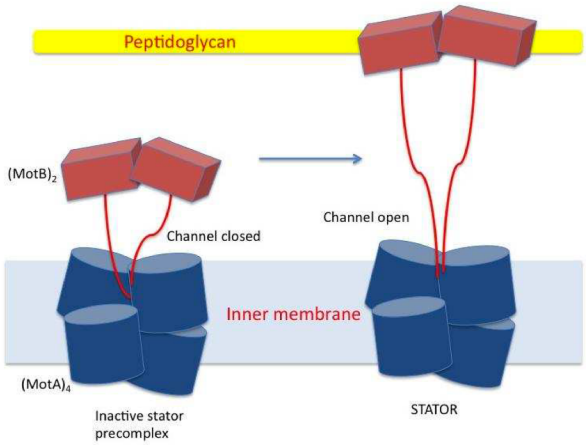

Figure 3.2: Hypothetical activation mechanism of $MotA_4MotB_2$ complex inside the membrane. The mechanism is specific for H. pylori (Scheme based on [100]).

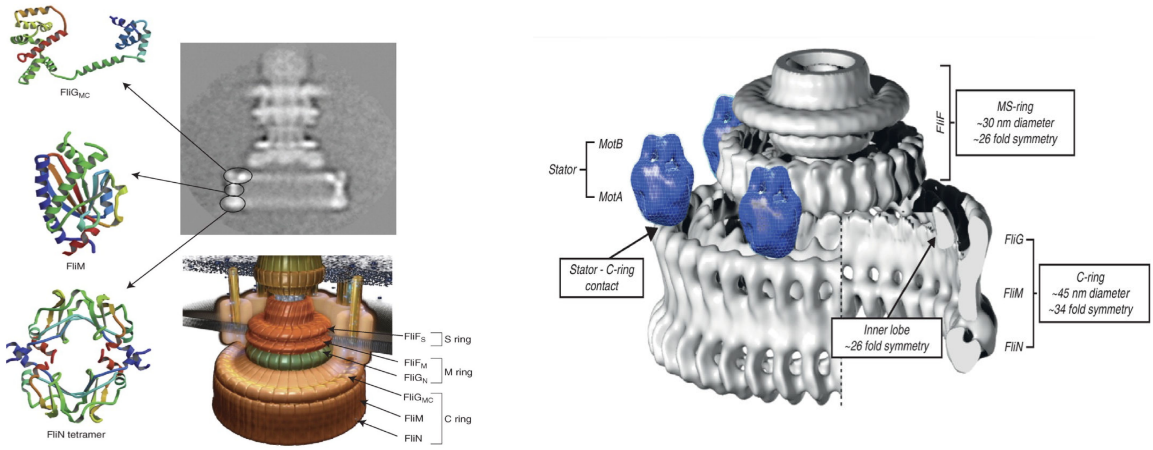

Figure 3.3: Schematic representatin of the flagellar basal body. On the left, the structure recorded by cryo-electron microscopy of the basal body is reported (adapted from [68]). Three proteins are supposed to fit the electron density and are reported from the top to the bottom as following: FliG, FliM and FliN. On the right, the same picture has been supplemented of the stator, which is located over the C ring (adapted from [71]).

The proteins involved in the switching are FliG, FliM and FliN present in 26, 34 and 136 copies in *Salmonella typhimurium* rotor, respectively [104]. They are arranged in a ring at the base of flagellum. Differently from other bacteria, *H. pylori* includes in its genome one more protein involved in flagellum switch regulation. The four motor switch proteins are identified as FliG (HP0352), FliM (HP1031), FliN (HP0584) and FliY (HP1030) (Fig. 4). Chen *et al.* demonstrated by some cryo-EM studies that the main core of C ring in most of the bacteria is principally regulated by the presence of FliG, FliM and FliN, but not FliY [95]. According to Lowenthal *et al.*, FliY has an N-termianl domain that belongs to CheC/CheX/FliY/FliM family, so it may substitute FliM in FliG C-domain interaction [105]. FliN plays a structural role in building the flagellar architecture, too. It binds the flagellar export protein FliH and localizes itself with FliI and FliJ. In *H. pylori*, FliN C-terminal domain is fused with a phosphatase/CheC-like domain [106][107]. Moreover,

FliN is involved in the direct binding of the regulator protein CheY-P [107]. Lowenthal *et al.* performed some alignment of FliN and FliY sequences taken from well characterized bacteria and they noticed some highly conserved structural regions. Taking into account the differences, alignment data seems to suggest that both FliN and FliY are involved in export and protein-protein interaction, beyond the motility function [105].

The export apparatus

The general assembly of the export apparatus is reminiscent of the Type 3 Secretion System (T3SS), a virulence factor common to most Gram-negative bacteria. It consists of six membrane proteins (FliO, FliP, FliQ, FliR; FlhA and FlhB) and three soluble (FliH, FliI and FliJ) [108]. All of them are associated to the MS-ring and assemble together to build the export machinery. Unfortunately nothing is known about *H. pylori* ATPase proteins (FliH, FliI and FliJ) involved in the export process of hook and filament components, but some pieces of information can be recovered on the membrane protein from their orthologous from *Salmonella typhimurium*.
The cytoplasmic domain of FlhA from *H. pylori* is the best characterized protein, whose structure was solved in 2010 [109]. Biochemical studies confirm the membrane region from FlhA interacts with FliF (the only components of the MS ring) [110], whist the interaction with the transcriptional factor FlgM induces a modulation of the translocated proteins [111].
FlhB from *Salmonella typhimorium* was demonstrated to play a structural role in the rod and basal body formation and regulates the length of the hook [112][113].
HP1419 shows high similarity in sequence with FliQ from *Salmonella typhimurium*, which is demonstrated to be a membrane protein involved in the export flagellar apparatus [114]. *H. pylori fliQ* mutants confirm this hypothesis, since they present a reduced ability of the bacterium to adhere to the gastric cells and the development of a non-motile species [115].
HP0685 is the membrane protein orthologous to *Salmonella typhimorium* FliP protein. It is another flagellar export component, which is supposed to assemble and form a component containing five subunits [116].
FliO appears to be largely conserved in the export apparatus from several different organisms. In *H. pylori*, its role is played by HP0583. *fliO* mutants show a nonflagellate nature [117]. Since *fliO* knockout induces a decrease in FlhA and RpoN-regulon (protein involved in biogenesis of flagellar genes) expression in *H. pylori* cells, it was postulated that FliO is essential for the transcription of RpoN-regulated flagellar genes (like FlaA and FlgE).
The MS-ring consists of FliF, which in *H. pylori* is supposed to be protein HP0351. The orthologous protein from *Salmonella typhimurium* is well characterized: it assembles in a ring with a diameter of 25 nm. Moreover, an axial projection map of the ring reveals a central channel, involved in the export system of flagellins [118].
The flagellar assembly goes on with the formation of the periplasmic elements of the rod (FliE, FlgB, FlgC, FlgF and FlgG), the P-ring (FlgI) and the L-ring (FlgH). Among them, FliE has a role as a structural adapter between the MS-ring and the rod [119], whilst FlgG is placed in a distal part of the rod [120]. For FlgB, FlgC and FlgF it is not clear the order along the rod and the regulatory mechanisms that determines the rod length. Plus, FlgI and FlgH form around the rod the P and the L ring, respectively. The first is located into

the peptidoglycan layer and the second in the outer membrane of *H. pylori* flagellum [6].

The hook and the hook-associated proteins

The hook is a highly curved, tubular structure that bridges the basal body and the filament. It is composed of about 120 copies of a single protein, FlgE. Up to now, two crystal structures of FlgE was deposited in the PDB database; they are from *Salmonella enterica serovar Typhimurium* (PDB ID 1WLG; [121]) and *Campilobacter jejuni* (PDB ID 5AZ4). A model of the hook was proposed docking the crystal structure of *St*FlgE in the electron density map recorded by cryo-electron microscopy [121]. The hook seems to be composed of three main domains: the outermost domain at the surface (7.5 nm), the middle domain (5-6 nm), which allows the interaction between the hook flagellins, and the inner core domain that forms a channel for the protein translocation (1nm thick; 3 nm axial lumen) and where the C- and N-terminus are supposed to be located (in a similar way to those flagellins in the filament) [122].
The proper assembly of the hook requires the participation of FliK, FlgD and FlhB. FliK is supposed to be involved in the regulation of the hook length [123][124], whilst FlgD is demonstrated to modulate the proper number of FlgE monomers assembled during the hook growth [125]. Finally, FlhB participates to the formation of the T3SS and is located in the inner membrane of the MS ring. It interacts with the C-terminus of FliK, triggering the stop signal of the export proteins, involved in the hook formation [126].
Two more proteins are located in between the hook and the filament structure. They are called *hook-associated proteins* (HAP) and are identified as FlgK and FlgL. They form a short hook-filament junction zone, important for the adaptation of the mechanical different structures. Indeed, the hook has a quite flexible structure, whilst the filament works as a propeller, with a much more rigid assembly. For this reason, the two HAPs must share a similar structure to both the hook protein and the filament.

The filament

The biosynthesis of the flagellum ends with the assembly of a filament. It acts as a helical propeller and it is composed of the major flagellin FlaA, the minor flagellin FlaB, coexpressed in a different amount, and a filament-associated cap protein, FliD [68]. They are respectively composed of 510 and 514 amino acids, with a sequence similarity of 61%. Both of them are essential for the bacterial motility, but their ratio FlaA/FlaB is influenced by environmental factors (pH, viscosity, etc.), providing the best filaments properties for the motility in certain conditions [127].
Besides, the crystal structure of FliS (3IQC) and HP1076 (3K1H) have been co-crystallized and solved. They were identified as chaperons of the proteins of *H. pylori* filaments. Lam *et al.* hypothesizes a role in flagellar biosynthesis due to the type of their interaction, involving a helical stacking with a measured association constant of $1.5\text{x}10^7$ M^{-1}. The role of FliS was supposed to be involved in the polymerization control of the filament proteins. Indeed, the last hundred C-terminal residues were considered essential for the interaction between FliS and FlaB. Since the two flagellins are supposed to interact, the same complex is assumed to be formed with FlaA, too. Moreover, the role of FliS could be related with a hypothetical interaction with FlgK and the cap protein FliD [128].
Finally, FliD behaves as a chaperon associated with FlaA and FlaB in the control of the filament length [129].

The transcriptional regulatory cascade

The expression of flagellar genes takes place in a hierarchical manner so that proteins can be produced by stage, favoring an order in secretion and inducing the proper interaction after their expression. For this reason, the flagellar genes can be divided into classes, according to their expression order [130]. Class 1 includes master transcription regulon and is expressed as first. That includes the genes encoding the T3SS proteins, the FlgSR two-component system (TCS), the FlhF GTPase, and σ^{54} [131][132]. The regulation includes a second pathway where the class 1 proteins modulate the class 2 gene expression through the factor σ^{54} [132].
The first interaction seems to take place at the level of the TCS with the two protein FlgS (histidine kinase) and FlgR (a transcriptional regulator and member of NtrC family of regulators). The T3SS proteins induce a phosphotransfer to the components FlgSR, even if more proofs of the interaction need to be reported [133]. The second pathway is required for expression of σ^{54}-dependent class 2 genes, which involves the FlhF GTPase, but the mechanism is still unclear [134].
The transcriptional factor σ^{28}, encoded by the gene *fliA*, is at the base of the expression of the class 3 genes and is negatively controlled by the antisigma factor FlgM. The interaction between the two factors induces a repression of the HBB proteins expression and causes the expression of the flagellins from the filament, which concludes the transcriptional phase of the flagellum [111][131].

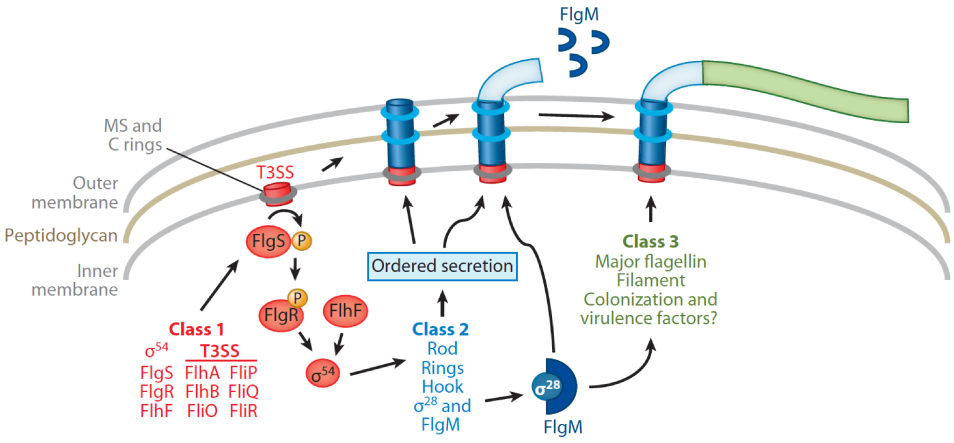

Figure 3.4: Molecular mechanism of the flagellar transcription. The flagellar genes and respective proteins included in the flagellar transcriptional cascade are shown. These factors are color-coded on the basis of their classification in the transcriptional regulatory cascade: red (class 1 genes), blue (class 2 genes), and green (class 3 genes), (adapted from [6]).

3.2 Results

3.2.1 The protein HP1119

The protein HP1119 is identified the hook-associated protein 1 (HAP1) FlgK, whilst the protein HP0295 is recognized as HAP3, also known as FlgL. The two proteins together provide a juction between the hook and the filament. The initialization of the filament assembly is guaranteed by the presence of a HAPs chaperon, identified as FlgN. The importance of its role was confirmed by some experiments performed on *Salmonella typhimurium* cell. Homma *et al.* noticed that *flgN* mutant phenotype closely resembles the flagellar assembly lacking HAPs [135], confirming that their role is paramount for the complete achievement of the flagellar structure.
The structural organization of the hook-filament junction can be gathered for two different organism: *Salmonella typhimurium*, whose FlgK and FlgL structures are known (PDB ID 2D4Y; 2D4X), and *Burkholderia pseudomallei* for whom FlgK structure has been recently solved ([136]; PDB ID 4UT1). From the analysis of the two FlgK sequences, two conserved domains are reported at the N-terminal and C-terminal, whilst their central regions are highly variable in length, likely as a consequence of the physiological environment of the two different organisms [136].

HP1119	
Number of residues	606
Molecular weight	68 kDa
Calculate molar extinction coefficient	46760 $cm^{-1}M^{-1}$

Table 3.1: Protein information and proprieties.

In analogy with the two already mentioned proteins, *Hp*FlgK is composed of 606 amino acids and presents a C-terminal domain recognized by Pfam as 'flagellar basal body rod FlgEFG protein C-terminal domain' (Pfam PF004609). The alignment of the sequences from *Hp*FlgK and *St*FlgK (Figure 3.5) shows also in this case that the two conserved domain are localized at the N- and C-terminus.
FlgK was successfully cloned, expressed and purified by affinity and size-exclusion chromatography (Superdex 12 10/300), as reported in Chapter 2, and a Histidine-tag was added at the C-terminus in order to selectively purify the protein from the bacterial supernatant.

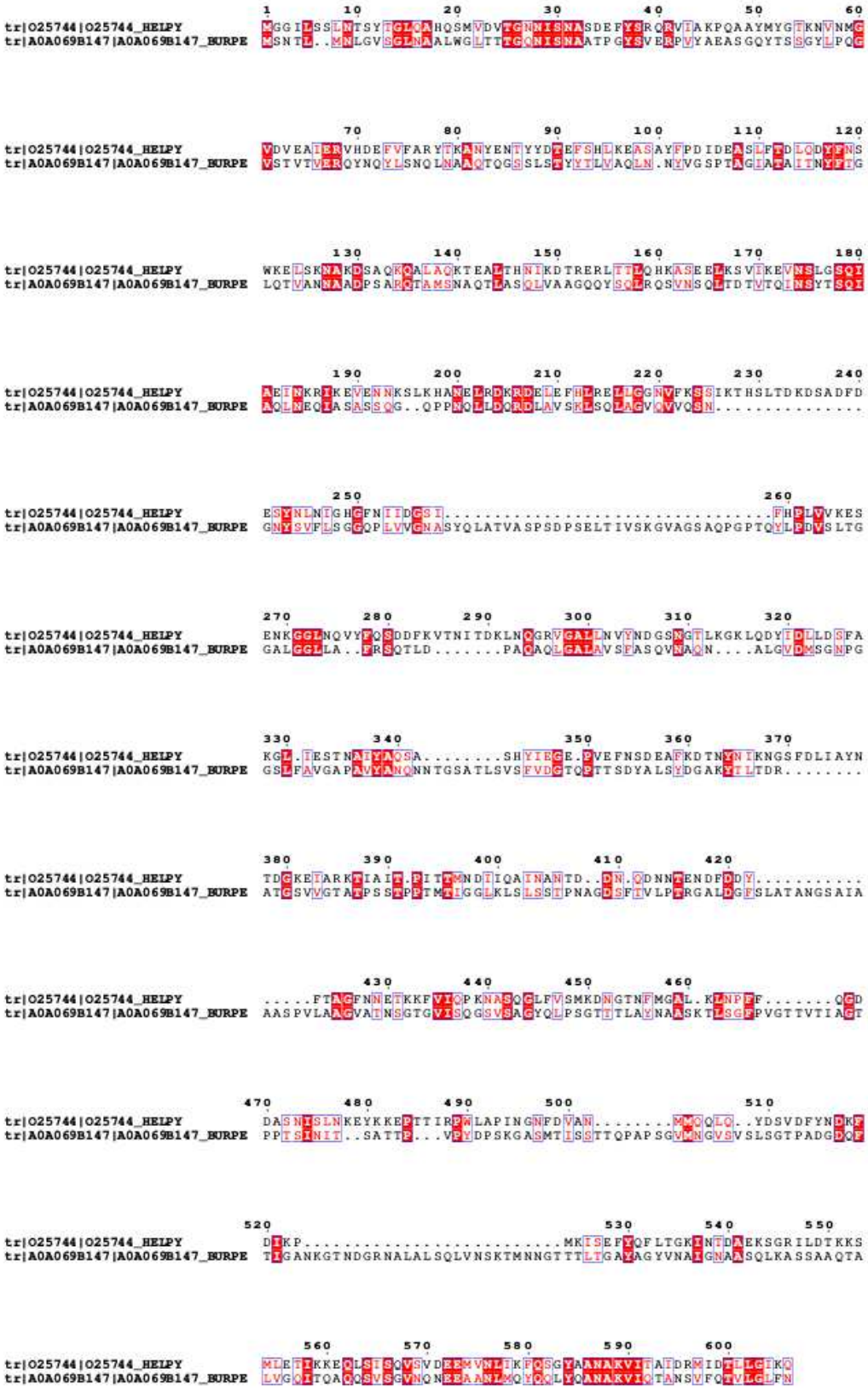

Figure 3.5: Sequence alignment of HpFlgK and BpFlgK. The two sequences share 22% of identity. The alignment was performed using the ClustalO server (https://www.ebi.ac.uk/Tools/msa/clustalo/) and designed by Espript3.0 (http://espript.ibcp.fr/ESPript/ESPript/).

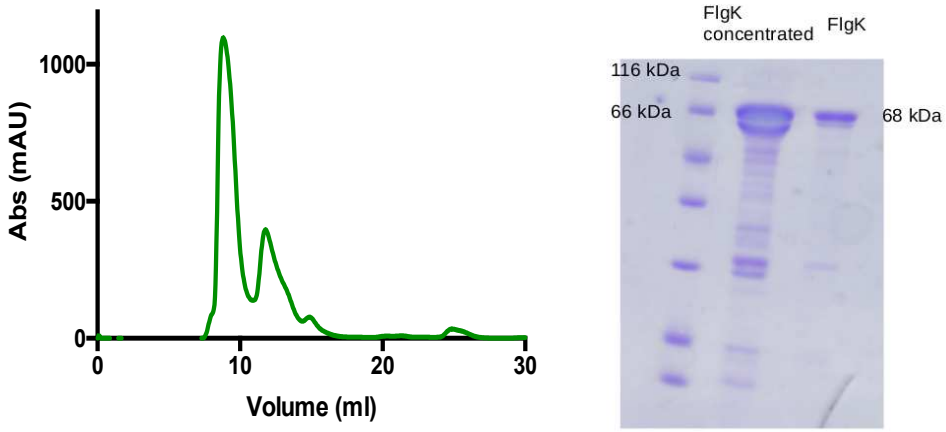

Figure 3.6: SDS-PAGE (right) of HP1119 purified after gel-filtration; elution profile of the purified protein (left). The elution profile shows two peaks and a shoulder: the first is eluted after 9 ml and corresponds to the aggregated form of the protein; the second peak is eluted after 12 ml and corresponds to the monomer and the shoulder eluted after 13.3 ml is a result of the partial degradation of the protein.

As it is evident form the gel-filtration profile, the protein is eluted in two main forms: the aggregate state, collected after 9 ml of elution, and the monomeric specie, after 12 ml of elution. The SDS-PAGE in Figure 3.6 represents the composition of one fraction under the first peak, highlighting the high level of degration of the protein. In order to see whether the aggregated form was properly folded, a CD spectrum was recorded and the analysis highlighted a strong contribute of α-helices due to the specific peaks located at 208 and 222 nm (Figure 3.7).

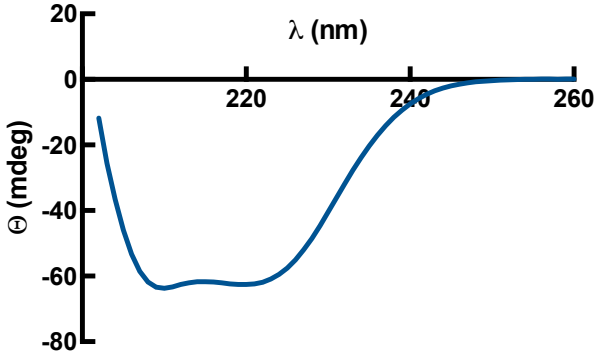

Figure 3.7: CD spectrum of the first peak fraction concentrated up to 12 mg/ml and diluted 1:20 in PBS.

