Università degli Studi di Padova

Department of Biological Chemistry

School of Bioscience and Biotechnology
Curriculum: Biotechnology
XXIV Cycle

STRUCTURAL AND FUNCTIONAL STUDY OF *HELICOBACTER PYLORI* PROTEINS CONTRIBUTING TO STOMACH COLONIZATION

Director of the PhD School: Ch.mo Prof. Giuseppe Zanotti
Coordinator of the Curriculum: Ch.mo Prof. Giorgio Valle
Supervisor: Ch.mo Prof.Giuseppe Zanotti

PhD Candidate Sandra Quarantini

January 2012

Università degli Studi di Padova

Department of Biological Chemistry

School of Bioscience and Biotechnology

Curriculum: Biotechnology

XXIV Cycle

STRUCTURAL AND FUNCTIONAL STUDY OF HELICOBACTER PYLORI PROTEINS CONTRIBUTING TO STOMACH COLONIZATION

Sandra Quarantini

To Reggie, Anna, Emanuele, and Jahan

“Fare i conti con la complessità significa accettare un’incertezza radicale, simile a quella che ognuno di noi sperimenta nella pluralità delle scelte possibili della vita.

...la miglior “narrazione” del mondo è l’evoluzione stessa dei processi naturali. Gli osservatori scientifici, invece, sono sempre situati, e l’incertezza ed i limiti dei loro modelli sono lo stimolo profondo per nuove esplorazioni e prospettive.”

“Deal with complexity means accepting radical uncertainty, similar to one that each of us experiences in the plurality of possible choices in life.

… the best "story" of the world is the very evolution of natural processes. Scientific observers, however, are always situated, and the uncertainty and limitations of their models are the stimulus for deeper exploration and new perspectives.”

I Principi Semplici della Complessità

-Ignazio Licata-

Contents

Summary

Helicobacter pylori is a gram negative, flagellated, spiral shaped bacterium able to establish a life-long chronic infection in the human gastric mucosa. *H. pylori* is spread world-wide; in fact more than half of the human population is affected. The infection is most of the times asymptomatic. Nonetheless, an important minority of cases experiences gastric diseases, including stomach and duodenal ulcers, adenocarcinomas, and stomach lymphomas. According to phylogenetics studies, *H. pylori*-human relationship began long time ago, thus suggesting a coevolution which could explain some positive effects recently postulated. The successful colonization of the human stomach by *H. pylori* is achieved through a combination of strategies that allow it to perfectly adapt to the gastric mucosa of its human host. Among these, the synthesis of urease is essential for the acidic resistance. Motility, which is mediated by its flagella, is also an important factor, allowing the bacterium to swim in response to a gradient of pH and to stay close to the mucus epithelial layer, where the pH is a generally higher with respect to the lumen. The adherence to the host is achieved by the expression of adhesins and surface components, which contact the host and are able to evade the immune recognition by displaying a high antigenic variation. *H. pylori* is characterized by high genetic variability, not only in gene sequence but also in gene content. One of the most striking differences in *H. pylori* strains is the presence or absence of a 40-kb DNA sequence named cag Pathogenicity Island, that encodes a Type IV Secretion System, causing the translocation of CagA toxin into epithelial gastric cells. After entering the host cell, CagA induces cellular modifications, including alteration of cell structure and motility, alteration of cell proliferation and of tight junctions. Most of *H. pylori* strains secrete VacA, a toxin that provokes vacuolization in epithelial cells. All *H. pylori* strains contain a vacA gene, but the gene sequence is highly variable, causing changes in VacA functional activity. This genetic variability is obviously essential in determining the level of pathogenic offense during colonization; however host polymorphism also contributes to the variety of outcomes from *H. p*ylori infection, making the relation with the host even more complex to understand. Even if the whole bacterial genome has been sequenced, several physiological aspects concerning the relation with the host must be elucidated. In fact, about 35% of *H. pylori* genome ORF's are still categorized as putative, since a certain function has not been assigned yet. Recently many research groups have moved their efforts towards the identification of new pharmacological targets against the bacterium and towards the assignment to all genes of specific roles within *H. pylori* physiology. In

this respect, secreted proteins are interesting, since they potentially come in direct contact with the host surface. Hp1454 (Chapter II) is a putative secreted protein without orthologues in other organisms. It has been chosen as a candidate to be structurally characterized due to the following interesting features: *i*) it is predicted to cross at least the cytoplasmic membrane, *ii*) it is totally uncharacterized, *iii*) its sequence is unique. The strategy adopted involved bioinformatic analyses, PCR amplification of the selected genes starting from purified *H. pylori* chromosomal DNA (strain CCUG 17874), cloning and expression of the protein in E. coli cells. The recombinant protein was then purified using two chromatographic steps and concentrated for crystallization trials. Crystals from both native and derivatized protein (heavy metals and selenomethionines) were obtained by vapor diffusion technique. The structure was solved by single-wavelength Anomalous Dispersion (SAD) at 2.8Å. The protein is composed by three domains, the C-terminal one being characterized by a large flexibility. A function for the protein is also proposed, based on the study of Hp1454 interaction network. Looking at the interactions of Hp1454 we can hypothesize that it is a periplasmic component of a TOL-PAL system not yet characterized in *H. pylori*. To better elucidate the function, experiments on *H. pylori* mutants are needed and they are currently in progress. Others two *H. pylori* proteins, essential for stomach colonization and currently better understood, have been studied: Hp-Fur, the ferric uptake regulator and Hp-HypB, a Nickel-containing GTPase involved in Urease and [Ni-Fe]- Hydrogenase maturation (Chapter III and IV). Fur is a transcription factor implicated in Iron homeostasis. In fact, iron binds to the protein triggering a conformational change that allows the binding of transcription factors to FUR-boxes on DNA. These are promoters of genes implicated in Iron uptake and homeostasis. Hp-Fur from *H. pylori* G27 strain was expressed and purified with different approaches, and tricks to avoid unspecific aggregation have been put in place. The GTPase Hp-HypB from *H. pylori* G27 strain was cloned, expressed and purified and crystallization trials are still in progress. Solving the structure of this protein could be very useful to elucidate the mechanism of Nickel delivery to enzymes whose function is crucial for the bacterium. Finally, in the Appendix a summary of the two projects I approached during my stay at UCI (University of California –Irvine) is presented.

Riassunto

Helicobacter pylori è un batterio gram-negativo, flagellato, spiraliforme, in grado di instaurare nella mucosa gastrica umana un'infezione cronica che può durare tutta la vita. *H. pylori* è diffuso in tutto il mondo e più della metà della popolazione umana risulta coinvolta. L'infezione è generalmente asintomatica, tuttavia un'importante minoranza dei casi sperimenta malattie gastriche varie, tra cui ulcere gastriche e duodenali, adenocarcinomi e linfomi dello stomaco. D'altra parte, secondo studi filogenetici, la relazione *H. pylori* - uomo sarebbe iniziata molto tempo fa, a suggerire una coevoluzione che potrebbe spiegare alcuni effetti positivi recentemente postulati. *H. pylori* colonizza efficacemente lo stomaco umano grazie ad una combinazione di strategie che gli permettono di adattarsi perfettamente alla mucosa gastrica del suo ospite. Tra questi, la sintesi di ureasi è essenziale per la resistenza all'acidità. La motilità, resa possibile dalla presenza di flagelli, è un altro elemento importante, permettendo al batterio di nuotare in risposta ad un gradiente di pH e di stare vicino allo strato di muco epiteliale, dove il pH è generalmente meno acido. Il contatto con l'ospite è ottenuto mediante l'espressione di adesine e componenti di superficie, che realizzano l'adesione eludendo il riconoscimento immunitario, grazie ad un alto tasso di variazione antigenica. *H. pylori* è caratterizzato da elevata variabilità genetica, non solo a livello di sequenza dei singoli geni, ma anche del contenuto di geni. Una delle differenze più evidenti tra ceppi di *H. pylori* è la presenza o l'assenza di una sequenza di DNA di 40-kb chiamata Isola di patogenicità cag, codificante un sistema di secrezione di tipo IV, che causa la traslocazione della tossina CagA nelle cellule epiteliali gastriche. Una volta iniettata nella cellula ospite, CagA induce modificazioni cellulari, alterazioni nella struttura, motilità e proliferazione cellulare, e nella struttura delle giunzioni strette. La maggior parte dei ceppi di *H. pylori* secerne VacA, una tossina che provoca vacuolizzazione nelle cellule epiteliali. Tutti i ceppi di *H. pylori* contengono il gene Vaca, ma la sequenza genica è molto variabile, causando cambiamenti nell'attività funzionale di VacA. Questa variabilità genetica è ovviamente essenziale nel determinare il grado di offesa patogena durante la colonizzazione, ma anche il polimorfismo dell'ospite contribuisce a rendere variabile l'esito di un'infezione da *H. pylori*, complicando ulteriormente il rapporto con l' ospite. Anche se l'intero genoma batterico è stato sequenziato, alcuni aspetti fisiologici riguardanti il rapporto con l'ospite restano da chiarire. Infatti, circa il 35% delle ORF del genoma di *H. pylori* è ancora classificato come putativo, dal momento che una funzione certa, non è stata ancora assegnata. Recentemente molti gruppi di ricerca si sono mossi verso l'identificazione di nuovi target e l'attribuzione dei ruoli specifici all'interno della fisiologia del batterio. A questo proposito,

le proteine di secrezione possono essere considerate interessanti targets, in quanto potenzialmente entrano in contatto diretto con la superficie ospitante. Hp1454 (capitolo II) è una proteina secreta putativa senza ortologhi in altri organismi. La proteina è stata scelta come candidato per la caratterizzazione strutturale in quanto presenta alcune interessanti caratteristiche: *i*) si prevede che attraversi almeno la membrana citoplasmatica, *ii*) è del tutto indefinita, *iii*) la sua sequenza è unica. La strategia di lavoro adottata ha compreso un'analisi bioinformatica, l'amplificazione dei geni selezionati a partire da DNA cromosomico purificato di *pylori* (ceppo CCUG 17.874), il clonaggio in un vettore con His-tag ed espressione della proteina in cellule di *E. coli*. La proteina ricombinante è stata poi purificata con due passaggi cromatografici e concentrata per le prove di cristallizzazione. Cristalli di proteina nativa e derivata (metalli pesanti e selenometionine) sono stati ottenuti con la tecnica di diffusione di vapore con il robot Oryx8. La struttura tridimensionale è stata risolta mediante Single-wavelength Anomalous Dispersion (SAD) a 2.8Å. La proteina è composta da tre domini, tra cui il dominio C-terminale risulta caratterizzato da un alto grado di flessibilità. Per Hp1454 è stata inoltre proposta una funzione, in base allo studio del network di interazioni. Dall' analisi delle interazioni possiamo infatti ipotizzare che la proteina sia un componente periplasmatico di un TOL-Pal system non ancora caratterizzato in *H. pylori.* Tuttavia, per chiarire meglio la funzione, esperimenti su mutanti di *H. pylori* sono necessari e sono attualmente in corso. Altre due proteine di *H. pylori*, essenziali per la colonizzazione dello stomaco e attualmente meglio comprese, sono state studiate: Hp-Fur, il regolatore di assorbimento di ferro, e Hp-HypB, una GTPasi contenente nichel coinvolta nella maturazione dell'ureasi e della [Ni-Fe]- idrogenasi (Capitolo III e IV). Fur è un fattore di trascrizione coinvolto nella omeostasi del ferro. Infatti, il legame del ferro alla proteina innesca un cambiamento conformazionale che permette il legame del fattore di trascrizione ai FUR-box sul DNA, al livello dei promotori di geni implicati nell'assorbimento e omeostasi del ferro. Hp-Fur (ceppo G27) è stata espressa e purificata con approcci diversi, e sono stati messi in atto accorgimenti per evitare aggregazioni aspecifiche. La GTPasi Hp-HypB (ceppo G27) è stata clonata, espressa e purificata, e prove di cristallizzazione sono ancora in corso. Risolvere la struttura di questa proteina potrebbe essere molto utile per chiarire il meccanismo di trasferimento del Nickel ad enzimi la cui funzione è cruciale per il batterio. Infine, in Appendice è presentata una sintesi dei due progetti che ho affrontato durante il mio soggiorno a UCI (University of California-Irvine,). Grazie a questa esperienza ho mosso i miei primi passi nel campo della caratterizzazione di proteine di membrana fotoattivate.

1- Introduction

1.1 Helicobacter pylori

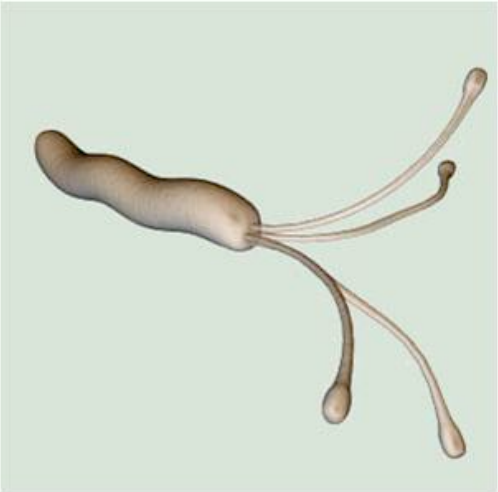


Fig. 1.1 The human gastric colonizer Helicobacter pylori.

Helicobacter pylori is a gram negative bacterium able to colonize the human stomach, a unique ecological niche. The very acidic pH, which makes the stomach an unfriendly environment for most microbes, does not prevent *H. pylori* colonization, in so much as it is the most successful human gastric bacterial parasite - more than half world's population is colonized. *H. pylori* is so well adapted to this exclusive environment that, after the first settlement which usually occurs early in life, it is able to establish a life-long chronic infection. In 1982 two Australian scientists, pathologist Barry Marshall and gastroenterologist Robin Warren, isolated an extracellular, curved bacillus lining the stomach epithelium of patients with gastritis and ulcer disease (Marshall, Warren 1984). The flagellate, spiral shaped bacterium was identified as *H. pylori* and since that time, favored by the Nobel Prize in Physiology or Medicine 2005 award, a relationship with peptic disorders, like gastritis and ulcer, has been established. Even today *H. pylori* is notorious as the major cause of stomach inflammation, which can lead in some infected individuals to gastritis, ulcer, mucosal-associated lymphoid tissue lymphoma (MALT) and even stomach cancer. Nonetheless, less than 15% of infected people display some symptoms. According to several phylogeographic studies, humans have been possibly colonized by *H. pylori* since their migration out of Africa, more than 60.000 years ago. Traveling around the world and populating new regions, they carried the bacterium with them. Since *H. pylori* is present in populations throughout the world as an ancestral host, its relation with humans looks more complex than simple parasitism, owing to their long co-evolution. This long association could have given rise not only to diseases, but also to possible protective effects, particularly with respect to esophageal disorders.

A more deep understanding of the human-bacterium interaction is needed to elucidate the role played by *H. pylori* in human life and to formulate an intelligent strategy of dealing with its most important clinical consequences.

1.2 Diffusion

From birth to death, humans are in contact with microbes, either transiently or persistently. Virtually, every mucosal and cutaneous surface in the human body is colonized by persistent residential microbes. Unlike several niches of the human body, including the oral cavity, esophagus, colon, and skin, where many bacterial species are present and no single species predominates, in the stomach *H. pylori*, when present, is the numerically dominant microorganism. *H. pylori* isolates from unrelated humans exhibit a high level of genetic diversity. Genetic variations are readily detectable by analyzing the nucleotide sequences of individual genes in different *H. pylori* strains. *H. pylori* allelic diversity is probably the consequence of multiple factors, including a high rate of mutation and intra-species genetic recombination, and a long evolutionary history of the species (McClain et al., 2009). A multi-locus sequence analysis of 370 strains isolated from humans in different part of the world identified 7 distinct populations, showing a geographic variation among *H. pylori* strains. The existence of distinct populations reflects human migrations from Africa to other parts of the earth in the Prehistory and could potentially contribute to explain the varying incidence of *H. pylori*-associated diseases in different parts of the world (Dominguez-Bello and Blaser, 2011). Unlike most other residential microbiota, *H. pylori* is becoming less common in human populations with socioeconomic development. Indeed, in Western countries the incidence of infections has dramatically fallen in the 20th century, due to environmental conditions that make the transmission more difficult, like improved sanitation, smaller family size, and frequent use of antibiotics during childhood. Asian countries, South America and Africa, on the other hand, are categorized as high risk (especially Japan, Korea and China) for incidence rate of gastric cancer.

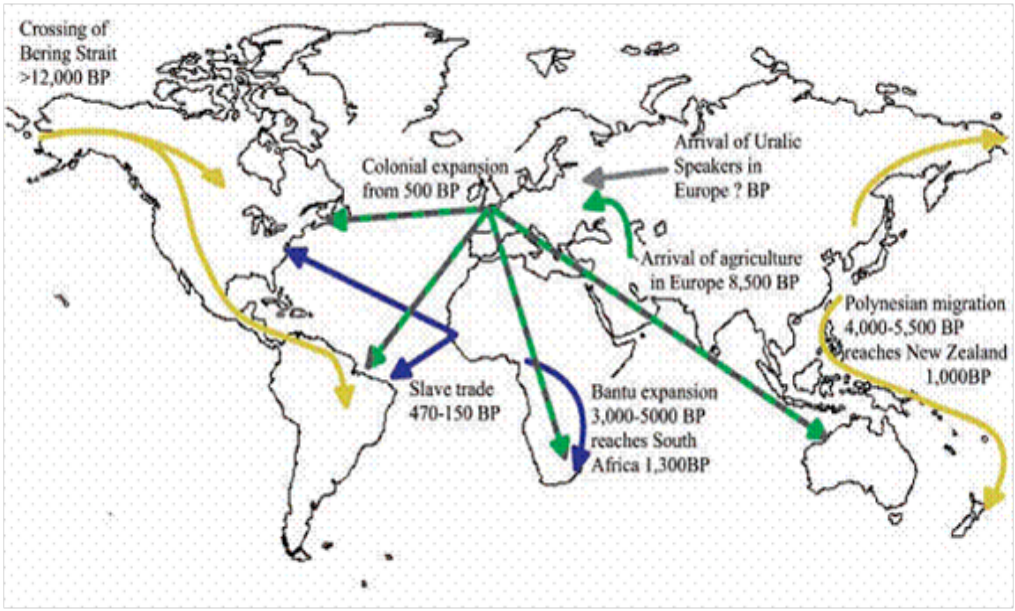


Fig. 1.2 One Interpretation of the ancient H. pylori migrations. Arrows indicate specific migrations of humans and H. pylori populations. BP, years before present. (Adapted from Falush et al., 2003)

One intriguing phenomenon is the "Asian paradox", whereby the *H. pylori* infection rate is very high in areas with little risk of gastric cancer such as South-Central Asia (India, Pakistan, Bangladesh) and South-Eastern Asia (Thailand, Indonesia), but lower in Eastern Asia, where the gastric cancer rate is the highest in the world (Japan, Korea) (Nguyen et al. , 2008). Thus, despite the undeniable connection between *H. pylori* and gastric cancer, this paradox seems to suggest that any explanation about gastric cancer onset must consider many factors, not only the bacterial incidence in some countries, but also the bacterial virulence, the environmental conditions, the host susceptibility, etc.

1.3 Genetic variability

The first sequenced *H. pylori* genome has been the 26695 strain. It consists of a circular chromosome with a size of 1,667,867 base pairs and average G+C content of 39%. 1590 predicted coding sequences were identified (Tomb et al., 1997) and among these, a putative biological role has been assigned to 1090 identified sequences. *H. pylori* strains display a high level of genetic diversity. There is a considerable variation among strains in the sequence of individual genes – corresponding alleles in different strains present an average identity of 92 to 99%, but several genes exhibit a much higher level of variability. In addition, a remarkable variation in genes content has been found. A study analyzing the genomic DNA from 56 different *H. pylor*i strains has identified 1150 genes, over a total of 1531 genes, that were present in all of the strains tested. They can be considered to represent a core genome. The remaining 25% were absent from at least

one of the 56 strains considered (MCclain et al., 2009), suggesting a unique and specific fingerprint for every strain. This extraordinary genetic heterogeneity is the result of the long evolution in different hosts and the high plasticity of *H. pylori* genome. Recombination events, including the presence of insertions elements, horizontally acquired genes (restriction recombination genes), mosaics and chromosomal rearrangements were frequently observed (Ahmed et al., 2005). Additionally, *H. pylor*i presents a natural transformation competence, since it is able to acquire plasmids and chromosomal DNA during infection. Multiple strains can colonize a host and interchange genetic material at the same time. The presence of non-randomly distributed repetitive chromosomal sequences and the high frequency of recombination between identical repeats ensure the variation within individual hosts. Moreover, during prolonged colonization of a single host, a micro-evolution phenomenon has been shown to occur; *H. pylori* isolates give rise to stable sub-species within a single individual, suggesting an adaptation of a *H. pylori* population to specific host niches, facilitated by unknown advantages conferred to them by selected plasticity regions (Ahmed et al., 2005).

1.4 Adaptation and gastric colonization

1.4.1 Acid adaptation

To colonize the gastroenteric tract, a bacterium has to survive in the acid lumen of the stomach (Montecucco, Rappuoli, 2001). *H. pylori* is able to avoid the bactericidal activity of the low pH in the lumen, since it has developed multiple mechanisms of acid adaptation. Among these, the synthesis of urease, a Ni^{2+}-containing enzyme capable of hydrolyzing urea into NH_3 and CO_2. Urease is crucial in buffering the periplasm and possibly in creating a neutral layer around the bacterial surface. Most of the urease is found in the bacterial cytoplasm, although up to 10% appears on the surface, owing to cell lysis during culture. Surface urease has a pH optimum between pH 7.5 and 8.0, but it is irreversibly inactivated below pH 4.0. The activity of cytoplasmic urease is low at neutral pH, but increases 10- to 20-fold as the external pH falls between 6.5 and 5.5, and its activity remains high down to pH 2.5 (Weeks et al., 2000). Urease is a dodecamer composed of two different subunits: six UreA (HP0073, 26.5 kDa) and six UreB (HP0072, 60.3 kDa) subunits form a huge complex, whose molecular mass is estimated to be ~600 kDa (Ha et al., 2001). Transmission electron microscopy (TEM) indicated that the enzyme is arranged as a double ring of about 13 nm diameter (Ha et al., 2001; PDB accession codes: 1E9Z). The biosynthesis of urease is governed by a seven-genes cluster, which includes *ureA* and *ureB* genes, encoding the 2

structural subunits, and *ureE, F, G, H* and *I*, which code for accessory proteins. *ureEFGH* products have been shown to be necessary for the production of an active enzyme, since they are implicated in urease assembly and nickel insertion in the active site of the apo-enzyme (Mobley et al., 1995). Among the urease genes cluster, *ureI* is the only one not involved in the biogenesis of an active urease (Skouloubris et al., 1998). In fact, ureI encodes an integral membrane protein with homology to putative amide transporters. Its expression in Xenopus oocytes demonstrated that it is an acid-activated urea transporter, able to regulate the rate of urea entry in the cytoplasm: it opens at acidic pH values to allow more urea in, and closes at neutral pH to avoid over-alkalinization. *H. pylori* pH-buffering mechanism is explained in detail in fig.1.3. Moreover, urease might also help to recruit neutrophils and monocytes in the inflamed mucosa and to activate production of proinflammatory cytokines (Montecucco, Rappuoli, 2001). As a consequence, urease is one of the main antigens recognized by the human immune response to *H. pylori*, (Ferrero et al., 1994) although the extent and the nature of this immune response after infection are not fully understood (Del Giudice et al., 2001).

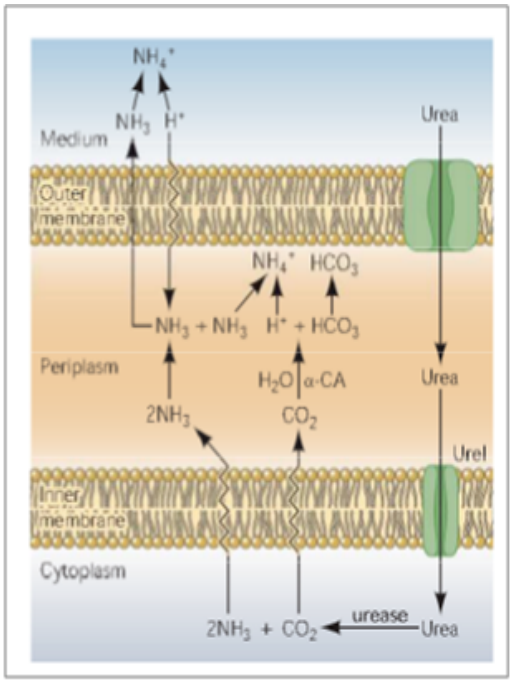


Fig.1.3 H. pylori pH-buffering mechanism. Following acidification, UreI opens and urea moves into the cytoplasm, increasing intrabacterial urease activity. Urease produces $2NH_3$ and CO_2, gases that readily diffuse into the periplasm. NH_3 and CO_2, both products from urease activity, enable acid acclimation. In fact NH_3 neutralizes entering protons by forming NH_4^+ while CO_2, is transformed into HCO_3^- and H^+ by the periplasmic carbonic anhydrase activity. NH_3/NH_4^+ pair neutralizes the entering acidity at a very acidic range of pH (pKa 9.2), while HCO_3^- is essential in buffering the periplasm at relatively neutral pH (pKa 6.1). (Adapted from Sachs et al., 2005)

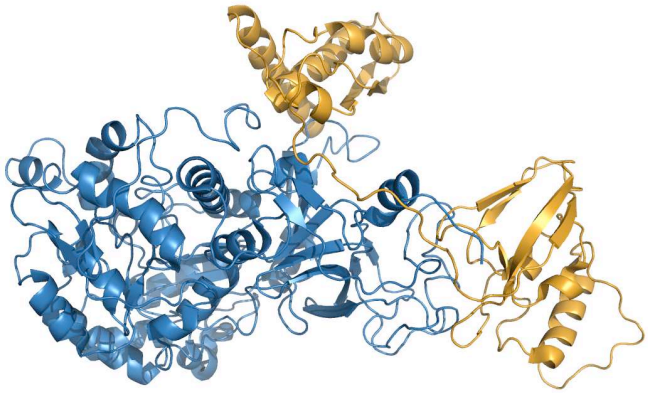


Fig. 1.4 Ribbon representation of the monomeric unit of H.pylori urease. Each unit consists of a α-subunit of 26.5 KDa (orange) and a β-subunit of 61.7 KDa (blu). (Ha et al., 2001) PDB accession code: 1E9Z.

In addition to urease, *H. pylor*i also possesses other ammonia-producing enzymes, including the Acylamide Amidohydrolase AmiE, and the Formamide Amidohydrolase AmiF (Skouloubris et al., 2001). These enzymes are aliphatic amidases specialized in the hydrolysis of short-chain amides to produce ammonia and the corresponding organic acid. AmiE and AmiF are assumed to have evolved separately from a duplication event of an ancestral gene. This suggests some enzymatic specialization: while AmiE hydrolyzes propionamide, acetamide and acrylamide in vitro, AmiF presents a unexpected substrate specificity: it only hydrolyzes formamide. Interestingly, *H. pylori* is the first pathogenic organism in which two amidases belonging to the aliphatic amidase family have been found, since these enzymes are usually found only in bacteria present in the environment, like *P. aeuruginosa* (Bury-Moné et al., 2003). According to phylogenetic analysis, AmiE and AmiF genes are predicted to have been acquired by horizontal transfer from a gram-positive organism. Due to the central role played by ammonia in the *H. pylori* survival as the major nitrogen source for acidic resistance and cytotoxic offense, the acquisition of these genes might have given to it a metabolic advantage when colonizing an unusual ecological niche, such as the stomach (Bury-Moné et al., 2003). AmiE and AmiF expression was found to be stimulated in *H. pylori* mutants deficient in urease and arginase (the latter is an enzyme which catalyzes the conversion of arginine to urea and ornithine). This observation suggests the presence of a complex regulatory network that allows the maintenance of an appropriate intracellular nitrogen balance. Urea or ammonia could serve as sensors of these regulations, avoiding toxic intracellular accumulation of ammonium and ensuring sufficient nitrogen supply in environmental conditions riches in urea, amino acids or amides (Bury-Moné et al., 2003). Compared to many other bacteria, *H. pylori* possesses relatively few transcriptional regulatory systems (Alm et al., 1999). Consistent

with this, many genes involved in the *H. pylori* acidic resistance and acclimation, including members of urease family and amidases, have been shown to be regulated in the same way by the two-component system ArsRS. ArsRS consist of a sensor kinase ArsS, which phosphorilates itself in response to pH changes and transfers the phosphoryl group, and a response regulator ArsR which, by accepting the phosphoryl group, functions as an activator or repressor of genes promoters.