All the fractions from the two peaks were used to perform several crystallization trials. The factions from the first peak were concentrated up to 12 mg/ml, the fractions from the second peak were concentrated up to 2.5 mg/ml, owing to the lower amount of purified protein. Although only the conditions F5 and E5 from the Structure Screen crystallization kit gave a reasonable result for the monomeric form of the protein (Figure 3.8), from the X-ray diffraction analysis the crystals turned out to be salts.

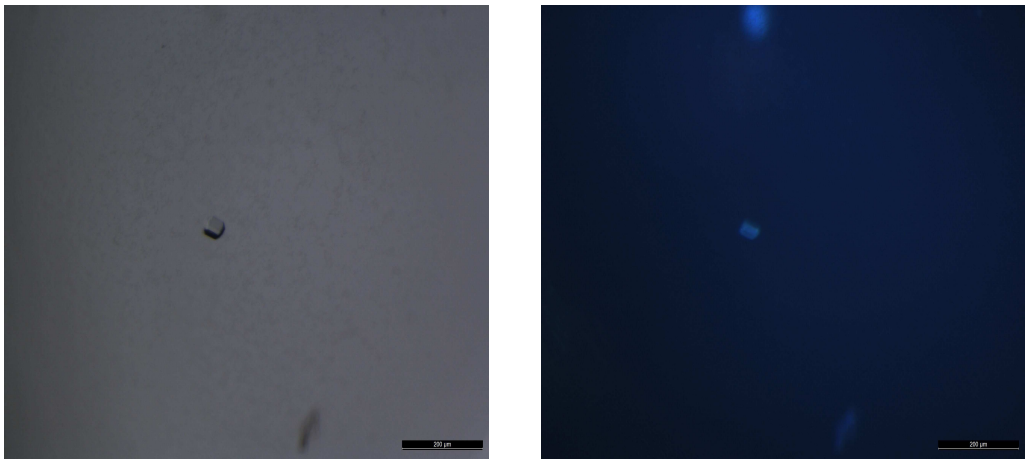

Figure 3.8: FlgK crystal from E5 (0.1 M Tris pH 8.5 and 25% v/v tert-Butanol) condition of the Structure Screen crystallization kit. On the left: stero image of the crystal recorder at whole spectrum light; on the right: the same image recorded using the fluorescence excitation filter.

3.2.2 The protein HP0870

In the case of *H. pylori* two different proteins HP0870 and HP0908, contribute to the hook assembly. They have been both recognized as FlgE and FlgE2 (Douillard 2009; Niehus, 2004), respectively, and share about 26% of sequence similarity (Figure 3.10). Moreover both of them are involved in the immunogenic response of the host against the bacterial colonization, so both can be treated as potential vaccine targets. In order to understand their role in the growth of the bacterial flagellum, both of them were expressed and characterized.

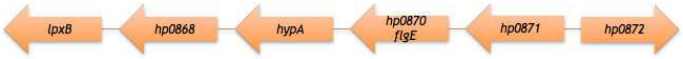

Figure 3.9: Location of the genes hp0870 in the bacterial operon.

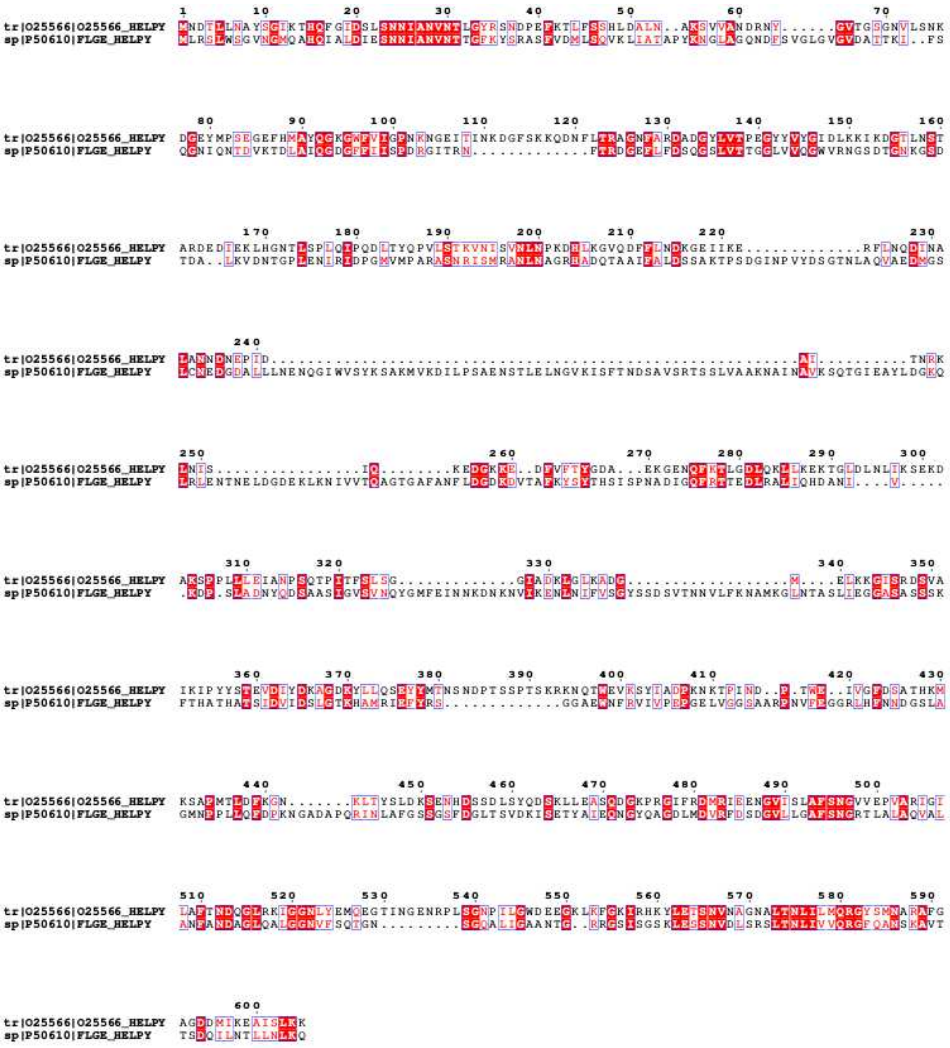

Figure 3.10: Sequence alignment between the proteins HP0908 and HP0870 respectively recognized as FlgE2 and FlgE. They share 26% of sequence identity. The alignment was performed using the ClustalO server (https://www.ebi.ac.uk/Tools/msa/clustalo/) and designed by Espript3.0 (http://espript.ibcp.fr/ESPript/ESPript/)

HP0870	
Number of residues	718
Molecular weight	76 kDa
Calculate molar extinction coefficient	41370 $cm^{-1}M^{-1}$

Table 3.2: Protein information and proprieties.

The protein HP0870 was successfully cloned, expressed and purified by affinity and size-exclusion chromatography (Superdex 12 10/300), as reported in Chapter 2, and a Histidine-tag was added at the C-terminus in order to selectively purify the protein from the bacterial supernatant.

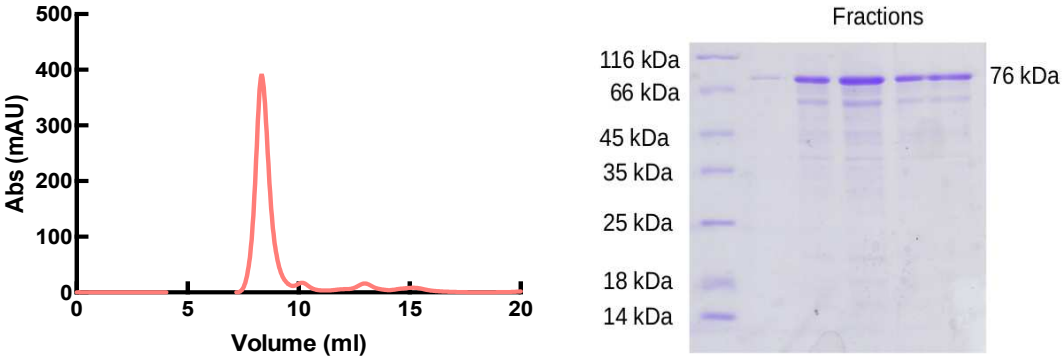

Figure 3.11: SDS-PAGE (right) of HP0870 purified after gel-filtration; elution profile of the purified protein (left). The elution profile shows one peak eluted after 9 ml and corresponding to the aggregated form of the protein.

The protein resulted to be completely eluted in an aggregated form (Figure 3.11) and to check its proper fold, a CD analysis was performed. The spectrum shows the protein is mainly constitute by β-strands, owing to its signal at 216 nm (Figure 3.12).

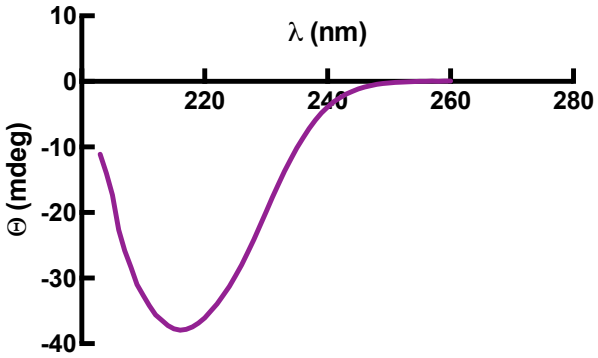

Figure 3.12: CD spectrum of the first peak fraction concentrated up to 14 mg/ml and diluted 1:20 in PBS.

Several crystallization tests have been performed utilizing the aggregate form of the protein, concentrated up to 14 mg/ml, but none of the conditions explored was successful.

3.2.3 The protein HP0908

The protein HP0908 contributes together with HP0870 to the hook assembly and is named FlgE2. In analogy with the domain observed in the protein HP1119, FlgE2 presents a flagellar basal body rod FlgEFG protein C-terminal domain. Comparing its sequence with the one from *Salmonella typhimurium*, it is noticeable that the two proteins share 27% of similarity, despite *Hp*FlgE2 is almost the double longer in sequence.

Figure 3.13: Location of the genes flgE2 and flgD in the bacterial operon.

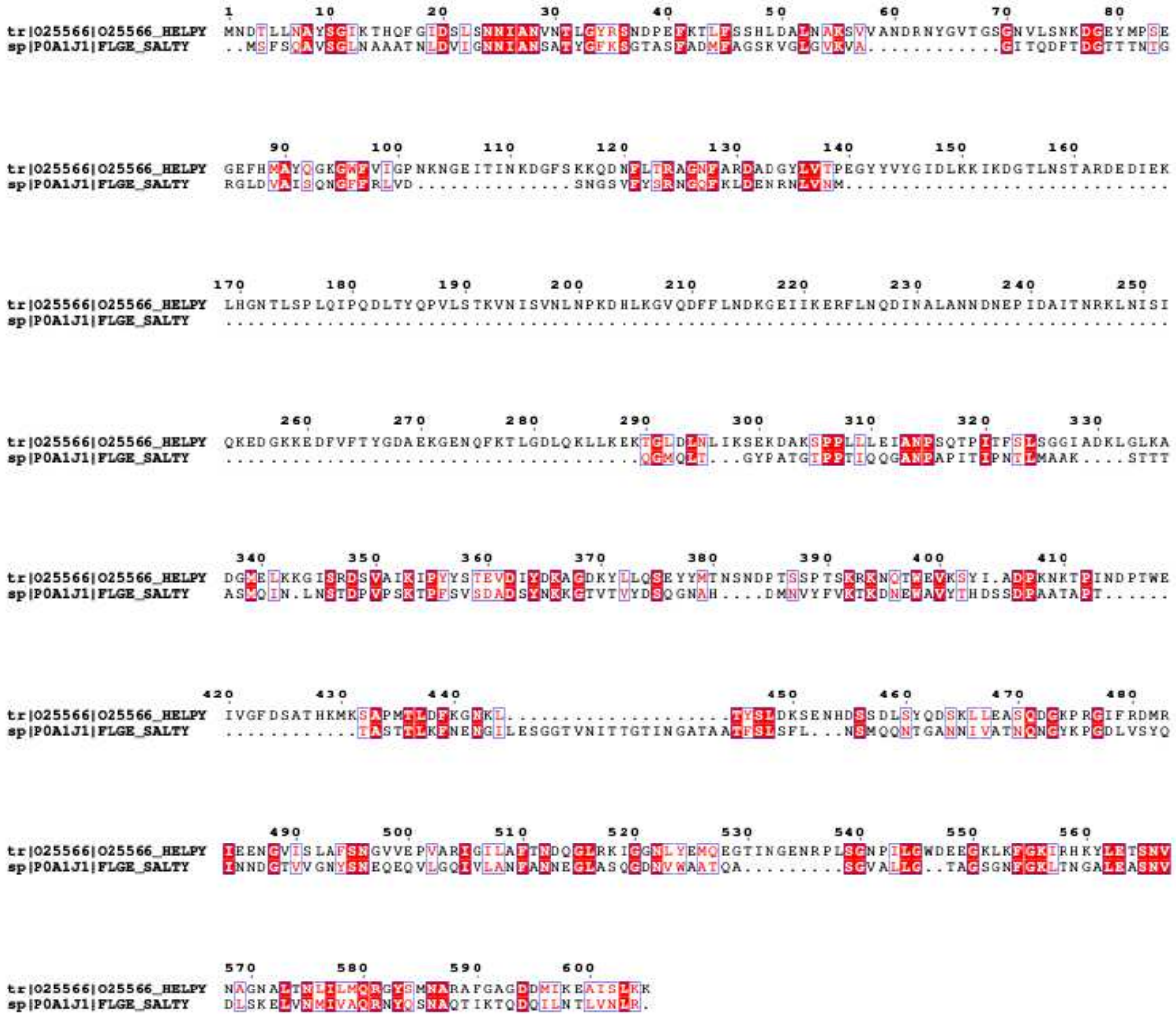

Figure 3.14: Sequence alignment between the proteins HP0908 and HP0870 respectively recognized as FlgE2 and FlgE. They share a sequence identity of about 27%. The alignment was performed using the ClustalO server (https://www.ebi.ac.uk/Tools/msa/clustalo/) and designed by Espript3.0 (http://espript.ibcp.fr/ESPript/ESPript/).

Moreover, differently from the other flagellar proteins, *hp0908* is located in a transcriptional locus adjacent to the gene *hp0907* coding for FlgD, the hook chaperon recently solved in prof. Zanotti's laboratory [9]. So, it is reasonable to hypothesize that the two proteins might interact during the polymerization process.

	Number of of amino acids	Molecular weight (kDa)	Molar Extinction coefficient ($cm^{-1}M^{-1}$
HP0908	605	68	56270
HP0908(85-570)	476	54	48820
FlgD	301	35	14440

Table 3.3: Protein information and proprieties.

Crystallization trials

The protein was cloned, expressed and purified as reported in Chapter 2. Figure 3.15 reports the profile size-exclution chromatography profile, performed on a Superdex 200 10/300 column and characterized by two peaks eluted after 11.5 ml and 13.3 ml of volume. From the calibration of the column, the first peak corresponds to the tetrameric form of the protein, whilst the second corresponds to the monomer. The SDS-PAGE profile demonstrates the protein is highly pure in all the fractions, but affected by a specific degradation.

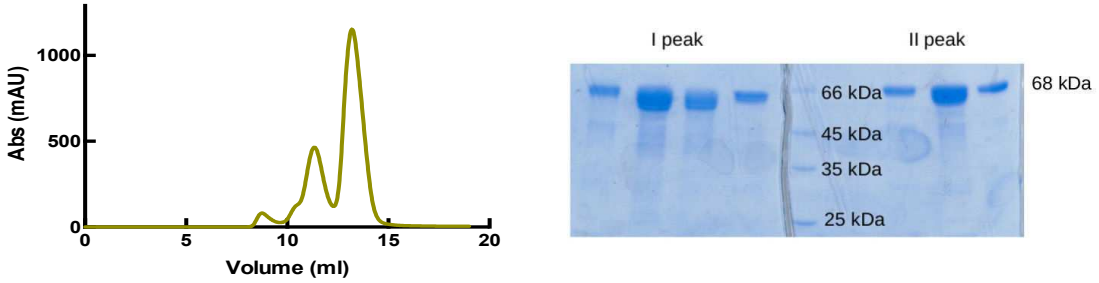

Figure 3.15: SDS-PAGE (right) of FlgE2 purified after gel-filtration; elution profile of the purified protein (left). The elution profile shows two peaks 11.5 ml corresponding to the tetrameric form of the protein and one at 13.3 ml corresponding to the monomer.

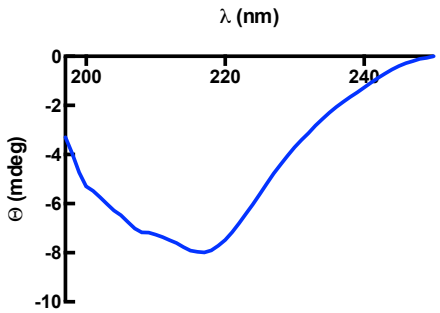

Figure 3.16: CD spectrum of the first peak fraction concentrated up to 11 mg/ml and diluted 1:20 in PBS.

All the factions under each peak were concentrated up to 9.5 mg/ml for the tetramer and 10.5 mg/ml for the monomer. In order to be sure that the protein eluted as tetramer was properly folded, a CD spectrum of the species was recorded. The spectrum reveals that β-strand components give the major contribution to the secondary structure (Figure 3.16).
Several crystallization conditions were explored but, the most successful were from PACT PREMIUM kit, giving crystals for both the species. Unfortunately, the first crystallization attempts gave rise only to small and fragile needles. To improve the quality of the crystals, the same conditions were explored again using the previous crystals as nucleation seeds. In the Figure 3.17, the crystals obtained after the seeding are shown.

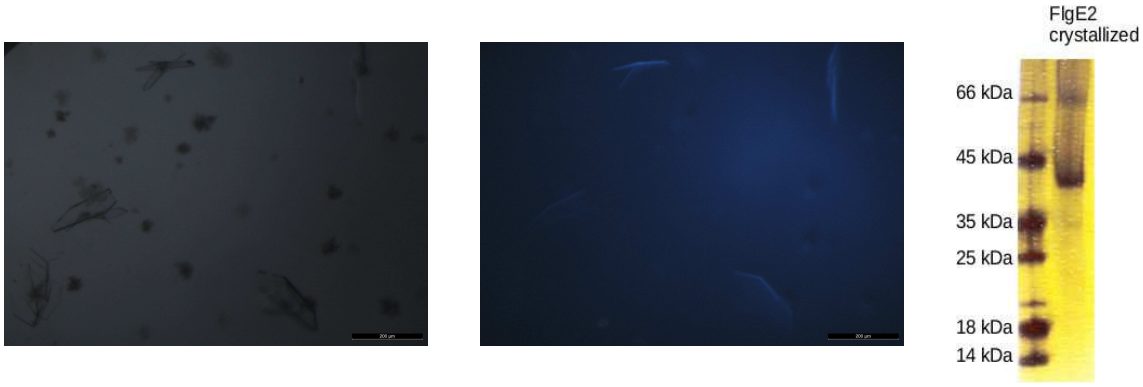

Figure 3.17: FlgE2 crystals from the condition H5 (0.2 M Sodium Nitrate, 0.1 M Bis-Tris Propane pH 8.5 and 20% w/v PEG 3350) of the PACT PREMIUM crystallization kit. In the left: stero image of the crystal recorder at whole spectrum light; in the middle: the same image recorded using the fluorescence excitation filter. On the left: Silver staining SDS-PAGE of the crystallized protein in condition H5.

$P2_12_12_1$		
a=43.50Å	b=50.19Å	c=239.62Å
α=90°	β=90°	γ=90°

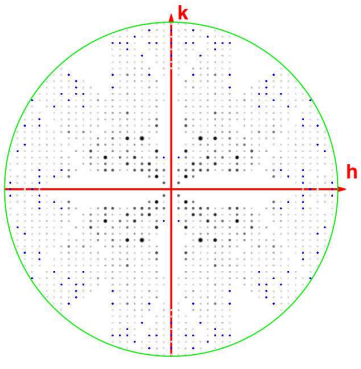

Figure 3.18: Diffraction pattern of the plane (0kl) of the protein FlgE2 reconstructed by HKLView software, included in CCP4 suite [84]. The crystals were obtained in condition H4 (0.2 M Potassium thiocyanate, 0.1 M Bis-Tris Propane pH 8.5, 20% w/v PEG 3350) with a protein concentration of 9.5 mg/ml. The space group and dimensions of the unit cell are reported in the table above the pattern.

	Overall	Inner Shell	Outer Shell
Low resolution limit (Å)	42.92	47.92	1.95
High resolution limit (Å)	1.91	8.96	1.91
R_{merge}	0.155	0.081	0.8
Number of observations	199353	2022	13050
Number of unique observation	41403	437	2743
$I/\sigma(I)$	6.4	14.4	1.7
Completeness	98.9	96.2	98.7
Multiplicity	4.8	4.3	4.8

Table 3.4: Data collection and processing statistics obtained from Aimless software [137].

All the crystals were tested by X-ray diffraction, but only the crystals from the tetrameric form were successful. Two diffraction data set were collected on the beamline ID23-2 at ESRF (Grenoble) with 1.9 Å and 2.1 Å of resolution. In this case, the cell is orthorombic and belongs to the space group $P2_12_12_1$. From the silver stainig of the SDS-PAGE of the analyzed crystal, the crystallized protein resulted to be a portion of the entire structure, with a molecular weight of 45 kDa (Figure 3.17). Later, the data were analyzed by the software Phenix [87] and the suite CCP4 [84]. Several attempts in solving the structure by molecular replacement, using the software Phaser [86] were performed, utilizing the structure of *St*FlgE (PDB ID 1WGL) as a template. Unfortunately the *Hp*FlgE2 structure has not been solved yet, but some considerations based on its preliminary structure can be done.
In order to facilitate the determination process of the structure, two parallel paths were pursued. In the first case seleniomethionine-labeled protein was produced according to the method described in the Chapter 2. The protein was expressed at 30°C after the induction overnight. The protein produced was purified as in the previous case. Figure 3.19 reports the elution profile (after gel-filtration) of the purified protein. Apparently, SeMet-protein is eluted only as a monomeric specie.

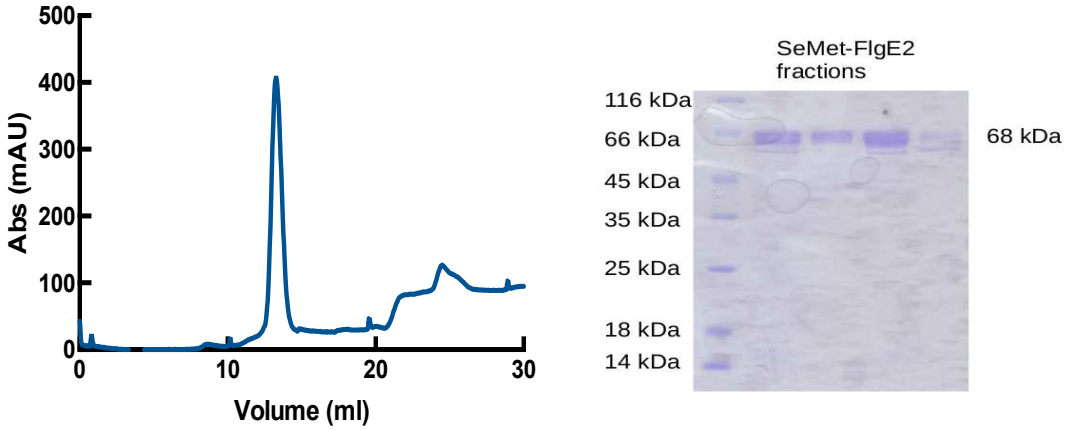

Figure 3.19: SDS-PAGE (right) of SeMet-FlgE2 purified after gel-filtration; Elution profile of the purified protein (left). The elution profile shows one peak after 13.3 ml of elution volume, corresponding to the monomer.

PACT PREMIUM kit was used as a crystallization matrix. Also in this case, the first attempt gave rise to thin and fragile crystals. The second trial was performed using crystals from the previous plates as seeds. In the last case, a set of diffraction data was collected, to a final resolution of 4 Å. Also in this case the cell is orthorombic and belongs to the space group $P2_12_12_1$. Although some attempts in solving the structure by SAD method using the pipeline SHARP/autoSHARP [85] have already been performed, the analysis of the last data is still ongoing

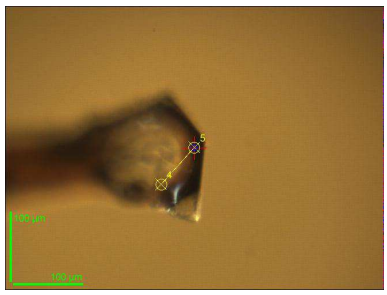

Figure 3.20: FlgE2 crystal mounted on a cryo-loop before the analysis.

$P2_12_12_1$		
a=44.87 Å	b=49.51 Å	c=244.14 Å
α=90°	β=90°	γ=90°

	Overall	Inner Shell	Outer Shell
Low resolution limit (Å)	81.36	81.36	4.47
High resolution limit (Å)	4.00	8.94	4.00
R_{merge}	0.294	0.087	1.088
Number of observations	357104	33659	106465
Number of unique observation	5073	528	1414
$I/\sigma(I)$	13.7	43.8	5.9
Completeness	100	99.8	100
Multiplicity	70.4	63.7	75.3
Anomalous completeness	100	100	100
Anomalous multiplicity	39.7	40.7	41.1
DelAnom correlation	0.682	0.858	0

Table 3.5: Data collection and processing statistics.

Another attempt to improve the quality of the crystals was to clone a mutant lacking both the N- and the C-terminus of the protein. According to the studies on *Salmonella typhimurium*, both the terminal are involved in the oligomerization process of the hook (Samatey, 2004). In analogy with the portion of the protein removed for *St*FlgE, only the domain from the amino acid 85 to 570 of *Hp*FlgE2 was used.
The same expression and purification methods of the native protein were used. Even in this case, the species purified by the gel-filtration was the monomeric form of FlgE2 (Figure 3.21).

The fractions eluted from gel-filtration were concentrated up to 40 mg/ml and two crystallization plates were prepared using PACT PREMIUM and PEG's crystallization kits. Protein crystals grew up in different conditions of PACT screen, but from the X-ray diffraction analysis none of them gave a signal at a higher resolution than 10 Å.

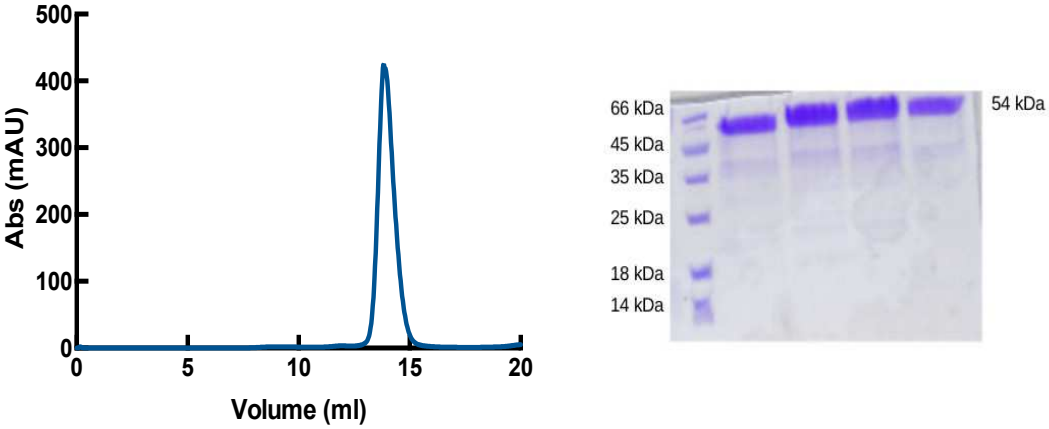

Figure 3.21: SDS-PAGE (right) of FlgE2(84-560) purified after gel-filtration; Elution profile of the purified protein (left). The elution profile shows one peak after 13.9 ml of elution volume, corresponding to the monomer.

Interaction tests

The role of the HP0908 is related to the lack of a flagellate phenotype in *H. pylori* [138], but its physiological role is still unclear. It is demonstrated nonflagellate bacteria are not able to colonize the host stomach, confirming its potential role as a pathogenic factor [5]. To functionally characterize the protein, the proof of the interaction with its cognate chaperon were sought. According to some recent studies on the flgellar growth from *Salmonella typhimurium*, the interaction between the two proteins FlgE2 and FlgD is supposed to be head-to-tail, favored by the protein unfolded state along the export route. Moreover, it seems to take place in a three-protein mechanism including the gate-protein FlhB [139]. The hypothesis is supported from the *Hp*FlgD structure too, where the C-terminal part represents a recognition motif for its partner [9].
In order to check the interaction, the two proteins were cloned separately in two different plasmids and co-expressed, as described in Chapter 2. The proteins FlgD and FlgE2 were expressed fused with a Strep-tag and a Histidine-tag, respectively. Using two different resins, the two proteins were successfully isolated from the bacterial supernatant. Firstly, the supernatant was incubated overnight with the Strep-tactin resin, in order to separate the protein FlgD from the *E. coli* proteins. Secondly, the purified elution was incubated overnight with the Ni-NTA resin, in order to distinguish the complex FlgE2-FlgD from the excess of chaperon in solution. To test unambiguously the assembly of the complex, a Western Blot was performed using both monoclonal antibodies for His-tag and Strep-tag. The last test confirmed that the interaction between the two proteins takes place during the expression process and involves both the full-length proteins and some degrades fragments. The thermodinac dissociation constant were estimated via a microscale thermophoresis analysis. The protocol used for the sample preparation is reported in Chapter 2. Merging the dissociation constant (K_D) for each curve in the plot (Figure 3.23), the

final average value of K_D was estimated to be (126.87±15.3) nM. The MST analysis was performed in collaboration with Dr. Ivana Pulić Kekez and Dr. Mario Kekez at the Dept. Biochemistry, University of Zagreb.

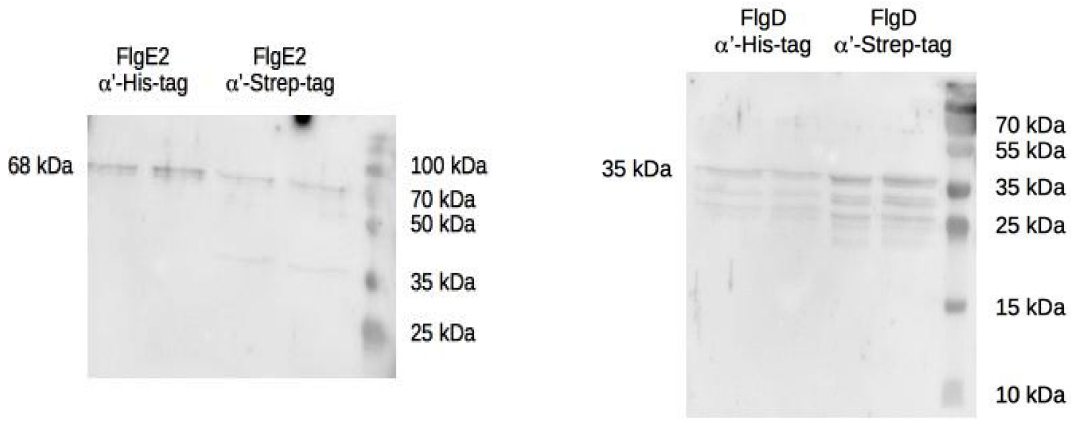

Figure 3.22: Western Blot of the co-purified protein FlgE2 (left) and FlgD (right).

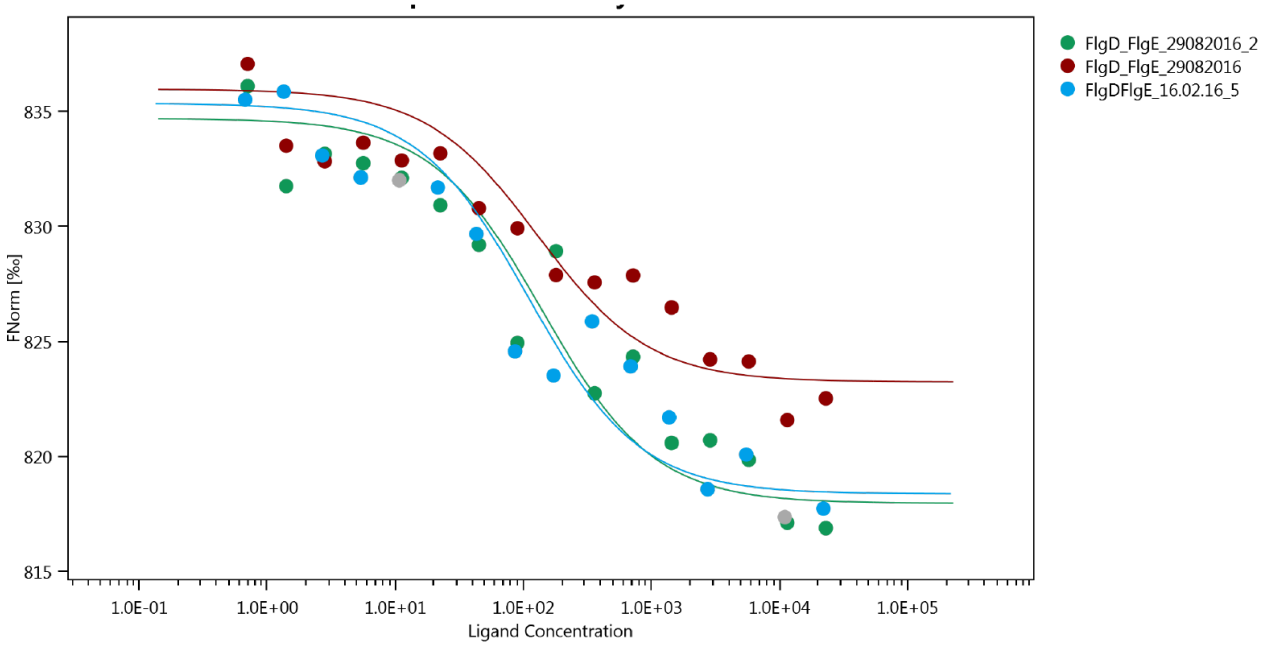

Figure 3.23: Profile of the thermophoresis analysis. The concentration of the titrant [FlgE2] is plotted vs the fluorescence intensity normalized (Fnorm).

3.3 Conclusions

3.3.1 The proteins FlgK and FlgE

The two proteins FlgK and FlgE were successfully purified and characterized. Both of them apprear to be soluble but aggregated in solution, as demonstrated by gel-filtration. From the CD spectra of the oligomeric species, it results quite evident FlgK is mainly characterized by α-helix motif, whilst FlgE is predominantly composed β-strands, in accordance with their roles in the flagellar architecture.

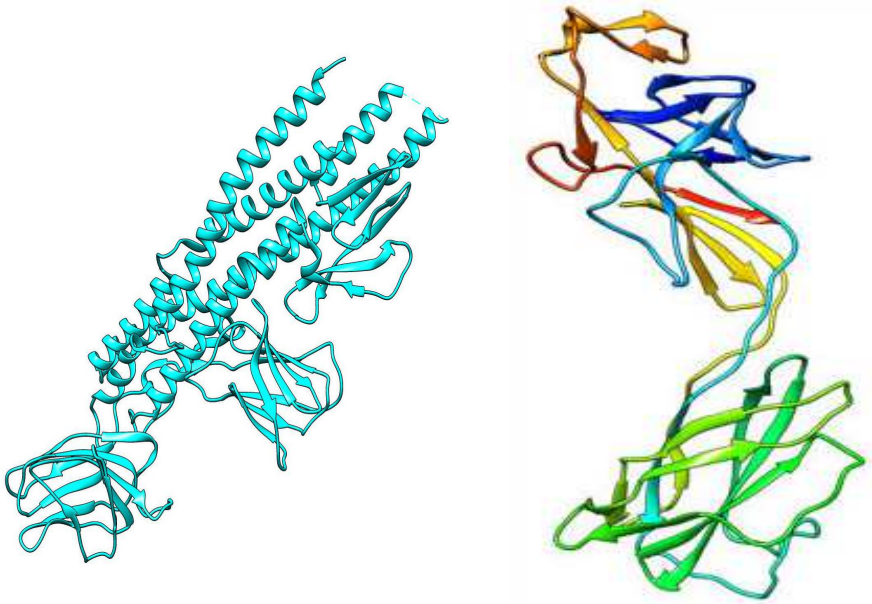

Figure 3.24: Structures of the proteins BpFlgK (PDB ID 4UT1) (left) and StFlgE (PDB ID 1WLG) (right).

In the case of *Bp*FlgK, three domains totally characterized by β-strand motives may be involved in protein-protein interactions with other FlgK subunits or HAPs, due to the high flexibility typical in the hook. On the other side, the helices may participate to the tight packing of the filament. The role of the helices was confirmed for *Hp*FlgK too. Thus, more attempts should be done its proper role in *H. pylori*.
The protein FlgE strongly confirms the presence of β-motives. In agreement with the role of the protein, the structure results highly flexible and adaptable to the conformation assumed by the hook during the motion of the flagellum.
Despite several crystallization attempts were performed on the aggregated state of the two proteins, no crystals have been obtained so far. In order to stabilize the protein in a more specific form in solution, a thermal shift assay will be performed in the near future.

3.3.2 The protein FlgE2

The structure of the protein FlgE2 from *H. pylori* has not been completely determined, yet. A preliminary model has been proposed solving the structure via molecular replacement and using as a template the structure from *St*FlgE. Although, the two structures

shares about 27% of sequence similarity, the proteins substantially diverges in several domains, probably due to the relevant contribution of the random coil motif, as highlighted by the CD spectrum.

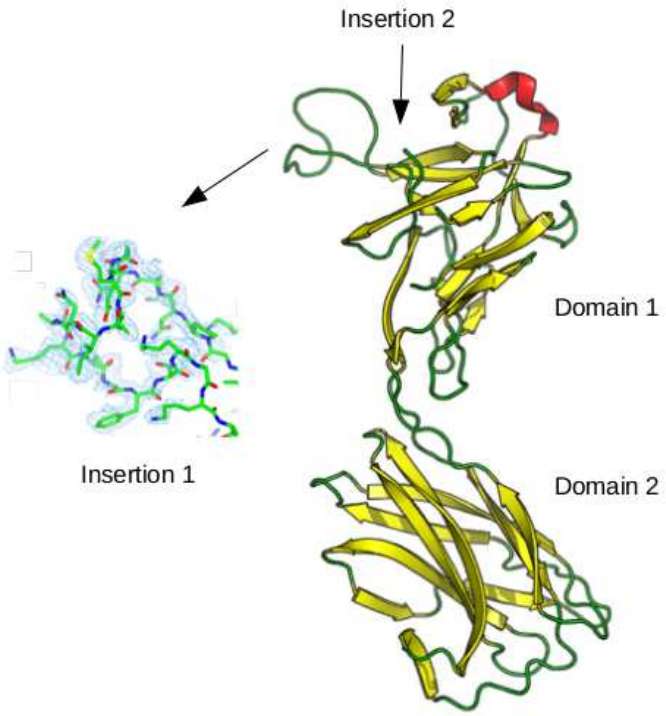

Figure 3.25: HpFlgE2 model solved by molecular replacement up to 1.9 Å of resolution and using as a template StFlgE (PDB ID 1WLG). It is remarkable to notice two main insertions were identified during the structural determination. Two main domains from StFlgE seem to be well conserved.

Figure 3.25 shows the preliminary model of *Hp*FlgE2. Two main domains from the ortholog from *Salmonella typhimurium* are conserved. According to the model proposed by Samatey *et al.*, the two domains, D1 and D2, are located in the middle part and the outer region of the hook, respectively. A third domain was identified by cryo-electron microscopy of the whole hook strcture. It has been named D0 and is supposed to provide a coiled-coil arrangement of the structure. Unfortunately, D0 is not visible in *Hp*FlgE2. D1 contributes to D1-D1 interactions in the assembly of the protofilament. The interactions are supposed to be essential for its rigidity and resistance to extension or compression. This is guaranteed by the interaction of β-structures. Moreover, the interaction between D1 and D2 allows an axial organization of the structure [121]. Two more insertions were highlighted for the *Hp*FlgE2. Nevertheless, they seem to be mainly characterized by the presence of disordered sequences.
Several refinement cycles were performed on the partial structure here proposed, but the final statistics proves the model is partially incompleted (R_{work}=0.43/R_{free}=0.45). The model used as a template is likely a poor representative of the orthologous protein from *H. pylori*. This idea is strongly supported by the large difference in the length of the sequence; indeed, *St*FlgE2 is composed of only 299 amino acids, whereas *Hp*FlgE2 counts

605 residues in the entire sequence. For this reason, the implementation of the prevoius structural information is ongoing, trying to get the location of the seleniomethionines in the density map, through SAD analysis.
Finally, recently the structure of a fragment of 79 kDa the orthologous protein FlgE2 from *Campilobacter jejuni* has been solved (PDB ID 5AZ4). The two protein are supposed to share about 22% of sequence identity and their length is comparable. So, one more attempt in solving the structure via molecular replacement can be simultaneously performed using the new structure from *C. jejuni* as a template.
Finally, the interaction assay confirms the idea that the two protein FlgD and FlgE2 are released together during the assembly of the hook. Evans *et al.* propose an interactive mechanism where the contact takes place in a head-to-tail fashion. According to their model, FlgD is docked at the gate of the export channel, bound to the protein FlhB. The protein FlgE (in the case of *Salmonella typhimurium*) captures its chaperon and is dragged outside the channel in an unfolded fashion, since the diameter of the exportary machinery can fit the dimension of the secondary structure only [139]. The experiment performed on the orthologous protein from *H. pylori* confirms that the interaction takes place in the second system too, confirming the hypothesis related to the binding motif located at the C-terminus of the FlgD protein [9]. But some more pieces of information concerning the position and the role of FlgE2 in the building of flagellum are still lacking. Indeed, from the solved structures of *Hp*FlgD and FlgE2 no details can be obtained, since both are lacking the N- and the C-terminus. Moreover, both show a strong degradation at the N-teminus. In order to clarify the type of interaction and further hypothesize a mechanism, the crystallization trials of the complex are ongoing.

4 | *Helicobacter pylori* adhesine A

4.1 Introduction

One of the first steps in bacterial infection is the attachment to the host cell surface. It is known that the bacteria and their toxins interact specifically with some region of the plasma membrane enriched in cholesterol, sphingomyelin, and glycosphingolipids [140]. These microdomains are referred to as lipid rafts and are attractive to a wide range of pathogens [141]. However, the molecular mechanism involved in bacteria-lipid interaction is still unclear. Some considerations can be proposed taking together some common structural features from the HIV-1 surface envelope glycoprotein gp120, Alzheimer's amyloid-β peptide and the cellular isoform of the prion protein (PrP^C). The structure consists in a hairpin motif containing a water-exposed aromatic residue and able to interact with the lipidic environment [142].
H. pylori adhesin A is a surface-located [143][144] lipoprotein [10] that was initially described as a sialic acid binding adhesin, expressed on the surface of gastrointestinal cells [145]. Experimentally, *H. pylori* interacts with selected glycolipids such as lactosylceramide (Galβ1-4Glcβ1-Cer; LacCer) [146] and its adherence can be inhibited by 3'-sialyllactose (NeuAcα2-3Galβ1-4Glc), a natural sialylated oligosaccharide, as well as by synthetic multivalent sialylated compounds [145]. HpaA is an outer membrane protein, totally exposed to the lumen and highly conserved among several strains from *H. pylori*. For all these reasons, it can be considered a potential vaccine antigen [147].
The three-dimensional structure of a typical glycolipid-binding protein has a linear motif which consists of a hairpin structure containing:

1. a solvent-exposed aromatic residue (Phe, Tyr, or Trp);

2. several charged residues (Asp, Glu, Arg, or Lys) with some of them oriented toward the solvent;

3. a Gly and/or a Pro residue inducing the turn in the backbone Cα-chain.