1.4.2 Flagellum and Chemiotassis

Not being an acidophilic bacterium, *H. pylori* has evolved several strategies to minimize exposure to the low pH in the stomach lumen by remaining in very close proximity to the surface of the epithelium, where the pH is near neutrality (Amieva et al., 2008). *H. pylor*i is a good swimmer, able to remain within a narrow band of the protective mucus gel that is constantly being secreted by epithelial cells of the stomach. This ability is made possible by its polar flagella, that are in fact considered an important colonization factor, since non-motile mutants cannot colonize the stomach (Montecucco, Rappuoli, 2001). Another ability important for colonization is the control of the direction of movement, which is achieved thanks to a chemotactic response to pH gradient. Indeed, from the stomach lumen to the apical surface of the stomach mucosa a large pH gradient generated by a different permeability of the mucus to different ions has been postulated (Montecucco, Rappuoli, 2001). *H. pylori* has been shown to increase its swimming speed when placed near an acidic gradient. It can also change its swimming paths to favor movement away from an acidic environment. In fact, as experiments on *H. pylori* mutants showed, the deletion on gene tlpB, which codes for a chemoreceptor, causes the loss of its ability to move away from acidic places (Amieva et al., 2008). *H. pylori* cells normally possess a bundle of 2 to 6 flagella about 3 µm long and sheathed probably to protect the flagellar structure from the attack of the stomach's acid. *H. pylori* flagella, like those of other enteric bacteria, are composed of three structural elements: a basal body, embedded in the cell wall containing the proteins complex essential for rotation and chemotaxis, an external helix-shaped filament as propeller extremity and a hook, functioning as a junction between these two portions. More than 50 putative proteins are predicted to be involved in expression, secretion, and assembly of this complex flagellar apparatus (Alm et al., 1999 and Tomb et al., 1997 in Spohn and Scarlato, 2001). Until now, just the proteins composing the filament and the hook have been characterized in details. FlaA and FlaB genes encode for the two principal constitutive units of filament, flagellin A and B. FlaA is the predominant subtype and FlaB the minor subtype, localized close to the basis of the filament (Kostrzynska et al., 1991). Both flagellins have similar molecular mass (53 kDa) and share

considerable amino acid homology (58% identity). They are essential for full motility and for the establishment of a persistent infection in animal model (Eaton et al., 1996). The hook is composed of FlgE subunits of 78 Kda, whereas the fliD gene encodes a hook-associated protein that promotes the incorporation of flagellin monomers into the growing flagellar filament. Sequence homology studies have identified many other genes implicated in the flagellum structure and function. The corresponding proteins, shown in the figure 1.5, are currently considered putative and they are not well characterized. The MS-ring is the first complex to be assembled during morphogenesis. It is composed of *FliF* subunits and its function is to provide a construction base for the other flagellar proteins (FlgB, FlgC and FlgG) and an anchor for the motor switch proteins FliM, FliN and FliG and the motor rotation proteins MotA and MotB. *FliI* and *FliH* genes encode the subunit of the P and L rings, which anchor the flagellum to the periplasm and to the outer membrane respectively (Spohn and Scarlato, 2001). Many other putative proteins seem to be implicated in the assembly of the complex. Since the flagellar apparatus is localized mainly beyond the cytoplasmic membrane, many flagellar proteins have to cross the membrane to reach their final destination (Spohn and Scarlato, 2001). The structural subunits of the filament and the hook are believed to be secreted by a specialized secretion pathway at the MS ring level. The members of this systems, actually not yet fully characterized, share homology to components of the type III secretion system found in other organisms. Other flagellar proteins are expected to be secreted *via* the traditional signal-peptide dependent pathway.

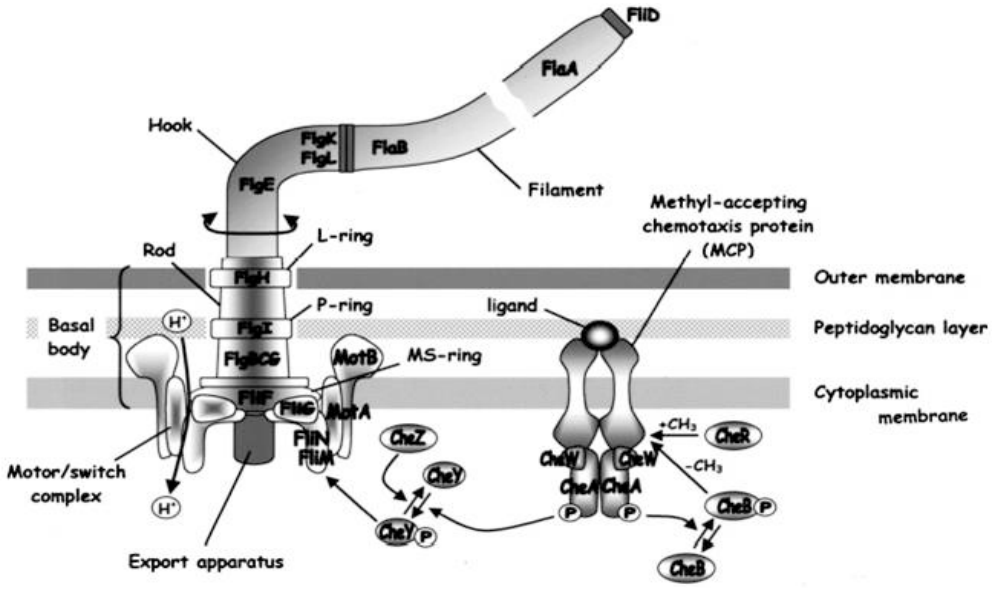


Fig.1.5 Flagellar and chemotaxis proteins and their putative locations and interactions. The protein components of the major elements of the flagellum, the external filament, the hook, the basal body, and the motor-switch complex are shown. (Adapted from Mobley , Mendz, and Hazell ,2001).

1.4.3 Adhesion

Approximately 20% of *H. pylori* bacteria in the stomach are found adhered to the surface of mucus epithelial cells (Hessey et al., 1990). Electron microscopy studies of gastric biopsy specimens have revealed extensive areas of adhesion to the host cells, and bacteria localize sometimes in the intracellular spaces, by involving a cytoskeleton adjustment at the adhesion site. Sometimes they are internalized into epithelial cells. The adhesion process involves specialized molecular interactions with the gastric mucosa that may lead to intimate attachment and modification of the cell surface and the underlying cytoskeleton (Amieva et al., 2008). More than 30 genes encoding for outer membrane proteins have been categorized as possible candidates for adhesion. Several of these have been identified as adhesins, suggesting multiple and perhaps redundant variable modes of attachment to the cell surface (Amieva et al., 2008). Moreover, different *H pylori* strains express different adhesins sets, and their expression can change within a single strain over time, thus suggesting an evolutional meaning in the redundancy of adhesive mechanisms. *H. pylori* has been shown to adhere to MG2 in the human salivary mucin (Andersen, 2007), making possible the first step of colonization. Adherence to sialic acid in the mucin seems to be a common feature in most *H. pylori* strains: there are at least six bacterial adhesins that bind sialic acid, and three of these genes (hpaA, nap, sapA) have been identified (Wadstrom, Hirmo & Borén, 1996). A well-known *H. pylori* adhesin is BabA, which binds the fucosylated blood group antigen Lewis-b present in the gastric secretory mucin MUC5AC. The biological role of the adhesion process is still controversial. On one hand, adhesive interactions contribute to inflammation and seem to be involved in disease progression. In particular, when transgenic mice expressing the Lewis-B antigen were infected with *H. pylori,* the mice showed increased bacterial attachment, more severe chronic gastritis, and parietal cell loss (Guruge et al., 1998). The delivery of the major *H. pylori* virulence factors CagA and VacA (see below) are intimately related to adhesion as well, suggesting a major role of adhesion in the delivery of toxins (Amieva et al., 2008). Some investigators have speculated that inflammation could have an adaptive role, since it promotes the release of nutrients into the gastric lumen. On the other hand, adhesion seems necessary to avoid mechanical clearance and allow the bacterium to hold firmly to its environment.

1.5 Virulence and pathogenesis

Stomach colonization by *H. pylori* is asymptomatic in many cases. This bacterium typically colonizes the human stomach for many decades without adverse consequences, but its presence is sometimes associated with an increased risk of various diseases, including peptic ulcers, noncardia gastric adenocarcinoma, and gastric mucosa-associated lymphoid tissue (MALT) lymphoma. *H. pylor*i has started colonizing humans so long ago that it is not surprising that its relationship is that of both a commensal bacterium and a pathogen, causing some diseases and possibly protecting against others (Dorer et al., 2009). The degree of pathogenicity, which varies depending on the strain, is called virulence. The ability of *H pylori* to survive and colonize the stomach, described in detail previously, is therefore essential for the development of virulence, but the latter does not necessarily occur. In the pathogenesis of *H. pylori* infection in humans, three steps can be distinguished: (1) entry to, adherence to, and colonization of the human gastric mucosa; (2) avoidance, subversion, or exploitation of the human immune system; and (3) multiplication, tissue damage, and transmission to a new susceptible host or spread to adjacent tissues (McGee et al., 2000). Virulence is mediated by several virulence factors that are defined by their involvement in one or more steps of the process. Among virulence factors, toxins secretion is directly connected to "intoxication" of host cells. Over the past century, more than hundred different toxins produced by various bacterial species have been characterized. The cellular effects of these toxins range from cell death to a variety of non-lethal changes, including permeabilization of membranes, blockade of exocytosis, and alterations of cellular signal transduction, cellular cytoskeletal properties and cellular proliferation (Cover and Blanke, 2005).

1.5.1 VacA

The toxin VacA is one of the best characterized *H. pylori* virulence factors. The vacuolating cytotoxin A was discovered in *H. pylori* broth culture filtrates for its propensity to form large intracellular vacuoles in cultured mammalian cells. The amino acid sequence of VacA does not show a significant similarity to any other known bacterial or eukaryotic protein (Cover and Blanke, 2005). As with most *H. pylori* factors, there is considerable genetic diversity in the VacA gene and, therefore, the activities of different alleles of the toxin differ in their cytotoxicity (Amieva et al., 2008). VacA shows toxic effects when it is secreted and delivered in an active form to host cell membranes. VacA is produced as a 140 kDa protoxin and is secreted via an auto-transporter pathway that involves the carboxylic end of the protoxin. The amino terminus contains a signal

sequence or “s” region that shows allelic variability and allows the toxin classification into different types. VacA s1 type is an active toxin strongly associated with strains causing ulcers and gastric cancer. The middle region of the gene, or “m” region, shows as well allelic variations, with m1 subtypes having stronger vacuolating activity. The mature toxin has a molecular mass of 88 Kda and is secreted as a soluble protein into the extracellular space. About 40-60% of the toxin has been shown to remain associated to the outer membrane of bacterium. Interestingly, the VacA molecules associated with the bacterial surface are still functional and are delivered to host cells by direct contact between adhered bacteria and the host cell membrane (Ilver D et al., 2004). The secreted toxin can assemble into water-soluble oligomeric flower-shaped structures (Fig. 1.7), and can be inserted into planar lipid bilayers to form anion-selective membrane channels. On exposure to either acidic or alkaline pH conditions, these oligomeric complexes dissociate into monomeric components (Cover et al., 1997). Water-soluble VacA oligomers have little vacuolating activity, but a short exposure to acid or alkaline media markedly increases the activity of these complexes. Accordingly, VacA cytotoxicity probably requires an initial interaction of monomeric forms of VacA with cells, and subsequent oligomerization contributes to the ability of VacA to form channels in cell membranes (Cover et al., 1997). The VacA monomer can undergo limited proteolytic cleavage to yield two fragments — p33 and p55 (Cover et al., 1997) that remain non-covalently associated. Several studies indicate that the p55 domain portion has an important role in the binding of VacA to host cells, while the p33 region presents the propensity to be inserted into membranes. Within the p33 region there is the only strongly hydrophobic region containing three tandem GXXXG motifs, characteristic of a transmembrane dimerization sequence. Mutagenesis on several residues within the N-terminal hydrophobic region abolishes the ability of VacA to form membrane channels in planar lipid bilayers and also abolishes vacuolating cytotoxic activity (McClain et al., 2003). These data provide evidence that the unique hydrophobic region at the N-terminus of VacA has a role in membrane channel formation, and that membrane channel formation is required for VacA-induced cell vacuolation (Cover and Blanke, 2005).

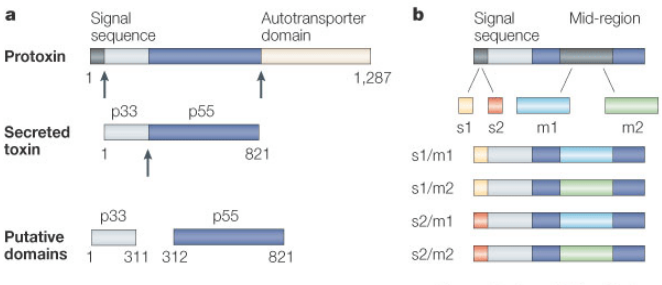


Fig. 1.6 a) The 140-kDa VacA protoxin is cleaved to yield an 88-kDa mature toxin that is secreted into the extracellular space via an autotransporter mechanism. The 33-kDa C-terminal domain of VacA is predicted to insert into the outer membrane and to form a channel, through which the mature VacA toxin is secreted. Arrows indicate sites of proteolytic cleavage. b) The high level of diversity among vacA alleles from different H. pylori strains is particularly striking near the 5' terminus of vacA (the s-region) and in the mid-region of the gene (the m-region) Adapted from Cover and Blanke 2005.

The acid-activated VacA toxin is internalized by the host and it has been localized in multiple intracellular sites, including endosomal compartments, the large intracellular vacuoles that form as a consequence of VacA intoxication, and the inner mitochondrial membrane. Many of the cellular effects of VacA, both in epithelial cells and T-lymphocytes, can be attributed to the ability of this toxin to be inserted into membranes and to form anion-selective channel (Cover and Blanke, 2005). It is believed that VacA inserts into the membranes of late endosomal vesicles induces the formation of pores with chloride channel activity, the alteration of the composition of anions within endosomes. It finally leads to osmotic swelling (Amieva et al., 2008). VacA-induced vacuoles are acidic because their limiting membrane contains the vacuolar-type ATPase proton pump (v-ATPase), whose function is essential for vacuole formation (Papini et al., 1996). Accordingly, vacuoles are promoted by the accumulation of membrane-permeable weak bases that are trapped by protonation in their lumen (Montecucco, Rappuoli, 2001). A proper function of endosomal trafficking is necessary for antigen presentation, which is shown to be perturbed by VacA essays *in vitro*. This inhibition of local antigen processing could be part of a strategy of the bacterium survival, that could significantly contribute to its chronic infection of the human stomach (Montecucco, Rappuoli 2001). In addition, VacA induces host's cell death through apoptosis by forming pores in mitochondrial membranes. A third reported effect of VacA is its ability to cause leakage of ions and small molecules, such iron, nickel, sugars, and amino acids, by disrupting the barrier function of tight junctions, without major disruptions in junction's integrity. This could be a mechanism by which *H pylori* acquires nutrients across an intact epithelial barrier (Amieva et al., 2008).

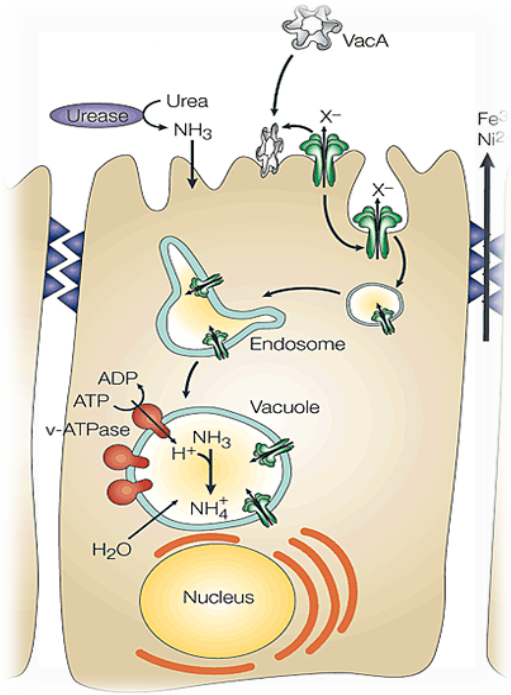


Fig.1.7 VacA vacuole formation mechanism. The toxin binds to the apical portion of epithelial cells and inserts itself into the plasma membrane, forming a hexameric anion-selective channel. These channels release bicarbonate and organic anions from the cell cytosol to support bacterial growth. The toxin channels are slowly endocytosed and, after reaching late endosomal compartments, they increase their permeability to anions with the enhancement of the electrogenic vacuolar proton pump. In the presence of weak bases, including the ammonia generated by the Helicobacter pylori urease, osmotically active acidotropic ions will accumulate in the endosomes. This leads to water influx and vesicle swelling, an essential step in vacuole formation (adapted from Montecucco et al. 2001).

1.5.2 Hp-NAP

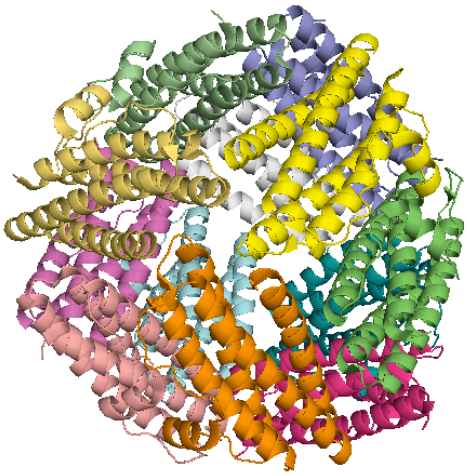


Fig. 1.8 Ribbon representation of Neutrophil activating protein (HP-NAP). HP-NAP is a dodecamer (PDB code 1JI4)

The *H. pylori* neutrophil-activating protein (Hp-NAP), is a 150 KDa oligomer composed of twelve identical 15-KDa subunits. The atomic structure reveals a four-helix bundle monomer that oligomerizes to form a dodecamer with a central internal iron-containing negative cavity (Zanotti

et al., 2002). Hp-NAP is similar to other Dps-like proteins (Zanotti et al. 2002; Papinutto et al., 2002). Dps-like proteins are a family of bacterial proteins whose expression is induced under stress conditions. They protect bacterial DNA from oxidizing radicals generated by the Fenton reaction and also from various other damaging agents. The napA gene is highly conserved among isolates of *H. pylori*, a property that may indicate a precise structurally-dependent function or a lack of immune selection for the diversification of Hp-NAP (De Bernard and D'Elios, 2010). Usually Dps-like proteins contain a positively charged N-terminus that promotes the binding to DNA and its condensation. Hp-NAP does not possess a positively charged N-terminus but, unlike the other members of the family, is characterized by a positively charged protein surface which has been proposed to be responsible for binding and condensing DNA (Ceci et al., 2007). Notably, HP-NAP is a major antigen in the human immune response; in fact, the majority of the infected patients have antibodies against it (Satin et al., 2000). Hp-Nap was shown to promote the adhesion of human neutrophils to endothelial cells and the production of reactive oxygen radicals (Yoshida et al., 1993; Evans et al., 1995). Thus, Hp-Nap is capable of crossing the epithelial layer and inducing the activation of several pro-inflammatory cytokines responsible for the inflammation process. Once accumulated within the tissue, Hp-NAP alone or together with other bacterial factors would recruit neutrophils and monocytes from the blood, thus amplifying the flogistic process (De Bernard and D'Elios, 2010). This inflammation could promote *H. pylori* growth by releasing of nutrients from the degraded epithelium, but the specific role of Hp-Nap inside the bacterium is still controversial.

1.5.3 CagPAI

H. pylori genome presents a 40Kb region of chromosomal DNA known as the cytotoxin-associated gene-pathogenicity island (cagPAI) containing about 30 genes coding for a type IV secretion system and the effector protein CagA. The latter is translocated into host cells. cagPAI is characterized by a different nucleotide composition with respect to the overall bacterial genome and is flanked by transposable elements, suggesting an acquisition by horizontal transfer. It can be hypothesized that early bacteria communities originating from crop plants, animals or rodent pests rampant in the vicinity of early human societies may have served as donors of some of the virulence cassettes (Ahmed et al., 2009). The presence or absence of this chromosomal region is one of the most striking differences among *H. pylori* strains, concurring in making more complex the genetic diversity of the bacterium. The pathogenicity island provides *H. pylori* with at least two unique properties: an increased transmission probability and the transformation of what would be

an almost commensal into a potential pathogen (Montecucco, Rappuoli, 2001). CagA does not present any sequence homology to other known proteins, while many of the genes adjacent have significant homology to genes coding for a type IV secretion system (T4SS). The T4SSs are a large group of transporter machines present in many gram-negative bacteria that are ancestrally related to conjugation systems (Cascales and Christie, 2003). The best studied among T4SS is that present in the plant pathogen *Agrobacterium tumefaciens.* These transporters are functionally diverse both in terms of the transported substrate (proteins or DNA or DNA–protein complexes) and the recipients, which can be either bacteria of the same or different species, or organisms from a different kingdom like fungi, plants or mammalian cells (Backert, Selbach, 2008). In *A. tumefaciens* the T4SS is composed of at least 12 proteins termed VirB1-11 and VirD4. They can be categorized into three groups: cytoplasmic or inner membrane proteins, a core complex located in the periplasm and in the membrane, and a surface structure or pilus, that projects beyond the outer membrane (Busler et al., 2006). *H. pylori* tipe IV secretion system has been characterized in some components: almost 11 of the 29 cag proteins have been shown to be part of the secretion machinery or have been proposed to represent functional homologues of VirB proteins (Cendron, Zanotti, 2011). Among these structural/functional homologs, the proteins forming the core complex, CagT/HP0532, CagX/HP0528, and CagY/HP0527 corresponding to VirB7, VirB9, VirB10 respectively, have been characterized; CagE/HP0544 (VirB4) and Cagα/HP0525 (VirB11) have been shown to have ATPase activity and Cagβ/HP0524 corresponds to the coupling protein VirD4, which binds DNA substrate in *A. tumefaciens*. CagL is an adhesin that works as an anchor of the apparatus to the host surface. It does so through binding to an integrin and promoting signal transduction. CagF is a chaperon-like protein crucial for CagA traslocation. After injection in the host, CagA is phosphorylated by a kinase of the host belonging to the Src family (SFKs). The phosphorylation event triggers several alterations in the gastric epithelial cells, including activation of SH2 domain-containing tyrosine phosphatase 2, alteration in cell structure and motility, alteration of tight junctions, alterations in cell scattering and proliferation, activation of β-catenin, and perturbation of epithelial cell differentiation and polarity (Cover and Blaser, 2009). Furthermore, CagA has been shown to take part in the inflammatory process, due to its propensity to stimulate the synthesis of interleukin-8. Moreover, a systematic mutagenesis study, conducted by Fischer et al.(2001), revealed the dramatic contribution of many cag proteins, about 14, in IL-8 activation. CagA amino acid sequence is characterized by the presence of EPIYA motifs (Glu-Pro-Ile-Tyr-Ala) located at the C-terminus of the protein. These motifs are the target of many host

kinases that recognize and phosphorylate the tyrosine residues of the motifs. Four different EPIYA motifs (A,B,C, and D) have been found, according to the different amino acid sequences flanking each of them. Thus, CagA can be categorized based on the alignment of different EPIYA motifs on two types: Western type, prevalent in strains present in Europe, America, Australia, and Africa, presenting EPIYA-A and EPIYA-B followed by 1-5 EPIYA-C; Eastern type, containing as a third motif EPIYA-D instead of EPIYA-C and circulating in Japan, China and Korea. CagA motifs EPIYA-A and B have shown to be much weakly susceptible of phosphorylation compared to their variant C and D. Thus, both number and type of the motifs, seem to be related to different level of pathogenicity.

1.6 Gastro-duodenal diseases

H. pylori typically colonizes the human stomach without causing adverse consequences for many years. Since the bacterium is so well adapted to its environment, in most cases the predominant symptom of a prolonged infection is a mild gastritis resulting from the inflammation process linked to nutrients acquisition. Moreover, *H. pylori* is versed in evading immune recognition by the host, by displaying a phase variation and an antigenic variation of its surface components, like outer membrane proteins (omp) and lipopolysaccharide (LPS) antigens. LPS from most bacterial organisms serve as a potent signal for the development of an inflammatory response. *H. pylori* strains commonly express LPS 0 antigens that are structurally related to Lewis blood group antigens present in human cells, showing an ability of molecular mimicry leading to long term colonization (Algood and Cover, 2006). Flagellar proteins and many bacterial factors also appear to be designed to reduce inflammation or recognition by the immune system (Amieva et al., 2008). A discrete incidence (~10%) of several diseases has been correlated with its long-term stomach colonization, including peptic ulcers, non-cardia gastric adenocarcinoma, and gastric mucosa-associated lymphoid tissue (MALT) lymphoma (Cover and Blaser, 2009). These various and divergent clinical outcomes deriving from the *H. pylori* infection can be elucidated considering more factors. First of all, bacterial polymorphism plays a role crucial in planning the consequences of the colonization. Thus, *H. pylori* strains expressing multiple “interaction factors” (CagA+, s1VacA+, BabA+ strains) are predicted to be highly interactive with the host, whereas strains lacking these factors would be relatively non-interactive. Consistent with these predictions, CagA+, s1VacA+, BabA+ strains are associated with increased gastric epithelial diseases, compared with strains that do not express these factors (Cover and Blaser, 2009; Hennig et al., 2004). Host polymorphism is also a determinant factor of *H. pylori*-associated disease risk. In effect, genetic

polymorphisms directly influence inter-individual variation in the magnitude of cytokine response (Amieva et al., 2008). Particularly some interleukin-1β and interleukin-10 genotypes have been correlated to a higher risk of gastric cancer in response to *H. pylor*i infection (Amieva et al., 2008). Various other proinflammatory gene polymorphisms have been associated whit an increased risk of noncardia adenocarcinoma (Cover and Blaser, 2009). Thus the individual-specific synergy between bacterial and host polymorphisms is crucial in determining disease risk during the colonization. Additionally, environmental factors, like dietary salt intake or, more generally, alimentary habits and sanitary conditions can influence the risk of gastric cancer.

1.7 Potential Benefits

H. pylori colonization can lead many biological costs to the host; conversely, the absence of *H. pylori* - caused by full eradication - is suggested to be associated with a potentially increased risk of onset of other disorders. According to this, an association has been observed between the presence of *H. pylori* and the absence of esophageal disorders, like gastro-esophageal reflux, Barrett's esophagus, and esophageal adenocarcinoma. Especially CagA+ strains, that are the most interactive and virulent, have been correlated to this protective effect. One hypothesis to explain these correlations is that *H. pylori* presence diminishes gastric acidity and protects the esophageal environment. Another negative correlation has been shown with childhood asthma, were the absence of *H. pylori* is associated with an increased risk of developing allergy syndromes, including asthma. The gastric recruitment of T-cells in the stomach due to *H. pylori* cagA+ presence could ultimately affect the activities of T cells present in other mucosal and cutaneous sites (Cover and Blaser, 2009). Other potential benefits on metabolism and on defense against various pathogens, like those triggering diarrheal diseases, have been also postulated (Cover and Blaser, 2009; Rothenbacher et al., 2000; Bode et al., 2001). A general question can be raised here: are these apparent benefits directly dependent on bacterium colonization or is *H. pylori* presence just a defense against the exposure to other bacteria or foreign antigens? If *H. pylori* eradication could cause the parallel disappearance of other microorganisms of the human microbioma, with the consequent exposure to other threats, a better approach in treatment would be needed. The recourse to antibiotics in the past seems to have not considered this aspect and has dramatically increased the resistance of the bacterium. Over the last years, several approaches have been used to develop effective *H. pylori* vaccines. Whole bacterial cell sonicates (first-generation vaccines) and individual *H. pylori* proteins (second-generation vaccines) have been used as antigens to

stimulate immunity in the host. Vaccination appears to be a promising approach, since it could allow to prevent colonization. More in general, a strong preliminary genetic analysis of host polymorphisms and interactions with bacterial variability is needed to identify high-risk sub-populations.

2- ***Structural Characterization of Hp1454, a novel secreted protein from H. pylori.***

2.1 Introduction

After the complete sequencing of *H. pylori* genomes of 26695 and J99 strains (Tomb et al., 1997; Alm et al.,1999), although several genes were easily identified and categorized into functional classes, a conspicuous amount of genetic products remained of uncertain classification. In fact, most proteins are still defined as hypothetical. Several aspects concerning *H. pylori* physiology, in particular its relationship with the host, still has to be further characterized in terms of molecular processes. As a result, in the past few years many research groups have moved their efforts towards the identification of new targets and the assignment of specific roles in an effort to "filling the gaps". In this respect, secreted proteins are interesting targets, since they potentially come in direct contact with the host surface. Hp1454 (strain 26695) is a totally unknown protein presenting an N-terminal signal peptide with a cleavage site between Ser18 and Ala19. The latter could indicate its affiliation to an *H. pylori* Sec system. The genome of strain 26695 is estimated to contain 517 putative signal sequences, among 1.590 orfs, indicating that these proteins could be secreted via the General Secretory Pathway (GSP). This figure may be an overestimated, since it is based on sequence similarities to a peptide motif derived from the characteristics of known signal peptides. Nevertheless, it indicates that a substantial portion of proteins synthesized by *H. pylori* are destined to cross at least the inner membrane (Ilver, Rappuoli, and Telford, 2001). Hp1454 has been identified first time during a comprehensive analysis of the secretome (Bumann et al. 2002), where it was found in the extracellular milieu. Hp1454 protein is up to now categorized as hypothetical and is totally uncharacterized. The corresponding genomic region does not present orthologues in other organisms. Despite its absence in others organisms, Hp1454 is well conserved among different pylori strains (100-98% identity), while it shows a lower grade of homology between *Helicobacter* species (*H. felis* homologue presents 35% identity with a coverage of 95%). Interestingly, Hp1454 displays a low similarity (28% identity on 40% of coverage) with some lipoproteins from others gram-negative bacteria, like *Nautilia profundicula*, an epsilon sulfur reducing proteobacteria, and the pathogenic *Francisella novicida*. Hp1454 is predicted to contain a short lipobox (KKIILAC) at the N-terminus according with the motif {DERK}(6)-[LIVMFWSTAG}(2)-[LIVMFYSTAGCQ]-[AGS]-C (Lippred server).