In the case of HpaA, Fantini *et al.* have identified this combination in the sequence (139)KKSEPGLLFSTGLDK(153), where F147 corresponds to the aromatic residue, whilst a turn in the chain is induced by P143 and G144 and K139, K140, E142, D152, K153 correspond to the charged amino acids in the vicinity of F147 [12].
In order to check the potential glycolipid-binding domain, the crystal structure of the protein was investigated. The native form of HpaA was previously expressed in prof. Zanotti's laboratory, but it seemed to easily aggregate in solution giving back different olimeric species (data not shown). In order to stabilized a single oligomeric form, a new construct of the HpaA fused together with a sfGFP1-10 was designed and new crystallization attempts were led on the fused protein.

4.1.1 Crystallization purpose

The intrinsic stability of GFP, together with its utility as a fluorescent reporter when correctly folded, has made it a potential target for reconstitution of separately expressed protein fragments [148]. Following this idea and fusing a terminal segment of GFP to a crystallization target, the resulting construct might recombine with the complementary fragment of GFP to create a new complex useful for crystallization [13]. In our case, a superfolder GFP [GFP(1-10)+GFP(11)] was used for the purpose. The two fragments were fused at the N- and C-terminus, keeping tight together the flexible portions of the protein, in order to improve its folding, increase its solubility and mediate the formation of a crystal lattice. Following this idea, the fused protein was designed as described in the Figure 4.1: HpaA sequence is missing the first twenty-one amino acids, identified by SignalP (http://www.cbs.dtu.dk/services/SignalP/) as a signal peptide; two glycines and serines sequences were designed as flexible linkers between the two proteins, to allow the target to easily rearrange in the proper conformation.

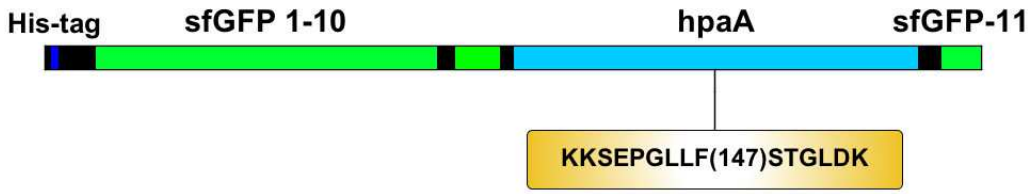

Figure 4.1: Scheme of the fused proteins. The target protein is inserted in between the 10th and the 11th β-stands from sfGFP.

HpaA-sfGFP	
Number of residues	510
Molecular weight	56 kDa
Calculate molar extinction coefficient	29465 $cm^{-1}M^{-1}$

Table 4.1: Protein information and proprieties.

4.2 Results

4.2.1 Protein Characterization

The HP0797 was successfully expressed and purified by Ni-NTA chromatography and size-exclusion chromatography (on Superdex 200 10/300), as previously described in Chapter 2. The fused protein counts 510 amino acids with a molecular weight of 56 kDa.

```
        10          20          30          40          50          60
MRGSHHHHHH  GMASMRKGEE  LFTGVVPILI  EKDGDVNGHK  FFVRGEGEGD  ATIGKLSLKF

        70          80          90         100         110         120
ICTTGKLPVP  WPTLVTTLTY  GVQCFSRYPD  HMKRHDFFKS  AMPEGYVQER  TISFKDDGTY

       130         140         150         160         170         180
KTRAEVKFEG  DTLVNRIELK  GIDFKEDGNI  LGHKLEYNFN  SHKVYITADK  QTNGIKANFT

       190         200         210         220         230         240
IRHNVEDGSV  QLADHYQHNT  PIGDGPVLLP  GGSLEGGGYT  MDLPDDHYLS  TQTILSKDLN

       250         260         270         280         290         300
GTDVGSGGGS  VDETNEVALK  LNYHPASEKV  QALDEKILLL  RPAFQYSDNI  AKEYENKFKN

       310         320         330         340         350         360
QTALKVEQIL  QNQGYKVISV  DSSDKDDLSF  SQKKEGYLAV  AMNGEIVLRP  DPKRTIQKKS

       370         380         390         400         410         420
EPGLLFSTGL  DKMEGVLIPA  GFVKVTILEP  MSGESLDSFT  MDLSELDIQE  KFLKTTHSSH

       430         440         450         460         470         480
SGGLVSTMVK  GTDNSNDAIK  SALNKIFANI  MQEIDKKLTQ  KNLESYQKDA  KELKGKRNRR

       490         500         510
SGSGGGSGGG  STSEKRDHMV  LLEYVTAAGI
```

Figure 4.2: Sequence of the fused protein

As shown in Figure 4.3, the protein is eluted as a trimer and concentrated by ultrafiltration up to 30 mg/ml. The purified protein was finally used for crystallization purposes.

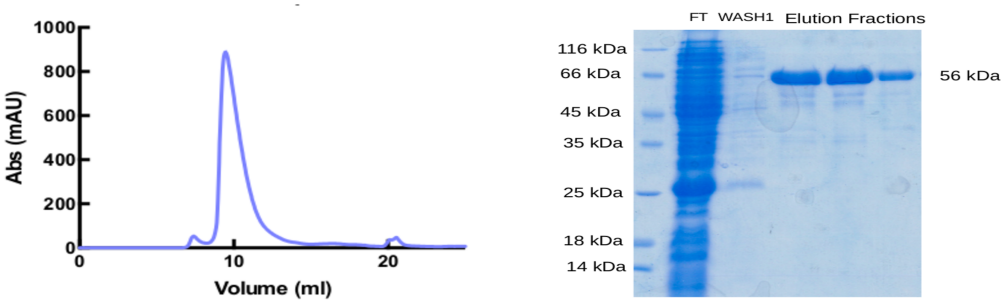

Figure 4.3: On the left: SDS-PAGE 12% acrylammide of the purified HpaA-sfGFP. On the right: gel-filtration profile of the protein HpaA-sfGFP. The peak is eluted after 10.6 ml of elution volume, corresponding to the trimeric form of the protein.

The homogeneity of the protein solution was confirmed by DLS analysis, demonstrating a single monodisperse species in solution (Figure 4.4).

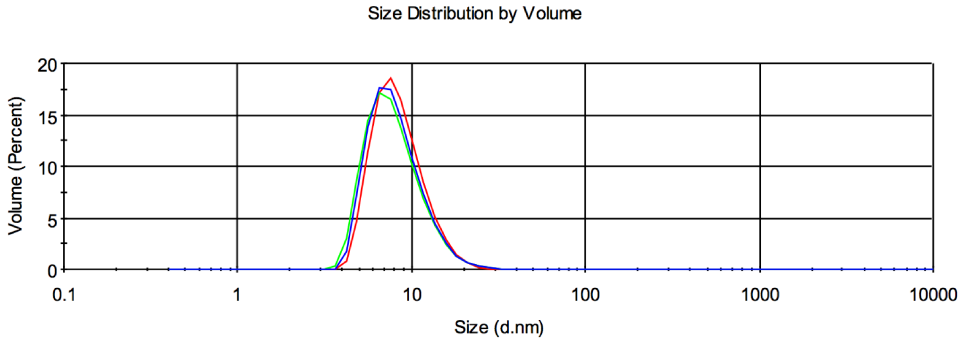

Figure 4.4: DLS spectrum of the purified fused protein. The three curves correspond to three repeated measures of the size distribution of the single particles versus the volume distribution, i.e. the distribution describing the relative proportion of multiple components in the sample based on their mass or volume rather than based on their scattering intensity.

Although we can assume the fused protein is properly folded thanks to its greenish appearance, a CD analysis was performed on the fused protein and compared with both the sfGFP and the wtHpaA (gently produced by Prof. Alessandro Negro, Dep. Biomedical Sciences, Univ. of Padua).

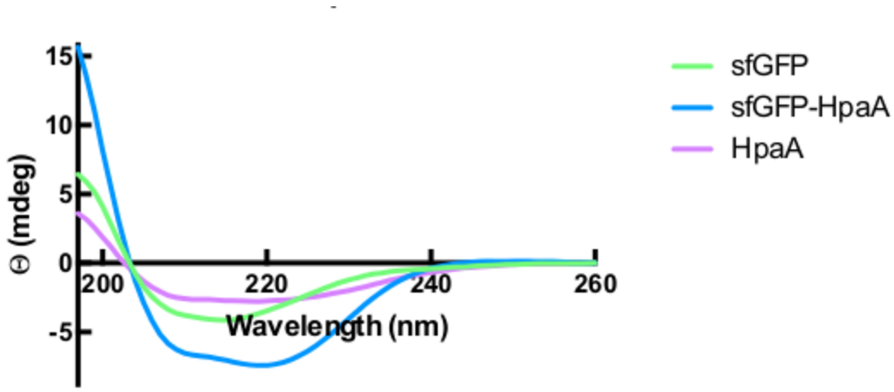

Figure 4.5: CD spectra of HpaA-sfGFP compared with sfGFP and HpaA alone. The curve related to the sfGFP (green line) shows an intense signal at 216 nm, representative of major composition in β-strands; on the contrary, the curve recorded for the protein HpaA (malve line) shows two strong signals at about 208 nm and 222 nnm. The last signals are recorded for the fused protein HpaA-sfGFP (blue line), too.

Comparing HpaA curve and HpaA-sfGFP is evident how a set of α-helices belonging to the pathogenic protein is present, with a more pronounced peak of absorption around 222 nm and 208 nm. The result confirms the adhesine fused together with the sfGFP assumes a correct fold in solution.

4.2.2 HpaA interaction with AGS cells

To confirm Fantini's hypothesis, the functional role of HpaA was checked in the fused form with sfGFP. AGS cells were treated as described in Chapter 2.2.12 and visualized by confocal microscope.

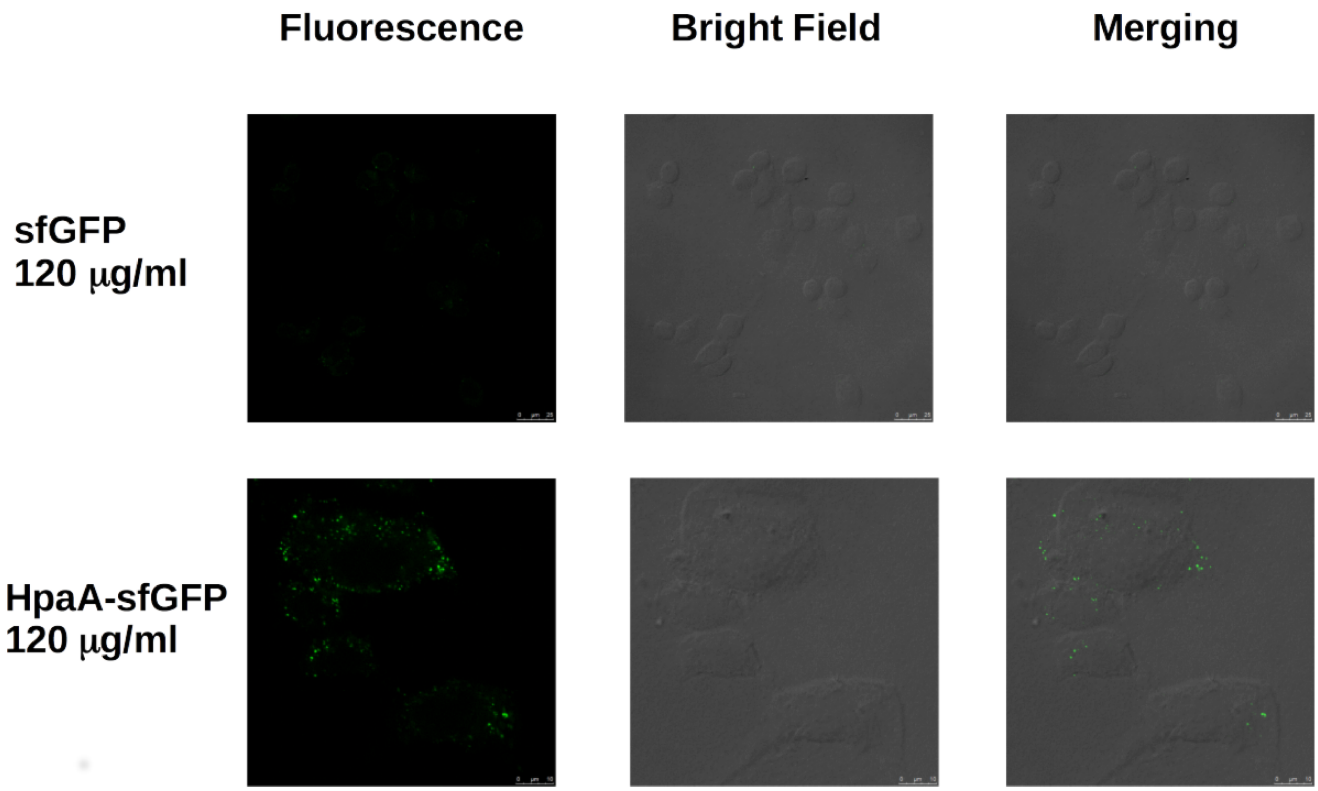

Figure 4.6: Comparison between AGS cells treated with HpaA-sfGFP and sfGFP alone.

From comparison between the cells treated both with sfGFP and HpaA-sfGFP at a concentration of 120 μg/ml, the physiological role of *H. pylori* protein is still maintained, confirming the protein is both properly folded and physiologically active. The last *in vitro* experiment gives support to the crystallization strategy of the fused protein, confirming the crystallized form will correspond with the active protein.

4.2.3 Crystallization and Data Collection

Purified samples of HpaA-sfGFP were concentrated up to 30 mg/ml and several crystallization kits were test in order to obtain good quality crystals. The first crystals were observed four month later the plate preparation. In Table 4.2 the positive conditions for the crystal growth are reported.

PACT Premium			
A5	0.1M SPG buffer, pH 8.0	/	25% (w/v) PEG 1500
A6	0.1M SPG buffer, pH 9.0	/	25% (w/v) PEG 1500
B7	0.1M MES, pH 6.0	0.1 M NaCl	20% (w/v) PEG 6000
C8	0.1 HEPES, pH 7.0	0.2 M NH_4Cl	20% (w/v) PEG 6000
C9	0.1M HEPES, pH 7.0	0.2M LiCl	20% (w/v) PEG 6000
D2	0.1M MMT buffer, pH 4.0	/	20% (w/v) PEG 1500
D8	0.1M Tris, pH 8.0	0.2 M NH_4Cl	20% (w/v) PEG 6000
G4	0.1M Bis Tris Propane, pH 7.5	0.2M KSCN	20% (w/v) PEG 3350
PEG's			
F2	0.1M HEPES, pH 7.5	/	25% (w/v) PEG 4000
G1	0.1M Tris, pH 8.5	0.2M Li_2SO_4	25% (w/v) PEG 5000 MME

Table 4.2: Successful crystallization conditions for the protein HpaA-sfGFP, obtained from two different crystallization kits after five months of incubation in the crystallization matrix.

The crystals were check for the GFP fluorescence, in order to be sure the entire complex was composing the crystal. The Figure 4.7 reports the results at the absorption wavelength of GFP.

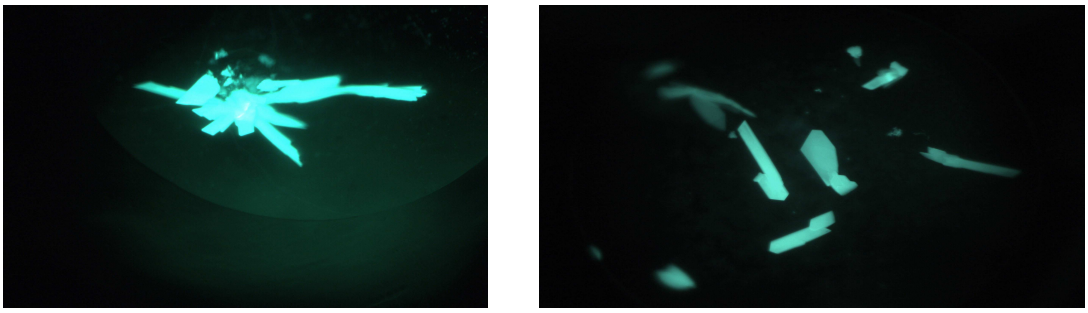

Figure 4.7: Crystals from HpaA-sfGFP precipitation, obtained in different conditions: A5 from PACT PREMIUM kit (left); F2 from PEG's kit (right). Fluorescence was recorded for the GFP emission peak at 505 nm.

Preliminary Statics	
Space Group	$P2_12_12$
Resolution	1.4 A
R_{work}	0.30

Table 4.3: Data collection and processing statics obtained by Phenix [87].

Later, crystals were analyzed by diffraction experiment at the Swiss Light Source synchrotron (Paul Scherrer Institute, Villigen, CH), on PXIII beamline. The structure was solved by molecular replacement, using the structure of the sfGFP as template (PDB ID 4JFG). The preliminary results, obtained by the software Phaser [86], are reported in Figure 4.8. Unfortunately, the crystals resulted composed only by the sfGFP.

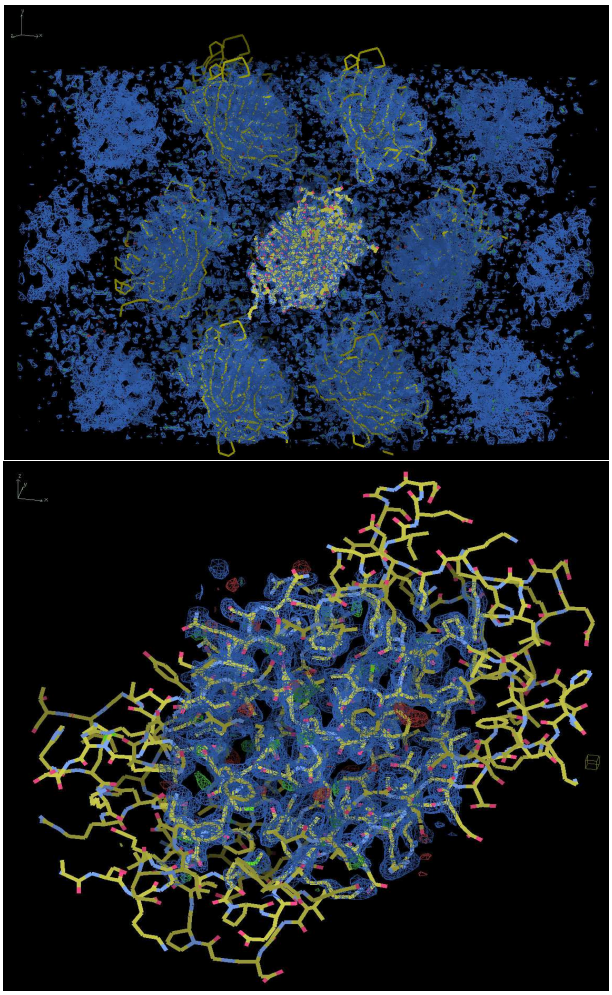

Figure 4.8: On the top: crystal packing of sfGFP; on the bottom: detailed view of electron density on a single sfGFP molecule (pictures prepared by Coot [88]).

A silver staining of the SDS-PAGE from the crystallizantion drop was performed to check whereas HpaA was still present in the crystallization drop, demonstrating that the hydrolysis took place during the crystallization process.

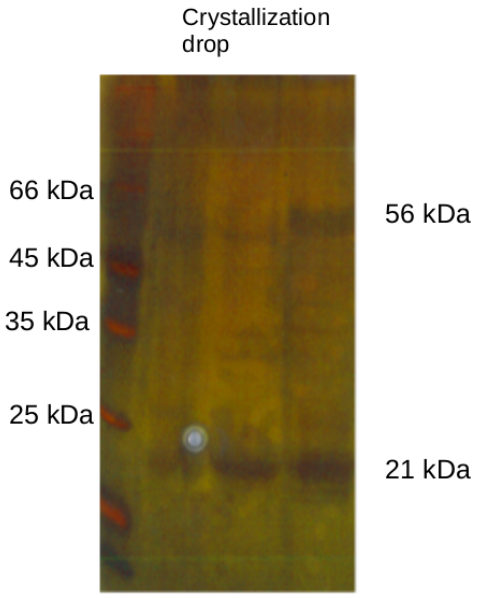

Figure 4.9: Silver staining of HpaA-sfGFP crystallization drop.

More crystallization attempts were performed in different conditions to prevent the enzymatic cut between HpaA and sfGFP. As a first trial, a Protease Inhibitor Cocktail (solution prepared according to the Product Datasheet) was added to a volume of 350 μl at a final concentration of protein of 20 mg/ml, diluted in a ratio 1:1000. Several crystallization conditions were later explored, but none of them was successful. As a second trial, a cross-linking reaction was proposed to stabilize the fused protein. The protein buffer was exchanged from 30 mM Tris pH 8.0 to 30 mM HEPES pH 8.0 and the solution was concentrated up to 30 mg/ml of protein in a final volume of 100 μl. The final solution was treated with 5 μl of 2.3% freshly prepared solution of glutaraldehyde for 15 minutes at 37°C. The reaction was terminated by addition of 10 μl of 1 M Tris-HCl, pH 8.0. Crystallization trials were performed with the fresh cross-linked solution, but even in this case, no crystals were grown up.

4.3 Conclutions

The recombinant protein was expressed in *E. coli* cells in a soluble form and fused together with a superfolder GFP. After gel-filtration purification, the protein results to be present in solution as a trimer. The result was confirmed from the DLS spectrum, indicating a monodisperse composition of the sample. The fused protein has its solubility increased compared to the HpaA protein alone (Sisinni, PhD thesis 2010), confirming the hypothesis proposed by Waldo *et al.* [149]. Moreover, the function of the protein is conserved as demonstrated from the AGS cells treatment.

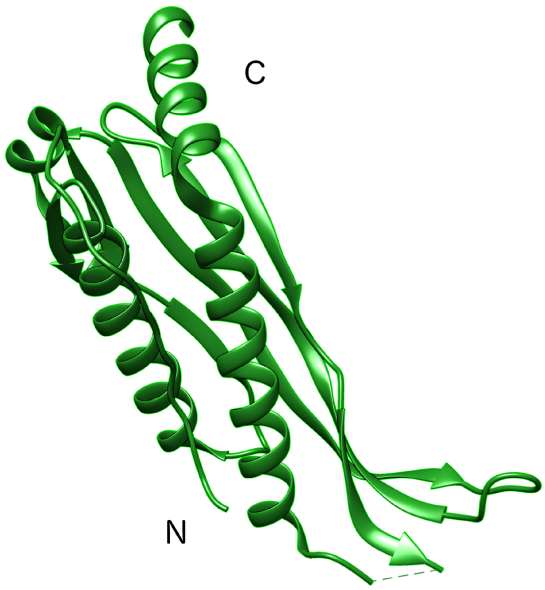

Figure 4.10: Crystal structure of HP0410 (PDB 3BGH).

Since the fused protein was produced at high concentration, more than 1500 crystallization conditions were tested, but none of them was informative. Indeed, the crystals obtained and analyzed in many measurement sessions at ESRF and SLS, provided information about the crystallographic form of sfGFP alone. As a matter of fact, the crystal structure was missing the protein sequence belonging to HpaA. From the silver staining of the SDS-PAGE of the crystal drop solution, this is likely due to an enzymatic cut at the level of the flexible linkers, hypothetically induced by a conformational stress. Indeed, from the model built using the crystal structure of the homologous protein HP0410 (PDB 3BGH), the N- and the C-termini are located at opposite positions (Figure 4.10), so the interaction between the two GFP fragments is likely to induce a tension at the level of the linkers. In the crystallization drop it is reasonable to hypothesize that the tension could be enhanced during to the crystallization process. Besides, considering the linkers as the cleavage sites and inducing a crystalline state on the protein, the linkers will result to be more exposed and prone for the enzymatic cut. To conclude, we can suppose that the system proposed by Waldo is susceptible of enzymatic cleavage in a constrained environment. In order to obtain a successful result, the location of the N- and C-terminus should be checked, in advance, *via* homology modeling.

5 | The putative lipoprotein HP1457

5.1 Introduction

Despite the whole genomes of the *H. pylori* 26695 and J99 were completely sequenced along the past years, several genetic loci have been identified, but not related with a specific function, so their role is still hypothetical. As a result of this lack, several unknown proteins have been explored from a structural point of view in order to identify a hypothetical role in the physiological activity of the bacterium.
Protein HP1457 belongs to this class of proteins and only recently has been recognized to act as a outer-membrane protein that contributes to the stimulation of the peptidoglycan synthase PBP1B [14]. It is characterized by the presence of a signal export peptide in between T21 and D22, recognized by LipoP server ($http://www.cbs.dtu.dk/services/LipoP/$) to belong to lipoprotein export peptide class. HP1457 is well conserved among the other *Helicobacter* species, sharing between 98-45% of sequence similarity, with a coverage of 99%. Moreover it shares a lower similarity (about 22%) with other Gram-negative bacteria, for instance *Escherichia coli*.

	Number of of amino acids	Molecular weight (kDa)	Molar Extinction coefficient ($cm^{-1}M^{-1}$
HP1457$\Delta(1-21)$	190	21	9970
HP1457$\Delta(1-49)$	160	18	8480

Table 5.1: Protein information and proprieties.

Figure 5.1: Location of the gene hp1457 in the bacterial operon. The protein is preceded by hp1454, a gene coding for a pathogenic protein with unkown activity, hp1455, another gene coding for a secreted protein with unknown function, and lpp20, a gene coding for a lipoprotein located outside the cell membrane.

Looking at the chromosomal locus corresponding to the *hp1457*, it is evident that it belongs to a transcription region poorly characterized. This area corresponds to an operon which includes the genes *hp1454*, *hp1455*, *1456* or *lpp20*, too.

5.2 Bioinformatic analysis

The protein sequence was aligned against the entire Uniprot database (http://www.uniprot.org/) in order to find some specific conserved portions in different bacteria. From the alignment it is evident that the protein is mainly conserved in several *Helicobacter* species, especially the C-terminal domain (Figure 5.2), here the proteins present about 95% of sequence similarity, with an average value of coverage of about 98%. In particular, the protein shares

about 22% sequence similarity with the already solved structure of LpoB (penicillin-binding protein activator B, PDB-ID: 2MII) from *E. coli*, suggesting the protein could play a similar role in *H. pylori* physiology (Figure 5.3). Indeed, the LpoB protein is known to be involved in the peptidoglycan regulation and synthesis, interacting with its cognate pennicellin-binding protein (PBP-1B) [14].

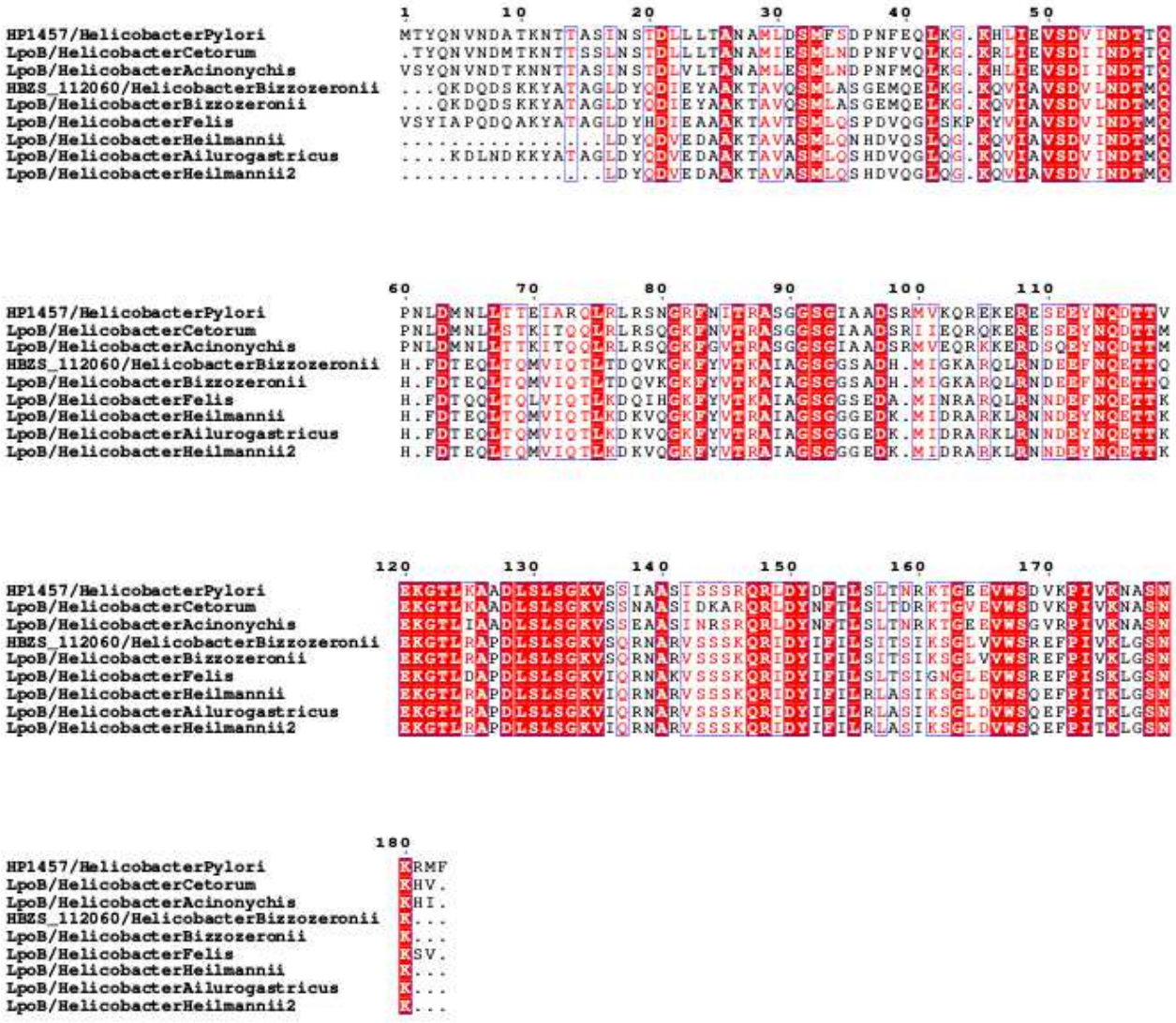

Figure 5.2: Multiple sequence alignment of the same protein from different species of Helicobacter. The proteins present about 95% of sequence identity. The alignment was performed using the ClustalO server (https://www.ebi.ac.uk/Tools/msa/clustalo/) and designed by Espript3.0 (http://espript.ibcp.fr/ESPript/ESPript/).

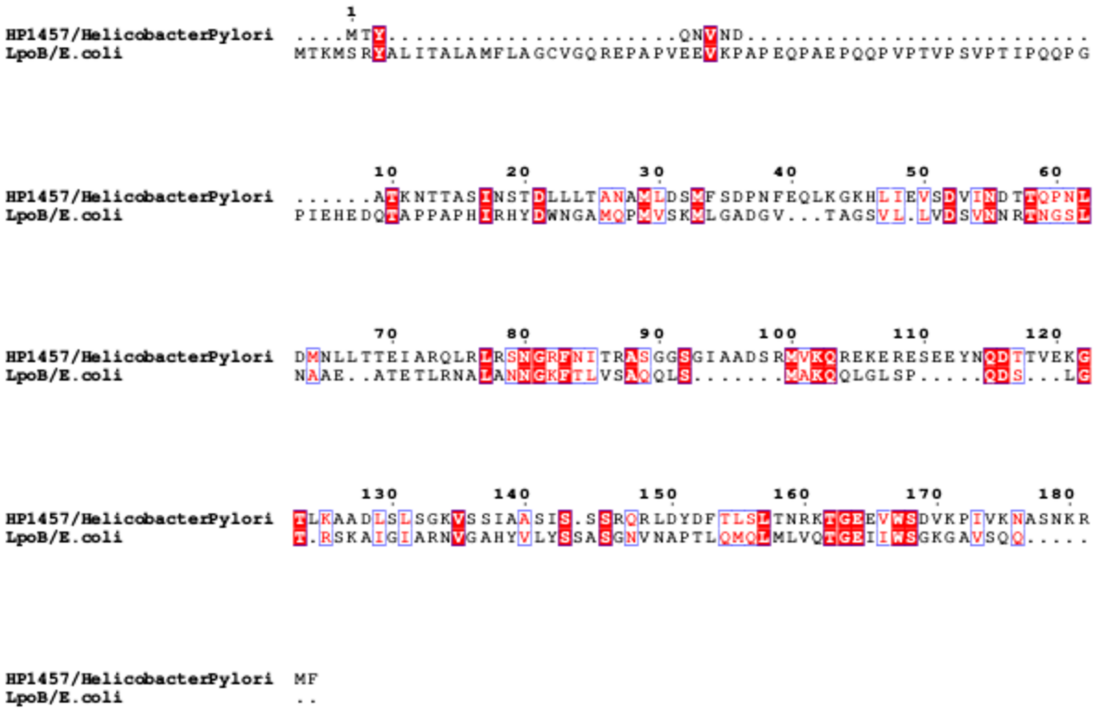

Figure 5.3: Sequence alignment between HP1457 protein from H. pylori and LpoB protein from E. coli. The alignment was performed using the ClustalO server (https://www.ebi.ac.uk/Tools/msa/clustalo/) and designed by Espript3.0 (http://espript.ibcp.fr/ESPript/ESPript/).

The hypothetical role of the protein in *H. pylori* is strongly supported by STRING server analysis ($http://string-db.org$), where a relationship between PBP1-B (pennicellin-binding protein 1B) from the bacterium with HP1457 was highlighted (Figure 5.4).

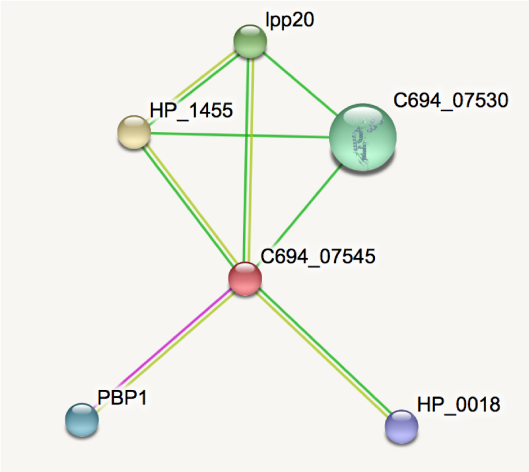

Figure 5.4: Map of the hypothetical interactions between the protein HP1457 (red circle in the middle) and known and hypothetical protein from H. pylori genome, obtained by STRING server with a high confidence high score of 0.7. The green lines indicates an interaction for proximity in chromosomal loci, the red lines indicate experimental evidences of interaction, whilst the yellow lines indicate the co-mention in PubMed abstracts.

Moreover, the protein is proposed to interact with the hypothetical protein HP1455, of unknown structure and function, and the protein HP1456, recognized to be a Lpp20 lipoprotein and whose structure has been recently solved in Prof. Zanotti's group (unpublished data).
Finally, Pfam has identified the amino acids going from 4 to 203 to belong the family of TolB proteins, a periplasmic protein involved in a translocation system in Gram-negative bacteria.

5.2.1 The role of LpoB-PBP1B interaction in peptidoglycan synthesis regulation

The synthesis of peptidoglycan takes place in three stages: 1) the nucleotides precursors (UDP-N-acetylglucosamine and UDP-N-acetylmuramylpentapeptide) are activated in the cytoplasm 2) and then assembled with undecaprenyl phosphate to form the lipid II and are flipped inside the membrane; 3) the lipid II polymerases, resulting in glycan chains and building up the peptidoglycan (PG) sacculus, essential for the maintainance of the shape and osmotic stability of the cell undergoing stress conditions [150].
In Gram-negative bacteria the sacculus layer is connected to the outer membrane through both covalent and noncovalent interactions by outer-membrane proteins. The PG role is essential for the structural integrity of the cell envelope, going through the insertion of new glycan strands and the release of old material from the cytosolic environment. Moreover its synthesis and turnover must be spatially controlled, in order to maintain the cell shape, and temporally, in order to be coordinated with the cell life cycle [151].
To generate and maintain a proper morphology, the cell undergoes to an elongation process that increases the lateral length of the cell, followed by a high concentration of PG synthesis at midcell, and allowing cell division [151]. Bacterial cytoskeletal proteins guide these processes. For instance, the bacterial actin homolog MreB is essential for elongation [152], whilst FtsZ (the bacterial structural homolog of tubulin) is required for PG synthesis at the septum [153]. So, MreB and FtsZ nucleate the assembly of inner-membrane enzymes that make the PG building block and controls PG synthesis. Peptidoglycan synthesis requires glycosyltransferases (GTases) to polymerize the glycan chains and transpeptidases (TPases) to crosslink the peptides (Typas, 2012). The pennicellin-binding protein (PBP) 1B is a bifunctional glycosil-transferase (GTase)-transpeptidase (TPase). It is recruited at the divisome[154], whilst its homologous PBP1A is located on the lateral wall for cell elongation. Both PBP1A and PBP1B are partially redundant, so the cell requires one of them for viability [155]. Both of them are associated with LpoA and LpoB, two lipoproteins anchored to the outer membrane. They do not share any similarity in sequence but they have the same role in docking the PBP1A and PBP1B respectively. The LpoB structure from *E. coli* has been recently reported from Egan *et al.*. It is composed by a flexible N-terminal region, which spans about 150 A of the periplasm, followed by 30-Å-long globular domain [14]. The Lpo are both involved in the TPase domain stimulation. Indeed, Typas *et al.* conducted some *in vitro* experiments, demonstrating the LpoA/B proteins stimulate the PBP1A/B to produce dimeric, trimeric and tetrameric structures of side peptides. The same authors noticed that LpoB can be chemically cross-linked to PBP1B *in vivo*, but not to a mutant that lacks PBP1B-UB2H domain, a noncatalytic region required for the proper system function (Typas, 2010). Moreover, Egan *et al.* grasped that

after the complex formation, a truncation of the N-terminal domain in the LpoB protein hinders the activation of the catalytic mechanism [14].
According to their theory, the PG synthesis takes place thanks the to outer-membrane LpoA/B proteins interacting with the TPase domain of each PBP1A/B. In particular, the LpoB-PBP1B complex localizes itself on the edge of the growing septum and, when the LpoB gets in contact with the UB2H domain from PBP1B, likely a conformational change induces a conformational change in PBP1B that impacts on the GTase domain. Secondly, the TPase activity takes place subsequently, so that the interaction could induce a conformational switch that repositions the domain and affects acceptor peptide binding, attachment to the PG, or the TPase activity itself [14].
Untill few years ago, the Tol-Pal complex was assumed to be the only system involved in the regulation of the outer membrane invagination, but it does not seem to be the only regulatory system [156]. As a member of the Tol-Pal complex, TolB participates in outer membrane construction during the cell division and its complex, together with LpoB-PBP1B, was colocalized at the constriction sites, underlining an overlapping or a link during the cellular process [157].This is supported by the similarity between the LpoB structure and the N-terminus from TolB protein, suggesting the two proteins may work concertly in the process of cell division. Moreover the PG remodeling appears to be part of a cellular strategy to adapt to the environmental changes. For instance, the regulation of the PG hydrolysis is proposed to twist and facilitate the *H. pylori* membrane curvature [158].

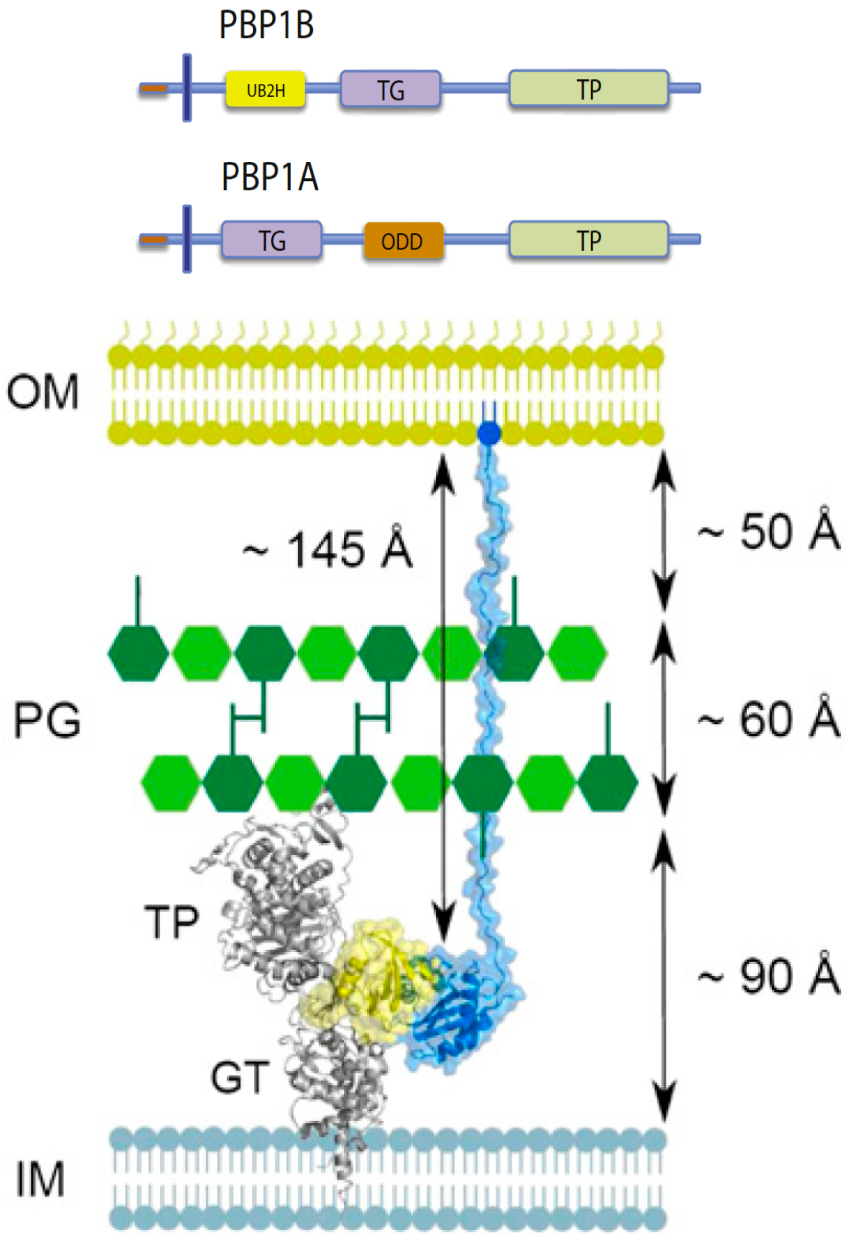

Figure 5.5: 1) On the top: structural domains in the two pennicellin-binding domains involved in the PG growth modulation; 2) On the bottom: model of the stimulation of PBP1B by LpoB in the bacterial envelope. LpoB (blue) uses its flexible N-terminal region to span the periplasm and places its globular domain in position to interact with the UB2H domain of PBP1B (yellow). Pictures adapted form [14].