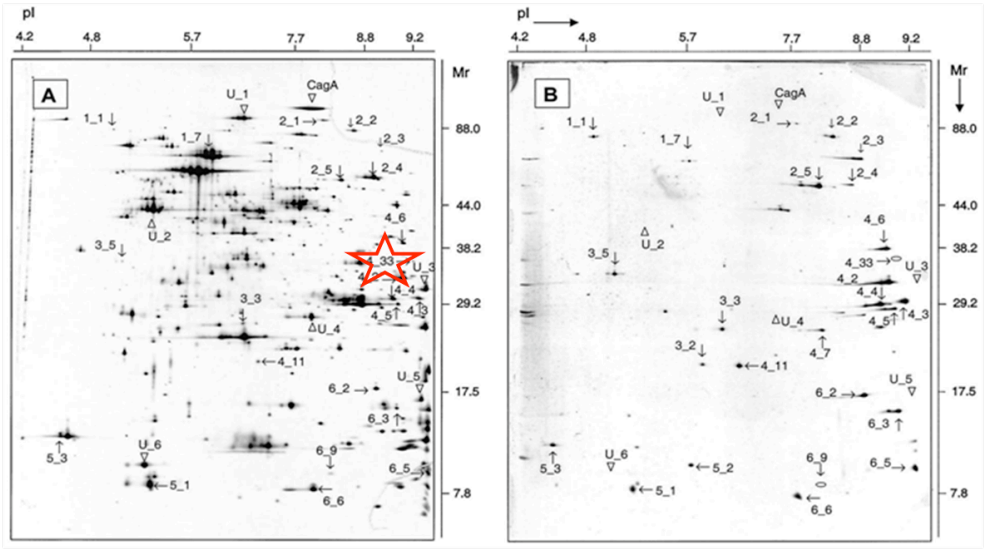


Fig2.1 2D-PAGE .Protein compositions of whole-cell proteins (A) and secreted proteins (B) of H. pylori 26695. Two-dimensional gel electrophoresis was performed with H. pylori harvested during exponential growth in liquid cultures (Adapted from Bumann et al., 2002) 4.33 spot has been identified as the product of Hp1454 gene.

```
HP1454/1-303          1 ----------MKKIILACLMAFVGANLSAEPKWYSKAYNKTN---- 32
jhp1347/1-303         1 ----------MKKIILACLVAFVGANLSAEPKWYSKAYNKTN---- 32
H._bilis_/1-299       1 ----MQKYLRSLCIFIYTLICIMLFGLHGCSHAKDTSFHKDK---- 38
H._felis/1-309        1 ----------MLTKSMRLALVFAPLLLGATPPAWYQNFNGN----- 31
lip_F.novicida/1-342  1 MRHLYFNLVSTIILVLATSSCSTGIPENSKNIKNQVVINQQSQKDT 46

HP1454/1-303         33 ---------------TQKGYLYGSGSATSKEASKQKALADLVASISV 64
jhp1347/1-303        33 ---------------AQKGYLYGSGSATSKEASKQKALADLVASISV 64
H._bilis_/1-299      39 ---------------TNKDYIYGEGSGPNFDIAKQNALQDLATNLQV 70
H._felis/1-309       32 ---------------PSAYFYGYGSGETKQAAKERALNDLASSIVV 62
lip_F.novicida/1-342 47 PNDTVPVWFSSGQPNTDQLLYGFGAAKSLEKATQKALADMVQKLQV 92

HP1454/1-303         65 VVNSQIHIQKSRVDNKLKSSDSQTINLKTDDLELNNVEIVNQEVQK 110
jhp1347/1-303        65 VVNSQIHIQKSRVDNKLKSSDSQTINLKTDDLELNNVEIVNQEAQK 110
H._bilis_/1-299      71 SIKYSTQQNINQKNDTLQTSGMTNTFLESKIKDIPSVEVEKTTKQG 116
H._felis/1-309       63 SVSSTTSSQTTRVNKTLDKRSSQNINLVVDSIKFVNVKVDKQECAQ 108
lip_F.novicida/1-342 93 TVSTTTNFTNVTTNDKVSQKLTQQITTTTTQITIPNYKVINQSELN 138

HP1454/1-303        111 GIYYTRVRINQNLFLQGLRDKYNALYGQFSTLMP---KVCKGVFLQ 153
jhp1347/1-303       111 GIYYTRVRINQNLFLQGLRDKYNALYGQFSTLMP---KVCKGVFLQ 153
H._bilis_/1-299     117 NKVNVRVRVSREILEGSIANRITQAQSELSTLLK----TCNQVSFS 158
H._felis/1-309      109 NICYTRLKLEKNMFLDTLAKRYNNLYNNLSALHPL--EGCRGIFLK 152
lip_F.novicida/1-342 139 QIFYVEVQTNKEQTINDLKELINSNINQAQQLLTTTNNKSSLYRFA 184

HP1454/1-303        154 QSKSMGDLLAKAMPIERILKAYSVPVGSLENYEKIYYQNAFKPKVQ 199
jhp1347/1-303       154 QSKSMGDLLAKAAPMERILKAYSVPVSSLENYEKIYYQNAFKPKVR 199
H._bilis_/1-299     159 QYKKFKEILHALKEDIVLYQALTKNMS-YGNIMLLEFQNSISTLPN 203
H._felis/1-309      153 EMQKLVVGVNKALPLQELLNAYGRNVASLTPYQNLLKENSPHPKAR 198
lip_F.novicida/1-342 185 IAQQVTKNIQLINSSLRTLIILVPDSNINSQIEILNQIDNELLNLK 230

HP1454/1-303        200 ITFDNNGDAEIKSALISAYARVLTPSDEEKLYQIKNEVFTDSANGI 245
jhp1347/1-303       200 IAFDDNSDTEIKNALMSAYARVLTPSDEEKLYQIKNEVFTDSANGI 245
H._bilis_/1-299     204 YALHWELDSLHSHKQEAQAILTAELSKFIKIDLKAKQTLHIKVSNN 249
H._felis/1-309      199 IIYAPGSDEEIADALTGQYARFVQQDDGPGLYSIENKIYQSSSNGQ 244
lip_F.novicida/1-342 231 RGLQIYVDKQNSGFFYNSLEKFLQVNNYNITDKKDLANININITLKLK 276

HP1454/1-303        246 TRIRVVVSASDCQGTPVLNRSLEVDEKNKNFAITRLQSLLYKELKD 291
jhp1347/1-303       246 TRIRVIISASDCQGTPVLNRSLEVDEKNKNFAITRLQSLLYKELKG 291
H._bilis_/1-299     250 NPLQFLLEFYDCKNNLESTIQINTHANQHDILQGSKKSRFGAIVYK 295
H._felis/1-309      245 FHIGLNLKIKNCQGNIIYSKLLSADQSSRSAAIERIKAQAFKQLRN 290
lip_F.novicida/1-342 277 DYDNKFNNNEYCIETKVELQALDDSSNQLSPKQYTIKVCSKQGRLA 322

HP1454/1-303        292 YANKEGQGNTGL--------                          303
jhp1347/1-303       292 YANKEGQGNTGL--------                          303
H._bilis_/1-299     296 AIDE----------------                          299
H._felis/1-309      291 YQQGVGAGTADVNSNKIDF-                          309
lip_F.novicida/1-342 323 AIDKAVEIFYSQLNDAESIY                         342
```

Fig. 2.2 Sequence alignment of Hp1454 protein. Sequences from H. pylori 26695 strain (Hp1454) and J99 strain (jhp1347) -100% of identity- are compared with sequences from two different Helicobacter spp. (H. bilis and H. felis) -respectively 51% coverage, 25%identity and 95% coverage, 35% identity- and with a putative lipoprotein from the bacterium F. novicida -40% coverage, 28% identity-. The % of total identity is indicated in shades of blue. Alignment and editing: Clustal & Jalview vers. 2.7 (Waterhouse et al., 2009)

2.2 Materials and Methods

The hp1454 gene cloned in our lab, belongs to the *H. pylori* CCUG17874 strain. The sequence below corresponds to the protein (after removal of the signal peptide) with the addition of a N-

terminal His-tag flanked by a cleavage sequence for TEV protease, coming from the construct codified by pET151 expression vector, in which the gene was cloned.

```
        -18        -8           3          13          23
MHHHHHHGKP IPNPLLGLDS TENLYFQ\GID PFTAEPKWYS  KAYNKTNTQK

        33         43         53         63         73
GYLYGSGSAT SKEASKQKAL ADLVASISVV VNSQIHIQKS RVDNKLKSSD

        83         93        103        113        123
SQTINLKTDD LELNNVEIVN QEAQKGIYYT RVRINQNLFL QGLRDKYNAL

       133        143        153        163        173
YGQFSTLMPK VCKGVFLQQS KNMGDLLAKA MPIERILKAY SVPVGSLENY

       183        193        203        213        223
EKIYYQNAFK PKVQITFDNN SDAEIKAALI SAYARVLTPS DEEKLYQIKN

       233        243        253        263        273
EVFTDSANGI TRIRVVVSAS DCQGTPVLNR SLEVDEKNKN FAITRLQSLL

       283      291
YKELKDYANK EGQGNTGL
```

\ corresponds to the cleavage site for TEV protease

Property	**Value**
Number of amino acids	291
Molecular Weigth	32.609
Theoretical pI	9,17
Abs (1g/l)	0,9

Table2.1 Some properties of Hp1454 referred to the construct after cleavage with TEV

2.2.1 Cloning

The coding sequence of Hp1454 was amplified from genomic *H. pylori* DNA (strain CCUG 17874) using the primers 5'-CACCGCAGAAGCCTAAGTGGTA-3'(forward) and 5'-TTATTATAACCCCGTATTGCC-3' (reverse). Primers were extended by the Phusion High-fidelity Dna polymerase (New England Biolabs) and the PCR product was then cloned into a pET151 vector

(Invitrogen) in frame with an N-terminal His-tag flanked by a TEV proteolysis site using a TOPO® Directional Cloning kit (Invitrogen).

2.2.3 Expression

E. coli BL21(DE3) competent cells were transformed with HP1454-pET151 plasmid by heat shock, than grown in a selective LB medium supplemented with 100 µg/mL ampicillin. A 3 liters culture was performed starting from a 25 ml overnight pre-inoculation. The cells were grown aerobically at 37°C until the absorption A_{600} reached value 0.6, at which point 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) was added to the medium to induce the expression. After an additional 4 h shacking at 30°C, cells were harvested by centrifugation. In order to produce the seleno-methionine derivative, the same transformation was carried out in *E. coli* B834(DE3), a methionine auxotroph strain. Cells were grown in a 2L minimal medium composed by M9, essential amino acids, 2% glucose, B1 vitamin, $CaCl_2$, trace of metals and seleno-methionines. The expression was performed at 28°C for 5 h after induction via IPTG.

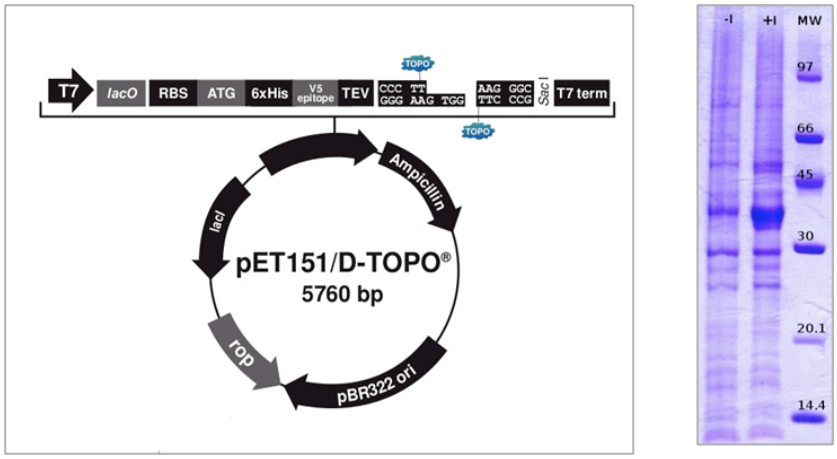


Fig. 2.3 Cloning and Expression of Hp1454. On the left: the expression plasmid pet151 (N-terminal His-tag) used. Right: SDS-PAGE of culture medium before (-I) and after(+I) induction via IPTG.

The bacterial pellet was resuspended in Buffer A (200 mM NaCl, 20 mM MOPS pH7.2, mM DTT 0.5), than a protease inhibitor cocktail tablet and DNase enzyme were added. The lysis was carried out by first treating the cells with 1 mg/ml lysozyme for 1h at 4°C, and then sonicating the extract 5 times for 45”, keeping the sample on ice. After sonication the cells were centrifuged at 25000g for 40' in order to separate the supernatant from the insoluble fraction.

2.4.3 Purification

As a first purification step, an IMAC -Immobilized Metal Affinity Chromatography- was performed. The soluble fraction was loaded onto a 1ml His-trap Ni-chelated column (GE Healtcare) equilibrated with Buffer A. After some washing steps using Buffer A supplemented with 10mM Imidazole, the protein was eluted in a gradient ranging from 20mM to 500mM imidazole. In the first purification trial of HP1454 the His-tag was removed by using TEV protease and the cleaved protein was recovered in the flow-through coming from a second affinity. The Imidazole concentration was diluted until a maximum of 10 mM before cutting the protein, to avoid any interference with proteolysis reaction. The protein was then concentrated by ultrafiltration until the desired volume was obtained. The gel filtration was performed on a Sephacryl S-200 prepgrade 16/60 size-exclusion column, using as eluent a Buffer containing 200mM NaCl, 30mM MES, pH 6. The protein eluted as single pick corresponding to the monomer and fractions were collected, concentrated to 9 mg/ml and stored at -80°C for crystallization trials. Since stability slightly decreases after removing the tag, some aliquots of the His-tag protein were also stored.

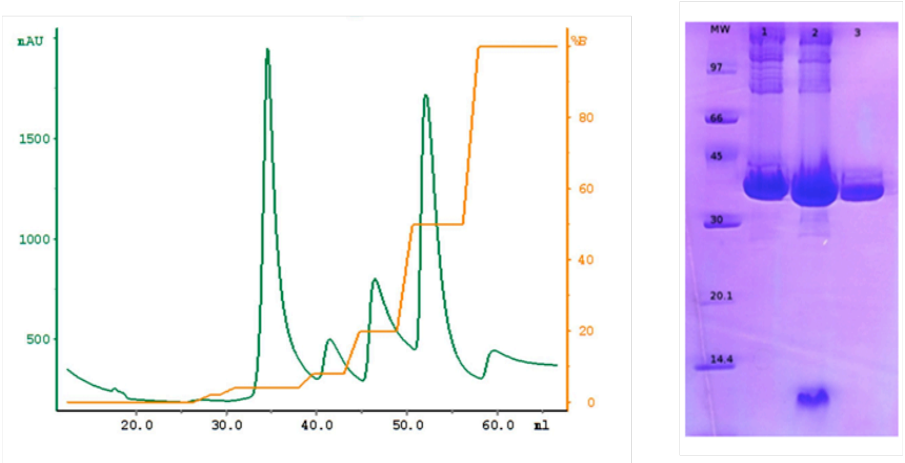


Fig.2.4 Left: Imac Affinity and elution of Hp1454 in imidazole steps gradient (0-500 mM) Right:SDS-PAGE Hp1454 after the affinity step; 1) His-tagHp1454; 2) TEV protease cleavage; 3) HP1454

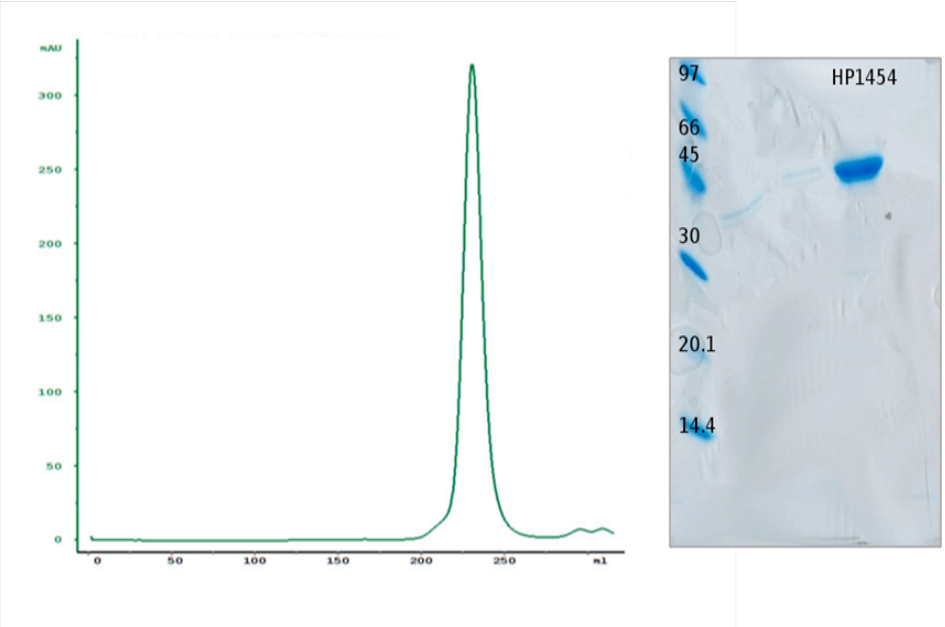


Fig.2.5 Left: Elution profile of Hp1454. Right: SDS-PAGE of Hp1454 at the end of the purification

2.4.4 UV-CD Analysis

Some preliminary information about Hp1454 secondary structure was obtained using Circular Dichroism analysis. CD measurements of the protein diluted to 1 mg/ml were performed using a JASCO-J715 Spectropolarimeter, in a 0.05 cm path length cell. The spectra in the far-UV (190-260 nm) were recorded at a scanning speed of 50 nm/min. Spectra were accumulated and averaged, followed by baseline correction by subtraction of the buffer. Mean residue weight ellipticity was calculated and expressed in units of degree cm^2 $dmol^{-1}$. The circular dichroism spectrum was deconvoluted using software program CDNN version 2.1 (Bohm, Muhr & Jaenicke, 1992). The spectrum deconvolution shows that secondary structure is predicted to contain mainly α-helix (more than 55%) while the content of beta sheet and random is 25% and 20% respectively.

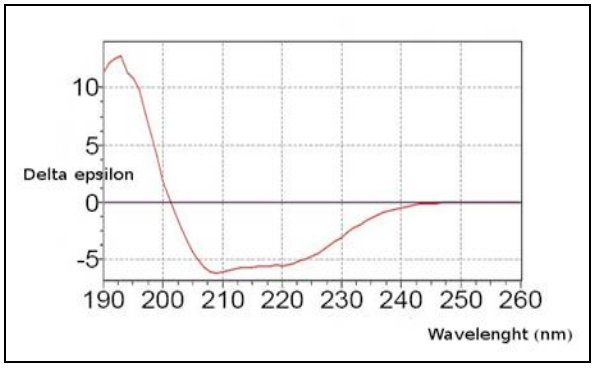


Fig. 2.6 CD Spectrum of Hp1454.

2.4.5 HPLC and Mass Spectroscopy

In order to check the integrity of the purified protein after His-tag cleavage, a portion of the protein used for crystallization trials was submitted to an additional analytical separation by high-pressure liquid chromatography (HPLC) and the mass determined by MALDI-TOF Mass Spectrometry. A 100 µl sample corresponding to 3.5 mg/ml was injected into the column and the resulting profile corresponded to a single species. Mass Spectroscopy analysis confirmed the purity and homogeneity of the sample through the identification of a prevalent type A (MW 32616.18 Da), corresponding to the whole Hp1454 (MW 32.609, 291 amino acids) in agreement with that expected from the sequencing of the construct. However, MS results showed the presence in solution of another very weak species B, which seems to correspond to the protein added of 2-4 additional residues (MW= 32.943.35). This contribution has been considered an irrelevant impurity, due to the weekness of the signal.

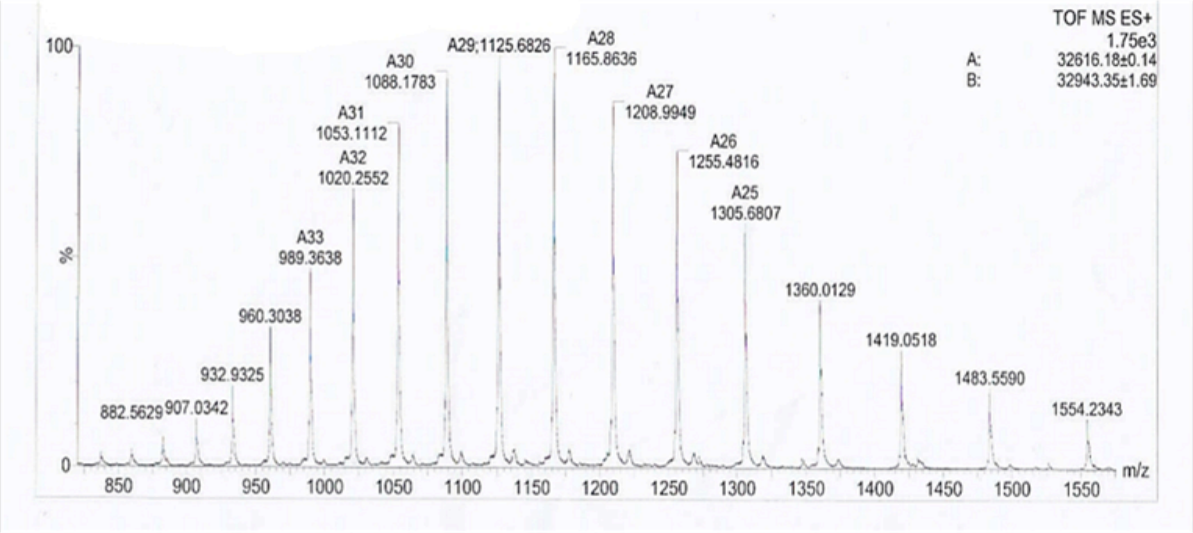


Fig2.7 Mass calculated by MALDI-TOF MS. Type A weight (32616.18 Da) corresponds to the expected for Hp1454 after His-tag cleavage.

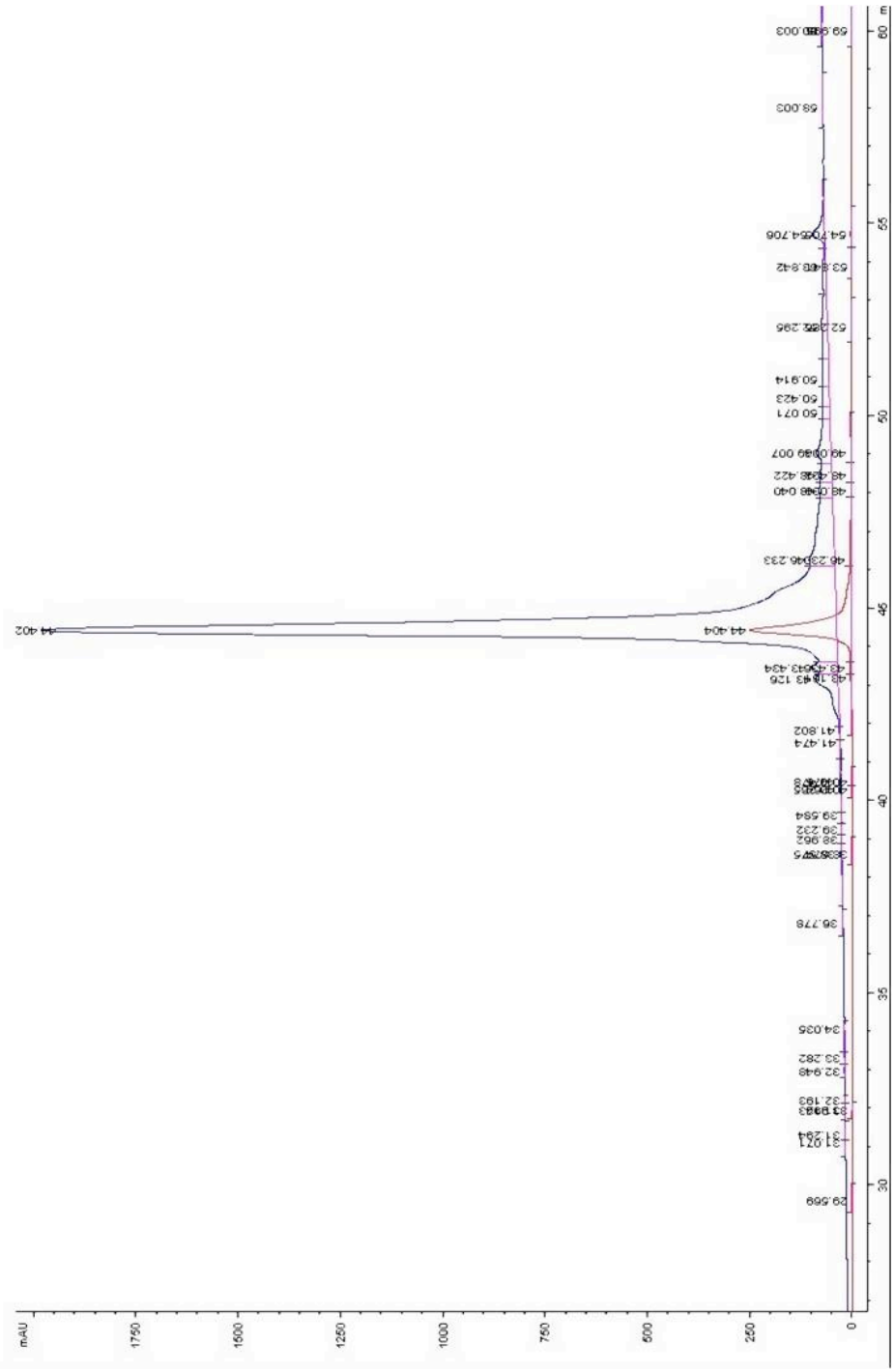


Fig2.8 HPLC profile elution for Hp1454

2.4.6 Bridge or not bridge?

Hp1454 sequence contains two cysteines at positions 135 and 245. In order to correctly interpret the electron density map, it has been very helpful to test the propensity of these residues in forming an intermolecular (or intramolecular) disulfide bridge. The intermolecular bridge option was examined by performing an SDS-PAGE in reducing and non-reducing conditions (in the presence or absence of 1 mM DTT). As shown in Fig. 2.9, in the presence of a reducing agent the apparent molecular mass is about 37,000 Da (lane B), in its absence 35,000 Da (lane A). They both correspond to a monomer, whose theoretical value is 32,609 Da. The fact that in the absence of DTT the protein migrates with a slightly lower apparent molecular mass could be explained in terms of conformational change imposed to the molecule by the reducing agent. The gel definitely exclude that an intermolecular bridge is present between two monomers, in agreement with gel-filtration data, that indicate the presence of a single monomeric species. In order to confirm the presence of an intramolecular disulfide bond, a titration of free thiols with the Elman's reagent was performed (Elman 1959). The test is based on the ability of DTNB 5,5'-Dithio-bis(2-nitrobenzoic acid) to bind free thiols in proteins, forming a disulfide bond and releasing a thiolate ion which is colored whit a maximal absorbance at 412 nm. In order to build a calibration line, thiols from DTT were used as standard and reactions were prepared with increasing concentrations of substrate from 10 µM to 10 mM.

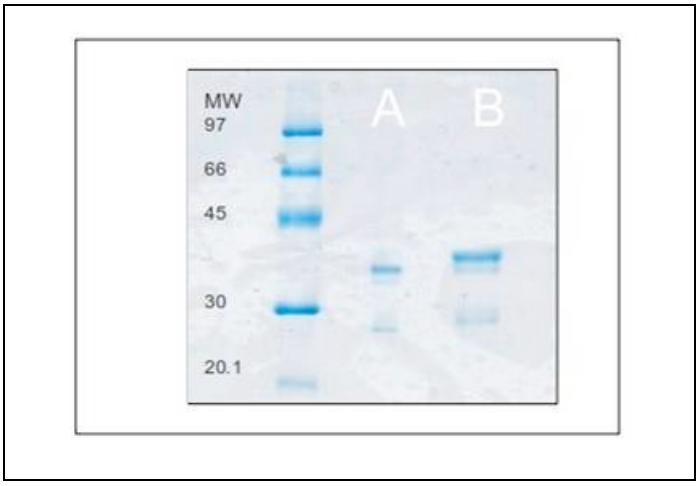


Fig2.9 SDS-PAGE from HP1454 crystals in A) non-reducing and B) reducing conditions.

In the case of a substrate like DTT, which has two free thiols/molecule, the molar ratio between the concentration of the colored product and substrate source of thiols is around 2, as expected. Every reaction was prepared in 1ml cuvette mixing the reagents at 4°C: 50µl of DTNB solution (50mM sodium acetate, 2mM DTNB), 100 µl Tris 1M pH 8 were mixed with 10 µl of

increasing concentrations of substrate. Water was added to fill the 1ml cuvette. After incubation at 37°C for 5 minutes, the A_{412} was measured and the molar ratio between product and substrate calculated. Since Hp1454 initial concentration was quite low for this experiment, the protein was concentrated until 1mM and the final concentration was measured by reading the absorbance A_{280}. The molar ratio DTNB/HP1454 was calculated for increasing concentrations of HP1454, ranging from 0.2 mM to 1 mM. The number of free thiols groups was calculated in two distinct reaction sets, by using as substrate the protein in native conditions and the protein denatured after incubation with 6 M GndHCl at 37°C.

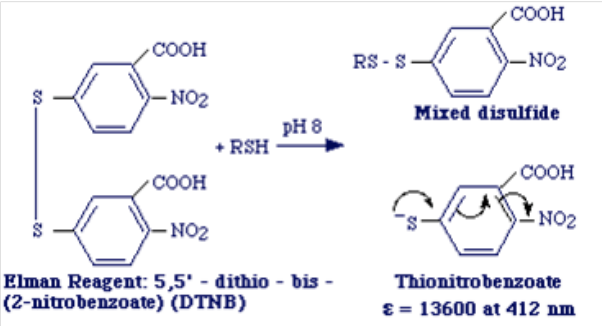


Fig 2.10 DTNB reaction. Elman Reagent reacts with a SH group forming a mixed disulfide and the colored product Thionitrobenzoate, which has an absorption peak at 412 nm.