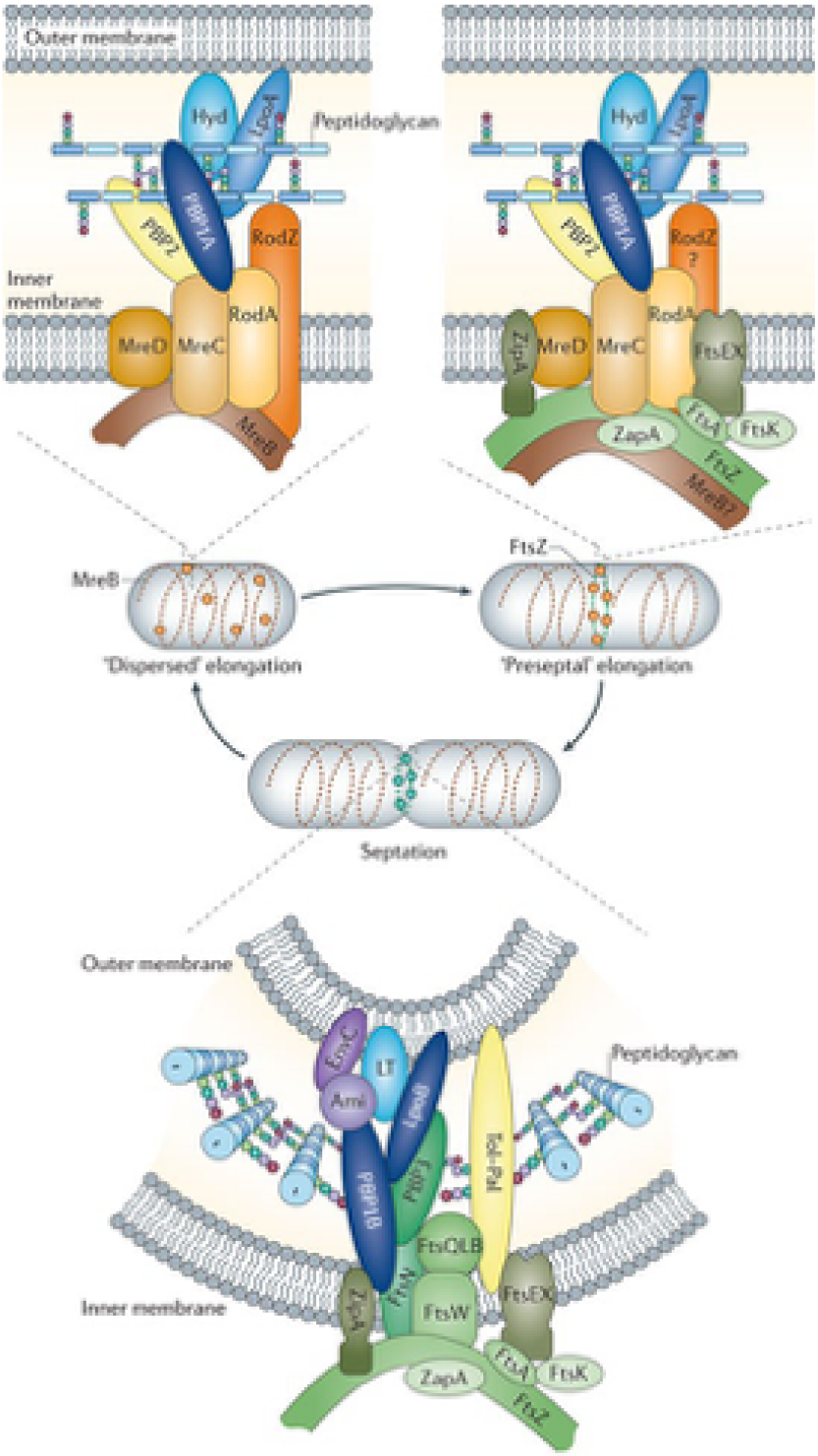

Figure 5.6: Activation mechanism of PBP1B in PG synthesis and regulation. The regulatory protein MreB and the associated membrane proteins control the position of the peptidoglycan syntheses penicillin-binding protein 1A (PBP1A). I the right panel, FtsZ and other early cell division proteins control the elongation-specific peptidoglycan synthesis complex during a 'preseptal' mode of the elongation. In the lower panel, cell division complex is represented with the inner membrane-localized cell division proteins, the peptidoglycan syntheses PBP1B and PBP3. The activity of the PBPs is regulated in part by the outer membrane-anchored lipoproteins such as LpoA and LpoB (taken from [159]).

5.3 Results

5.3.1 Protein Characterization of the construct HP1457Δ(1-21)

HP1457 was successfully cloned, expressed and purified by affinity and size-exclusion chromatography (Superose 12 10/300), as reported in Chapter 2. Its entire sequence counts 183 amino acids with a molecular weight of 19.6 kDa. Since a signal peptide was identified at the N-terminus between the amino acids T21 and D22, a construct missing a N-terminal portion of 21 amino acids was designed. A Histidine-tag at the C-terminus was added in order to selectively purify the protein from the bacterial supernatant.

```
        30          40          50          60          70
DLLLTANAML  DSMFSDPNFE  QLKGKHLIEV  SDVINDTTQP  NLDMNLLTTE

        80          90         100         110         120
IARQLRLRSN  GRFNITRASG  GSGIAADSRM  VKQREKERES  EEYNQDTTVE

       130         140         150         160         170
KGTLKAADLS  LSGKVSSIAA  SISSSRQRLD  YDFTLSLTNR  KTGEEVWSDV

       180         190    195
KPIVKNASNK  RMFLVPRGSH  HHHHH
```

Figure 5.7: HP1457 sequence, lacking the first 21 amino acids and supplement of a Histidine-tag and a cleavage-site for the Thrombin at the C-teminus.

The protein was separated from the supernatant and eluted in gel-filtration as both a dimeric and monomeric species (Figure 5.8).

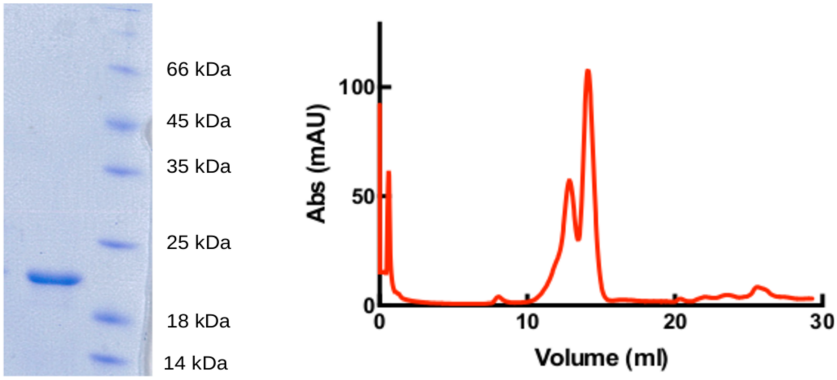

Figure 5.8: SDS-PAGE (left) of HP1457Δ(1-21) purified after gel-filtration. The sample correnspons to the concentrated fractions from the second peak of gel-filtration. The elution profile of the purified protein: the first corresponds to the dimeric form of the protein; the second corresponds to the monomer (right).

The monomeric fraction (purified in highest yields) was lately concentrated by ultra-filtration untill different concentrations were reached. Several crystallization trials were attempts at about 10 mg/ml, 20 mg/ml, 30 mg/ml, 40 mg/ml and 60 mg/ml. The same

fractions were treated with Thrombin, performing the cleavage overnight at 16°C. The protein solubility decreased after the lost of the Histidine-tag. In the last case, the protein was concentrated up to 5 mg/ml to perform some crystallization trials. None of the mentioned cases was successful.

5.3.2 HPLC-MS and CD analysis of the construct HP1457Δ(1-21)

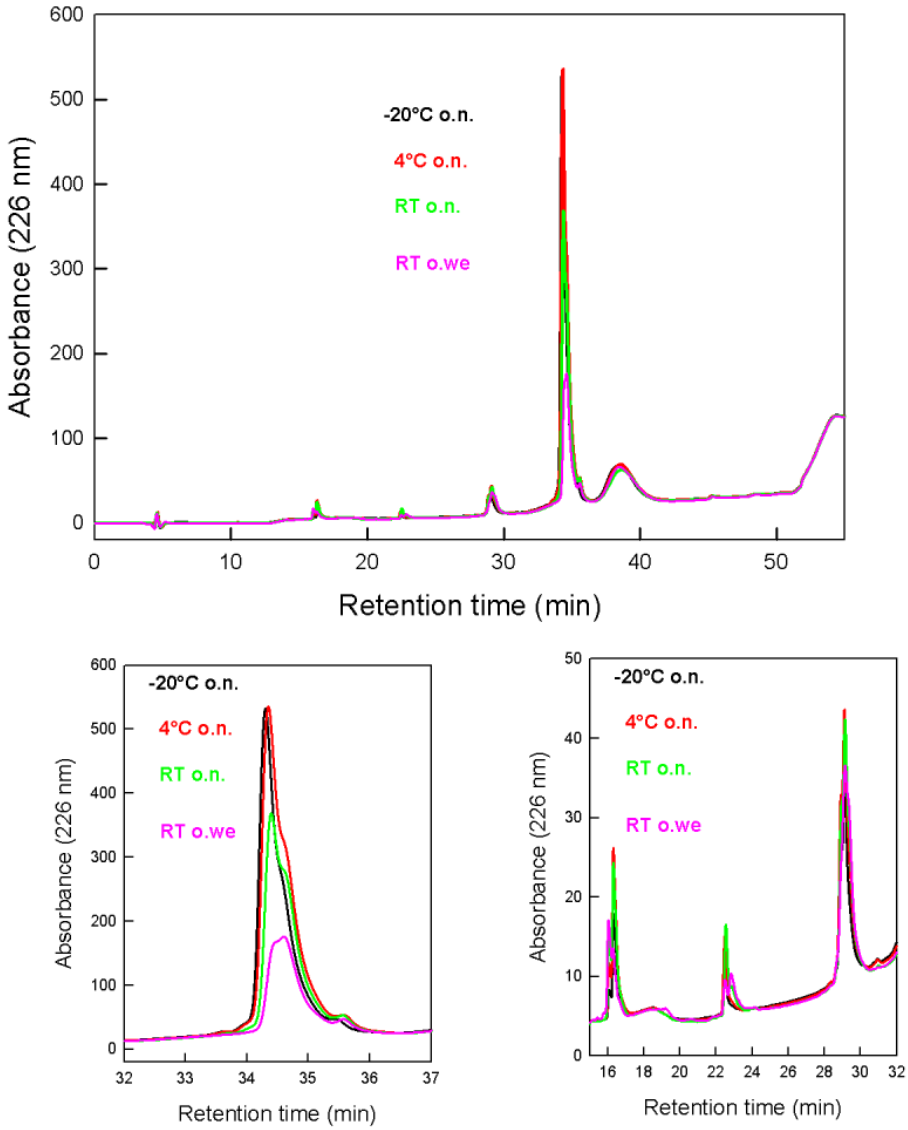

Figure 5.9: The HPLC chromatographic profile of the HP1457Δ(1-21) protein stability. Its tendency to degrade was checked at different temperatures, as indicated form the different colored lines: 1) the black line represents the purified protein stored overnight at -20° C; 2) the red line represents the purified protein stored overnight at 4° C; 3) the green line represents the purified protein stored overnight at RT; 4) the magenta line represents the purified protein stored over-weekend at RT.

In order to understand the causes of the low tendency to crystallize, a HPLC analysis was performed on the sample left overnight at different temperatures. As shown in Figure 5.9, the HPLC analysis highlights a degradation pattern that includes the production of a major species with a retention time of elution of 28.6 min. Even if their contribution to degradation is relatively low compared with the major peak, the loss of some fragments at earliest retention times (16 min, 22.5 min and 28.6 min) highlights a hypothetical fragmentation of the protein at some specific level of the sequence.
The composition of each specific peak obtained from the HPLC purification was analyzed by a ESI/q-TOF Mass Spectrometry and revealed the molecular weight of the components, as reported in Table 5.2.

RT (min)	Molecular weight (Da)		Associate peptide
	Experimental	Theoretical	
16.0	1119	1119	182-189
	1274	1275	181-189
	1290	1291	181-189 (oxidised)
22.5	473		4-7 (21-24)
	601	602	23-28
	649	649	8-13
	689	690	88-96
	905	906	9-17
	1033	1034	6-15
	1121	1121	6-16
28.6	8852	8851	102-180
	9440	9440	97-180
	10110	10111	88-180
	10699	10697	97-189
34.0	21041	21036	2-189
34.5	19784	19779	2-180

Table 5.2: Schematic summary of the Mass spectrometry led on the each concentrated fraction from the HPLC analysis. The position of the peaks is indicated in the first column, whilst in the second and in the third the molecular weight and the hypothetical peptide related to the fragment are reported.

From Table 5.2 each peak can be related to a specific fragment from the amino adic sequence: 1) the first peak corresponds to the histidine-tag fragment located at the C-terminus; 2) both the second and the third peak corresponds to fragments from the N-terminus, specifically going from the residue 8 to the residue 180; 3) the peak at 34 minutes of retention time corresponds to the entire protein sequence; 4) finally, the shoulder eluted after 34.5 minutes is representative of the protein lacking the Histidine-tag. Therefore, the lacking of a relevant degradation, with the exception of the Histine-tag, demonstrates that the protein is quite stable in solution. As a consequence, some structural features should influence the crystallization process.
In order to check the role of the protein fold in the crystal packing, three different CD spectra at different temperatures was recorded (Figure 5.10).

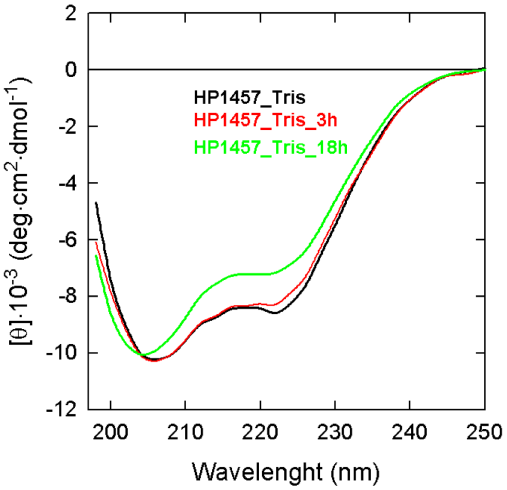

Figure 5.10: CD spectra of HP1457Δ(1-21) protein in its purification buffer. Each line represents the degradation process at room temperature after: 1) black bold line, 0 hours; 2) red bold line, after 3 hours; 3) green bold line, 18 hours.

CD spectra clearly show a typical profile of α-helix, with two peaks fixed around 208 nm and 222 nm. The protein seems to be mostly composed by a series (or a bundle) of helices, with a small contribution of random-coil that increases along the exposition of the protein at room temperature, with a signal around 200 nm.

5.3.3 Protein Characterization of the construct HP1457Δ(1-49)

Comparing the sequence of HP1457 with the protein LpoB from *E. coli*, a portion at the N-terminus (afflicted by degradation, as shown in the HPLC analysis) seems to belong to a hypothetical disordered motif, supposed to span across the peptydoglycan layer. To reduce the influence of disordered portions, a new construct lacking the first 49 amino acids was designed (Figure 5.11). Its entire sequence presents a molecular weight of 16.5 kDa. A Histidine-tag and cleavage site for Thrombin were added at the C-terminus in order to selectively purify the protein from the bacterial supernatant and reduce the disorder due to the presence of a tag.

```
         60         70         80         90        100        110        120
EVSDVINDTTQ PNLDMNLLTT EIARQLRLRS NGRFNITRAS GGSGIAADSR MVKQREKERE SEEYNQDTTV

        130        140        150        160        170        180        190
EKGTLKAADL SLSGKVSSIA ASISSSRQRL DYDFTLSLTN RKTGEEVWSD VKPIVKNASN KRMFLVPRGS

    196
HHHHHH
```

Figure 5.11: Amino acid sequence of HP1457Δ(1-49) protein.

Both the gel-filtration profiles, for mutant and wildtype proteins, show a preferential oligomeric form of the protein. In the case of the Histidine-tag, the protein is also organized as a dimer in solution. The oligomerization equilibrium is shifted to the monomeric form after Thrombin cleavage (Figure 5.12). The reaction was performed overnight at 16°C. To be sure the tag was totally removed, the sample was checked by a Western Blot essay against an anti-polyHistidine antibody (Sigma, Germany).

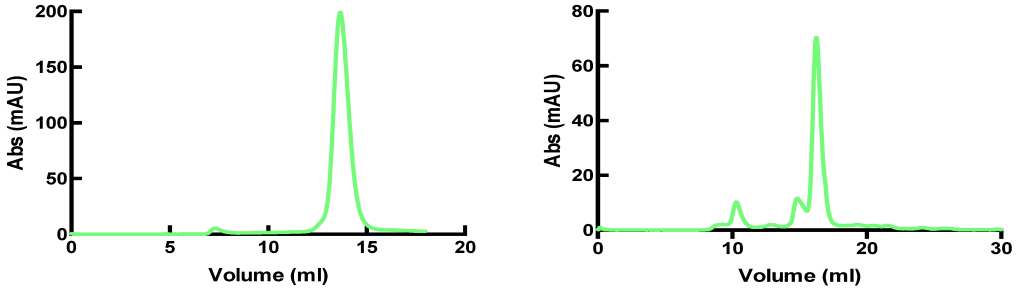

Figure 5.12: HP1457Δ(1-49) purification profiles: before the Thrombin cleavage (on the left); after the enzymatic cleavage (on the right).

As a consequence of the Histidine-tag cleavage, the protein strongly decreased its solubility. Hence, it was concentrated only up to 7 mg/ml. A 96-crystallization plate was prepared, without any relevant result. Likely, the protein was not enough concentrated for crystallization purposes.

5.3.4 Interaction between the proteins HP1457 and Lpp20

In order to overcome the crystallization issue, the *in vitro* interaction between the proteins HP1457 and Lpp20 (HP1456) was checked. The experiment was performed by Dr. Francesca Vallese and Dr. Nigam Mishra in prof. Zanotti's laboratory. The proteins were separately expressed and, after the cleavage of the His-tag from HP1456, were joined together and co-eluted in a gel-filtration purification procedure. The SDS-PAGE of the purified sample is reported in the Figure 5.13 and, as shown in the last lane of the acrylammide gel, the two proteins seem to interact in solution.

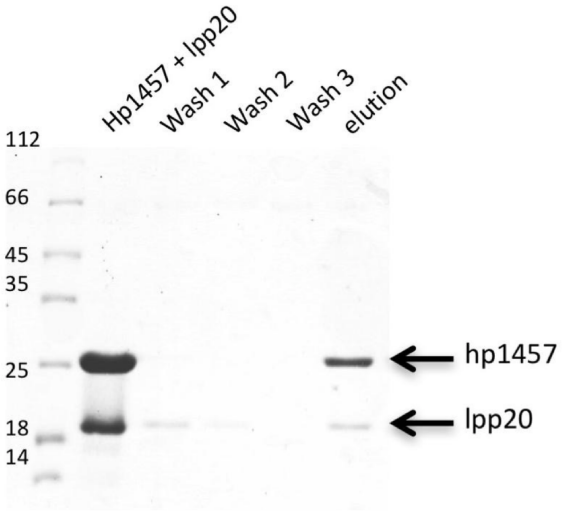

Figure 5.13: Western Blot of the gel-fitration purification, demonstrating the interaction between HP1457 and Lpp20 (HP1456).

5.4 Conclusions

The structure of HP1457 is regrettably still unknown. Even if several crystallization attempts were performed, none of them led to X-ray quality crystals. As predicted by the bioinformatic analysis of the sequence, its structural organization, resembling the LpoB protein from *E. coli*, could explain how its high flexibility could affect the organization of the system (Figure 5.14).

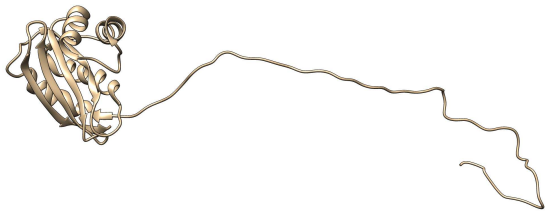

Figure 5.14: Structure of the protein LpoB from E. coli (PDB ID 2MII). Two main domains are clearly distinguished in the structure: i) a flexible domain (N-terminus), which is suppose to span across the peptidoglycan layer; ii) a globular domain (C-terminus), which is supposed to interact with the PBP1B protein.

Moreover, its role might be justified looking at the genes in the same operon, coding for the proteins going from HP1454 to HP1456. All of them are characterized by the presence of a SEC peptide at the N-terminus, which indicates their probable location could be in the periplasm, plus their secreted form offers a role as vaccine targets with high immunogenic response [15]. Specifically, the protein HP1454 has been identified as a component of *H. pylori* vesicles, where outer membrane proteins are mostly contained [160]. Interestingly, its structure shows a three-domain organization with: 1) Domain I similar to tumor necrosis factor (TNF)-a inducer protein, Tip-α; 2) Domain II consists of a three-helix bundle with a good similarity with epsin-3, which contributes to the membrane deformation and 3) Domain III which is similar to a putative lipoprotein from *Shewanella oneidensis* [15]. Furthermore, the protein HP1456, recognized to be a Lipoprotein 20 (Lpp20), was identified as a candidate vaccine antigen for protective immunization [161], whilst its structure is similar to tumor necrosis factor (TNF)-α inducer protein, Tip-α, as already seen in the Domain II of HP1454 (unpublished data). Finally, HP1455 is a protein of 130 amino acids of unknown function.
Their proximity in the bacterial genome suggests their functions could be related and likely involved in the same physiological mechanism. Indeed, the three proteins, along with HP1455, could participate in the assembly of a complex involved in regulation and/or modification of the bacterial outer membrane. To conclude, the future purpose of this work is to crystallize the HP1456-HP1457 complex in oder to clarify the proper nature of the interaction and try to understand the role of HP1457 in the physiological activity of bacteria.

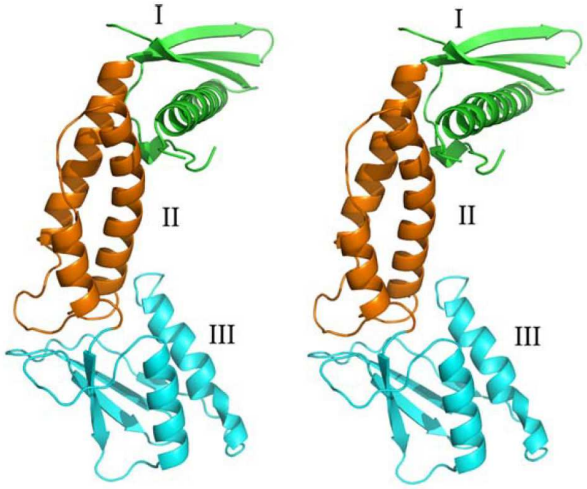

Figure 5.15: Stereo view of the structure of the protein HP1454 (PDB ID 4KZS), with the three functional domains highlighted in different colors. The structure has been adapted from the paper [15].

Bibliography

[1] Julie Parsonnet, Gary D Friedman, Daniel P Vandersteen, Yuan Chang, Joseph H Vogelman, Norman Orentreich, and Richard K Sibley. Helicobacter pylori infection and the risk of gastric carcinoma. *New England Journal of Medicine*, 325(16):1127–1131, 1991.

[2] Cesare Montecucco and Rino Rappuoli. Living dangerously: how helicobacter pylori survives in the human stomach. *Nature Reviews Molecular Cell Biology*, 2(6):457–466, 2001.

[3] SJ Hessey, J Spencer, JI Wyatt, G Sobala, BJ Rathbone, AT Axon, and MF Dixon. Bacterial adhesion and disease activity in helicobacter associated chronic gastritis. *Gut*, 31(2):134–138, 1990.

[4] Harry LT Mobley. Helicobacter pylori factors associated with disease development. *Gastroenterology*, 113(6):S21–S28, 1997.

[5] WO Paul, Michael C Lane, Steffen Porwollik, et al. Helicobacter pylori motility. *Microbes and infection*, 2(10):1207–1214, 2000.

[6] Paphavee Lertsethtakarn, Karen M Ottemann, and David R Hendrixson. Motility and chemotaxis in campylobacter and helicobacter. *Annual review of microbiology*, 65:389–410, 2011.

[7] Jiunn-Jong Wu, Bor-Shyang Sheu, Ay-Huey Huang, Shin-Ting Lin, Hsiao-Bai Yang, et al. Characterization of flgk gene and flgk protein required for h pylori colonization-from cloning to clinical relevance. *World journal of gastroenterology*, 12(25):3989, 2006.

[8] Niamh Kinsella, Patricia Guerry, Jakki Cooney, et al. The flge gene of campylobacter coli is under the control of the alternative sigma factor sigma54. *Journal of bacteriology*, 179(15):4647–4653, 1997.

[9] Ivana Pulić, Laura Cendron, Marco Salamina, Patrizia Polverino de Laureto, Dubravka Matković-Čalogović, and Giuseppe Zanotti. Crystal structure of truncated flgd from the human pathogen helicobacter pylori. *Journal of structural biology*, 194(2):147–155, 2016.

[10] Paul W O'Toole, Lars Janzon, Peter Doig, Jinzhou Huang, Magdalena Kostrzynska, et al. The putative neuraminyllactose-binding hemagglutinin hpaa of helicobacter pylori ccug 17874 is a lipoprotein. *Journal of bacteriology*, 177(21):6049–6057, 1995.

[11] Angela C Jones, RP Logan, Susan Foynes, Alan Cockayne, Brendan W Wren, and Charles W Penn. A flagellar sheath protein of helicobacter pylori is identical to hpaa, a putative n-acetylneuraminyllactose-binding hemagglutinin, but is not an adhesin for ags cells. *Journal of bacteriology*, 179(17):5643–5647, 1997.

[12] Jacques Fantini, Nicolas Garmy, and Nouara Yahi. Prediction of glycolipid-binding domains from the amino acid sequence of lipid raft-associated proteins: application to hpaa, a protein involved in the adhesion of helicobacter pylori to gastrointestinal cells. *Biochemistry*, 45(36):10957–10962, 2006.

[13] Hau B Nguyen, L-W Hung, Todd O Yeates, Thomas C Terwilliger, and Geoffrey S Waldo. Split green fluorescent protein as a modular binding partner for protein crystallization. *Acta Crystallographica Section D: Biological Crystallography*, 69(12):2513–2523, 2013.

[14] Alexander JF Egan, Nicolas L Jean, Alexandra Koumoutsi, Catherine M Bougault, Jacob Biboy, Jad Sassine, Alexandra S Solovyova, Eefjan Breukink, Athanasios Typas, Waldemar Vollmer, et al. Outer-membrane lipoprotein lpob spans the periplasm to stimulate the peptidoglycan synthase pbp1b. *Proceedings of the National Academy of Sciences*, 111(22):8197–8202, 2014.

[15] Sandra Quarantini, Laura Cendron, and Giuseppe Zanotti. Crystal structure of the secreted protein hp1454 from the human pathogen helicobacter pylori. *Proteins: Structure, Function, and Bioinformatics*, 82(10):2868–2873, 2014.

[16] Ping Cao, Mark S McClain, Mark H Forsyth, and Timothy L Cover. Extracellular release of antigenic proteins byhelicobacter pylori. *Infection and immunity*, 66(6):2984–2986, 1998.

[17] R. MACNAB. Flagella and motility. *Escherichia coli and Salmonella*, 1996.

[18] Jean-F Tomb, Owen White, Anthony R Kerlavage, Rebecca A Clayton, Granger G Sutton, Robert D Fleischmann, Karen A Ketchum, Hans Peter Klenk, Steven Gill, Brian A Dougherty, et al. The complete genome sequence of the gastric pathogen helicobacter pylori. *Nature*, 388(6642):539–547, 1997.

[19] Frank Maixner, Ben Krause-Kyora, Dmitrij Turaev, Alexander Herbig, Michael R Hoopmann, Janice L Hallows, Ulrike Kusebauch, Eduard Egarter Vigl, Peter Malfertheiner, Francis Megraud, et al. The 5300-year-old helicobacter pylori genome of the iceman. *Science*, 351(6269):162–165, 2016.

[20] Daniel Falush, Thierry Wirth, Bodo Linz, Jonathan K Pritchard, Matthew Stephens, Mark Kidd, Martin J Blaser, David Y Graham, Sylvie Vacher, Guillermo I Perez-Perez, et al. Traces of human migrations in helicobacter pylori populations. *science*, 299(5612):1582–1585, 2003.

[21] Martin J Blaser. Helicobacter pylori and gastric diseases. *Bmj*, 316(7143):1507–1510, 1998.

[22] Sebastian Suerbaum and Christine Josenhans. Helicobacter pylori evolution and phenotypic diversification in a changing host. *Nature reviews microbiology*, 5(6):441–452, 2007.

[23] Arnoud HM van Vliet, Ernst J Kuipers, Barbara Waidner, Beverly J Davies, Nicolette de Vries, Charles W Penn, Christina MJE Vandenbroucke-Grauls, Manfred Kist, Stefan Bereswill, and Johannes G Kusters. Nickel-responsive induction of urease expression inhelicobacter pylori is mediated at the transcriptional level. *Infection and immunity*, 69(8):4891–4897, 2001.

[24] Cara L Cooke, Jennifer L Huff, and Jay V Solnick. The role of genome diversity and immune evasion in persistent infection with helicobacter pylori. *FEMS Immunology & Medical Microbiology*, 45(1):11–23, 2005.

[25] Qin Jiang, Koji Hiratsuka, and Diane E Taylor. Variability of gene order in different helicobacter pylori strains contributes to genome diversity. *Molecular microbiology*, 20(4):833–842, 1996.

[26] Edgar Eduardo Lara-Ramírez, Aldo Segura-Cabrera, Xianwu Guo, Gongxin Yu, Carlos Armando García-Pérez, and Mario A Rodríguez-Pérez. New implications on genomic adaptation derived from the helicobacter pylori genome comparison. *PLoS One*, 6(2):e17300, 2011.

[27] G Sachs, JM Shin, K Munson, O Vagin, N Lambrecht, DR Scott, DL Weeks, and K Melchers. The control of gastric acid and helicobacter pylori eradication. *Alimentary pharmacology & therapeutics*, 14(11):1383–1401, 2000.

[28] Nam-Chul Ha, Sang-Taek Oh, Jae Young Sung, Kyeung Ah Cha, Mann Hyung Lee, and Byung-Ha Oh. Supramolecular assembly and acid resistance of helicobacter pylori urease. *Nature Structural & Molecular Biology*, 8(6):505–509, 2001.

[29] Bruce E Dunn, Nimish B Vakil, Barbara G Schneider, Margaret M Miller, Jason B Zitzer, Thomas Peutz, and Suhas H Phadnis. Localization of helicobacter pylori urease and heat shock protein in human gastric biopsies. *Infection and immunity*, 65(4):1181–1188, 1997.

[30] David L Weeks, Sepehr Eskandari, David R Scott, and George Sachs. A h+-gated urea channel: the link between helicobacter pylori urease and gastric colonization. *Science*, 287(5452):482–485, 2000.

[31] TL Cover and MJ Blaser. Purification and characterization of the vacuolating toxin from helicobacter pylori. *Journal of Biological Chemistry*, 267(15):10570–10575, 1992.

[32] Xaver Sewald, Wolfgang Fischer, and Rainer Haas. Sticky socks: Helicobacter pylori vaca takes shape. *Trends in microbiology*, 16(3):89–92, 2008.

[33] Nathalie Dautin and Harris D Bernstein. Protein secretion in gram-negative bacteria via the autotransporter pathway*. *Annu. Rev. Microbiol.*, 61:89–112, 2007.

[34] John L Telford, Antonello Covacci, Paolo Ghiara, Cesare Montecucco, and Rino Rappuoli. Unravelling the pathogenic role of helicobacter pylori in peptic ulcer: potential new therapies and vaccines. *Trends in biotechnology*, 12(10):420–426, 1994.

[35] Antoine Galmiche, Joachim Rassow, Anne Doye, Sebastien Cagnol, Jean-Claude Chambard, Stephanette Contamin, Virginie de Thillot, Ingo Just, Vittorio Ricci, Enrico Solcia, et al. The n-terminal 34 kda fragment of helicobacter pylori vacuolating cytotoxin targets mitochondria and induces cytochrome c release. *The EMBO journal*, 19(23):6361–6370, 2000.

[36] Maurizio Molinari, Mariolina Salio, Carmela Galli, Nathalie Norais, Rino Rappuoli, Antonio Lanzavecchia, and Cesare Montecucco. Selective inhibition of ii-dependent antigen presentation by helicobacter pylori toxin vaca. *The Journal of experimental medicine*, 187(1):135–140, 1998.

[37] Mark S Sundrud, Victor J Torres, Derya Unutmaz, and Timothy L Cover. Inhibition of primary human t cell proliferation by helicobacter pylori vacuolating toxin (vaca) is independent of vaca effects on il-2 secretion. *Proceedings of the National Academy of Sciences of the United States of America*, 101(20):7727–7732, 2004.

[38] Kelly A Gangwer, Darren J Mushrush, Devin L Stauff, Ben Spiller, Mark S McClain, Timothy L Cover, and D Borden Lacy. Crystal structure of the helicobacter pylori vacuolating toxin p55 domain. *Proceedings of the National Academy of Sciences*, 104(41):16293–16298, 2007.

[39] Timothy L Cover and Steven R Blanke. Helicobacter pylori vaca, a paradigm for toxin multifunctionality. *Nature Reviews Microbiology*, 3(4):320–332, 2005.

[40] Barbara Satin, Giuseppe Del Giudice, Vittorina Della Bianca, Stefano Dusi, Carlo Laudanna, Fiorella Tonello, Dermot Kelleher, Rino Rappuoli, Cesare Montecucco, and Filippo Rossi. The neutrophil-activating protein (hp-nap) of helicobacter pylori is a protective antigen and a major virulence factor. *The Journal of experimental medicine*, 191(9):1467–1476, 2000.

[41] DJ Evans, Dolores G Evans, Toshiki Takemura, Hironobu Nakano, Heather C Lampert, David Y Graham, D Neil Granger, and Peter R Kvietys. Characterization of a helicobacter pylori neutrophil-activating protein. *Infection and immunity*, 63(6):2213–2220, 1995.

[42] Mario Milco D'Elios, Amedeo Amedei, Andrea Cappon, Gianfranco Del Prete, and Marina De Bernard. The neutrophil-activating protein of helicobacter pylori (hp-nap) as an immune modulating agent. *FEMS Immunology & Medical Microbiology*, 50(2):157–164, 2007.

[43] Pasqualina Montemurro, Hiroaki Nishioka, William G Dundon, Marina de Bernard, Giuseppe Del Giudice, Rino Rappuoli, and Cesare Montecucco. The neutrophil-activating protein (hp-nap) of helicobacter pylori is a potent stimulant of mast cells. *European journal of immunology*, 32(3):671–676, 2002.

[44] Amedeo Amedei, Andrea Cappon, Gaia Codolo, Anna Cabrelle, Alessandra Polenghi, Marisa Benagiano, Elisabetta Tasca, Annalisa Azzurri, Mario Milco D'Elios, Gianfranco Del Prete, et al. The neutrophil-activating protein of helicobacter pylori promotes th1 immune responses. *The Journal of clinical investigation*, 116(4):1092–1101, 2006.

[45] Giuseppe Zanotti, Elena Papinutto, William G Dundon, Roberto Battistutta, Michela Seveso, Giuseppe Del Giudice, Rino Rappuoli, and Cesare Montecucco. Structure of the neutrophil-activating protein from helicobacter pylori. *Journal of molecular biology*, 323(1):125–130, 2002.

[46] Fiorella Tonello, William G Dundon, Barbara Satin, Maurizio Molinari, Giuseppe Tognon, Guido Grandi, Giuseppe Del Giudice, Rino Rappuoli, and Cesare Montecucco. The helicobacter pylori neutrophil-activating protein is an iron-binding protein with dodecameric structure. *Molecular microbiology*, 34(2):238–246, 1999.

[47] Johannes G Kusters, Arnoud HM van Vliet, and Ernst J Kuipers. Pathogenesis of helicobacter pylori infection. *Clinical microbiology reviews*, 19(3):449–490, 2006.

[48] Qiao Zhong, Shi-He Shao, Lei-Lei Cui, Run-Hong Mu, Xiao-li Ju, and Su-rong Dong. Type iv secretion system in helicobacter pylori: a new insight into pathogenicity. *Chinese medical journal*, 120(23):2138–2142, 2007.

[49] Manfred Rohde, Jürgen Püls, Renate Buhrdorf, Wolfgang Fischer, and Rainer Haas. A novel sheathed surface organelle of the helicobacter pylori cag type iv secretion system. *Molecular microbiology*, 49(1):219–234, 2003.

[50] Wolfgang Fischer, Jürgen Püls, Renate Buhrdorf, Bettina Gebert, Stefan Odenbreit, and Rainer Haas. 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(5):1337–1348, 2001.

[51] Jiro Tanaka, Toshihiko Suzuki, Hitomi Mimuro, and Chihiro Sasakawa. Structural definition on the surface of helicobacter pylori type iv secretion apparatus. *Cellular microbiology*, 5(6):395–404, 2003.

[52] Terry Kwok, Dana Zabler, Sylwia Urman, Manfred Rohde, Roland Hartig, Silja Wessler, Rolf Misselwitz, Jürgen Berger, Norbert Sewald, Wolfgang König, et al. Helicobacter exploits integrin for type iv secretion and kinase activation. *Nature*, 449(7164):862–866, 2007.

[53] Isabelle Pattis, Evelyn Weiss, Romy Laugks, Rainer Haas, and Wolfgang Fischer. The helicobacter pylori cagf protein is a type iv secretion chaperone-like molecule that binds close to the c-terminal secretion signal of the caga effector protein. *Microbiology*, 153(9):2896–2909, 2007.

[54] Hideaki Higashi, Ryouhei Tsutsumi, Syuichi Muto, Toshiro Sugiyama, Takeshi Azuma, Masahiro Asaka, and Masanori Hatakeyama. Shp-2 tyrosine phosphatase as an intracellular target of helicobacter pylori caga protein. *Science*, 295(5555):683–686, 2002.

[55] Osamu Handa, Yuji Naito, and Toshikazu Yoshikawa. Caga protein of helicobacter pylori: a hijacker of gastric epithelial cell signaling. *Biochemical pharmacology*, 73(11):1697–1702, 2007.

[56] Steffen Backert, Marguerite Clyne, and Nicole Tegtmeyer. Molecular mechanisms of gastric epithelial cell adhesion and injection of caga by helicobacter pylori. *Cell Communication and Signaling*, 9(1):1, 2011.

[57] Anna Muotiala, ILKKA M Helander, L Pyhälä, TIMO U Kosunen, and AP Moran. Low biological activity of helicobacter pylori lipopolysaccharide. *Infection and immunity*, 60(4):1714–1716, 1992.

[58] MA Monteiro, KH Chan, DA Rasko, DE Taylor, PY Zheng, BJ Appelmelk, HP Wirth, M Yang, MJ Blaser, SO Hynes, et al. Simultaneous expression of type 1 and type 2 lewis blood group antigens by. *Helicobacter pylori*, pages 11533–11543, 1998.

[59] Holly M Scott Algood and Timothy L Cover. Helicobacter pylori persistence: an overview of interactions between h. pylori and host immune defenses. *Clinical microbiology reviews*, 19(4):597–613, 2006.

[60] Hironori Yoshiyama, Hiroki Nakamura, Mitsuaki Kimoro, Kiwamu Okita, and Teruko Nakazawa. Chemotaxis and motility of helicobacter pylori in a viscous environment. *Journal of gastroenterology*, 34, 1999.

[61] Jafar Mahdavi, Berit Sondén, Marina Hurtig, Farzad O Olfat, Lina Forsberg, Niamh Roche, Jonas Ångström, Thomas Larsson, Susann Teneberg, Karl-Anders Karlsson, et al. Helicobacter pylori saba adhesin in persistent infection and chronic inflammation. *Science*, 297(5581):573–578, 2002.

[62] Elisabet Carlsohn, Johanna Nyström, Ingrid Bölin, Carol L Nilsson, and Ann-Mari Svennerholm. Hpaa is essential for helicobacter pylori colonization in mice. *Infection and immunity*, 74(2):920–926, 2006.

[63] Jeroen HB Van de Bovenkamp, Jafar Mahdavi, Korteland-Van Male, M Anita, Hans A Büller, Alexandra WC Einerhand, Thomas Borén, and Jan Dekker. The muc5ac glycoprotein is the primary receptor for helicobacter pylori in the human stomach. *Helicobacter*, 8(5):521–532, 2003.

[64] Anke Walz, Stefan Odenbreit, Kai Stühler, Andreas Wattenberg, Helmut E Meyer, Jafar Mahdavi, Thomas Borén, and Stefan Ruhl. Identification of glycoprotein receptors within the human salivary proteome for the lectin-like baba and saba adhesins of helicobacter pylori by fluorescence-based 2-d bacterial overlay. *Proteomics*, 9(6):1582–1592, 2009.

[65] Hilde De Reuse and Stefan Bereswill. Ten years after the first helicobacter pylori genome: comparative and functional genomics provide new insights in the variability and adaptability of a persistent pathogen. *FEMS Immunology & Medical Microbiology*, 50(2):165–176, 2007.

[66] Stefan Odenbreit, Markus Till, Dirk Hofreuter, Gerhard Faller, and Rainer Haas. Genetic and functional characterization of the alpab gene locus essential for the adhesion of helicobacter pylori to human gastric tissue. *Molecular microbiology*, 31(5):1537–1548, 1999.

[67] Alberto Danielli and Vincenzo Scarlato. Regulatory circuits in helicobacter pylori: network motifs and regulators involved in metal-dependent responses. *FEMS microbiology reviews*, 34(5):738–752, 2010.

[68] Ivana Pulic, Valentina Loconte, and Giuseppe Zanotti. Structural characterization at the atomic level of a molecular nano-machine: The state of the art of helicobacter pylori flagellum organization. *American Journal of Biochemistry & Biotechnology*, 10(3):143, 2014.

[69] Koushik Paul, Gabriela Gonzalez-Bonet, Alexandrine M Bilwes, Brian R Crane, and David Blair. Architecture of the flagellar rotor. *The EMBO journal*, 30(14):2962–2971, 2011.

[70] Dennis R Thomas, Noreen R Francis, Chen Xu, and David J DeRosier. The three-dimensional structure of the flagellar rotor from a clockwise-locked mutant of salmonella enterica serovar typhimurium. *Journal of bacteriology*, 188(20):7039–7048, 2006.

[71] Daniela Stock, Keiichi Namba, and Lawrence K Lee. Nanorotors and self-assembling macromolecular machines: the torque ring of the bacterial flagellar motor. *Current opinion in biotechnology*, 23(4):545–554, 2012.

[72] Jorge E Galán. Common themes in the design and function of bacterial effectors. *Cell host & microbe*, 5(6):571–579, 2009.

[73] Daniela Büttner and Sheng Yang He. Type iii protein secretion in plant pathogenic bacteria. *Plant physiology*, 150(4):1656–1664, 2009.

[74] Mark J Pallen, Scott A Beatson, and Christopher M Bailey. Bioinformatics, genomics and evolution of non-flagellar type-iii secretion systems: a darwinian perpective. *FEMS microbiology reviews*, 29(2):201–229, 2005.

[75] Jorge E Galán, Maria Lara-Tejero, Thomas C Marlovits, and Samuel Wagner. Bacterial type iii secretion systems: specialized nanomachines for protein delivery into target cells. *Annual review of microbiology*, 68:415, 2014.

[76] Christoph J Hueck. Type iii protein secretion systems in bacterial pathogens of animals and plants. *Microbiology and molecular biology reviews*, 62(2):379–433, 1998.

[77] Sophie Woestyn, Abdelmounaaim Allaoui, Pierre Wattiau, and Guy R Cornelis. Yscn, the putative energizer of the yersinia yop secretion machinery. *Journal of bacteriology*, 176(6):1561–1569, 1994.

[78] Tohru Minamino, Yusuke V Morimoto, Noritaka Hara, and Keiichi Namba. An energy transduction mechanism used in bacterial flagellar type iii protein export. *Nature communications*, 2:475, 2011.

[79] Julia Radics, Lisa Königsmaier, and Thomas C Marlovits. Structure of a pathogenic type 3 secretion system in action. *Nature structural & molecular biology*, 21(1):82–87, 2014.

[80] Chelsea R Epler, Nicholas E Dickenson, Esther Bullitt, and Wendy L Picking. Ultrastructural analysis of ipad at the tip of the nascent mxih type iii secretion apparatus of shigella flexneri. *Journal of molecular biology*, 420(1):29–39, 2012.

[81] Samuel Wagner, Lisa Königsmaier, María Lara-Tejero, Matthew Lefebre, Thomas C Marlovits, and Jorge E Galán. Organization and coordinated assembly of the type iii secretion export apparatus. *Proceedings of the National Academy of Sciences*, 107(41):17745–17750, 2010.

[82] Andrew GW Leslie and Harold R Powell. Processing diffraction data with mosflm. In *Evolving methods for macromolecular crystallography*, pages 41–51. Springer, 2007.

[83] Wolfgang Kabsch. Xds. *Acta Crystallographica Section D: Biological Crystallography*, 66(2):125–132, 2010.