As expected (and confirmed by the calculated calibration line), the molar ratio between the colored thiolate ion (product) and DTT (substrate) is ~2 (fig. 2.11A). In fact, SH groups present in each DTT molecule are 2. The same ratio is expected from Hp1454, if Cys135 and Cys245 are reduced and not involved in a S-S bridge. The average of experimental molar ratio calculated using native Hp1454 as substrate was about 0.2 as showed in fig.2.11B. This value strongly indicates the presence of an intra-molecular bridge. However, a limited fraction of Hp1454 molecules without S-S bridge seems to be present in solution, since a colored product is formed, even if at very small concentration. The molar ratio calculated for reactions with Hp1454 denatured slightly increased at the value of 0.6/0.7.

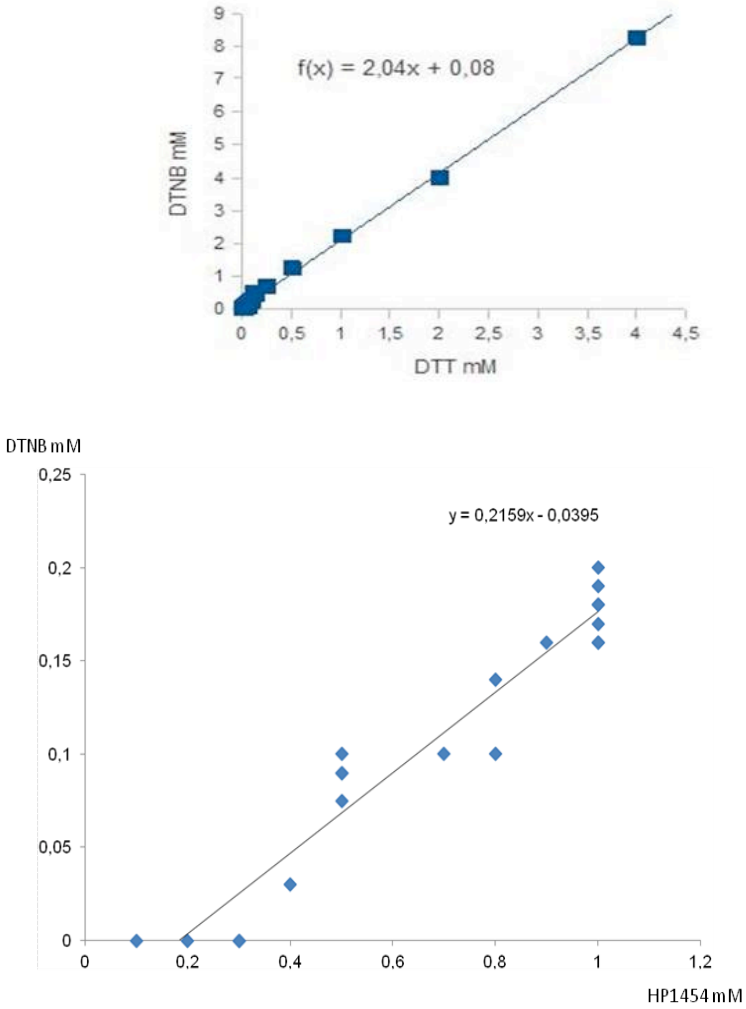


Fig.2.11 A (top) Calibration Line showing the molar ratio (~2) between product (DTNB) and substrate (DTT). DTNB concentration was calculated measuring absorption peak at 412 nm. B (bottom) Ratio of molar concentrations between product (DTNB) and substrate (Hp1454 in native condition). The value is ~ 0.2

2.4.7 Study of protein-protein Interactions

Interactions have been inferred by querying several platforms dedicated to the exploration of protein pathways available in the network, like PIMRider® , String 9.0, and IRefIndex. IRefIndex consolidates protein interactions data from a number of primary interaction databases including BIND, BioGRID, CORUM, DIP, HPRD, IntAct, MINT, MPact, MPPI and OPHID (Razick et al., 2008). Part of these databases are integrated as well in the String 9.0 Platform, which additionally provides an in-house prediction collection. STRING quantitatively integrates interaction data from different sources (genomic context, yeast two hybrid system, previous knowledge) for a large number of organisms, and transfers information between these organisms where applicable (Szklarczyc et al., 2011). Finally PIMRider® use as source its ultimate Y2H platform Hybrigenics.

Particularly, Pimrider platform contains a section dedicted to *H. pylori* where 1,200 interactions connecting 46.6% of the proteome are available. (Rain et al., 2001). These interactions have been identified by screening 261 *H. pylori* bait proteins against a highly complex library using Hybrigenics high-throughput proprietary yeast two-hybrid-based technology (Rain et al., 2001).

Fig. 2. 12 Scheme of the yeast two hybrid system screens for interacting proteins based on trascriptional activation. The protein of interest is fused to a DNA-binding domain while proteins that bind to "bait" are fused to a transcription activation domain. Any protein that binds to the bait, "the fish", will activate the transcription of a reporter gene.

2.4.8 Crystallization

Crystallization trials were carried out using the vapor diffusion technique whit an Oryx8 crystallization device (Douglas Instruments). Different precipitant solutions from several crystallization kits (Structure Screen I and II by Molecular Dimensions, PACT Suite, JCSG Suite, Anions and Cations Suites, PEGs and PEGs II, AmSO4, MPD Suite, MbClass I and II Suites by Qiagen) were screened by using the sitting drop technique. Dilutions of precipitant solutions at different percentages were also tested in order to find the metastable zone. In all cases (native, heavy atoms derivatives and Se-Met derivatives, see below), crystals were obtained in the same crystallization condition, corresponding to number 5 of the SSI kit (Molecular Dimensions): 0.1 M Na acetate, 2.0 M ammonium sulfate, pH 4.6. Crystals were grown at the temperature of 291K. Se-Met crystals were obtained only by microseeding from crystals of the native protein. The quality of crystals was improved by performing microseeding and using small Hp1454 crystals as seeds to grow bigger crystals in the same precipitation condition. Crystals were crushed by pipetting up and down several times and the resulting suspension was added to an Eppendorf tube containing 100 µl of the precipitant solution. The tube was centrifuged for five minutes at about 100g. A series of dilutions from 1 to 10^{-4} seed solutions were performed and new drops with different concentrations of seeds were prepared. Some drops were prepared by adding H_2O_2 inside the reservoir in order to test any difference in the crystallization process by increasing the oxidative power of the precipitant and to evaluate the grade of improvement in diffraction analysis.

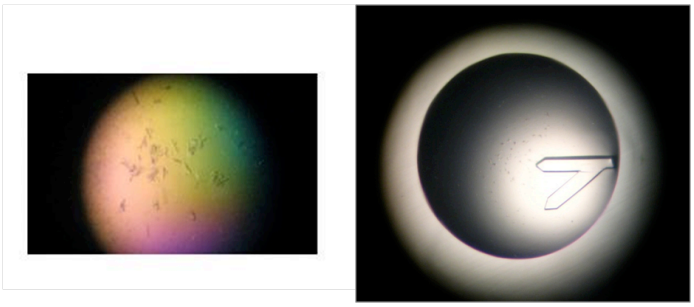


Fig. 2. 13 Crystals of native Hp1454 before and after improvement by microseeding. Crystals were grown as large rods ~0.2 mm long.

2.4.9 MIR and SAD

Fig2. 14 Crystals obtained through the co-crystallization of Hp1454 with Ytterbium chloride hydrate 10mM. (Best Resolution 2.9 Å)

Potential derivatives crystals were obtained through the co-crystallization of Hp1454 with heavy atoms in order to obtain phase information by the multiple isomorphous replacement method (MIR). Co-crystallization was apparently successful with different heavy atoms, which were added into the drop before crystallization at the concentration of 10mM. Among several heavy atoms tested, data were collected from crystals formed in the presence of Ytterbium(III) chloride hydrate, Gadolinium(III) chloride hydrate, Potassium hexachlororhenate(IV), Samarium(III) chloride hexahydrate and Europium(III) nitrate hexahydrate. To obtain initial phases from single anomalous diffraction (SAD), selenomethionine derivatized protein was expressed, purified and crystallized. Se-Met derivatives were obtained from two different targets. The first was HP1454, the recombinant protein expressed from CCUG strain, while the second was the homologue protein (jhp1347) coming from J99 strain. The second form apparently seemed particularly useful, since the amino acid sequence has four methionines instead of the three present in the original target. Crystals were produced from both of these, but definitely just those corresponding to *H. pylori* CCUG strain were found to diffract at a useful resolution.

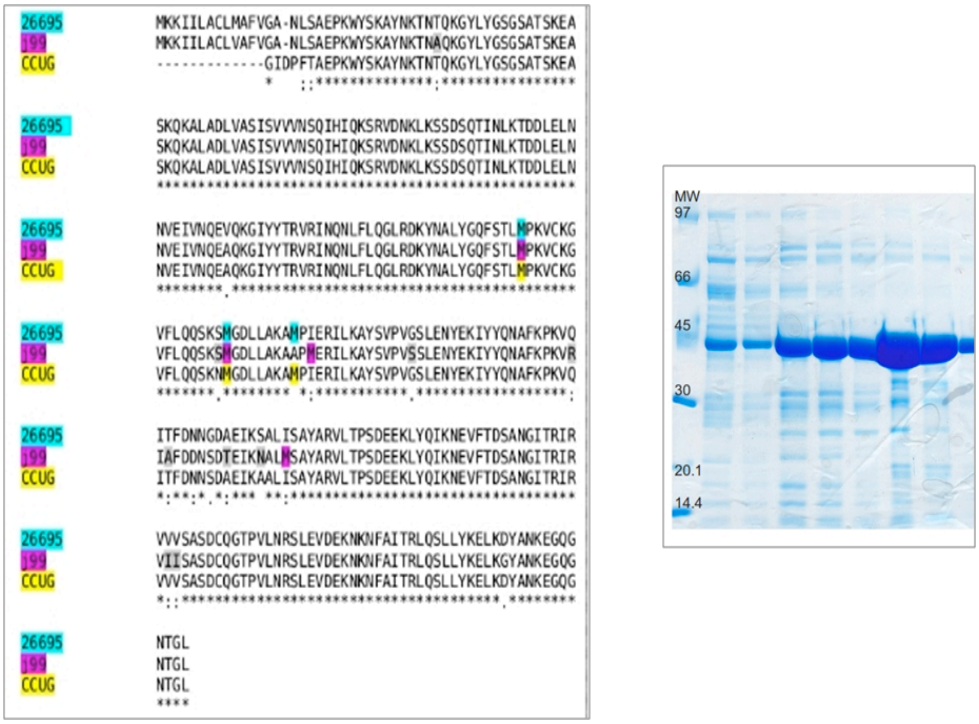


Fig.2. 15 Left: Sequence alignment for Hp1454 from three different strains: 26695, J99, and CCUG. 4 Methionines are present in the protein from j99, strain (3 in the protein from CCUG strain) Right: SDS-PAGE of jhp1347 purification

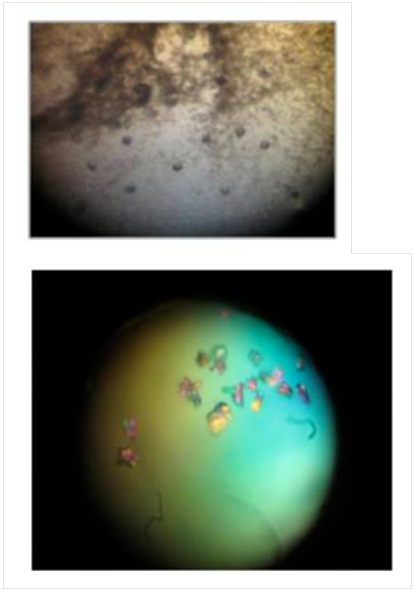


Fig.2. 16 Crystals from SE-Met jp1347 (top) vs Crystals from Se-Met Hp1454 (bottom).

2.4.10 Data collection

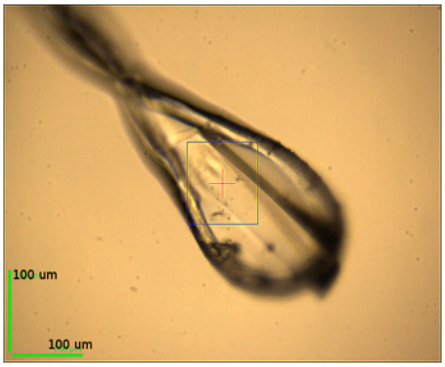


Fig.2. 17 Hp1454 crystal mounted in a loop

Just before data collection, crystals were soaked for few seconds with ethylene glycol as cryoprotectant. The soaked crystals were then mounted on a cryo-loop and immediately frozen under a nitrogen stream at 100 K (fig. 2.17). The best diffraction data of the native protein were collected at the beamline ID23-1 of the European Synchrotron Radiation Facility (ESRF) in Grenoble. 400 frames of 1° oscillation each were measured. Crystals of derivatives (Se-met and heavy atoms) were measured at the BM14 and ID14-4 of ESRF, respectively. All the data sets were integrated using MOSFLM (Leslie 1992) and merged with SCALA (Collaborative Computational Project 1994) as implemented in CCP4 (Collaborative Computational Project, Number 4, 1994).

Hp1454 crystals belong apparently to the orthorhombic space group $C222_1$, with cell parameters a = 73.90 Å, b = 101.50 Å, c = 121.68 Å. Assuming a molecular mass of 32,609 Da and 1 monomer in the asymmetric unit, this corresponds to a VM of 3.44 $Å^3$/Da and a percentage of solvent of about 67%. For the SAD experiment, data at the wavelengths corresponding to Se peak (Fig.2.18) were collected with an oscillation angle of 1° per image from five crystals. Data-collection statistics for native and Se-Met are shown in Table 2.2. To determine the positions of the three selenium atoms and to calculate the phases, the SHARP/autoSHARP suite was used (Bricogne et al. 2003).

Data set	Native	Anomalous
Beamline	ID23	BM14
Wavelengths	0.9835	0.9782
Space Group	$C222_1$	$C222_1$
Unit Cell parameters(Å)	a = 73.90 b = 101.50 c = 121.68	a=73.29 b=101.75 c=120.42
Resolution (Å)	61.17-2.9 (**3.06-2.9**)	59.47-3.2(**3.37-3.2**)
Tot. No. of reflections	463534	126462
No. of unique reflections	10354(1487)	7737(1099)
Multiplicity	9(9.1)	16.2 (16.7)
Anomalous Multiplicity		8.8(8.8)
Completeness(%)	99.1 (99.7)	100(100)
Anomalous completeness		100(100)
Rmerge	0.1 (0.46)	0.11(0.62)
I/σI	0.92(0.99)	5.6(1.1)

Table 2.2 Data-collection and processing statistics relative to Hp1454. Values in parentheses are for the highest resolution shell.

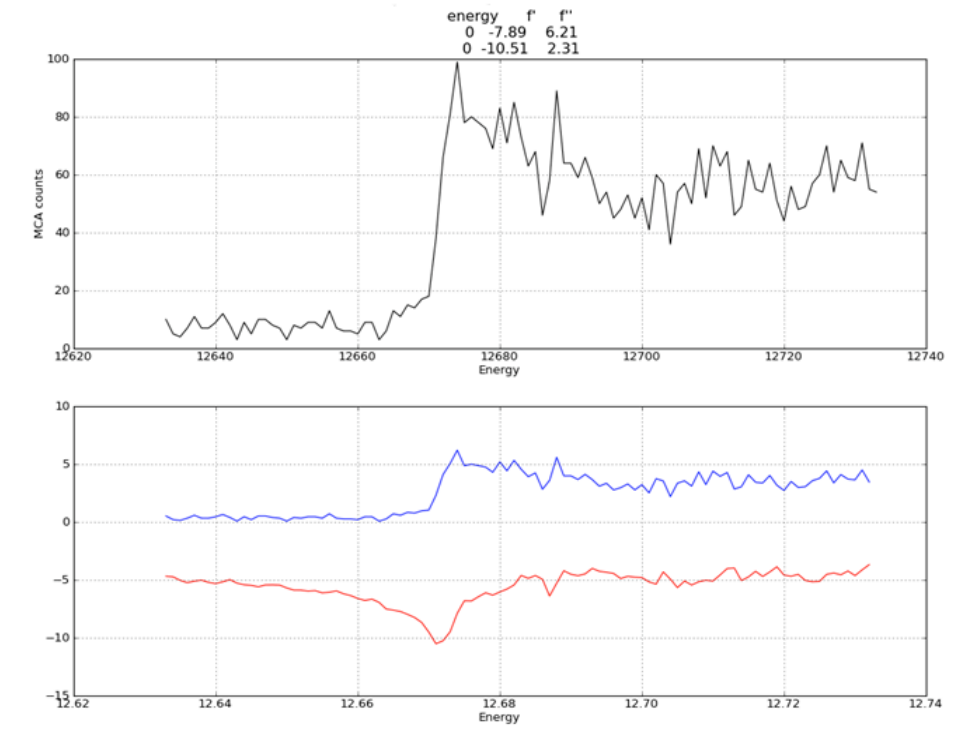


Fig 2.18 A fluorescence scan was performed to verify the presence and location of the absorption edge for Se-Met Hp1454 crystals. The X-ray Energy is indicated in eV. The maximum value of the f" anomalous coefficient (blue), corresponding to the peak wavelength, was measured (6.21); f' (red) was calculated from f" using the Kramers-Kronig equation.

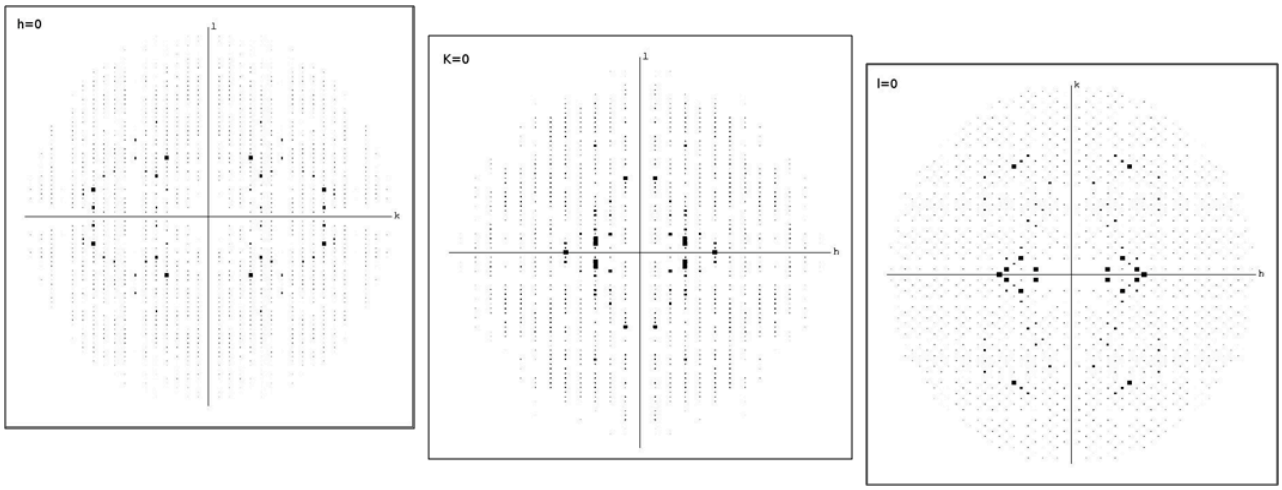


Fig. 2.19 Reciprocal space planes of HP1454 crystal processed as $C222_1$: 0kl, h0l, and hk0

2.5 Results and Discussion

2.5.1 Structure of Hp1454

Hp1454 from *H. pylori* CCUG strain was successfully expressed, purified, and crystallized. Despite their relatively large size, obtained by microseeding, crystals showed a modest diffracting power. This may be ascribed to the loose packing of the protein in the crystal cell, which leaves a large amount of empty space, ~ 65% of the volume, filled with solvent. Phase information was obtained through SAD experiments, while MIR experiments turned out insufficient in order to calculate the electron density map (very low occupancy). Phases, after application of density modification procedures, were good enough to allow a preliminary tracing of the polypeptide chain for most of the monomer. In particular, α-helices were clearly visible, specially for the first two domains (see below). The connection between the second and the third domain was on the contrary not clear, and the electron density for the third domain was less defined. The structure was solved in the $C222_1$ orthorhombic space group, with one monomer in the asymmetric unit, consistent with the characterization in solution. However, the interpretation of a part of the monomer - the last 80 C-terminal residues - proved to be very hard, due to the weakness of the electron density in this portion. Hp1454 model consists in fact of three-domain. The N-terminal

domain (domain I, from 21 to 110) is composed of 3 β-sheets (labeled A1, A2 and A3) with antiparallel orientation, one long α-helix (residues 34-58) and a long stretch (residues 60-90) connecting the helix to β-strand A2. Residues from 1 to 20 and from 58 to 75, are not visibile in the elctron density map and they are not included in the model. The central domain (domain II, residues 111-182) is mostly alpha-helical: 3 α-helices are arranged in an anti-parallel way. The interpretation of the C- terminal domain (domain III) presents several uncertainty: it is organized in a four-stranded anti-parallel β-sheet flanked by two α-helices (Fig. 2.20). Whilst the position of every amino acid in domains I and II is well defined, side chains of domain III are hard to identify. In addition, the crystal packing relative to domain III is quite unusual, since the only contact between two asymmetric units is confined to two β-strands, one for each monomer, belonging to domain III. Unfortunately the symmetry operation relating these two strands is a two-fold axis oriented nearly parallel to the strands. This hinders the formation of correct H-bond interactions between the two strands.

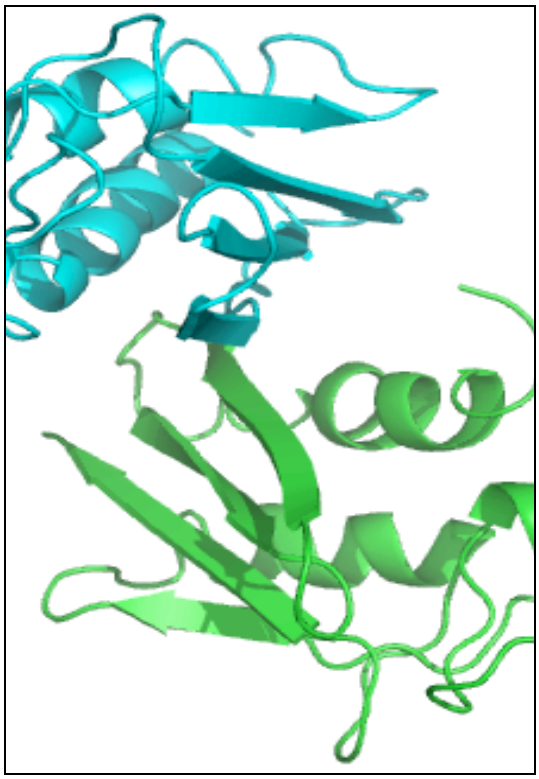


Fig.2.20 A cartoon view of the Hp1454 C-terminal domain (domain III) consisting of four-stranded anti-parallel β-sheet flanked by two α-helices. A local two-fold axis oriented nearly parallel to the two β-strands of each monomer has been observed.

Since all these problems seem to be relative only to domain III, whilst the electron density of the remaining largest portion of the model is well defined, we concluded that the orientation of domain III could be flexible with respect to the other two domains. This is also justified by the quite loose crystal contacts in this area. A possible hypothesis is that domain III does not respect the orthoromibic symmetry of the rest of the molecule and we tested the possibility that the correct space group has a lower symmetry. Data were re-processed as C2, with two monomers in the asymmetric unit (Table 2.3). Statistics are only sligtly better, in particular for the Rmerge, which drops from 0.10 for the native data set to 0.091 (Table 2.2 and 2.3). Phases were also re-calculated in the new space group.

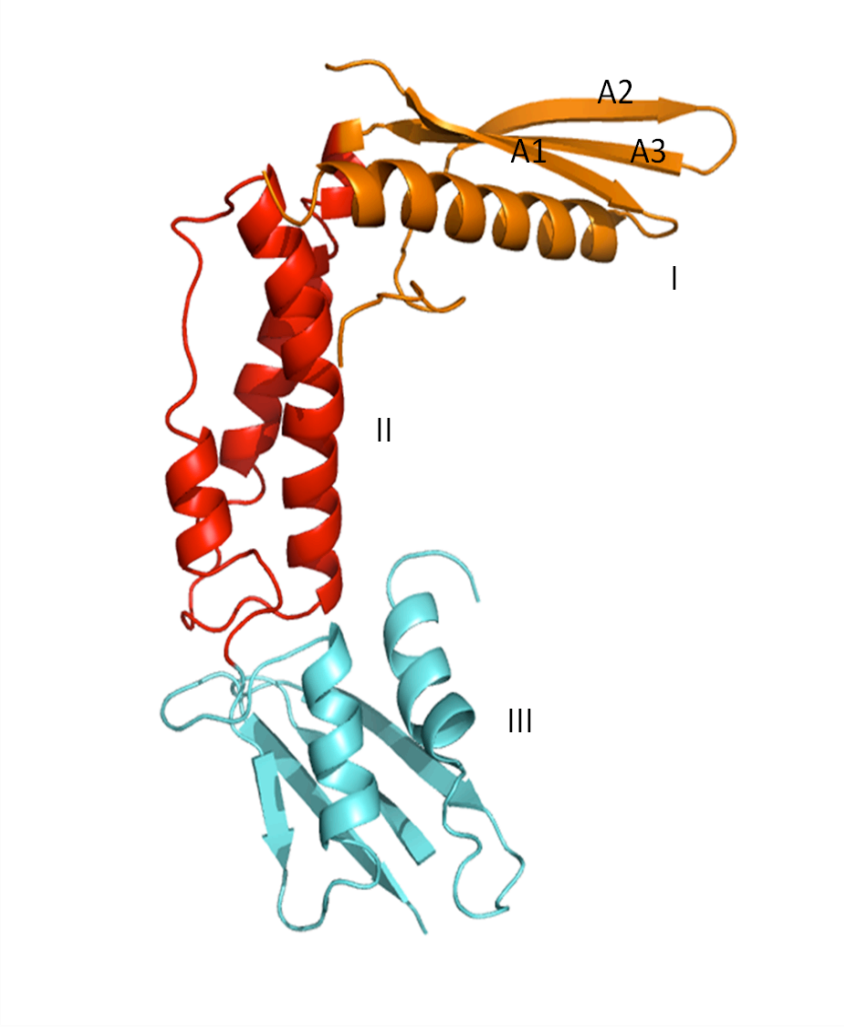


Fig. 2.21 Cartoon model of HP1454. The Three domains are shown: The N-terminal domain (domain I, orange), composed of A1, A2 and A3 β-sheets with antiparallel orientation, one long α-helix and a stretch connecting the helix to A2, the central domain (domain II, red) mostly alpha-helical (3 α-helices arranged in an anti-parallel way), and the flexible C terminal domain (domain III , blue) consisting in four-stranded anti-parallel β-sheet flanked by two α-helices.

The new maps are similar for domains I and II and perhaps slightly better for domain III. The refinement is not fully completed, and statistics regarding the current model are reported in Table 2.4. Form the chemical point of view, the advantage in using space group C2 is that now the two β-strands involved in the inter-molecular contacts in the crystal are not any more constrained by the crystallographic symmetry. A cartoon view of the two monomers present in the asymmetric unit is illustrated in Fig. 2.22.

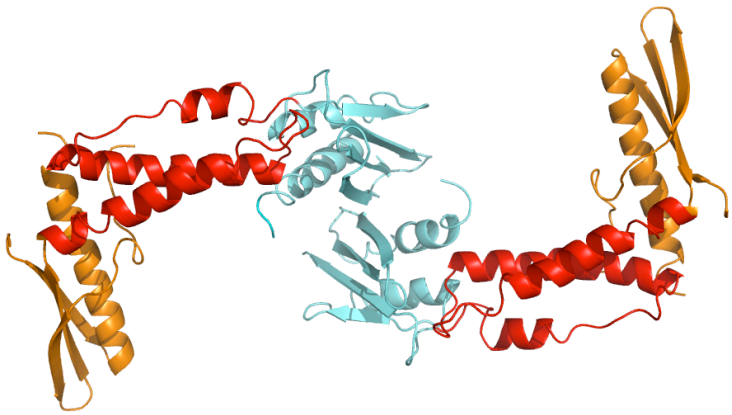


Fig. 2. 22 A cartoon view of the asymmetric unit in the monoclinic space group containing two monomers. According with the C2 Space group symmetry, the overall Hp1454 architecture is substantially the same with respect to the model built in the $C222_1$ space group while some uncertainty about the flexible C-terminal domain building was solved .

Space group	C2	C2
***a*, *b*, *c* (Å), β(°)**	73.54 102.07 121.64, 90.25	73.911 102.10 122.337 90.20
Resolution (Å)	60.82-2.90 (3.06-2.90) *	61.17 (2.98-2.83)
R_{sym} or R_{merge}	0.091 (0.452)	0.104 (0.571)
$\langle I / \sigma I \rangle$	7.3 (2.0)	9.8 (2.5)
Completeness (%)	98.8 (98.9)	97.0 (98.7)
Redundancy	2.9 (2.8)	5.6 (5.4)
No. reflections	19734 (2873)	21101 (3143)

Table 2.3 Data in the second column refer to 1 crystal, in the third column data from two crystals were merged together and the resolution was cut to a slightly higher value.

Rwork / Rfree	0.266 / 0.360
No. protein atoms	3662
R.m.s. deviations	
Bond lengths (Å)	0.008
Bond angles (°)	1.54
Ramachandran plot (%)	
Favored or allowed	87.3
Generously allowed	8.4
Disallowed	4.2
G-factor	0.2

Table 2.4 Refinement statistics for the model solved in the C2 space group.