[84] Computational Project Collaborative et al. The ccp4 suite: programs for protein crystallography. *Acta crystallographica. Section D, Biological crystallography*, 50(Pt 5):760, 1994.

[85] Clemens Vonrhein, Eric Blanc, Pietro Roversi, and Gérard Bricogne. Automated structure solution with autosharp. *Macromolecular Crystallography Protocols: Volume 2: Structure Determination*, pages 215–230, 2007.

[86] Airlie J McCoy, Ralf W Grosse-Kunstleve, Paul D Adams, Martyn D Winn, Laurent C Storoni, and Randy J Read. Phaser crystallographic software. *Journal of applied crystallography*, 40(4):658–674, 2007.

[87] Paul D Adams, Pavel V Afonine, Gábor Bunkóczi, Vincent B Chen, Ian W Davis, Nathaniel Echols, Jeffrey J Headd, L-W Hung, Gary J Kapral, Ralf W Grosse-Kunstleve, et al. Phenix: a comprehensive python-based system for macromolecular structure solution. *Acta Crystallographica Section D: Biological Crystallography*, 66(2):213–221, 2010.

[88] Paul Emsley and Kevin Cowtan. Coot: model-building tools for molecular graphics. *Acta Crystallographica Section D: Biological Crystallography*, 60(12):2126–2132, 2004.

[89] Karen M Ottemann and Andrew C Lowenthal. Helicobacter pylori uses motility for initial colonization and to attain robust infection. *Infection and immunity*, 70(4):1984–1990, 2002.

[90] QN Karim, RP Logan, J Puels, A Karnholz, and ML Worku. Measurement of motility of helicobacter pylori, campylobacter jejuni, and escherichia coli by real time computer tracking using the hobson bactracker. *Journal of clinical pathology*, 51(8):623–628, 1998.

[91] Robert M Macnab. How bacteria assemble flagella. *Annual Reviews in Microbiology*, 57(1):77–100, 2003.

[92] Ronan O'Toole, Debra L Milton, and Hans Wolf-Watz. Chemotactic motility is required for invasion of the host by the fish pathogen vibrio anguillarum. *Molecular microbiology*, 19(3):625–637, 1996.

[93] Hirofumi Suzuki, Koji Yonekura, and Keiichi Namba. Structure of the rotor of the bacterial flagellar motor revealed by electron cryomicroscopy and single-particle image analysis. *Journal of molecular biology*, 337(1):105–113, 2004.

[94] Ferenc Vonderviszt, Péter Závodszky, Miyuki Ishimura, Hatsuho Uedaira, and Keiichi Namba. Structural organization and assembly of flagellar hook protein fromsalmonella typhimurium. *Journal of molecular biology*, 251(4):520–532, 1995.

[95] Songye Chen, Morgan Beeby, Gavin E Murphy, Jared R Leadbetter, David R Hendrixson, Ariane Briegel, Zhuo Li, Jian Shi, Elitza I Tocheva, Axel Müller, et al. Structural diversity of bacterial flagellar motors. *The EMBO journal*, 30(14):2972–2981, 2011.

[96] Hiroyuki Terashima, Seiji Kojima, and Michio Homma. Flagellar motility in bacteria: structure and function of flagellar motor. *International review of cell and molecular biology*, 270:39–85, 2008.

[97] Jiadong Zhou, Scott A Lloyd, and David F Blair. Electrostatic interactions between rotor and stator in the bacterial flagellar motor. *Proceedings of the National Academy of Sciences*, 95(11):6436–6441, 1998.

[98] David F Blair and Howard C Berg. Mutations in the mota protein of escherichia coli reveal domains critical for proton conduction. *Journal of molecular biology*, 221(4):1433–1442, 1991.

[99] Sophie Roure, Mathilde Bonis, Catherine Chaput, Chantal Ecobichon, Austin Mattox, Charlotte Barrière, Nina Geldmacher, Stéphanie Guadagnini, Christine Schmitt, Marie-Christine Prévost, et al. Peptidoglycan maturation enzymes affect flagellar functionality in bacteria. *Molecular microbiology*, 86(4):845–856, 2012.

[100] Jenna O'Neill, Meng Xie, Marcel Hijnen, and Anna Roujeinikova. Role of the motb linker in the assembly and activation of the bacterial flagellar motor. *Acta Crystallographica Section D: Biological Crystallography*, 67(12):1009–1016, 2011.

[101] Fumiko Togashi, Shigeru Yamaguchi, May Kihara, Shin-Ichi Aizawa, and Robert M Macnab. An extreme clockwise switch bias mutation in flig of salmonella typhimurium and its suppression by slow-motile mutations in mota and motb. *Journal of bacteriology*, 179(9):2994–3003, 1997.

[102] Lawrence K Lee, Michael A Ginsburg, Claudia Crovace, Mhairi Donohoe, and Daniela Stock. Structure of the torque ring of the flagellar motor and the molecular basis for rotational switching. *Nature*, 466(7309):996–1000, 2010.

[103] Mayukh K Sarkar, Koushik Paul, and David Blair. Chemotaxis signaling protein chey binds to the rotor protein flin to control the direction of flagellar rotation in escherichia coli. *Proceedings of the National Academy of Sciences*, 107(20):9370–9375, 2010.

[104] Dennis Thomas, David Gene Morgan, and David J DeRosier. Structures of bacterial flagellar motors from two flif-flig gene fusion mutants. *Journal of bacteriology*, 183(21):6404–6412, 2001.

[105] Andrew C Lowenthal, Marla Hill, Laura K Sycuro, Khalid Mehmood, Nina R Salama, and Karen M Ottemann. Functional analysis of the helicobacter pylori flagellar switch proteins. *Journal of bacteriology*, 191(23):7147–7156, 2009.

[106] Jonathan L McMurry, James W Murphy, and Bertha González-Pedrajo. The flin-flih interaction mediates localization of flagellar export atpase flii to the c ring complex. *Biochemistry*, 45(39):11790–11798, 2006.

[107] Koushik Paul and David F Blair. Organization of flin subunits in the flagellar motor of escherichia coli. *Journal of bacteriology*, 188(7):2502–2511, 2006.

[108] Tohru Minamino and Keiichi Namba. Self-assembly and type iii protein export of the bacterial flagellum. *Journal of molecular microbiology and biotechnology*, 7(1-2):5–17, 2004.

[109] Stanley A Moore and Yunhua Jia. Structure of the cytoplasmic domain of the flagellar secretion apparatus component flha from helicobacter pylori. *Journal of Biological Chemistry*, 285(27):21060–21069, 2010.

[110] Jonathan L McMurry, John S Van Arnam, May Kihara, and Robert M Macnab. Analysis of the cytoplasmic domains of salmonella flha and interactions with components of the flagellar export machinery. *Journal of bacteriology*, 186(22):7586–7592, 2004.

[111] Frédéric Colland, Jean-Christophe Rain, Pierre Gounon, Agnes Labigne, Pierre Legrain, and Hilde De Reuse. Identification of the helicobacter pylori anti-σ28 factor. *Molecular microbiology*, 41(2):477–487, 2001.

[112] Tohru Minamino, Tetsuo Iino, and K Kutuskake. Molecular characterization of the salmonella typhimurium flhb operon and its protein products. *Journal of bacteriology*, 176(24):7630–7637, 1994.

[113] Tohru Minamino and Robert M Macnab. Components of the salmonella flagellar export apparatus and classification of export substrates. *Journal of Bacteriology*, 181(5):1388–1394, 1999.

[114] Kouhei Ohnishi, Fan Fan, Gary J Schoenhals, May Kihara, and Robert M Macnab. The flio, flip, fliq, and flir proteins of salmonella typhimurium: putative components for flagellar assembly. *Journal of bacteriology*, 179(19):6092–6099, 1997.

[115] Susan Foynes, Nick Dorrell, Stephen J Ward, Zun W Zhang, Andy A McColm, Mike JG Farthing, and Brendan W Wren. Functional analysis of the roles of fliq and flhb in flagellar expression in helicobacter pylori. *FEMS microbiology letters*, 174(1):33–39, 1999.

[116] Fan Fan, Kouhei Ohnishi, Noreen R Francis, and Robert M Macnab. 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 microbiology*, 26(5):1035–1046, 1997.

[117] Jennifer Tsang and Timothy R Hoover. Requirement of the flagellar protein export apparatus component flio for optimal expression of flagellar genes in helicobacter pylori. *Journal of bacteriology*, 196(15):2709–2717, 2014.

[118] Hirofumi Suzuki, Koji Yonekura, Kazuyoshi Murata, Teruhisa Hirai, Kenji Oosawa, and Keiichi Namba. 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(2):104–114, 1998.

[119] Tohru Minamino, Shigeru Yamaguchi, and Robert M Macnab. Interaction between flie and flgb, a proximal rod component of the flagellar basal body ofsalmonella. *Journal of bacteriology*, 182(11):3029–3036, 2000.

[120] H Okino, M Isomura, S Yamaguchi, Y Magariyama, S Kudo, and SI Aizawa. Release of flagellar filament-hook-rod complex by a salmonella typhimurium mutant defective in the m ring of the basal body. *Journal of bacteriology*, 171(4):2075–2082, 1989.

[121] Fadel A Samatey, Hideyuki Matsunami, Katsumi Imada, Shigehiro Nagashima, Tanvir R Shaikh, Dennis R Thomas, James Z Chen, David J DeRosier, Akio Kitao, and Keiichi Namba. Structure of the bacterial flagellar hook and implication for the molecular universal joint mechanism. *Nature*, 431(7012):1062–1068, 2004.

[122] Koji Yonekura, Saori Maki-Yonekura, and Keiichi Namba. Complete atomic model of the bacterial flagellar filament by electron cryomicroscopy. *Nature*, 424(6949):643–650, 2003.

[123] Kieran A Ryan, Najma Karim, Mulugeta Worku, Charles W Penn, and Paul W O'Toole. Helicobacter pylori flagellar hook-filament transition is controlled by a flik functional homolog encoded by the gene hp0906. *Journal of bacteriology*, 187(16):5742–5750, 2005.

[124] Nahid Kamal, Nick Dorrell, Aparna Jagannathan, Susan M Turner, Chrystala Constantinidou, David J Studholme, Gemma Marsden, Jason Hinds, Ken G Laing,

Brendan W Wren, et al. Deletion of a previously uncharacterized flagellar-hook-length control gene flik modulates the σ54-dependent regulon in campylobacter jejuni. *Microbiology*, 153(9):3099–3111, 2007.

[125] Tomoko Kubori, Nobuo Shimamoto, Shigeru Yamaguchi, Keiichi Namba, and Shin-Ichi Aizawa. Morphological pathway of flagellar assembly in salmonella typhimurium. *Journal of molecular biology*, 226(2):433–446, 1992.

[126] Tohru Minamino, Nao Moriya, Takanori Hirano, Kelly T Hughes, and Keiichi Namba. Interaction of flik with the bacterial flagellar hook is required for efficient export specificity switching. *Molecular microbiology*, 74(1):239–251, 2009.

[127] Christine Josenhans, Agnes Labigne, and Sebastian Suerbaum. Comparative ultrastructural and functional studies of helicobacter pylori and helicobacter mustelae flagellin mutants: both flagellin subunits, flaa and flab, are necessary for full motility in helicobacter species. *Journal of bacteriology*, 177(11):3010–3020, 1995.

[128] Wendy Wai Ling Lam, Eui Jeon Woo, Masayo Kotaka, Wai Kwan Tam, Yun Chung Leung, Thomas Kin Wah Ling, and Shannon Wing Ngor Au. Molecular interaction of flagellar export chaperone flis and cochaperone hp1076 in helicobacter pylori. *The FASEB Journal*, 24(10):4020–4032, 2010.

[129] Mohammad Khalifeh Gholi, Behnam Kalali, Luca Formichella, Gereon Göttner, Fereshteh Shamsipour, Amir hassan Zarnani, Mostafa Hosseini, Dirk H Busch, Mohammad Hasan Shirazi, and Markus Gerhard. Helicobacter pylori flid protein is a highly sensitive and specific marker for serologic diagnosis of h. pylori infection. *International Journal of Medical Microbiology*, 303(8):618–623, 2013.

[130] Gunther Spohn and Vincenzo Scarlato. Motility of helicobacter pylori is coordinately regulated by the transcriptional activator flgr, an ntrc homolog. *Journal of bacteriology*, 181(2):593–599, 1999.

[131] Catherine D Carrillo, Eduardo Taboada, John HE Nash, Patricia Lanthier, John Kelly, Peter C Lau, Rachel Verhulp, Oksana Mykytczuk, Jonathan Sy, Wendy A Findlay, et al. Genome-wide expression analyses of campylobacter jejuni nctc11168 reveals coordinate regulation of motility and virulence by flha. *Journal of Biological Chemistry*, 279(19):20327–20338, 2004.

[132] David R Hendrixson and Victor J DiRita. Transcription of σ54-dependent but not σ28-dependent flagellar genes in campylobacter jejuni is associated with formation of the flagellar secretory apparatus. *Molecular microbiology*, 50(2):687–702, 2003.

[133] Stephanie N Joslin and David R Hendrixson. Activation of the campylobacter jejuni flgsr two-component system is linked to the flagellar export apparatus. *Journal of bacteriology*, 191(8):2656–2667, 2009.

[134] Murat Balaban, Stephanie N Joslin, and David R Hendrixson. Flhf and its gtpase activity are required for distinct processes in flagellar gene regulation and biosynthesis in campylobacter jejuni. *Journal of bacteriology*, 191(21):6602–6611, 2009.

[135] MICHIO Homma, K Kutsukake, T Iino, and S Yamaguchi. Hook-associated proteins essential for flagellar filament formation in salmonella typhimurium. *Journal of bacteriology*, 157(1):100–108, 1984.

[136] Louise J Gourlay, Rachael J Thomas, Claudio Peri, Oscar Conchillo-Solé, Mario Ferrer-Navarro, Arnone Nithichanon, Jordi Vila, Xavier Daura, Ganjana Lertmemongkolchai, Richard Titball, et al. From crystal structure to in silico epitope discovery in the burkholderia pseudomallei flagellar hook-associated protein flgk. *FEBS journal*, 282(7):1319–1333, 2015.

[137] Philip R Evans and Garib N Murshudov. How good are my data and what is the resolution? *Acta Crystallographica Section D: Biological Crystallography*, 69(7):1204–1214, 2013.

[138] Andre Schmitz, Christine Josenhans, and Sebastian Suerbaum. Cloning and characterization of the helicobacter pylori flba gene, which codes for a membrane protein involved in coordinated expression of flagellar genes. *Journal of bacteriology*, 179(4):987–997, 1997.

[139] Lewis DB Evans, Simon Poulter, Eugene M Terentjev, Colin Hughes, and Gillian M Fraser. A chain mechanism for flagellum growth. *Nature*, 504(7479):287–290, 2013.

[140] Matthew J Duncan, Jeoung-Sook Shin, and Soman N Abraham. Microbial entry through caveolae: variations on a theme. *Cellular microbiology*, 4(12):783–791, 2002.

[141] Jacques Fantini. Interaction of proteins with lipid rafts through glycolipid-binding domains: biochemical background and potential therapeutic applications. *Current medicinal chemistry*, 14(27):2911–2917, 2007.

[142] Radhia Mahfoud, Murugesapillai Mylvaganam, Clifford A Lingwood, and Jacques Fantini. A novel soluble analog of the hiv-1 fusion cofactor, globotriaosylceramide (gb3), eliminates the cholesterol requirement for high affinity gp120/gb3 interaction. *Journal of lipid research*, 43(10):1670–1679, 2002.

[143] Kristina Blom, B Samuel Lundin, Ingrid Bölin, and Ann-Mari Svennerholm. Flow cytometric analysis of the localization of helicobacter pylori antigens during different growth phases. *FEMS Immunology & Medical Microbiology*, 30(3):173–179, 2001.

[144] Anneli M Lundström, Kristina Blom, Vivianne Sundaeus, and Ingrid Bölin. Hpaa shows variable surface localization but the gene expression is similar in different helicobacter pylori strains. *Microbial pathogenesis*, 31(5):243–253, 2001.

[145] Doyle J Evans and Dolores G Evans. Helicobacter pylori adhesins: review and perspectives. *Helicobacter*, 5(4):183–195, 2000.

[146] J Angström, S Teneberg, M Abul Milh, T Larsson, I Leonardsson, BM Olsson, M Olwegård Halvarsson, D Danielsson, I Näslund, Aî Ljungh, et al. The lactosylceramide binding specificity of helicobacter pylori. *Glycobiology*, 8(4):297–309, 1998.

[147] ANNELI M LUNDSTRÖM, INGRID BÖLIN, MONA BYSTRÖM, and SUSANNE NYSTRÖM. Recombinant hpaa purified from escherichia coli has biological properties similar to those of native helicobacter pylori hpaa. *Apmis*, 111(3):389–397, 2003.

[148] Stéphanie Cabantous, Thomas C Terwilliger, and Geoffrey S Waldo. Protein tagging and detection with engineered self-assembling fragments of green fluorescent protein. *Nature biotechnology*, 23(1):102–107, 2005.

[149] Pawel Listwan, Thomas C Terwilliger, and Geoffrey S Waldo. Automated, high-throughput platform for protein solubility screening using a split-gfp system. *Journal of structural and functional genomics*, 10(1):47–55, 2009.

[150] Waldemar Vollmer, Didier Blanot, and Miguel A De Pedro. Peptidoglycan structure and architecture. *FEMS microbiology reviews*, 32(2):149–167, 2008.

[151] James T Park and Tsuyoshi Uehara. How bacteria consume their own exoskeletons (turnover and recycling of cell wall peptidoglycan). *Microbiology and Molecular Biology Reviews*, 72(2):211–227, 2008.

[152] Yu-Ling Shih and Lawrence Rothfield. The bacterial cytoskeleton. *Microbiology and Molecular Biology Reviews*, 70(3):729–754, 2006.

[153] Michelle Aaron, Godefroid Charbon, Hubert Lam, Heinz Schwarz, Waldemar Vollmer, and Christine Jacobs-Wagner. The tubulin homologue ftsz contributes to cell elongation by guiding cell wall precursor synthesis in caulobacter crescentus. *Molecular microbiology*, 64(4):938–952, 2007.

[154] Ute Bertsche, Thomas Kast, Benoît Wolf, Claudine Fraipont, Mirjam EG Aarsman, Kai Kannenberg, Moritz Von Rechenberg, Martine Nguyen-Distèche, Tanneke Den Blaauwen, Joachim-Volker Höltje, et al. Interaction between two murein (peptidoglycan) synthases, pbp3 and pbp1b, in escherichia coli. *Molecular microbiology*, 61(3):675–690, 2006.

[155] SAMIRA Y YOUSIF, JENNY K BROOME-SMITH, and BRIAN G SPRATT. Lysis of escherichia coli by.-lactam antibiotics: Deletion analysis of the role of penicillin-binding proteins 1a and 1b. *Microbiology*, 131(10):2839–2845, 1985.

[156] Matthew A Gerding, Yasuyuki Ogata, Nicole D Pecora, Hironori Niki, and Piet AJ De Boer. The trans-envelope tol–pal complex is part of the cell division machinery and required for proper outer-membrane invagination during cell constriction in e. coli. *Molecular microbiology*, 63(4):1008–1025, 2007.

[157] Lisandro H Otero, Alzoray Rojas-Altuve, Leticia I Llarrull, Cesar Carrasco-López, Malika Kumarasiri, Elena Lastochkin, Jennifer Fishovitz, Matthew Dawley, Dusan Hesek, Mijoon Lee, et al. How allosteric control of staphylococcus aureus penicillin binding protein 2a enables methicillin resistance and physiological function. *Proceedings of the National Academy of Sciences*, 110(42):16808–16813, 2013.

[158] Laura K Sycuro, Zachary Pincus, Kimberley D Gutierrez, Jacob Biboy, Chelsea A Stern, Waldemar Vollmer, and Nina R Salama. Peptidoglycan crosslinking relaxation promotes helicobacter pylori's helical shape and stomach colonization. *Cell*, 141(5):822–833, 2010.

[159] Athanasios Typas, Manuel Banzhaf, Carol A Gross, and Waldemar Vollmer. From the regulation of peptidoglycan synthesis to bacterial growth and morphology. *Nature Reviews Microbiology*, 10(2):123–136, 2012.

[160] Annelie Olofsson, Anna Vallström, Katja Petzold, Nicole Tegtmeyer, Jürgen Schleucher, Sven Carlsson, Rainer Haas, Steffen Backert, Sun Nyunt Wai, Gerhard Gröbner, et al. Biochemical and functional characterization of helicobacter pylori vesicles. *Molecular microbiology*, 77(6):1539–1555, 2010.

[161] Jacqueline Keenan, Jane Oliaro, Neil Domigan, Howard Potter, Geoff Aitken, Randall Allardyce, and Justin Roake. Immune response to an 18-kilodalton outer membrane antigen identifies lipoprotein 20 as a helicobacter pylorivaccine candidate. *Infection and immunity*, 68(6):3337–3343, 2000.

Acknowledgements

I must express my sincere gratitude to Prof. Giuseppe Zanotti, because he has always welcomed new ideas and points of view, for his knowledge and his ability in handing it down, for his proper advices and his patient and gentle manners. I will never thank him enough for the chance he gave me to learn new biochemical methods and, mainly, crystallography.
I would like to thank the entire laboratory (old and new people), especially Dr. Paola Berto and Dr. Francesca Vallese, for their fruitful help in planning and performing experiments, and Dr. Tito Calí, who introduced me to the fascinating world of confocal microscopy.
An enourmus thank to Dr. Ivana Pulić Kekez and Prof. Dubravka Matković-Calogović for their sincere and deep interest in H. pylori pathological effects, for the meaningful discussions and their availability in using their instruments. Finally, my sincere thanks to Dr. Laura Cendron for her trust and patient during the nights spent collecting data; I learnt so much.
I would like to thank my family, who has always been in the shadow during this three year, but not far, allowing me to take my time to do my best.
In the end, thanks to my second family, a bunch of yelling, happy and super-supportive people who have made my whole permanence in Padua. Wherever we will be in a short while, I hardly think we will be apart.