2.5.2 Possible roles of Hp1454.

In complete absence of sequence homology, succeeding in assigning a physiological function to Hp1454 is not a simple task. We can at least try to speculate about a possible role for the protein looking at its interactions with other proteins**.** In Fig.2.23, the results from predicted or experimentally determined interactions are shown. Querying String 9.0 with a medium confidence score (0.4), the program identified ten possible interactions. The network map decreased to five interactions by setting a higher confidence score (0.7) (Fig.2.23 ,bottom). Among these five interactions, three are predicted to occur with products of nearby genes, like HP1455 and HP1457, both hypothetical proteins and HP1456. The latter has been identified as Lpp20, a membrane-associated lipoprotein with a strong antigenic activity (Keenan et al., 2000). Any putative

conserved domain has been detected from Hp1455 blast, while HP1457 is classified as a lipoprotein and shows a TolB amino-terminal domain homology. Other possible interactions (String9.0) have been found with HP1230, which corresponds to HobA (PDB code 2UVP), a protein implicated in the organization of the first step of DnaA oligomerization and attachment to oriC during the initiation of replication, and HP0592 a type III restriction enzyme R protein.

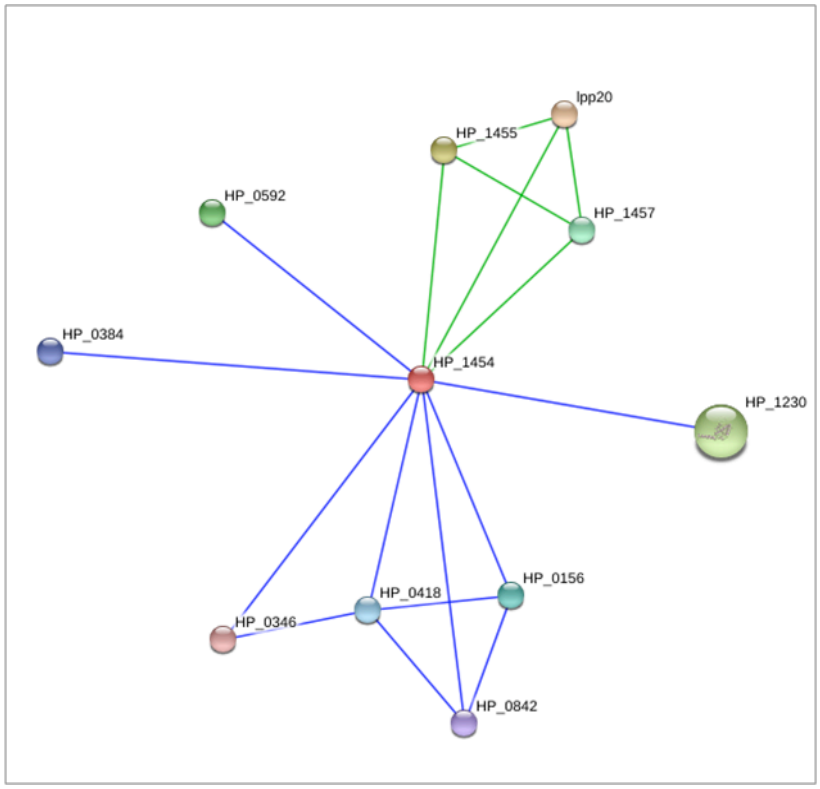
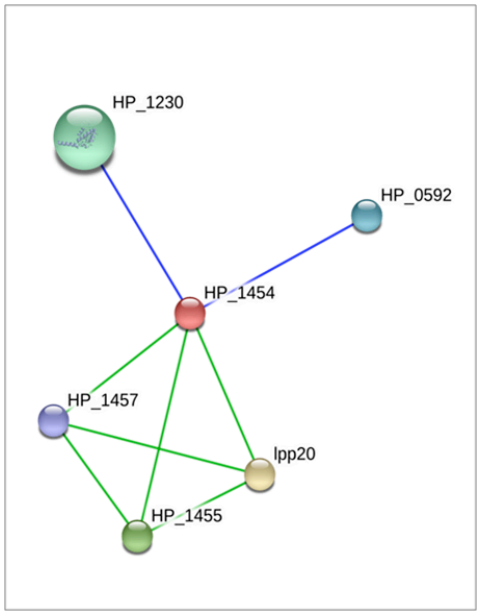


Fig.2.23 String 9.0 interactions map for Hp1454. Green connections: neighborhood; Blue connections: co-occurrence. TOP: Prediction score range 0.8 /0.4: 10 interactions detected (0.150 low confidence, 0.9 highest confidence) BOTTOM: Prediction score range 0.8/0.7: 5 interactions detected (Jensen et al. 2009).



Both these proteins are cytoplasmic and, even though a relationship has been predicted, a functional connection is unlikely. The other two platforms queried, PIMrider and IrefIndex, have found two additional interactions. The first is a strong connection with Hp0241 (E value 0.7/1), an hypothetical 131 amino acids protein well conserved inside Helicobacter genus and without putative conserved domains. A second, weaker interaction (E value 0.3/1) was found with fliS (Hp0753), a flagellar export chaperone, crucial in preventing the premature polymerization of flagellins essential in the formation of functional flagella. FliS in addition to 2 flagellin subunits (FlaA and FlaB), was found to interact with about 50 different proteins, including hook-associated proteins, basal-body proteins, and motor switch proteins. All components of this complex interactome work together in the assembly of the flagellum from the cytoplasmic membrane through the periplasmic space to the outer membrane (Lam et al., 2010). Despite the apparent lack of interaction with other components of the complex, the idea that Hp1454 could take part in this functional network has been taken into account. To test this hypothesis, knockout experiments of *H. pylori* mutants are in progress at Munich University by Prof. Wolfgang Fischer.

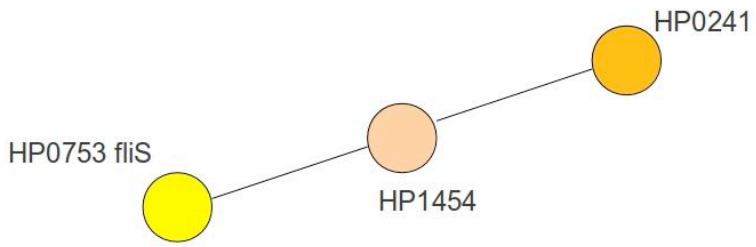


Fig. 2.24 Interactions found by Pimrider and IrefIndex platforms. Hp1454-fliS (E value=0.3) interaction is predicted as weak, while Hp0241-HP1454 (E value 0.7)seems a stronger connection.

2.5.3 A new gene-cluster

Looking at the chromosomal locus hosting the hp1454 gene, we find a region transcriptionally poor characterized. This area was found to correspond to an operon which includes the genes hp1454, hp1455, hp1456, and hp1457 (source : MicrobesOnline Regulon Prediction, Price et al., 2005). In agreement with the interactions previously found, this additional evidence allowed us to speculate about a possible functional role for Hp1454. In fact since hp1454 gene is transcribed

together with three others genes and their respective products have been found to interact, this is apparently the most promising clue to follow.

Fig. 2.25 The gene-cluster that contains hp1454 consists of three additional components: hp1455 (not characterized yet), hp1456 (lpp20), and hp1457 that presents sequence homology with the N-terminal domain of the Tol-B family.

The downstream presence inside the cluster of the hp1457 gene, that encodes a 210 amino acids protein containing a region (4-203) of homology with the amino-terminal domain of TolB proteins family, led us to investigate the role of TolB proteins in bacteria. TolB is an essential periplasmic component of the Tol-PAL (peptidoglycan-associated lipoprotein) translocation system present in many gram-negative bacteria where it plays a role in cell-envelope integrity. An important role in group A colicin translocation has also been described (Lazzaroni et al., 1999) In *E. coli*, unlike *H. pylori*, the Tol–PAL system has been characterized in detail. It consists of several proteins encoded by two adjacent operons in which genes are positioned in the following order: (ybgC–tolQ–tolR–tolA)(tolB–pal–ybgF). Their transcription mainly occurs from two promoters located upstream of the 5' ends of genes ybgC and tolB, although the presence of a long ybgC–ybgF transcript has been observed (Vianney et al., 1996; Muller & Webster, 1997). However, the Tol–PAL gene cluster is not universal and its components display a low level of similarity between different gram negative bacteria. Proteins encoded by these operons form two complexes. One is located in the cytoplasmic membrane and involves the TolA, TolQ and TolR proteins interacting with each other by their transmembrane segments. The second complex, associated with the outer membrane, is composed of TolB in the periplasm and Pal in the outer membrane. *E. coli* Tol-PAL cluster includes YbgC and YbgF genes as well; YbgC is a cytoplasmic protein with thioesterase activity (Zhuang et al., 2002; Gully and Bouveret, 2006), while YbgF is a protein of unknown function located in the periplasm .In the cytoplasmic membrane complex, TolA is composed of three structural domains: the transmembrane TolA I domain (amino acids 1-47) , which anchors the protein in the cytoplasmic membrane, and two globular periplasmic domains TolA II (48–310) and TolA III (311–421). These three domains are separated by segments rich in glycine residues, which provide a certain level of flexibility (Levengood et al., 1991). TolA I domain interacts with three transmembrane domains of TolQ (TolQI – amino acids 9–36, TolQII 127–159, and TolQIII 162–191), which additionally presents two large cytoplasmic loops, while TolR has a N-terminal

anchor and most of its residues are localized in the periplasm. At the second complex, Pal is anchored in the OM and interacts with the peptidoglycan and the C-terminal domain of the periplasmic TolB, which possesses a β-propeller structure (PDB code 1CRZ) (Bouveret et al., 1995; Abergel et al.1999; Ray et al., 2000). In addition, Pal interacts with TolA and other cell envelope proteins such as OmpA and Lpp. The N-terminal domain of Tol B possess a five-stranded mixed β-sheet structure that sandwiches two major α-helices and it was found to interact with TolA domain in the periplasm. The two Tol–Pal complexes interact with each other through the periplasm via TolB and the C-terminal domain of TolA. Actually TolB can be considered as an adapter protein that coordinates the association of OM/Pal with the peptidoglycan and the IM.

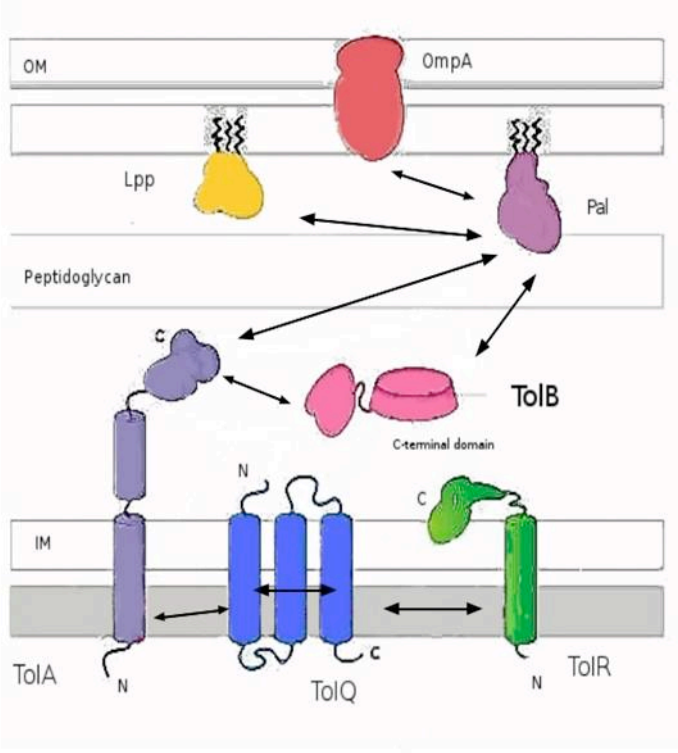


Fig. 2.26 Location and interactions of Tol–Pal proteins in the cell envelope of Gram-negative bacteria (E. coli). The various interactions are indicated by arrows (Adapted from Godlewska et al.,2009)

Homologues of the Tol-Pal proteins have been found in *H. pylori* strain 26695 (Angelini et al., 2008). However in this case a cluster comparable to that of *E. coli* is not present. In fact, while Tol B (hp1126) and Pal (hp1125) have been recognized as well as TolQ (hp1113) and TolR (hp1129), the identification of TolA and YbgF remain still uncertain, owing to the lack of an evident homology. Additionally, YbgC gene has been found to be located in a totally different region of the chromosome (hp0496) with respect to the others components. This alternative organization of the genetic information in *H. pylori* could offer new openings of interpretation.We are tempted to

speculate that Hp1454 could be a component of a *H. pylori* Tol-pal system or an equivalent system due to its co-occurrence with a tolB homologue protein (Hp1457) and the lipoprotein Lpp (Hp1456). This scenario would require the localization for Hp1454 inside the periplasmic space or in the proximity of the inner or outer bacterial membrane. The presence of a lipobox at the N-terminus of the sequence and the homology -weak indeed- with some lipoproteins, seem to be consistent with this prediction. The only known physiological role of the Tol proteins is to maintain the integrity of the outer membrane. Therefore these proteins are likely to be involved in the biogenesis of some outer membrane components (Lazzaroni et al., 1999). One potential role of the Tol proteins could be to participate in the transport of an outer membrane component crucial for its stability, such as LPS, Lpp or Pal. This synthesis should need the participation of several actors that could act in coordination with TolA and Tol B. The lipoprotein Hp1454 could be a component of this periplasmic network. The previous idea will be used to direct the next experiments on *H. pylori* mutants. Phenotype of mutants lacking the hp1454 gene will be tested, looking for possible damages in terms of cell envelope integrity.

3- Expression and Purification of Fur, the Ferric Uptake Regulator from H. pylori.

3.1 Iron Requirement

Iron is one of the most abundant elements on the planet and is a basic requirement for life (Gilbreath et al., 2011). Despite its abundance in nature, it is not easily available to living organisms because, in the presence of oxygen, the oxidized ferric form (Fe^{3+}), which is very insoluble, is the prevalent one. Therefore, bacteria have developed high-affinity iron uptake systems in order to acquire the iron necessary for growth (Tomb et al., 1997). *H. pylori* genome suggests the presence of several uptake systems for iron acquisition in both ferrous and ferric forms. However, under anaerobic or acidic conditions, iron is found in the more physiologically accessible form Fe^{2+}. In bacteria this element plays a critical role in processes such as metabolism, electron transport, oxidative stress responses and regulation of genes expression (Gilbreath et al., 2011). The physiological role of iron inside organisms is intimately connected to oxidative metabolic pathways. Enzymes necessary for the oxidative metabolism of the tricarboxylic acid cycle, like aconitase, and the cytochromes and non-heme iron electron carriers of the electron transport chain require iron for their activity. Since glycolytic enzymes do not require iron, aerobic and facultative organisms can overcome any deficiency of iron by shifting their metabolism in favor of glycolysis rather than the tri-carboxylic acid cycle. *H. pylori* is a microaerophilic organism. It requires oxygen to survive, but at a lower level than that present in the atmosphere. The metabolism of pyruvate reflects the microaerophilic character of this organism (Tomb et al., 1997). Despite the reduced need for oxygen, which is consistent with the environment chosen to establish the infection, the stomach, *H. pylori* cannot survive without iron. Iron in fact participates in electron transfer reactions and it is present in the active site of enzymes in coordination with oxygen, sulfur and nitrogen ligands. Iron-containing proteins are mainly involved in basic cell metabolism. Bacterial respiration and peroxide reduction need heme-containing enzymes, while iron-sulfur proteins are important mediators of electron transport, anaerobic respiration, amino acids metabolism and energy metabolism. Finally iron containing non-heme proteins are involved in DNA synthesis and protection from superoxide (Van Vliet, Bereswill, and Kusters , 2001). In bacteria, iron possesses a regulatory role as well. In particular, pathogenic bacteria depend on iron availability to modulate their interaction with the host. One condition that has been clearly demonstrated to regulate bacterial toxins synthesis is, together with other important variables like nutrients availability, ions concentration and especially iron concentration. Bacteria seem to respond to low iron concentration by increasing expression of virulence factors that may promote

their growth and survival and ultimately cause damage to the host (Payne et al. 2006). In *H. pylori,* cytotoxin VacA and several outer membrane proteins exhibit changes in expression in response to varying concentration of iron (Merrel et al., 2003). Particularly, VacA synthesis is negatively regulated by iron concentration.

3.2 Iron homeostasis, uptake and storage

Whereas iron is essential for growth of most pathogens, free iron is toxic. In fact it is known that Fe^{2+} catalyzes free radical formation through Fenton reaction by producing hydroxyl radicals, that results in cellular damage. Therefore iron homeostasis between iron uptake and storage is crucial in maintaining a normal activity and to avoid ferric toxic effects. The nature and number of iron acquisition systems in bacteria usually reflect the type of environment in which they live (Gilbreath et al., 2011). In fact, depending on the type of iron source, bacteria develop different and specialized uptake systems. In the human stomach iron is available as ferrous ion (Fe^{2+}), and coordinated to organic molecules such as lactoferrin, hemoglobin, transferrin and heme groups. As with other pathogens, *H. pylori* require an iron-scavenging system for survival in the host (Tomb et al., 1997). The major ferrous iron transport protein known in *H. pylori* is FeoB. Although the mechanism of ferrous iron uptake through FeoB remains unclear, this transporter is required for colonization in a mouse model of infection. FeoB is present in *E. coli* as well, together whit two others auxiliary components, FeoA and FeoC, that are lacking in *H. pylori*. As seen in *E. coli* FeoB, the N-terminal region of *H. pylori* FeoB contains two regions that show similarity to nucleotide-binding domains, thus suggesting a possible ATPase activity (Gilbreath et al. 2011). Furthermore, iron uptake could be mediated by a FecA ferric citrate uptake system, since *H. pylori* genome carries three putative fecA genes (fecA1, fecA2, and fecA3) homologues to iron transport proteins in *E. coli.* However, *H. pylori* looks unsupplied with others components of the system (FecBCDE) and the specific function of fecA gene remains to be elucidated. Inside *Helicobacter* species, different iron uptake systems have been described, according with different nutritional niches they occupy. Nongastric *Helicobacter* spp, such as *Helicobacter cinaedi*, *Helicobacter fennelliae*, and *Helicobacter bilis,* inhabit nutritionally competitive niches and have developed an iron uptake system based on siderophores biosynthesis. Siderophores are small organic molecules that specifically bind, solubilize and aid the delivery of ferric iron in bacteria (Gilbreath et al., 2011). In contrast, a mechanism for acquiring ferric ion through siderophores has never been described in *Helicobacter pylori,* as well as others gastric pathogens like *H. felis* and *H. mustelae.* This aspect is

interesting and could be explained considering the presence in the stomach of the prevalent reduced ferric ion (Fe^{2+}), and the lacking of a strong competition for nutrients from other bacterial species that require the ability to steal sources by a specialized capture system. Instead, gastric *Helicobacter* spp. have evolved the ability to acquire iron from their only major competitor, the mammalian host (Gilbreath et al., 2011). In mammalian hosts, extracellular iron is stored inside chelating molecules such as heme, transferrin and lactoferrin. Lactoferrin is the major iron-chelating molecule present at mucosal surfaces, even if the ultimate iron carrier is transferrin. During colonization of the gastric mucosa *H. pylori* utilizes host lactoferrin and transferrin, which represent a readily available iron source for the bacterium. *H. pylori* is able to use iron bound to heme group as well. The mechanism of binding remains still unclear, even if three iron-repressible outer membrane proteins ($IROMP_S$) that specifically could bind heme have been described (Worst et al. 1999). Interestingly, *H. pylori* expresses an heme oxygenase homolog, HugZ, thus suggesting an ability to utilize heme in terms of metabolic process. Another important iron source is hemoglobin: to utilize it, *H. pylori* expresses a specific receptor, FrpB2, similar to ChuA. The latter is an hemoglobin-binding protein present in *E. coli*. There are evidences that the bacterium is able to use hemoglobin directly. This is yet another proof of the long adaptation of *H. pylori* to its environment. In the absence of dedicated export systems, in order to maintain an adequate intracellular iron concentration avoiding both accumulation and shortage, *H. pylori* as well as others bacteria developed a system to store its iron content. The major *H. pylori* iron storage protein is the ferritin Pfr, a non-heme protein regulated in response to iron. The importance of Pfr is illustrated by the fact that this iron storage protein is necessary for survival during iron starvation and it is required for efficient colonization in an animal model of infection (Gilbreath et al., 2011).

3.3 Maintaining iron balance: Fur

In bacteria both iron acquisition and storage are governed at the higher level of the regulation of genes transcription. Fur, the ferric uptake regulator, is a global transcription factor responsible for this primary regulatory checkpoint. In numerous bacterial species, Fur binds to promoter regions of iron-regulated genes at level of conserved regions named fur-box. The binding to fur-boxes is iron-dependent and target genes, which can be both repressed and induced, are involved in iron homeostasis or storage. Hp-Fur belongs to the large Fur and Fur-like family composed of metal uptake regulators that control the homeostasis of iron (Fur from *E. coli* was the first

member described), zinc (Zur), nickel (Nur) and manganese (Mur) (Dian et al., 2011). A similar architecture is detectable in all these proteins, although they widely differ in metal selectivity and biological functions. It consists of a N-terminal DNA-binding domain (about 80 residues) containing a winged-helix motif and a C-terminal dimerization domain (about 70 residues) which are connected by a hinge loop (Dian et al., 2011). It was found that *Hp*-Fur regulates more than 50 genes in response to iron (Ernst et al., 2005; Danielli et al. 2006) and it is considered to play an important role in the ability of *H. pylori* to colonize model animals (Bury-Moné et al., 2004). Iron-bound Fur repression in *H. pylori* is similar to that in other bacteria like *E. coli*: Fur is thought to bind to fur-boxes as a dimer and to repress transcription in an iron-dependent manner. In *H.pylori* Fur dimerization is ferric dependent, since it has been demonstrated that the ion bound in S1 is crucial to form and stabilize the dimer. Under conditions of high concentrations of iron, Fur binds to the promoters of iron uptake genes resulting in an inhibition of transcription. However, unlike Fur in other organisms, *H. pylori* Fur has also been shown to act as an apo-repressor (Gilbreath et al., 2011). Currently the apo-fur regulon is predicted to contain 16 genes targets, among which only sodB and pfr have been definitively shown to be directly regulated. In iron-depleted conditions, apo-fur binds to iron-storage genes blocking their transcription. Iron homeostasis seems further regulated in *H. pylori*, as shown by the presence of a mechanism of auto-regulation of Hp-Fur transcription. In fact fur gene is transcribed from its own promoter, which has three distinct fur boxes. Fe^{2+}-Fur bound to box I and box II under iron-repleted conditions or iron excess, respectively, result in a repression of fur expression. In the absence of iron or under iron-restricted conditions, apo-Fur binds to box III, activating fur transcription.

3.4 Hp-Fur features

Hp-Fur is a small protein of 17681 Da and 150 amino acids. It corresponds to hp1027 gene in *H. pylori* strain 26695 and it shares a high sequence homology (between 59%-49% similarity and 37%-23% identity) with other Fur proteins: it share 32.4% and 28.9% identity and 70% similarity with *E. coli* Fur (Ec-Fur) and P. *aeruginosa* Fur (Pa-Fur) respectively. The first X-ray structure of a Fur protein (from *P. aeruginosa*) has been reported in a zinc-substituted form (Pohl et al.2003).

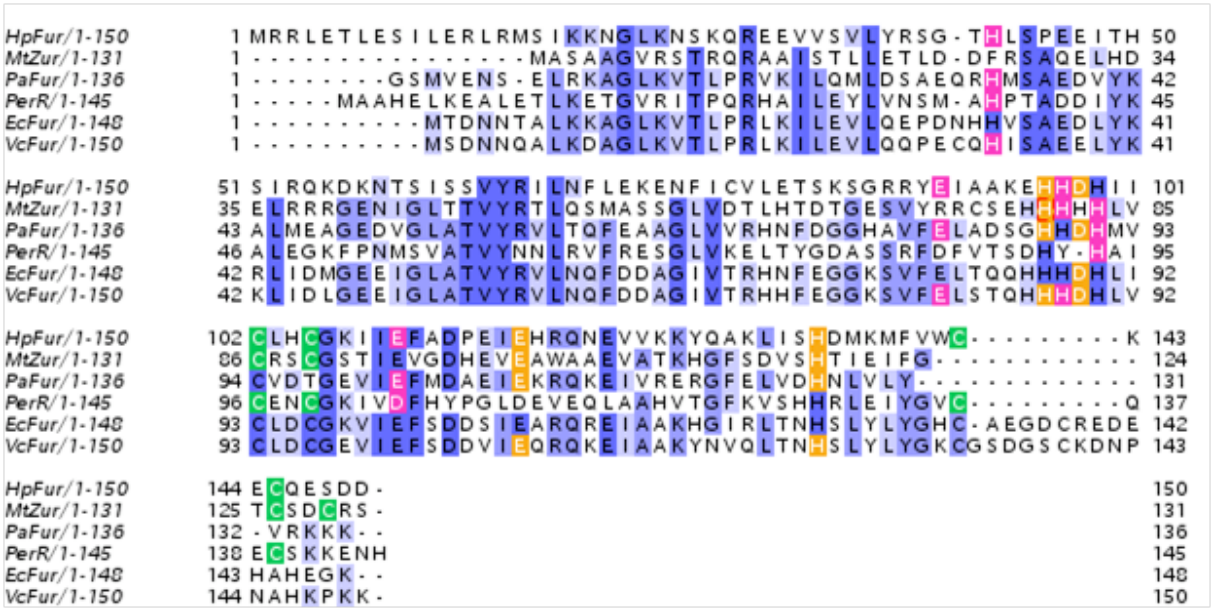


Fig. 3.1 Sequence alignment of Fur and Fur-like proteins. The amino acids sequences of Fur proteins from H. pylori 26695(Hp-Fur), P. aeruginosa(PaFur), E. coli (EcFur), V. cholerae (VcFur) and of the Fur-like Zur from M. tuberculosis (MtZur), and B. subtilis PerR (PerR) were aligned using Clustal. The % of identity is indicated in shades of blue; residues involved in S1 (or equivalent) in green, S2 in pink, S3 in orange. Alignment editing: Jalview vers. 2.7 (Waterhouse et al., 2009).

The original goal of this work was trying to solve the structure of Hp-Fur by X-ray crystallography. Unfortunately, the work has been stopped after the deposition of the structure of HPfur2M (PDB 2XIG) in January 2011. Those that will be presented here are mostly preliminary purification studies.

3.5 Materials and Methods

3.5.1 Expression and Lysis

E. coli BL21(DE3) competent cells were transformed with a plasmid pet22b containing the Fur coding sequence from *H. pylori* G27 strain (*NdeI BamHI double restriction*). The plasmid was kindly sent by A. Danielli from the University of Bologna. After transformation, cells were grown in a selective LB medium supplemented with 100 µg/mL ampicillin. An overnight pre-inoculation was the starting point for a 1L culture performed at 37°C until the absorbance A_{600} reached the value 0.7. Before the induction with 1 mM IPTG, in the first expression trial, 0.1 mM $ZnSO_4$ was added to the medium in order to possibly stabilize the dimeric state. The overexpression was performed at 27°C for 4 hours. Cells were harvested from the LB by centrifugation and subsequently resuspended in a lysis Buffer (100mM NaCl, 20mM Hepes pH 7.4 and 5 mM DTT) at 4°C and disrupted by French Press. In order to avoid any DNA interactions, not only a protease inhibitor

cocktail tablet but DNases as well were added to the lysis buffer. In order to separate the soluble fraction from the pellet, a centrifugation at 25000g for 40' was performed.

3.5.2 Purification

The crude protein mixture was loaded on a DEAE column for anionic exchange equilibrated with Buffer A (10mM Hepes pH 8.5, 50 mM NaCl, 5 mM DTT). Before the injection the sample was been diluted 1:2 in order to decrease salt concentration and the pH adjusted to 8.5. The protein was eluted on a salt gradient from 0 to 60% with Buffer B (10mM Hepes pH8.5, 1 M NaCl, 5mM DTT). The fractions containing HpFur were concentrated and further purified on a gel filtration column (Superdex 75 prep grade HR16/60; GE Healthcare) equilibrated with 100mM NaCl, 20mM Hepes pH 7.4, 3mM DTT buffer. Since an additional purification step was necessary, fractions were again concentrated and injected in a Superdex 200 HiLoad 26/60 GE Healthcare.

3.5.3 One-step Purification and Crystallization trials

A different purification step was performed in order to reduce the amount of unspecific aggregation found in the previous trials (Pellicer et al. 2009). Cell lysis was carried out using buffer A (100mM NaH_2PO_4 , 10 mM Tris pH 8 and 2MGdnHCl) in the presence of protease inhibitors. After centrifugation at 25000g for 40', the resulting supernatant was applied to a Chelating Sepharose Fast Flow column charged with $ZnSO_4$ and equilibrated with the same buffer A. Two subsequent washing steps were performed, first with 5 volumes of 0.5 M $(NH_4)_2SO_4$ in buffer A and then with 35 mM glycine in buffer A. HpFur was eluted on a gradient of Imidazole from 0% to 60% with Buffer B (100mM NaH_2PO_4 , 10 mM Tris pH 8, 2MGdnHCl, 500mM Imidazole). In order to remove GdnHCl and imidazole, the purified protein was dialyzed against 20 mM Acetate, 100mM NaCl buffer at pH 4.5. A second purification step was performed on the soluble portion after an unexpected precipitation during the dialysis step (column Superdex 200 HiLoad 26/60 GE Healthcare). The purified protein was concentrated to a final concentration of 10 mg/ml and stored at-80°C. Crystallization trials have been performed using the vapor diffusion technique and an Oryx8 Robot (Douglas Instruments). Different crystallization conditions (Structure Screen I and II by Molecular Dimensions, PACT Suite, JCSG Suite) have been tested on drops of 0.6 µl volume.

3.6 Results and Discussion

Two different approaches have been put in place in order to purify Hp-Fur for crystallization purposes. The first attempt was aimed at obtaining the protein in the dimeric state. For this

reason, purification was performed in reducing conditions and in the presence of 0.1 mM $ZnS0_4$, added in the medium during the expression of the recombinant protein. After anionic exchange, fractions containing HpFur were easy detectable, but they presented a strong contamination from others species. In order to eliminate the noise, purification was extended with different gel filtration steps that, even so, did not allow the recovery of the protein at a sufficient level of purity.

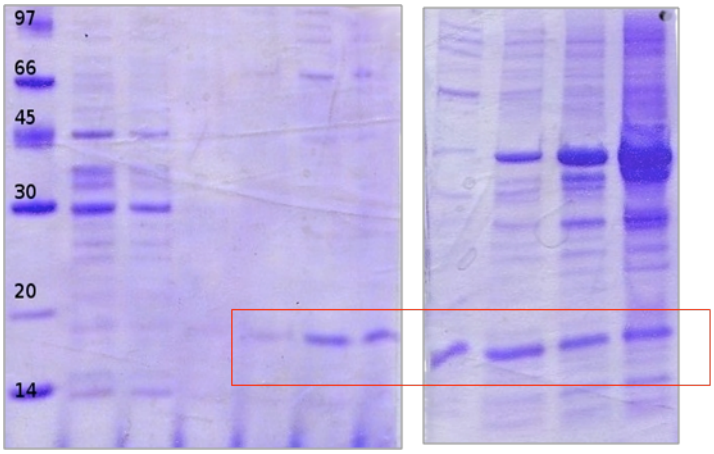


Fig.3.2 SDS-Page results: HP-Fur is present in many bands. However, bands corresponding to higher [NaCl] of gradient look strongly contaminated.

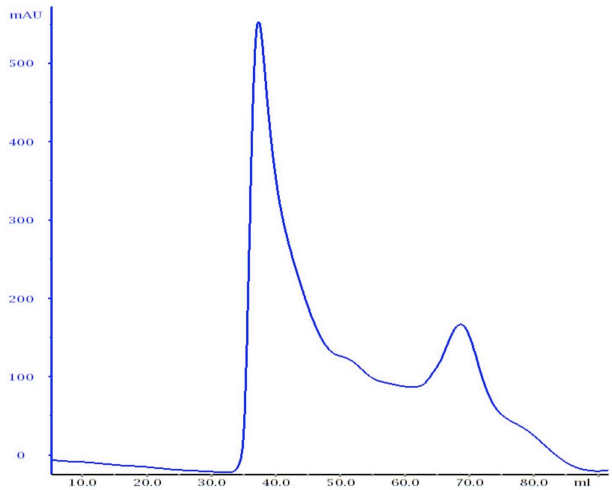


Fig. 3.3A

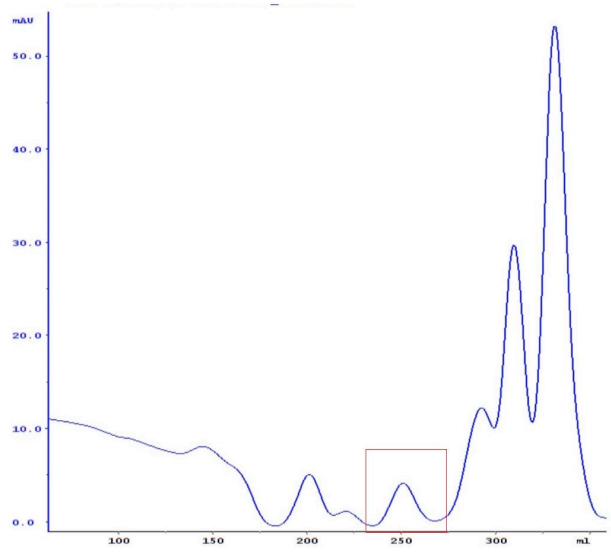


Fig. 3.3B

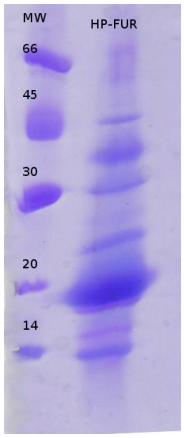


Fig. 3.3C

Fig.3.3 Two steps of Gel filtration: Superdex 75 prepgrade HR16/60 (A) and Superdex 200 HiLoad 26/60 (B). Hp-Fur was detectable in the peak at about 250mL of retention V (highlighted).(C) SDS-PAGE of HP-fur at the end of purification. Unspecific Aggregation has been reported.

The second purification strategy adopted was based on the use of chaotropic agents under mild denaturing conditions, according to a previous protocol (Pellicer et al. 2009). In solution Hp-Fur showed a strong tendency to form oligomers, due to the presence of six cysteines and the occurrence of hydrophobic interactions, which lead to unspecific aggregation. The problem of solubility was solved by adding GdnHCl 2M and 24 mg of almost pure protein were purified in this unique step.

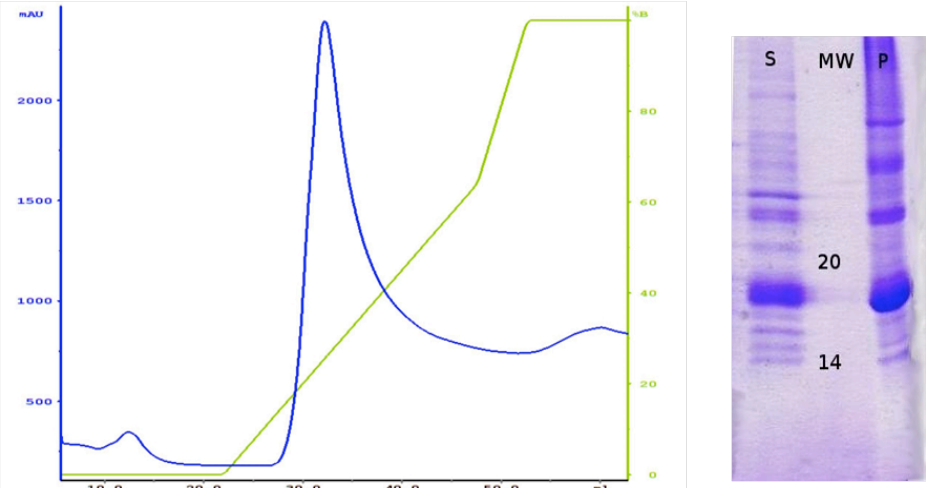


Fig. 3.4 Left: One step purification:IMAC affinity (Zinc-chelated column . HpFur was eluted on a gradient of Imidazole (until 300mM) in presence of GdnHCl. Right: SDS-PAGE Hp-Fur after dialysis. S = soluble fraction, P =precipitated fraction

Unfortunately, solubility gained during the affinity was almost completely lost in the subsequent step. In fact, a visible precipitate was observed after removal of the chaotropic agent during dialysis (Fig. 3.4 right). Starting from 2L of culture, about 2 mg of total protein were purified by this method and used for crystallization trials. A this point of the work the new goal should have been the improvement of the purification protocol in order to increase the protein yield (and the grade of purity) and avoid the precipitation of the sample. New dialysis buffer should have been tested and the effect of arginine in limiting the precipitation, evaluated. Unfortunately, the work was stopped due to the publication of the structure of Hp-Fur2M (PDB 2XIG) by another group of researchers. This inconvenient is an unpleasant aspect of making science in a very competitive field.

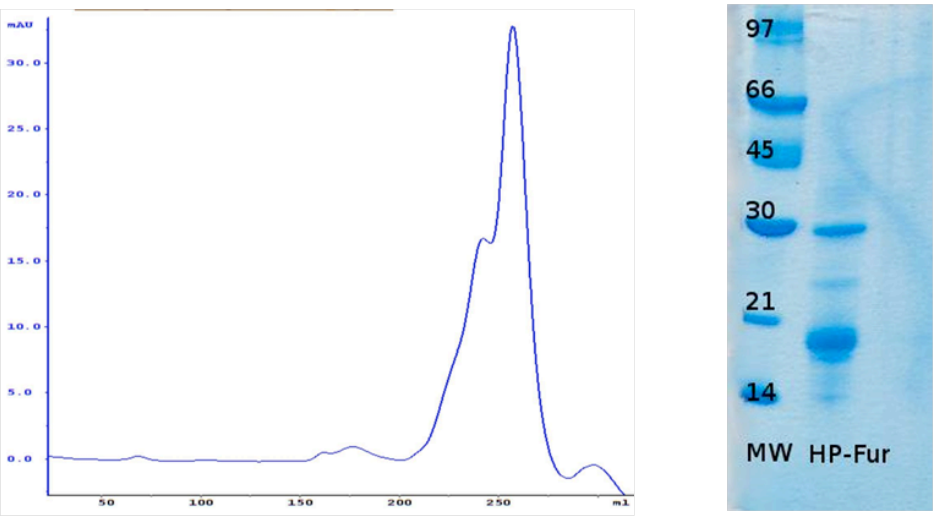


Fig.3.5 Left: Gel filtration on the soluble fraction. (Superdex 200 HiLoad 26/60). Right: SDS-PAGE of Hp-Fur at the end of purification. The band corresponds to peak at 250ml of retention volume.

The structure of Hp-Fur with 2 mutations (C78S and C150S) has been solved in complex with zinc ions. As structural studies on HpFur2M mutant showed (Dian et al.2011), Hp-Fur folding is similar to that of other Fur and Fur-like proteins consisting in a N-terminal domain (residues 1-92) connected to the C-terminal dimerization domain(residues 98-150) by a short hinge region (residues 93-97). The structure of double mutant HpFur2M revealed three different binding sites named S1, S2, and S3. S1 is a structural ZnS_4 site which binds Zn ion in a tetra-co-ordinated manner by two CX_2C motifs containing Cys 102, Cys 105 and Cys 142, Cys 145 respectively (Dian et al., 2011). S2 is required for DNA-binding activity and shows a variable and flexible coordination geometry in different HpFur subunits (between four to six atoms), while S3 has been demonstrate to be a non-essential site for DNA-binding, although it could be involved in determining the strength of HpFur response to excess of the metal (Dian et al., 2011).

4- Cloning, Expression and Purification of HypB, a GTPase from H. pylori.

4.1 Introduction

Successful colonization of the human gastric mucosa by *H. pylori* requires the activity of two nickel-containing enzymes: urease and hydrogenase. Urease contains a dinuclear nickel cluster and is crucial for acid resistance (see Chap.1-introduction), while the [Ni-Fe]-hydrogenase enzyme provides a means for *H. pylor*i to utilize hydrogen. Molecular hydrogen is produced in the large intestine of animals due to the fermentation reactions of carbohydrates catabolism. *H. pylori* is able to use molecular hydrogen as a respiratory substrate. The amount of the gas within tissues colonized by these pathogens is ample, and use of H_2 significantly increases the stomach colonization ability of *H. pylori*. In fact an hydrogenase mutant strain of *H. pylori* has been demonstrated to be much less efficient in colonization of mice (Olson and Maier, 2002); thus, H_2 is an energy-yielding substrate that can facilitate the maintenance of the gastric pathogen and the importance of H_2 use to *H. pylori* colonization capability represents a new aspect toward our understanding of host–pathogen relationships. *H. pylori* hydrogenase , a membrane-associated nickel-containing enzyme, is able to couple H_2 oxidation with reduction of a variety of positive redox-potential acceptors but it is not yet known whether H_2 oxidation by *Helicobacter* provides ATP, reducing agents, or a transmembrane potential (Maier, 2003). A direct coupling of H_2 oxidation to O_2 reduction would be the most energetically favorable way for harvesting electrons, but in the case of *H. pylori*, which colonize a low free-O_2 environment, the presence of a terminal oxidase able to efficiently operate in this conditions, binding and reducing O_2, has still to be clarified (Maier, 2003). However, H_2 activation via hydrogenase in membranes results in cytochrome reduction indicating that H_2 oxidation is linked to a respiratory energy-generating electron-transport chain (Benoit and Maier 2008). While the minimum content of [Ni -Fe]-hydrogenase in all organisms is two subunits (a large subunit containing the [Ni-Fe] site and a small subunit containing multiple [Fe-S] clusters), *H. pylori* [Ni-Fe]-hydrogenase appears to be organized as an heterotrimeric complex (Benoit and Maier, 2008). The [Ni-Fe]-hydrogenase gene cluster (HydABCDE) includes genes HydA , encoding the small subunit, HydB, the large subunit, HydC the membrane-anchored cytochrome b subunit, while HydD and HydE are annotated as hypothetical genes of accessory, not stuctural components. The metallocenters of both urease and [Ni-Fe]-hydrogenase are complex and need a battery of accessory proteins for their assembly. Two gene complexes have been identified to be implicated in urease and hydrogenase maturation, UreEFGH and HypABCDEF respectively. HypABCDEF genes are clustered together in one operon in most H_2

utilizing bacteria, while they are separated into three different locations on *H. pylori* genome (HypA - HypBCD and HypEF). UreEFGH components and in particular the better characterized ureE- a dimer binding nickel - and ureG - a GTPase- only participate in the maturation of urease and do not appear to be involved in hydrogenase maturation, while Hyp B and HypA have been shown to aid in the maturation of both the nickel-containing enzymes. In fact, gene deletion mutants in *H. pylori* revealed that interruption of either HypA or HypB not only abrogated hydrogenase activity, but also disrupted urease activity (Olson et al., 2001). Moreover, since addition of nickel can restore urease activity in HypAB mutants, this is not the case for either UreE and UreG mutants, which seems to suggest a secondary, despite important role for HypA and HypB in maturation of urease with respect to Ure accessory proteins (Maier et al., 2009). HypA is a dimer that is currently believed to act as a nickel chaperone, each dimer binding two nickel ions with micromolar affinity (Gasper et al., 2006). HypB is a GTPase regulator enzyme that has been proposed to controls the donation of the metal to the hydrogenase apoprotein or to release of a nickel free HypA (Hube et al., 2002). *In vitro* experiments suggest that HypB forms a complex with HypA (Atanassova et al.,2005). According with the actual model based on *E. coli* Hyp homologues, both the enzymes cooperate in hydrogenase maturation to insert the nickel ion into the large subunit of the hydrogenase after the iron is incorporated (that process is mediated by Hyp CDEF components). The exact mechanism of this nickel delivery step is still unknown (Sydor et al. 2011).

4.2 Hp-HypB characterization

The *H. pylori* metal-binding guanine nucleotide-binding protein Hp-HypB displays several unusual features if compared with its homologues proteins present in other bacteria. A few HypB proteins have sequences rich in histidine residues that are implicated in the binding of multiple nickel ions and are thought to function in nickel storage (Fig.4.1, Bj-HYpB sequence). Additionally, other HypB homologues bind nickel to a three- cysteine motif at the N terminus (Fig.4.1, Ec-HypB and Bj-HypB sequences). Finally Ec-HypB has been shown to bind either nickel or zinc with a micromolar affinity in an additional metal-binding site in the G-domain (CH motif) and a crystal structure of HypB from *Methanocaldococcus jannaschii* (PDB code 2HF8/9) revealed a corresponding dinuclear zinc site (Cys-95, His-96, and Cys-127) bridging two monomers of an homodimer (Sydor et al 2011). Hp-HypB lacks both the N terminal high affinity binding site and the polihistidine region, but interestingly it presents a CH motif in the G-domain.

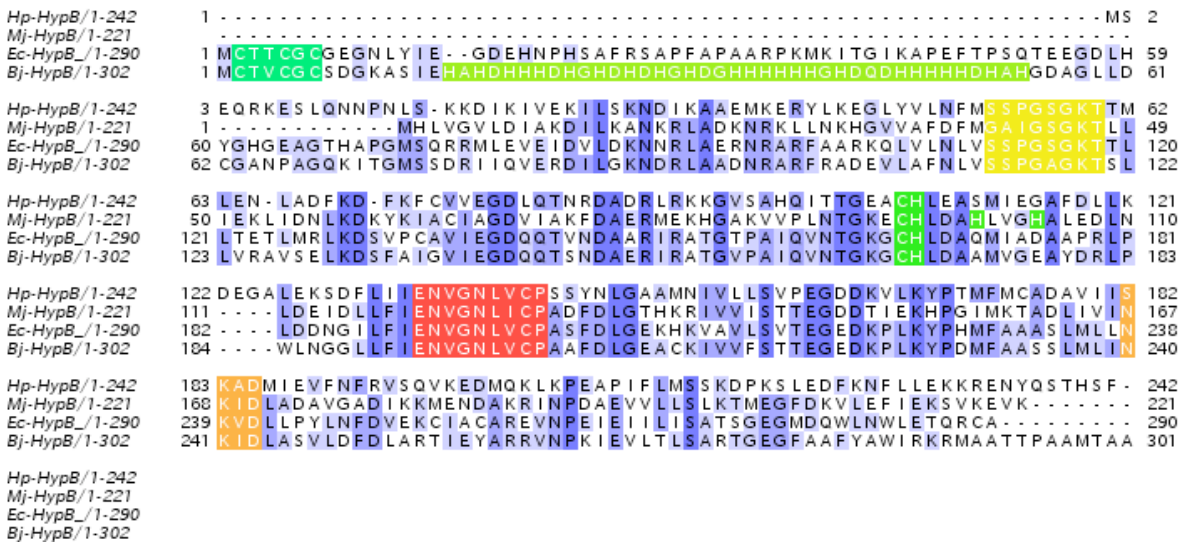


Fig.4.1 Sequences alignment of HypB proteins from H. Pylori (Hp-HypB), M.jannaschii (Mj-HypB), E.coli (Ec-HypB), and B.japonicum(Bj-HypB).In shades of blue is shown the identity above 50%. In shades of green the metal binding sites (the three-cysteine motif and the polyHis region (Nickel binding sites); CH motif (Nickel and Zinc binding site).In red, yellow and orange tipical GTPase motifs: switch motif ENV/IGNLV/ICP, P-loop (GXXXXGK(S/T)) and guanine recognition site (NKXD). Alignment and editing: Clustal & Jalview vers. 2.7 (Waterhouse et al., 2009)

While previous studies have suggested that Hp-HypB is incapable of binding Nickel (Maier et al. 1995), very recently it has been demonstrated that HypB when fully reduced binds an equivalent of either Nickel or Zinc and that metal modulates the GTPase activity and the dimeritazion state (Sydor et al., 2011). Cys 106 and His 107 residues (the CH motif) have been proposed to be involved in metal (zinck or nickel) binding according to an homology model created by using Mj-HypB crystal structure as a template (identity 30% and similarity 51%) (Sydor et al., 2011).

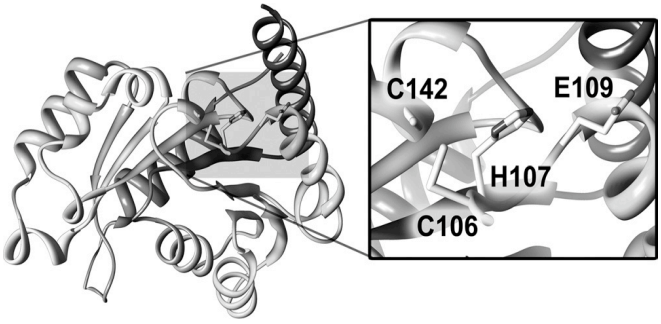


Fig.4.2 Homology model of HypB (Modeler 9v7).C106 and H107 located on the surface of the protein are proposed to bind Zinc or Nickel. C142 and E109 are shown as other candidate metal-binding residues (Sydor et al., 2001).

The aim of this project is the structural characterization of Hp-HypB in order to elucidate the Nickel step delivery and the maturation process. The work is still in progress and a parallel structural study of Hp-HypA is also scheduled.

4.3 Materials and Methods

4.3.1 Cloning

Hyp B gene from genomic Dna of *H. pylori* G27 strain was amplified by using primer CACCATGGCTAGCTGGAGCCACCCGCAGTTCGAAAAAGGCAGCGAACAACGAAAAGAATCT (fw), AGTTTAAAACGAATGCGTGGATTGGT (rv) with Dna polymerase Eurotaq (Euroclone). The forward primer contained the proper sequence for a TOPO® Directional Cloning (Invitrogen) in pET 101 expression vector and the signal sequence coding for a Strep-tag II (SA-WSHPQFEK). The reverse primer was designed with a stop sequence corresponding to the C-terminus (before the 6His-tag included in the vector).

Property	**Value**
Number of amino acids	242+11
Molecular Weight	28.607
Theoretical pI	5.6
Abs (1g/l)	0.28

Table 4.1 Some properties of the construct added of the Strep-tag II

4.3.2 Expression

HypB-pET101 plasmid was transformed into *E.coli* BL21(DE3) competent cells which were grown in a selective LB medium supplemented with 100 µg/mL ampicillin. A 3 L culture was performed at 37°C shacking until 0.7 OD, than expression was induced via 0.5 mM IPTG and bacteria were grown for additional 4 hours at 25°C. A 3L LB culture supplemented with 1.5 mM $NiS0_4$ was performed in parallel in order to test the best expression conditions. Cells were harvested and medium eliminated by centrifugation. Pellet was resuspended in a lysis buffer containing 100mM NaCl, 20mM Tris pH 7.5 and in the presence of $NiS0_4$, 4mM TCEP (tris(2-carboxyethyl)phosphine). In order to perform the lysis, lisozime was added at the concentration of 1mg/ml to the resuspended pellet and shacked 1h at 4°C. Then, cells were disrupted by sonication (8 cycles of 45'' at 4°C) after adding a protease inhibitor cocktail tablet.

4.3.3 Purification

HypB was purified by a Strep-Tactin affinity purification, which looks as the best system for metal-binding proteins, allowing also one-step purification (higher protein purity (> 95%) than 6xHistidine-tag). Crude lysate was loaded on a pre-packed strep-tactin (an engineered streptavidin) superflow column (IBA) at a flow rate of 1 ml/minute. The 5ml column was previously equilibrated with 2CV of buffer W (100mM Tris pH8, 150mM NaCl). WSHPQFEK-HypB bound to streptavidin was eluted from the column after 5 washings with Buffer W, by applying to the column Buffer E (Buffer W containing 2.5 mM desthiobiotin as a reversibly binding specific competitor). Purified protein was eluted and collected in three clean fractions corresponding to 0.5 CV each, as showed in the SDS-PAGE (Fig.4.3 left). Interestingly the protein migrate with an apparent molecular weight of 32,000 Da. The affinity was performed on the crude lysate coming from both cultures, with and without $NiS0_4$ supplementation. In order to characterize HypB in solution, size exclusion chromatography steps were carried out in three different conditions :1) HypB coming from expression without $NiS0_4$ and without previous TCEP treatment; 2) HypB coming from Nickel supplemented expression and incubated previously overnight in a glovebox with TCEP 10mM; 3) Hyp B expressed in presence of $NiS0_4$ and, after reduction with TCEP 10mM, added of Nickel 0.5 mM. A Superdex 75 prep grade HR16/60 (GE Healthcare) column was equilibrated with the following Buffer 25 mM Hepes pH 7.6, 200mM NaCl, and 1 mM TCEP. Elution profiles of HypB in these different conditions were not defined, indicating the presence of aggregates (data not shown). Therefore, the protocol was changed and desthiobiotin was eliminated from the solution soon after affinity, in order to possibly reduce the noise. Unfortunately and despite the addition of 4mM TCEP before injection in size exclusion chromatography, the protein largely precipitated. The soluble portion was still analyzed by gel filtration (Buffer 25 mM Hepes pH 7.6, 200mM NaCl, and 1 mM TCEP) and the resulting elution profile seem to correspond to a dimer. A yield of 2.5 mg was obtained and the protein was than concentrated to 10mg/ml and frozen at -80°C.

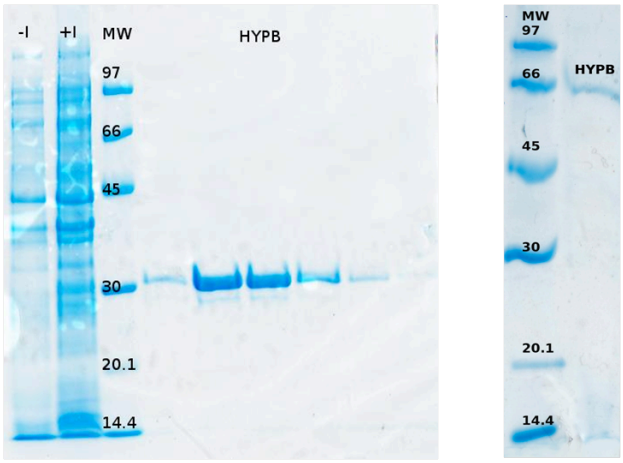


Fig. 4.3 SDS-Page for Strep-tag HypB after affinity (left) and at the end of the purification step (right).

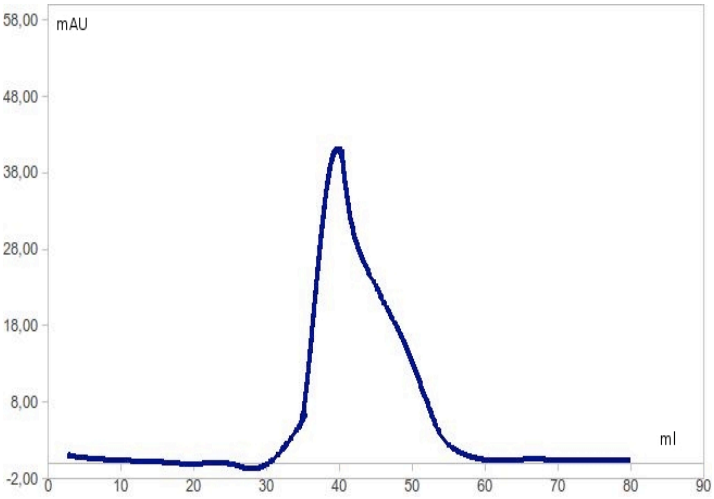


Fig. 4.4 Elution profile for HypB.

4.3.4 Crystallization

Crystallization trials were performed by exploring several precipitant conditions (Structure Screen I and II by Molecular Dimensions, PACT Suite, JCSG Suite). Unfortunately crystals diffracting well have not been obtained yet. However, some small crystals formed in presence of 150mM Maleic acid disodium salt, 42.5 mM HCl and 20% PEG 3350. They need further tests.

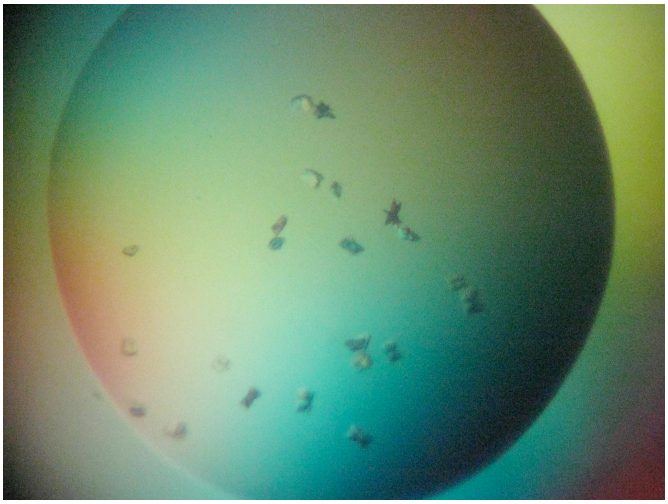


Fig. 4.5 HypB crystals obtained by vapor diffusion technique.

4.4 Results and Discussion

HypB was cloned, purified and some preliminary crystallization trials were performed. The characterization in solution was attempted, considering what is already known from the literature (Sydor et. al 2011). While the dimeric form of HypB was expected only in the presence of Nickel and in reducing conditions, trials of purification in these conditions did not led to isolate HypB as a dimer. On the contrary a HypB dimer seemed to result from a simpler purification in the absence of both Ni^{2+} and high concentration of reducing agent TCEP. The presence of the dimer looks confirmed by the SDS-PAGE, where indeed a band corresponding to the monomer was also expected. Interestingly, according with (Mehta et al., 2003) HypB was found to exist as homodimer with a molecular mass of 52.4 Kda in the native solution. Further studies are needed to optimize the purification step and the yield in order to explore more crystallization conditions. New efforts will be needed to clarify the dimerization metal-dependent process of Hp-HypB and, possibly, to elucidate HypB-HypA interaction.

Appendix

During the Phd period I had the special opportunity to spend six months at UCI, University of California – Irvine US - in the laboratory of Prof. Hartmut Luecke. This appendix will describe two projects I worked on during that time. Both projects are not about *H. pylori* proteins. They consist in the characterization of the role of the sensory rhodopsin transducer (ASRT) from *Anabaena* sp. and in the structural study of rhodopsin constructs from alga (Chr2). The short period that I have spent in the lab has not allowed to complete the project, nevertheless has allowed me to make some experience on the exciting field of membrane proteins.

Project I: ASRT characterization

Introduction

Anabaena sensory rhodopsin transducer (ASRT) is a soluble protein of 14.7 Kda from *Anabaena sp.* (Nostoc), a freshwater cyanobacterium. ASRT is part of the same operon of ASR, the light-sensing transmembrane protein *Anabaena* sensory rhodopsin (Vogeley et al., 2004) and they are co-transcribed. ASR is a member of the microbial rhodopsin family, membrane-embedded proteins that share some structural features, like the presence of seven transmembrane helices and of a covalent bond to a retinal chromophore. The latter is attached via a Schiff-base linkage to a lysine in the last of the seven helices. The absorption of one photon by the retinal (λmax 543 nm for ASR) results in a configurational change at the C_{13}= C_{14} double bond in the chromophore, which in turn triggers a number of conformational changes in the protein that eventually return the photoreceptor to its initial ground state (Lanyi, 2004). Inside the microbial rhodopsin family, ion-pumping rhodopsins exploit the photon energy to transport ions (H^+ in bacteriorhodopsins and Cl^- in halorhodopsin) across the membrane against electrochemical gradients, while sensory rhodopsins, like ASR, use the photocicle for signaling pathways. In the latter case the structural changes associated with photon absorption are detected by transducer molecules that connect the sensory rhodopsin to the signaling cascades involved in the regulation of light-dependent processes, like photosynthesis, phototaxis and circadian rhythm control (Vogeley, and Luecke 2006). ASRT has been demonstrated to bind to ASR, consistently with their co-transcription; experiments of ASR-ASRT co-expression in *E. coli* showed an increase of about 20% in the rate of photocycle, suggesting a physical interaction between the two proteins (Jung et al., 2003). These findings together with the lacking of a detectable proton transport activity when expressed in *E. coli*, seem to suggest that ASRT works as a transducer for ASR. The best sensory rhodopsin characterized in structural terms is the haloarcheal phototaxis receptor sensory rhodopsin II (SRII) from *Natronomonas pharaonis*. Bacteriorhodopsin (BR), halorhodopsin (HR) and SRII are all archeal in origin and have been well characterized, functionally and structurally (Vogeley et al. 2007). *Halobacterium salinarum* has been the first organism considered, but several variants have been documented in related halophilic archaea, such as *N. pharaonis* and *Haloarcula vallismortis*. ASR-ASRT pair represents the first sensory rhodopsin and transducer system detected in eubacteria. The eubacterial ASR-ASRT system uses a completely different mechanism of transduction with respect to archaeal SRII (*N. pharaonis)* and its cognate transducer HtrII. In fact,

an extensive membrane-embedded transducer-binding region has been observed in the X-ray structure of *N. pharaonis* SRII (Luecke et al., 2001) co-crystallized with its taxis transducer fragment (Gordeliy et al., 2002; Spudich 2002). Thus, while archaeal sensory rhodopsins transmit signals by transmembrane helix–helix interactions with integral membrane transducers, the interaction of the soluble ASRT with ASR seems to suggest a new transduction mechanisms used by microbial sensory rhodopsin (Vogeley et al. 2006). Atomic resolution structures for the *Anabaen*a photoreceptor and its putative transducer have been obtained (PDB code 1XIO and 2II9, 2II7, 2II8, 2IIA respectively) (Vogeley et Luecke, 2004), but the physiological function of *Anabaena* SR has not been established yet. Asr is the first structurally characterized eubacterial retinylidene photoreceptor. The structure, obtained by X-ray diffraction of crystals grown in a cubic liquid phase, showed similarity in the overall architecture with archaeal rhodopsins (Vogeley et al.2004). However ASR exhibits some peculiar differences in the chromophore and the cytoplasmic-side portion that is believed to interact with ASRT. Unlike other microbial rhodopsins, ASR photoactive site displays light-induced shifts of isomeric configuration. While retinal 13-cis state predominates in orange illumination, in blue light the all-trans content increase. This mixture of retinal isomers, where cis and trans chromophore configurations depend on the color of illumination (λ), is a peculiar property providing a possible mechanism for single-pigment color sensing (Vogeley et al. 2004). Indeed, *Anabaena* is a photosynthetic organism and chromatic adaptation has been recognized as an essential process to regulate the synthesis of light-harvesting pigments like phycocyanin (orange light stimulated) and phycoerythrin (blue light stimulated). Moreover, the cytoplasmic face of ASR, rich in hydrophilic residues, introduces an other important structural difference between *Anabaena* and archaeal rhodopsins, the latter being highly hydrophobic in this region. The numerous ordered water molecules forming hydrogen bonds with polar and charged residues in the cytoplasmic-side of ASR are thought to be involved in the interaction with the soluble ASRT .ASRT structure was solved in three different space groups, P4 (PDB code 2IIA), C2 (2II9), and $P2_12_12_1$ (2II7, 2II8) in all of which the protein was present as a planar tetramer. While residues 3-18 and 31-104 showed clear density in all four crystal forms consisting of seven β-strands arranged in a β-sandwich (β-face), the 20 C-terminal residues and 19-30 residues looked quite disordered corresponding to a very flexible an variable α-helical portion (α-face) (Vogeley et al., 2007). The tetrameric state of ASRT, which is consistent with the characterization in solution, is stabilized by interactions provided by the completely ordered β-face, while the flexible helical in the α-face is supposed to interact with the membrane or ASR. Isothermal titration calorimetric

(ITC) measurement suggested that ASRT binds ASR with a stoichiometric value near 4, thus the binding could occur between ASR and one tetramer of ASRT.

Interestingly, the tetrameric structure of ASRT has been found to superficially resemble to the G_β protein of eukaryotic heterotrimeric G-proteins, which functions as a transducer for G protein-coupled receptors ($GPCR_s$) (Vogeley et al.2007). Little information is known about the mechanism of signal transduction operated by ASRT. The structure seems to suggest that the variable and flexible α-face could be a major key in the signaling mechanism. In fact, the assumption of alternative conformations on the cytoplasmic surface of ASR could induce the adoption of different signalling conformations by ASRT. In turn, ASRT, due to its solubility may diffuse and interact with other components or directly as transcription factor. The general aim of the project was to elucidate the role of ASRT, considering the possibility that it is a transcription factor working as a direct downstream signaling component. The specific goal was to test the propensity of ASRT to bind to DNA carrying out crystallization trials in complex with suitable DNA probes.

Experimental procedure and Results

ASRT was overexpressed in *E.coli* BL21 in a 4L LB culture. The plasmid, containing the ASRT coding sequence with an additional 6-His tag at the N-terminus, under IPTG-inducible promoter, was been given by Prof. Elena Spudich of Texas University. Protein was then purified by IMAC affinity with a talon resin column (Clontech Laboratories) and by an additional anionic exchange chromatography step (Fig. *II* and *III*). After purification the protein was concentrated in a Buffer 20mM Tris, 80mM NaCl pH 7.9 until 16.8 mg/ml (1mM) and stored at -80°C. In order to test different DNA sequences and to detect the substrate with higher affinity for ASRT for co-crystallization purposes, an Electrophoretic Mobility Shift Assay (EMSA) was performed. Four different double strand DNA constructs were tested, according with previous isothermal titration calorimetric ASRT-DNA experiments. Single strand oligonucleotides (from IDT) were annealed by heating and slow cooling in a Buffer 10 mM Tris pH 8, 50 mM NaCl, 1 mM EDTA. Sequences of the constructs are shown below. Constructs 1 (20 bp long, blunt ends) 2 and 3 (21 and 22 bp long sticky ends) corresponded to phycoerythrocyanin (pec) operon promoter sequence (according to Wang et al., 2001) a micromolar affinity is expected), while construct 4 was a sticky ends tata-box like sequence.

1 ATG ACC TTT AGG AGGAAAGA

TAC TGG AAA TCC TCC TTT CT

2 **T** ATG ACC TTT AGG AGG AAA G A

TAC TGG AAA TCC TCC TTT CT **A**

3 **AA** ATG ACC TTT AGG AGG AAA GA

TAC TGG AAA TCC TCC TTT CT**TT**

4 **A** ATATATATATATATATATAT

TATATATATATATATATATA **T**

EMSA was performed in agarose gel 1% in Binding buffer (100mM Tris, 500mM KCl, 10 mM DTT pH 7.5) , 2.5% glycerol, 5 mM MgCl according with Pierce protocol. Concentrated DNA solution was added stepwise to ASRT sample to reach a *dsDNA:ASRT tetramer* molar ratio of 0.25, 0.5, 0.75, 1.0, 2.0, and 4.0. Unfortunately no evidence of binding have been obtained, possibly due to the conditions used for the essay (salt, pH, temperature,). Nonetheless, some crystallization trials were performed with ASRT and the blunt construct (1) by testing several crystallization kits (Structure Screen I and II by Molecular Dimensions, PACT Suite, JCSG Suite by Qiagen).

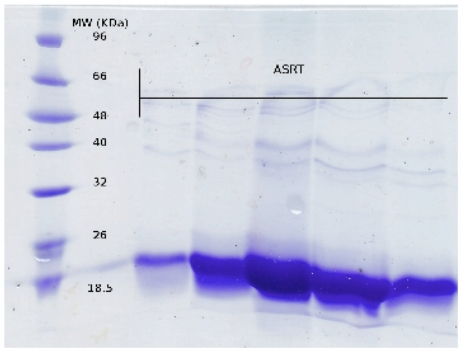
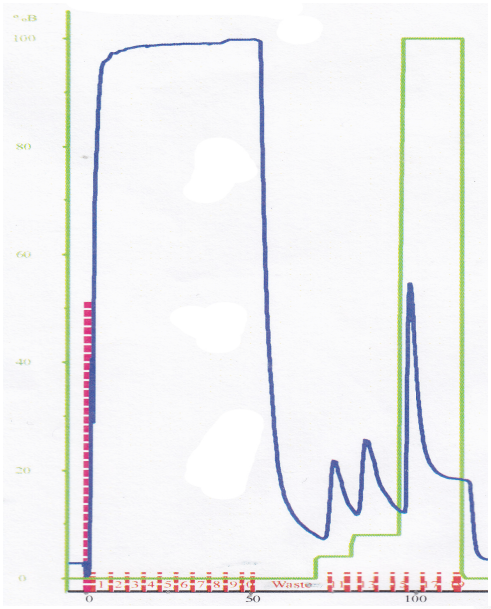


Fig. I SDS-Page for ASRT at the end of the two purification steps.The different runs correspond to different portion of the elution peak shown in Fig. III.



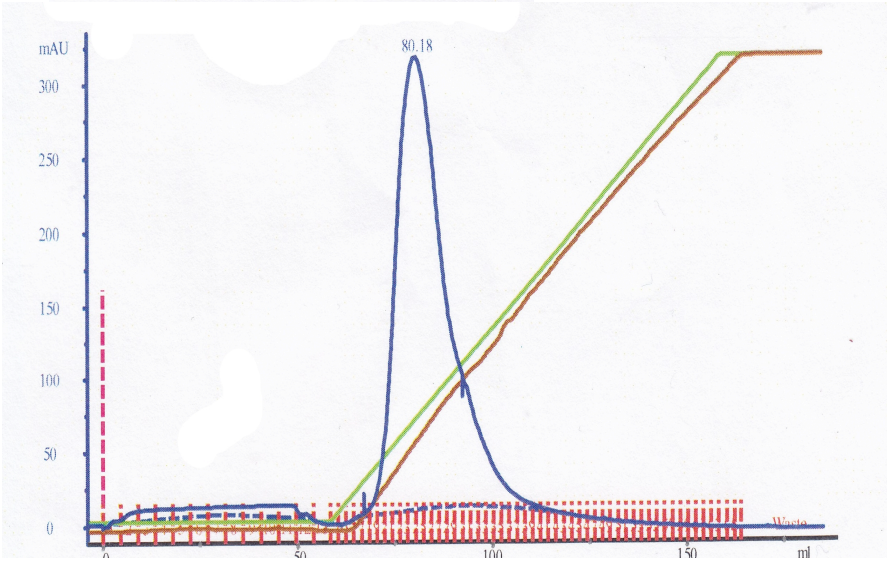


Fig II (top) and Fig. III The IMAC affinity elution step for ASRT and the following anionic exchange chromatography (the % of Imidazole 500mM and NaCl 1M are indicated in green).

Project II: Channelrhodopsins Constructs Expression

Introduction

More than three hundred photochemically reactive proteins that use retinal as their chromophore have been described in both prokaryotic and eukaryotic organisms. A common feature of these proteins is the presence of seven membrane-embedded α-helices that form an internal pocket in which the retinal is bound (Spudich et al., 2000). Rhodopsins are classified in two families or types, according to the sequence conservation. The Type 1 rhodopsins are found in archaeal organisms, eubacteria , fungi and algae and they function as light-driven ion transporters (BR and HL), phototaxis receptors (SR I and II), and light activated channels (Channelrhodopsins). The type 2 rhodopsins consist of the photosensitive receptor proteins in animal eyes, including human rod and cone visual pigments, receptor proteins in the pineal gland, hypothalamus, and other tissues of lower vertebrates (Spudich et al., 2000). Among type I rhodopsins, Channelrhodopsins, are currently the subject of intense study due to their growing importance for applications in neuro- and cell biology. Channelrhodopsins are a completely new class of light-gated ion channels present in green algae and to date four have been found (ChR1 and ChR2 from *Chlamydomonas reinhardtii and* VChR1 and VChR2 from *Volvox carteri* (Hegemann et al. 2010). ChR1 and ChR2 are expressed in the eye-spot of *C. reinhardtii* and are responsible for phototactic and photophobic responses. While ChR1 predominantly mediates proton conductance over the cell membrane, ChR2 allows the permeation of different monovalent and divalent cations (Bamann et al 2007). After absorption of a photon, ChR2 opens rapidly to form a pore permeable to monovalent (Na^+, K^+ and H^+) and divalent cations (Ca^{2+}). The light-induced cation conductance leads to strong and rapid depolarization of the membrane (Nagel et al. 2003). According with the current model, the photoisomerization of the all-trans-retinal to the 13-cis configuration induces the protein to change its conformation and to open the channel (Hegemann, 2008).

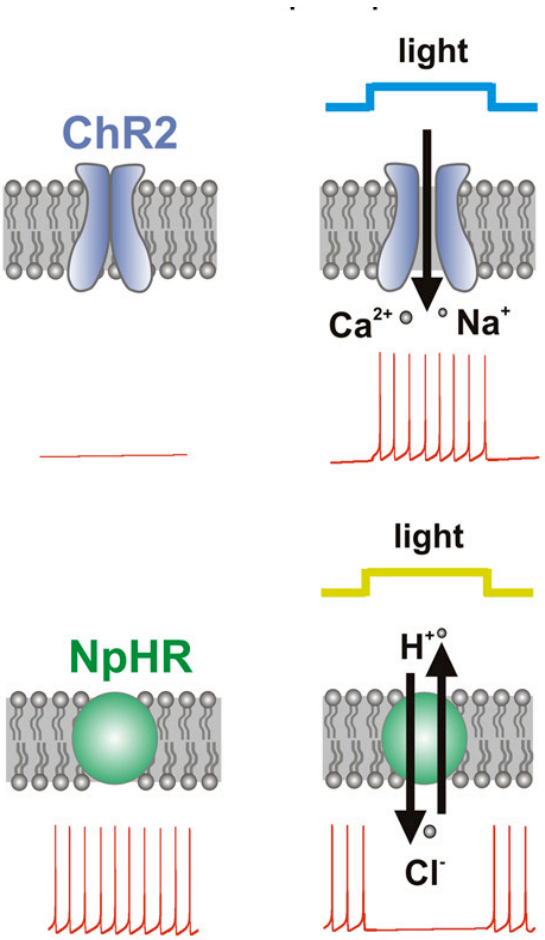


Fig.IV CHR2 and NpHR (Halorhodopsin from N. pharaonis) are useful optogenetics tools for modulating membrane voltage potential. When expressed in neural cells, ChR2 opens in response to blue light (470 nm) producing a depolarization of the expressing neuron and generating action potentials. The chloride pump HR when activated (by yellow light at 580 nm) produces in neurons hyperpolarization, thus inhibiting the spikes generation. ChR2 and NpHR can be used togheter in the same cell to activate and quiet neural activity with two different colors of light (Adapted from Kno¨pfel et al. 2010).

A major question about light-gated ion channels concerns the coupling between the light activation and the protein action (Bamann et al 2007). In the light-driven proton pump bacteriorhodopsin (BR), which is one of the best known structures among membrane proteins, the outwardly directed proton transfer is strictly coupled to the photocycle. Therefore, the turnover for the proton transport is limited to the reaction time of the photocycle, that is, one proton per cycle time (Bamann et al. 2007). After illumination, the *trans–cis* isomerization of the retinal triggers the deprotonation of the Schiff base and the consequent release of the proton to the extracellular side. Reprotonation of the Schiff base occurs from the intracellular side until a new isomeration cycle of the retinal starts. For channel rhodopsins the situation is different. Light-induced modifications at the chromophore site are transmitted to a conformational change leading to the open state of the channel, which allows a high number of cations to passively cross the membrane, driven only by their electrochemical gradient. The lifetime of the open state of the

channel was determined to be approximately 10 ms by electrophysiological methods, and it exceeds by far the estimated cation transfer time. Therefore, the passage of the cations is uncoupled from the supposed photocycle (Bamann et al.2007).

Because of its unique properties as a light-gated, inwardly directed cation channel, ChR2 has become an extremely useful tool for neural applications. In fact, it has opened the way of controlling neural activity simply by virtue of the light-induced depolarization of the cell (Boyden et al., 2005). ChR2 was expressed in cultivated neural cells, in living animals like *C. elegans* (in muscle cells) and in the photoreceptor-deficient retina of rodents. Boyden et al.(2005) showed that pulses of blue light at approximately 470 nm wavelengths induce spikes of action potentials in cells expressing the ChR2 protein. Studies in mice with retinal photoreceptor degeneration have demonstrated that virally delivered ChR2 in retinal cells restores the ability of the neurons to transmit information to the visual cortex in response to light in a safe, stable long-term manner. Since the development of ChR2, other opsins have been identified or engineered and used to control neural activity (LaLumiere 2011). Various mutants of ChR2 are currently available (Lin, 2011), including ChR2(H134R) (Nagel et al. 2005). Unfortunately most of the channelrhodopsin mutations generated so far have been chosen on the basis of the homology of ChR2 to BR, which is quite low specially for helices 1 and 2 of the two proteins. In order to improve the application of ChR2, a better understanding of the molecular mechanism of this natural light-induced neural switch is needed. To date, high-resolution three-dimensional structures for the various dark and open states of channelrhodopsin are the most urgent requirements (Hegemann et al., 2011).

Experimental procedures

In order to structurally characterize new channelrhodopsin members, several constructs deriving from different algal species were cloned and the expression in various expression systems (*E. coli*, insect cells) was attempted. Constructs, sent from Dr. Ed Boyden lab, were previously tested according to their propensity to generate a photocurrent after illumination in human embryonic kidney cell line 293 (HEK293)cells. Among the four constructs I tried to express, three (from *Neochlosarcina sp, Scherffelia dubia,* and *Brachiomonas submarina*), were predicted to be new algal channelrhodopsins, and one (from *Cyanophora paradoxa*) was assumed to belong to a new class of inward pumps. A sequence allignment between the four new constructs sequences and BR, ChR2 and SRII sequences is shown in Fig.*V*. The overall sequence identity in the seven transmembrane sections is about 15% to 18%, but the homology percentage increases in some

regions (helix 3 and helix 7) as indicated by shades of blue. All the constructs were cloned with an N-term 6His-tag in a pET28(b) vector and transformed in C43(DE3) *E. coli*. Three liters of LB cultures supplemented with 50 µg/ml ampicillin were performed and the expression of the proteins was induced by 1mM IPTG added in the medium at A_{600} 0.5. 5 µM all-trans-retinal was then added and the expression was carried out in the dark shacking for 4 hours at 30°C. Cells were harvested by centrifugation at 5,000RPM, for 15min. After lysis by French press, membrane were extracted by ultracentrifugation at 150000g, 1h, 4°C repeated twice. Membrane solubilization was performed in a buffer containing β-DDM 0.2 mM. Unfortunately the protein was not detected in *E. coli* membranes, as revealed by an IMAC affinity performed later. Currently others trials to express these channels proteins are in progress. New constructs have been designed (with a C-term 6His-tag) and various expression systems are being testing (insect cells, yeast, cell free).

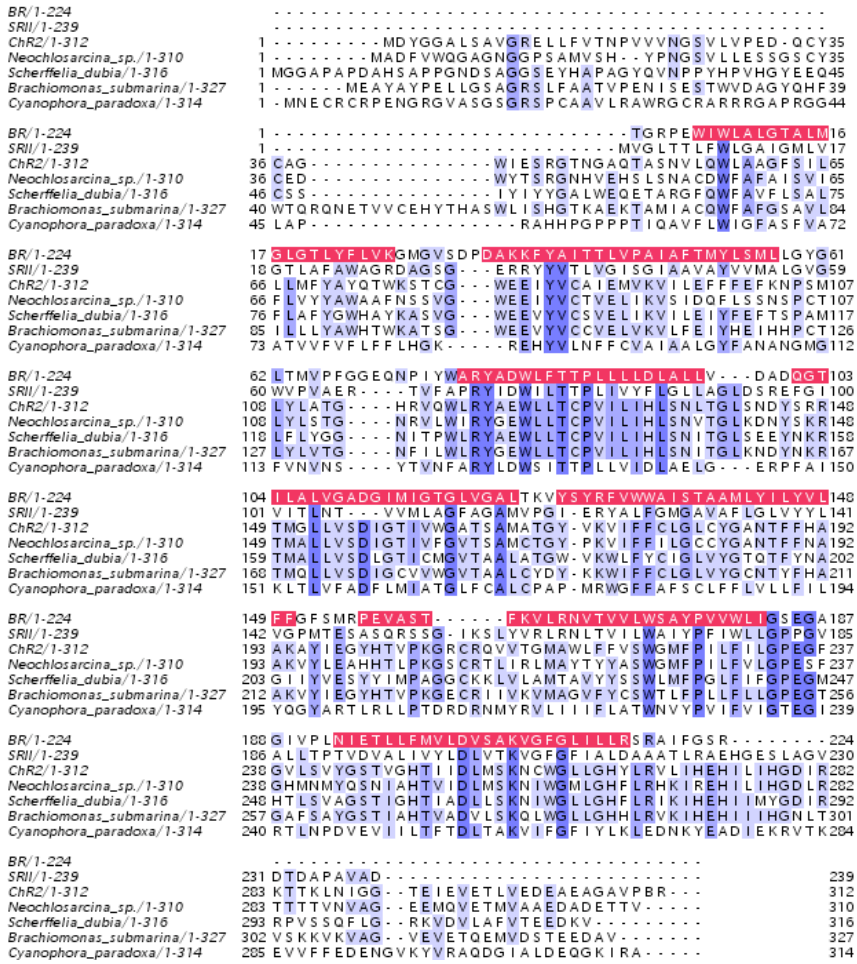


Fig. V Clustal sequence alignment. Sequences from BR, SRII, CHR2 and the four new candidates (Neochlosarcina sp, Scherffelia dubia, Brachiomonas submarina and Cyanophora paradoxa* are compared. Identity is shown in shades of blue ,the 7 transmembrane helices regions are *highlighted in red. Editing Jalview vers. 2.7 (Waterhouse et al., 2009).

References

CHAPTER I

1. Agarwal K, Agarwal S. *Helicobacter pylori* vaccine: from past to future.Mayo Clin Proc. 2008 Feb;83(2):169-75.

2. Ahmed N, Tenguria S, Nandanwar N. *Helicobacter pylori*--a seasoned pathogen by any other name.Gut Pathog. 2009 Dec 23;1:24.

3. Ahmed N, Sechi LA. *Helicobacter pylori* and gastroduodenal pathology: new threats of the old friend.Ann Clin Microbiol Antimicrob. 2005 Jan 5;4:1.

4. Algood HM, Cover TL. *Helicobacter pylori* persistence: an overview of interactions between H. pylori and host immune defenses.Clin Microbiol Rev. 2006 Oct;19(4):597-613.

5. Alm RA, Ling LS, Moir DT, King BL, Brown ED, Doig PC, Smith DR, Noonan B, Guild BC, deJonge BL, Carmel G, Tummino PJ, Caruso A, Uria-Nickelsen M, Mills DM, Ives C, Gibson R, Merberg D, Mills SD, Jiang Q, Taylor DE, Vovis GF, et al. Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori*.Nature. 1999 Jan 14;397(6715):176-80.

6. Alm RA, Trust TJ. Analysis of the genetic diversity of *Helicobacter pylori*: the tale of two genomes.J Mol Med (Berl). 1999 Dec;77(12):834-46.

7. Amieva MR, El-Omar EM. Host-bacterial interactions in *Helicobacter pylori* infection.Gastroenterology. 2008 Jan;134(1):306-23.

8. Andersen LP. Colonization and infection by *Helicobacter pylori* in humans.Helicobacter. 2007 Nov;12 Suppl 2:12-5.

9. Backert S, Clyne M, Tegtmeyer N. Molecular mechanisms of gastric epithelial cell adhesion and injection of CagA by *Helicobacter pylori*.Cell Commun Signal. 2011 Nov 1;9(1):28.

10. Backert S, Clyne M. Pathogenesis of *Helicobacter pylori* infection.Helicobacter. 2011 Sep;16 Suppl 1:19-25.

11. Backert S, Selbach M. Role of type IV secretion in *Helicobacter pylori* pathogenesis.Cell Microbiol. 2008 Aug;10(8):1573-81.

12. Basso D, Plebani M, Kusters JG. Pathogenesis of *Helicobacter pylori* infection.Helicobacter. 2010 Sep;15 Suppl 1:14-20.

13. Bode G, Rothenbacher D, Brenner H. *Helicobacter pylori* colonization and diarrhoeal illness: results of a population-based cross-sectional study in adults.Eur J Epidemiol. 2001;17(9):823-7.

14. Bourzac KM, Guillemin K. *Helicobacter pylori*-host cell interactions mediated by type IV secretion.Cell Microbiol. 2005 Jul;7(7):911-9.

15. Bury-Moné S, Skouloubris S, Dauga C, Thiberge JM, Dailidiene D, Berg DE, Labigne A, De Reuse H. Presence of active aliphatic amidases in Helicobacter species able to colonize the stomach.Infect Immun. 2003 Oct;71(10):5613-22.

16. Busler VJ, Torres VJ, McClain MS, Tirado O, Friedman DB, Cover TL. Protein-protein interactions among *Helicobacter pylori* cag proteins.J Bacteriol. 2006 Jul;188(13):4787-800.

17. Cascales E, Christie PJ. The versatile bacterial type IV secretion systems.Nat Rev Microbiol. 2003 Nov;1(2):137-49.

18. Ceci P, Mangiarotti L, Rivetti C, Chiancone E. The neutrophil-activating Dps protein of *Helicobacter pylori*, HP-NAP, adopts a mechanism different from *Escherichia coli* Dps to bind and condense DNA.Nucleic Acids Res.2007;35(7):2247-56.

19. Cendron L, Zanotti G. Structural and functional aspects of unique type IV secretory components in the *Helicobacter pylori* cag-pathogenicity island.FEBS J. 2011 Apr;278(8):1223-31.

20. Cooksley C, Jenks PJ, Green A, Cockayne A, Logan RP, Hardie KR. NapA protects *Helicobacter pylori* from oxidative stress damage, and its production is influenced by the ferric uptake regulator.J Med Microbiol. 2003 Jun;52(Pt 6):461-9.

21. Cover TL, Blaser MJ. *Helicobacter pylori* in health and disease.Gastroenterology.n2009 May;136(6):1863-73.

22. Cover TL. Genetic diversity in *Helicobacter pylori*. Helicobacter. 1997 Jun;2(2):108-9.

23. Cover TL.Commentary: *Helicobacter pylori* transmission, host factors, and bacterial factors.Gastroenterology. 1997 Dec;113(6 Suppl):S29-30.

24. Cover TL, Hanson PI, Heuser JE. Acid-induced dissociation of VacA, the *Helicobacter pylori* vacuolating cytotoxin, reveals its pattern of assembly.J Cell Biol. 1997 Aug 25;138(4):759-69.

25. Cover TL, Blanke SR. *Helicobacter pylori* VacA, a paradigm for toxin multifunctionality.Nat Rev Microbiol. 2005 Apr;3(4):320-32.

26. D'Elios MM, Andersen LP. *Helicobacter pylori* inflammation, immunity, and vaccines.Helicobacter. 2007 Oct;12 Suppl 1:15-9.

27. de Bernard M, D'Elios MM.The immune modulating activity of the *Helicobacter pylori* HP-NAP: Friend or foe?Toxicon. 2010 Dec 15;56(7):1186-92.

28. Del Giudice G, Covacci A, Telford JL, Montecucco C, Rappuoli R.The design of vaccines against *Helicobacter pylori* and their development.Annu Rev Immunol. 2001;19:523-63.

29. Dominguez-Bello MG, Blaser MJ. The human microbiota as a Marker for migrations of individuals and populations. Annu Rev Anthropology. 2011 Oct;40:451-474

30. Dorer MS, Talarico S, Salama NR. *Helicobacter pylori*'s unconventional role in health and disease.PLoS Pathog. 2009 Oct;5(10):e1000544.

31. Eaton KA, Suerbaum S, Josenhans C, Krakowka S.Colonization of gnotobiotic piglets by *Helicobacter pylori* deficient in two flagellin genes.Infect Immun. 1996 Jul;64(7):2445-8.

32. Evans DJ Jr, Evans DG, Takemura T, Nakano H, Lampert HC, Graham DY, Granger DN, Kvietys PR.Characterization of a *Helicobacter pylori* neutrophil-activating protein.Infect Immun. 1995 Jun;63(6):2213-20.

33. Falush D, Wirth T, Linz B, Pritchard JK, Stephens M, Kidd M, Blaser MJ, Graham DY, Vacher S, Perez-Perez GI, Yamaoka Y, Mégraud F, Otto K, Reichard U, Katzowitsch E, Wang X, Achtman M, Suerbaum S.Traces of human migrations in *Helicobacter pylori* populations.Science. 2003 Mar 7;299(5612):1582-5.

34. Ferrero RL, Thiberge JM, Huerre M, Labigne A.Recombinant antigens prepared from the urease subunits of Helicobacter spp.: evidence of protection in a mouse model of gastric infection.Infect Immun. 1994 Nov;62(11):4981-9.

35. Fischer W, Püls J, Buhrdorf R, Gebert B, Odenbreit S, Haas R.Systematic mutagenesis of the *Helicobacter pylori* cag pathogenicity island: essential genes for CagA translocation in host cells and induction of interleukin-8.Mol Microbiol. 2001 Dec;42(5):1337-48.

36. Guruge JL, Falk PG, Lorenz RG, Dans M, Wirth HP, Blaser MJ, Berg DE, Gordon JI.Epithelial attachment alters the outcome of *Helicobacter pylori* infection.Proc Natl Acad Sci U S A. 1998 Mar 31;95(7):3925-30.

37. Ha NC, Oh ST, Sung JY, Cha KA, Lee MH, Oh BH.Supramolecular assembly and acid resistance of *Helicobacter pylori* urease.Nat Struct Biol. 2001 Jun;8(6):505-9.

38. Hennig EE, Mernaugh R, Edl J, Cao P, Cover TL.Heterogeneity among *Helicobacter pylori* strains in expression of the outer membrane protein BabA.Infect Immun. 2004 Jun;72(6):3429-35.

39. Hessey SJ, Spencer J, Wyatt JI, Sobala G, Rathbone BJ, Axon AT, Dixon MF.Bacterial adhesion and disease activity in Helicobacter associated chronic gastritis.Gut. 1990 Feb;31(2):134-8.

40. Ilver D, Barone S, Mercati D, Lupetti P, Telford JL. *Helicobacter pylori* toxin VacA is transferred to host cells via a novel contact-dependent mechanism.Cell Microbiol. 2004 Feb;6(2):167-74.

41. Kostrzynska M, Betts JD, Austin JW, Trust TJ.Identification, characterization, and spatial localization of two flagellin species in *Helicobacter pylori* flagella.J Bacteriol. 1991 Feb;173(3):937-46.

42. Loh JT, Gupta SS, Friedman DB, Krezel AM, Cover TL.Analysis of protein expression regulated by the *Helicobacter pylori* ArsRS two-component signal transduction system.J Bacteriol. 2010 Apr;192(8):2034-43.

43. Marais A, Mendz GL, Hazell SL, Mégraud F.Metabolism and genetics of *Helicobacter pylori*: the genome era.Microbiol Mol Biol Rev. 1999 Sep;63(3):642-74. Review.PubMed [citation] PMID: 10477311, PMCID: PMC103749

44. Marshall BJ, Warren JR.Unidentified curved bacilli in the stomach of patients with gastritis and peptic ulceration.Lancet. 1984 Jun 16;1(8390):1311-5.

45. McClain MS, Shaffer CL, Israel DA, Peek RM Jr, Cover TL.Genome sequence analysis of *Helicobacter pylori* strains associated with gastric ulceration and gastric cancer.BMC Genomics. 2009 Jan 5;10:3.

46. McClain MS, Iwamoto H, Cao P, Vinion-Dubiel AD, Li Y, Szabo G, Shao Z, Cover TL.Essential role of a GXXXG motif for membrane channel formation by *Helicobacter pylori* vacuolating toxin.J Biol Chem. 2003 Apr 4;278(14):12101-8.

47. McGee DJ, Mobley HL.Pathogenesis of *Helicobacter pylori* infection.Curr Opin Gastroenterol. 2000 Jan;16(1):24-31.

48. Mobley HL, Island MD, Hausinger RP.Molecular biology of microbial ureases.Microbiol Rev. 1995 Sep;59(3):451-80.

49. Mobley HLT, Mendz GL, Hazell SL.Overview.In: Mobley HLT, Mendz GL, Hazell SL, editors. *Helicobacter pylori*: Physiology and Genetics. Washington (DC): ASM Press; 2001. Chapter 1.

50. Montecucco C, Rappuoli R.Living dangerously: how *Helicobacter pylori* survives in the human stomach.Nat Rev Mol Cell Biol. 2001 Jun;2(6):457-66.

51. Nguyen LT, Uchida T, Murakami K, Fujioka T, Moriyama M.*Helicobacter pylori* virulence and the diversity of gastric cancer in Asia.J Med Microbiol. 2008 Dec;57(Pt 12):1445-53.

52. Papini E, Gottardi E, Satin B, de Bernard M, Massari P, Telford J, Rappuoli R, Sato SB, Montecucco C.The vacuolar ATPase proton pump is present on intracellular vacuoles induced by *Helicobacter pylori*.J Med Microbiol. 1996 Aug;45(2):84-9.

53. Papinutto E, Dundon WG, Pitulis N, Battistutta R, Montecucco C, Zanotti G.Structure of two iron-binding proteins from Bacillus anthracis.J Biol Chem. 2002 Apr 26;277(17):15093-8.

54. Rektorschek M, Buhmann A, Weeks D, Schwan D, Bensch KW, Eskandari S, Scott D, Sachs G, Melchers K.Acid resistance of *Helicobacter pylori* depends on the UreI membrane protein and an inner membrane proton barrier.Mol Microbiol. 2000 Apr;36(1):141-52.

55. Rothenbacher D, Blaser MJ, Bode G, Brenner H.Inverse relationship between gastric colonization of *Helicobacter pylori* and diarrheal illnesses in children: results of a population-based cross-sectional study.J Infect Dis. 2000 Nov;182(5):1446-9.

56. Sachs G, Weeks DL, Wen Y, Marcus EA, Scott DR, Melchers K.Acid acclimation by *Helicobacter pylori*.Physiology (Bethesda). 2005 Dec;20:429-38.

57. Satin B, Del Giudice G, Della Bianca V, Dusi S, Laudanna C, Tonello F, Kelleher D, Rappuoli R, Montecucco C, Rossi F.The neutrophil-activating protein (HP-NAP) of *Helicobacter pylori* is a protective antigen and a major virulence factor.J Exp Med. 2000 May 1;191(9):1467-76.

58. Shaffer CL, Gaddy JA, Loh JT, Johnson EM, Hill S, Hennig EE, McClain MS, McDonald WH, Cover TL.*Helicobacter pylori* exploits a unique repertoire of type IV secretion system components for pilus assembly at the bacteria-host cell interface.PLoS Pathog. 2011 Sep;7(9):e1002237.

59. Skouloubris S, Thiberge JM, Labigne A, De Reuse H.The *Helicobacter pylori* UreI protein is not involved in urease activity but is essential for bacterial survival in vivo.Infect Immun. 1998 Sep;66(9):4517-21.

60. Skouloubris S, Labigne A, De Reuse H.The AmiE aliphatic amidase and AmiF formamidase of *Helicobacter pylori*: natural evolution of two enzyme paralogues.Mol Microbiol. 2001 May;40(3):596-609.

61. Spohn G, Scarlato V.Motility, Chemotaxis, and Flagella.In: Mobley HLT, Mendz GL, Hazell SL, editors. *Helicobacter pylori*: Physiology and Genetics. Washington (DC): ASM Press; 2001. Chapter 21.

62. Tomb JF, White O, Kerlavage AR, Clayton RA, Sutton GG, Fleischmann RD, Ketchum KA, Klenk HP, Gill S, Dougherty BA, Nelson K, Quackenbush J, Zhou L, Kirkness EF, Peterson S, Loftus B, Richardson D, Dodson R, Khalak HG, Glodek A, McKenney K, Fitzegerald LM, et al.The complete genome sequence of the gastric pathogen *Helicobacter pylori*.Nature. 1997 Aug 7;388(6642):539-47. Erratum in: Nature 1997 Sep 25;389(6649):412.

63. Wadström T, Hirmo S, Borén T.Biochemical aspects of *Helicobacter pylori* colonization of the human gastric mucosa.Aliment Pharmacol Ther. 1996 Apr;10 Suppl 1:17-27.

64. Yoshida N, Granger DN, Evans DJ Jr, Evans DG, Graham DY, Anderson DC, Wolf RE, Kvietys PR.Mechanisms involved in *Helicobacter pylori*-induced inflammation.Gastroenterology. 1993 Nov;105(5):1431-40.

65. Zanotti G, Papinutto E, Dundon W, Battistutta R, Seveso M, Giudice G, Rappuoli R, Montecucco C.Structure of the neutrophil-activating protein from *Helicobacter pylori*.J Mol Biol. 2002 Oct 11;323(1):125-30.

66. Zhang J, Zhang X, Wu C, Lu D, Guo G, Mao X, Zhang Y, Wang DC, Li D, Zou Q.Expression, purification and characterization of arginase from *Helicobacter pylori* in its apo form.PLoS One. 2011;6(10):e26205.

CHAPTER II

1. Abergel C, Bouveret E, Claverie JM, Brown K, Rigal A, Lazdunski C, Bénédetti H.Structure of the *Escherichia coli* TolB protein determined by MAD methods at 1.95 A resolution.Structure. 1999 Oct 15;7(10):1291-300.

2. Alm RA, Trust TJ.Analysis of the genetic diversity of *Helicobacter pylori*: the tale of two genomes.J Mol Med (Berl). 1999 Dec;77(12):834-46.

3. Angelini A, Cendron L, Goncalves S, Zanotti G, Terradot L.Structural and enzymatic characterization of HP0496, a YbgC thioesterase from *Helicobacter pylori*.Proteins. 2008 Sep;72(4):1212-21.

4. Böhm G, Muhr R, Jaenicke R.Quantitative analysis of protein far UV circular dichroism spectra by neural networks.Protein Eng. 1992 Apr;5(3):191-5.

5. Bonsor DA, Hecht O, Vankemmelbeke M, Sharma A, Krachler AM, Housden NG, Lilly KJ, James R, Moore GR, Kleanthous C.Allosteric beta-propeller signalling in TolB and its manipulation by translocating colicins.EMBO J. 2009 Sep 16;28(18):2846-57.

6. Bouveret E, Derouiche R, Rigal A, Lloubès R, Lazdunski C, Bénédetti H.Peptidoglycan-associated lipoprotein-TolB interaction. A possible key to explaining the formation of contact sites

between the inner and outer membranes of *Escherichia coli*.J Biol Chem. 1995 May 12;270(19):11071-7.

7. Bumann D, Aksu S, Wendland M, Janek K, Zimny-Arndt U, Sabarth N, Meyer TF, Jungblut PR.Proteome analysis of secreted proteins of the gastric pathogen *Helicobacter pylori*.Infect Immun. 2002 Jul;70(7):3396-403.

8. Cascales E, Buchanan SK, Duché D, Kleanthous C, Lloubès R, Postle K, Riley M, Slatin S, Cavard D.Colicin biology.Microbiol Mol Biol Rev. 2007 Mar;71(1):158-229.

9. Godlewska R, Wiśniewska K, Pietras Z, Jagusztyn-Krynicka EK.Peptidoglycan-associated lipoprotein (Pal) of Gram-negative bacteria: function, structure, role in pathogenesis and potential application in immunoprophylaxis.FEMS Microbiol Lett. 2009 Sep;298(1):1-11.

10. Gully D, Bouveret E.A protein network for phospholipid synthesis uncovered by a variant of the tandem affinity purification method in *Escherichia coli*.Proteomics. 2006 Jan;6(1):282-93.

11. Ilver D, Rappuoli R, Telford JL.Protein Export.In: Mobley HLT, Mendz GL, Hazell SL, editors. *Helicobacter pylori*: Physiology and Genetics. Washington (DC): ASM Press; 2001. Chapter 19.

12. Jensen LJ, Kuhn M, Stark M, Chaffron S, Creevey C, Muller J, Doerks T, Julien P, Roth A, Simonovic M, Bork P, von Mering C.STRING 8--a global view on proteins and their functional interactions in 630 organisms.Nucleic Acids Res. 2009 Jan;37(Database issue):D412-6.

13. Jungblut PR, Bumann D, Haas G, Zimny-Arndt U, Holland P, Lamer S, Siejak F, Aebischer A, Meyer TF.Comparative proteome analysis of *Helicobacter pylori*.Mol Microbiol. 2000 May;36(3):710-25.

14. Keenan J, Oliaro J, Domigan N, Potter H, Aitken G, Allardyce R, Roake J.Immune response to an 18-kilodalton outer membrane antigen identifies lipoprotein 20 as a *Helicobacter pylori* vaccine candidate.Infect Immun. 2000

Jun;68(6):3337-43.

15. Kostrzynska M, O'Toole PW, Taylor DE, Trust TJ.Molecular characterization of a conserved 20-kilodalton membrane-associated lipoprotein antigen of *Helicobacter pylori*.J Bacteriol. 1994 Oct;176(19):5938-48.

16. Lam WW, Woo EJ, Kotaka M, Tam WK, Leung YC, Ling TK, Au SW.Molecular interaction of flagellar export chaperone FliS and cochaperone HP1076 in *Helicobacter pylori*.FASEB J. 2010 Oct;24(10):4020-32.

17. Lazzaroni JC, Germon P, Ray MC, Vianney A.The Tol proteins of *Escherichia coli* and their involvement in the uptake of biomolecules and outer membrane stability.FEMS Microbiol Lett. 1999 Aug 15;177(2):191-7.

18. Levengood SK, Beyer WF Jr, Webster RE.TolA: a membrane protein involved in colicin uptake contains an extended helical region.Proc Natl Acad Sci U S A. 1991 Jul 15;88(14):5939-43.

19. Muller MM, Webster RE.Characterization of the tol-pal and cyd region of *Escherichia coli* K-12: transcript analysis and identification of two new proteins encoded by the cyd operon.J Bacteriol. 1997 Mar;179(6):2077-80.

20. Price MN, Huang KH, Alm EJ, Arkin AP.A novel method for accurate operon predictions in all sequenced prokaryotes.Nucleic Acids Res. 2005 Feb 8;33(3):880-92. Print 2005.

21. Rain JC, Selig L, De Reuse H, Battaglia V, Reverdy C, Simon S, Lenzen G, Petel F, Wojcik J, Schächter V, Chemama Y, Labigne A, Legrain P.The protein-protein interaction map of *Helicobacter pylori*.Nature. 2001 Jan 11;409(6817):211-5.

22. Ray MC, Germon P, Vianney A, Portalier R, Lazzaroni JC.Identification by genetic suppression of *Escherichia coli* TolB residues important for TolB-Pal interaction.J Bacteriol. 2000 Feb;182(3):821-4.

23. Razick S, Magklaras G, Donaldson IM.iRefIndex: a consolidated protein interaction database with provenance.BMC Bioinformatics. 2008 Sep 30;9:405.

24. Szklarczyk D, Franceschini A, Kuhn M, Simonovic M, Roth A, Minguez P, Doerks T, Stark M, Muller J, Bork P, Jensen LJ, von Mering C.The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored.Nucleic Acids Res. 2011 Jan;39(Database issue):D561-8.

25. Tokuda H, Matsuyama S.Sorting of lipoproteins to the outer membrane in *E. coli*.Biochim Biophys Acta. 2004 Nov 11;1694(1-3):IN1-9.

26. Vianney A, Muller MM, Clavel T, Lazzaroni JC, Portalier R, Webster RE.Characterization of the tol-pal region of *Escherichia coli* K-12: translational control of tolR expression by TolQ and identification of a new open reading frame downstream of pal encoding a periplasmic protein.J Bacteriol. 1996 Jul;178(14):4031-8.

27. Zhuang Z, Song F, Martin BM, Dunaway-Mariano D.The YbgC protein encoded by the ybgC gene of the tol-pal gene cluster of *Haemophilus influenzae* catalyzes acyl-coenzyme A thioester hydrolysis.FEBS Lett. 2002 Apr 10;516(1-3):161-3.

CHAPTER III

1. Bijlsma JJ, Waidner B, Vliet AH, Hughes NJ, Häg S, Bereswill S, Kelly DJ, Vandenbroucke-Grauls CM, Kist M, Kusters JG.The *Helicobacter pylori* homologue of the ferric uptake regulator is involved in acid resistance.Infect Immun. 2002 Feb;70(2):606-11.

2. Carpenter BM, Gancz H, Benoit SL, Evans S, Olsen CH, Michel SL, Maier RJ, Merrell DS.Mutagenesis of conserved amino acids of *Helicobacter pylori* fur reveals residues important for function.J Bacteriol. 2010 Oct;192(19):5037-52.

3. Danielli A, Roncarati D, Delany I, Chiarini V, Rappuoli R, Scarlato V.In vivo dissection of the *Helicobacter pylori* Fur regulatory circuit by genome-wide location analysis.J Bacteriol. 2006 Jul;188(13):4654-62.

4. Delany I, Spohn G, Rappuoli R, Scarlato V.The Fur repressor controls transcription of iron-activated and -repressed genes in *Helicobacter pylori*.Mol Microbiol. 2001 Dec;42(5):1297-309.

5. Dhaenens L, Szczebara F, Husson MO.Identification, characterization, and immunogenicity of the lactoferrin-binding protein from *Helicobacter pylori*.Infect Immun. 1997 Feb;65(2):514-8.

6. Dian C, Vitale S, Leonard GA, Bahlawane C, Fauquant C, Leduc D, Muller C, de Reuse H, Michaud-Soret I, Terradot L.The structure of the *Helicobacter pylori* ferric uptake regulator Fur reveals three functional metal binding sites.Mol Microbiol. 2011 Mar;79(5):1260-75. doi: 10.1111/j.1365-2958.2010.07517.x.

7. Ernst FD, Bereswill S, Waidner B, Stoof J, Mäder U, Kusters JG, Kuipers EJ, Kist M, van Vliet AH, Homuth G.Transcriptional profiling of *Helicobacter pylori* Fur-and iron-regulated gene expression.Microbiology. 2005 Feb;151(Pt 2):533-46.

8. Ernst FD, Homuth G, Stoof J, Mäder U, Waidner B, Kuipers EJ, Kist M, Kusters JG, Bereswill S, van Vliet AH.Iron-responsive regulation of the *Helicobacter pylori* iron-cofactored superoxide dismutase SodB is mediated by Fur.J Bacteriol. 2005 Jun;187(11):3687-92.

9. Gancz H, Merrell DS.The *Helicobacter pylori* Ferric Uptake Regulator (Fur) is essential for growth under sodium chloride stress.J Microbiol. 2011 Apr;49(2):294-8.

10. Gancz H, Censini S, Merrell DS.Iron and pH homeostasis intersect at the level of Fur regulation in the gastric pathogen *Helicobacter pylori*.Infect Immun. 2006 Jan;74(1):602-14.

11. Gilbreath JJ, Cody WL, Merrell DS, Hendrixson DR.Change is good: variations in common biological mechanisms in the epsilonproteobacterial genera Campylobacter and Helicobacter.Microbiol Mol Biol Rev. 2011 Mar;75(1):84-132.

12. Lee JS, Choe YH, Lee JH, Lee HJ, Lee JH, Choi YO.*Helicobacter pylori* urease activity is influenced by ferric uptake regulator.Yonsei Med J. 2010 Jan;51(1):39-44. Epub 2009 Dec 29.

13. Merrell DS, Thompson LJ, Kim CC, Mitchell H, Tompkins LS, Lee A, Falkow S.Growth phase-dependent response of *Helicobacter pylori* to iron starvation.Infect Immun. 2003 Nov;71(11):6510-25.

14. Miles S, Piazuelo MB, Semino-Mora C, Washington MK, Dubois A, Peek RM Jr, Correa P, Merrell DS.Detailed in vivo analysis of the role of *Helicobacter pylori* Fur in colonization and disease.Infect Immun. 2010 Jul;78(7):3073-82.

15. Miles S, Carpenter BM, Gancz H, Merrell DS.*Helicobacter pylori* apo-Fur regulation appears unconserved across species.J Microbiol. 2010 Jun;48(3):378-86

16. Payne SM, Wyckoff EE, Murphy ER, Oglesby AG, Boulette ML, Davies NM.Iron and pathogenesis of Shigella: iron acquisition in the intracellular environment.Biometals. 2006 Apr;19(2):173-80.

17. van Vliet AHM, Bereswill S, Kusters JG.Ion Metabolism and Transport.In: Mobley HLT, Mendz GL, Hazell SL, editors. *Helicobacter pylori*: Physiology and Genetics. Washington (DC): ASM Press; 2001. Chapter 17.

18. Vitale S, Fauquant C, Lascoux D, Schauer K, Saint-Pierre C, Michaud-Soret I.A ZnS(4) structural zinc site in the *Helicobacter pylori* ferric uptake regulator.Biochemistry. 2009 Jun 23;48(24):5582-91.

19. Wang G, Alamuri P, Maier RJ.The diverse antioxidant systems of *Helicobacter pylori*.Mol Microbiol. 2006 Aug;61(4):847-60.

20. Worst DJ, Maaskant J, Vandenbroucke-Grauls CM, Kusters JG.Multiple haem-utilization loci in *Helicobacter pylori*.Microbiology. 1999 Mar;145 (Pt 3):681-8. Erratum in: Microbiology 1999 Apr;145(Pt 4):985.

CHAPTER IV

1. Atanassova A, Zamble DB.*Escherichia coli* HypA is a zinc metalloprotein with a weak affinity for nickel.J Bacteriol. 2005 Jul;187(14):4689-97.

2. Benoit SL, Maier RJ.Hydrogen and nickel metabolism in helicobacter species.Ann N Y Acad Sci. 2008 Mar;1125:242-51.

3. Gasper R, Scrima A, Wittinghofer A.Structural insights into HypB, a GTP-binding protein that regulates metal binding.J Biol Chem. 2006 Sep 15;281(37):27492-502.

4. Hube M, Blokesch M, Böck A.Network of hydrogenase maturation in *Escherichia coli*: role of accessory proteins HypA and HybF.J Bacteriol. 2002 Jul;184(14):3879-85.

5. Maier RJ.Availability and use of molecular hydrogen as an energy substrate for Helicobacter species.Microbes Infect. 2003 Oct;5(12):1159-63.

6. Maier RJ, Benoit SL, Seshadri S.Nickel-binding and accessory proteins facilitating Ni-enzyme maturation in *Helicobacter pylori*.Biometals. 2007 Jun;20(3-4):655-64.

7. Mehta N, Olson JW, Maier RJ.Characterization of *Helicobacter pylori* nickel metabolism accessory proteins needed for maturation of both urease and hydrogenase.J Bacteriol. 2003 Feb;185(3):726-34.

8. Mehta N, Benoit S, Maier RJ.Roles of conserved nucleotide-binding domains in accessory proteins, HypB and UreG, in the maturation of nickel-enzymes required for efficient *Helicobacter pylori* colonization.Microb Pathog. 2003 Nov;35(5):229-34.

9. Olson JW, Mehta NS, Maier RJ.Requirement of nickel metabolism proteins HypA and HypB for full activity of both hydrogenase and urease in *Helicobacter pylori*.Mol Microbiol. 2001 Jan;39(1):176-82. Erratum in: Mol Microbiol 2001 Apr;40(1):270.

10. Olson JW, Maier RJ.Molecular hydrogen as an energy source for *Helicobacter pylori.*Science. 2002 Nov 29;298(5599):1788-90.

11. Sydor AM, Liu J, Zamble DB.Effects of metal on the biochemical properties of *Helicobacter pylori* HypB, a maturation factor of [NiFe]-hydrogenase and urease.J Bacteriol. 2011 Mar;193(6):1359-68.

12. Vignais PM, Billoud B.Occurrence, classification, and biological function of hydrogenases: an overview.Chem Rev. 2007 Oct;107(10):4206-72.

APPENDIX

1. Bamann C, Kirsch T, Nagel G, Bamberg E.Spectral characteristics of the photocycle of channelrhodopsin-2 and its implication for channel function.J Mol Biol. 2008 Jan 18;375(3):686-94.

2. Baylor D.How photons start vision.Proc Natl Acad Sci U S A. 1996 Jan 23;93(2):560-5.

3. Boyden ES, Zhang F, Bamberg E, Nagel G, Deisseroth K.Millisecond-timescale, genetically targeted optical control of neural activity.Nat Neurosci. 2005 Sep;8(9):1263-8.

4. Gordeliy VI, Labahn J, Moukhametzianov R, Efremov R, Granzin J, Schlesinger R, Büldt G, Savopol T, Scheidig AJ, Klare JP, Engelhard M.Molecular basis of transmembrane signalling by sensory rhodopsin II-transducer complex.Nature. 2002 Oct 3;419(6906):484-7.

5. Govorunova EG, Jung KH, Sineshchekov OA.[Phototaxis of the green algae: the new class of rhodopsin receptors].Biofizika. 2004 Mar-Apr;49(2):278-93.

6. Hegemann P.Algal sensory photoreceptors.Annu Rev Plant Biol. 2008;59:167-89.

7. Hegemann P, Möglich A.Channelrhodopsin engineering and exploration of new optogenetic tools.Nat Methods. 2011 Jan;8(1):39-42.

8. Jung KH, Trivedi VD, Spudich JL.Demonstration of a sensory rhodopsin in eubacteria.Mol Microbiol. 2003 Mar;47(6):1513-22.

9. Kaneko A.Our recent studies on sensory transduction: from vision to taste.Keio J Med. 2001 Mar;50(1):13-9.

10. Kosower EM.Sensory mechanisms on the molecular level.Endeavour. 1992;16(1):23-8.

11. LaLumiere RT.A new technique for controlling the brain: optogenetics and its potential for use in research and the clinic.Brain Stimul. 2011 Jan;4(1):1-6.

12. Lanyi JK.Bacteriorhodopsin.Annu Rev Physiol. 2004;66:665-88.

13. Lin JY.A user's guide to channelrhodopsin variants: features, limitations and future developments.Exp Physiol. 2011 Jan;96(1):19-25.

14. Lin JY, Lin MZ, Steinbach P, Tsien RY.Characterization of engineered channelrhodopsin variants with improved properties and kinetics.Biophys J. 2009 Mar 4;96(5):1803-14.

15. Luecke H, Schobert B, Lanyi JK, Spudich EN, Spudich JL. Crystal structure of sensory rhodopsin II at 2.4 angstroms: insights into color tuning and transducer interaction.Science. 2001 Aug 24;293(5534):1499-503.

16. Nagel G, Szellas T, Huhn W, Kateriya S, Adeishvili N, Berthold P, Ollig D, Hegemann P, Bamberg E. Channelrhodopsin-2, a directly light-gated cation-selective membrane channel.Proc Natl Acad Sci U S A. 2003 Nov 25;100(24):13940-5.

17. Nagel G, Szellas T, Kateriya S, Adeishvili N, Hegemann P, Bamberg E.Channelrhodopsins: directly light-gated cation channels.Biochem Soc Trans. 2005 Aug;33(Pt 4):863-6.

18 Kno¨pfel T, Lin MZ, Levskaya A, Tian L, Lin JY, Boyden ES Toward the Second Generation of Optogenetic Tools. J of Neurosc. 2010 Nov 10;30(45): 14998-15004.

19. Sasaki J, Spudich JL.Proton transport by sensory rhodopsins and its modulation by transducer-binding.Biochim Biophys Acta. 2000 Aug 30;1460(1):230-9.

20. Schoenenberger P, Schärer YP, Oertner TG.Channelrhodopsin as a tool to investigate synaptic transmission and plasticity.Exp Physiol. 2011 Jan;96(1):34-9. Epub 2010 Jun 18.

21. Spudich J.Spotlight on receptor/transducer interaction.Nat Struct Biol. 2002 Nov;9(11):797-9.

22. Spudich JL, Yang CS, Jung KH, Spudich EN.Retinylidene proteins: structures and functions from archaea to humans.Annu Rev Cell Dev Biol. 2000;16:365-92.

23. Spudich JL.Variations on a molecular switch: transport and sensory signalling by archaeal rhodopsins.Mol Microbiol. 1998 Jun;28(6):1051-8.

24. Vogeley L, Luecke H.Crystallization, X-ray diffraction analysis and SIRAS/molecular-replacenent phasing of three crystal forms of Anabaena sensory rhodopsin transducer.Acta Crystallogr Sect F Struct Biol Cryst Commun. 2006 Apr 1;62(Pt 4):388-91.

25. Vogeley L, Sineshchekov OA, Trivedi VD, Sasaki J, Spudich JL, Luecke H.Anabaena sensory rhodopsin: a photochromic color sensor at 2.0 A.Science. 2004 Nov 19;306(5700):1390-3.

26. Vogeley L, Trivedi VD, Sineshchekov OA, Spudich EN, Spudich JL, Luecke H.Crystal structure of the Anabaena sensory rhodopsin transducer.J Mol Biol. 2007 Mar 30;367(3):741-51.

27. Wang S, Kim SY, Jung KH, Ladizhansky V, Brown LS. A eukaryotic-like interaction of soluble cyanobacterial sensory rhodopsin transducer with DNA. J Mol Biol. 2011 Aug 12;411(2):449-62.

Acknowledgments

I would like to express my most sincere gratitude to my supervisor Prof. Giuseppe Zanotti for the knowledge that he has given to me and the many doors that he have always kept open. I will not thank him enough for having given me the opportunity to enter into the world of crystallography and for his rare ability to light the way with kindness and elegance.

Special thanks to Dr Laura Cendron, for constantly encouraging me with her presence and lively enthusiasm. Her wide-ranging advice has been invaluable to face the challenge of this work.

I would like to thank also Dr.Ilenia Bertipaglia for the great help and the good mood in working that she transmitted to the lab.

I am also grateful to Dr.Monica Santarosa for her divine ability to transform and create.

Thanks to all the colleagues with whom I had the opportunity to work: Nicola, Lorenza, Munan, Sara, Giorgia, and Marco.

Finally, I would like to express my special gratitude to Prof. Hartmut Luecke "Hudel" for the opportunity to visit his lab at UCI-University of California, Irvine. Despite having spent a brief period in Hudel's lab, I'm still glad to have approached a so exciting field like the characterization of channelrhodopsins and to have known good friends like Reggie, Anna, Emanuele, and Jahan to whom this dissertation is dedicated.