UNIVERSITA' DEGLI STUDI DI PADOVA

Dipartimento di Scienze Chimiche

Scuola di Dottorato di Ricerca in Scienze Molecolari

Indirizzo: Scienze Chimiche

XXII Ciclo

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

Direttore della Scuola: Ch.mo Prof. Maurizio Casarin

Co-Supervisore: Ch.mo Prof. Giuseppe Zanotti

Supervisore: Ch.mo Prof. Roberto Battistutta

Dottorando: Nicola Barison

31 Gennaio 2010

UNIVERSITY OF PADUA

Department of Biological Chemistry

Doctorate School of Molecular Sciences

Specialization: Chemical Sciences

Thesis

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

Nicola Barison

Graduate School director:	Professor Maurizio Casarin *(Department of Chemical Sciences, University of Padua, Padua, Italy)*
Co-Supervisor:	Professor Giuseppe Zanotti (*Department of Biological Chemistry, University of Padua and VIMM, Padua, Italy*)
Supervisor:	Professor Roberto Battistutta (*Department of Chemical Sciences, University of Padua and VIMM, Padua, Italy*)
Supervisor at MPIIB:	Professor Thomas F. Meyer (*Max Planck Institute of Infection Biology, Berlin, Germany*)

To Diane

Contents

Summary

Helicobacter pylori is a Gram-negative microaerophilic bacterium that is able to establish a life-long chronic infection in the stomach of more than half of the human population. The infection is most of the times asymptomatic but, in an important minority of cases, *H. pylori* causes gastroduodenal pathologies, including stomach and duodenal ulcers, adenocarcinomas and stomach lymphomas. *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 traslocation 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. Other properties affected by genetic variability are expression and binding ability of adhesins. All these features are involved in pathogenical variability of *H. pylori*, in addition to HP-NAP, a toxin able to activate immunological response (Chapter I).

Although *H. pylori* can be successfully eradicated by antibiotics in many patients, increasing antibiotic resistance in the bacterium remains a serious problem. New therapies are required to eradicate *H. pylori* infection and looking for new pharmaceutical targets could be useful for the development of new antibiotics in the future.

Recently, new genes involved in *H. pylori* stomach colonization were identified by mutagenesis. Of the genes analyzed, 29% had a predicted stomach colonization effect. These included known *H. pylori* virulence genes, genes implicated in virulence in other pathogenic bacteria, and genes codifying for hypothetical proteins.

The aim of this project was to determine the three-dimensional structure of *H. pylori* proteins important for stomach colonization and pathogenesis and to obtain information about their putative functions. In particular, we focalized our attention on the TenA homologue HP1287, on the putative adhesin HP1028, on HP0175, a toxin with PPIase activity, and on HP0421, a cholesterol-α-glucosyltransferase that promotes immune evasion.

The research described in this thesis was mostly carried out at the Department of Biological Chemistry, University of Padua and Venetian Institute of Molecular Medicine (VIMM), Padua.

The strategy adopted involved bioinformatic analyses, PCR amplification of the selected genes starting from purified *H. pylori* chromosomal DNA (strain CCUG 17874), cloning in an His-tag-containing vector and expression of the protein in *E. coli* cells. The soluble recombinant proteins were then purified using two chromatography steps and concentrated for crystallization trials. The proteins HP1028 and HP1287 were subsequently crystallized and, in the case of HP1287, the structure was determined by X-Ray diffraction. Techniques to make the protein samples more suitable for crystallization, as DLS (Dynamic light scattering), and to investigate the secondary structure, as CD (Circular dichroism), were also used. Mutagenesis techniques were included for functional studies and for crystallographic purposes.

After a general introduction, the structural and functional analysis of HP287 from *H. pylori* is described in Chapter II. HP1287 is a homologue of TenA from *B. subtilis*, an enzyme involved in thiamin precursors salvage. The protein was cloned, expressed, purified and crystallized. The structure was solved by Molecular Replacement and refined at 2.7 Å. The structure reveals a very high similarity with TenA from *B. subtilis*, with the only difference of the substitution of one amino acid in the active site (a tyrosine substituted with a phenylalanine in HP1287). Since HP1287 was not active on the same substrate of TenA from *B. subtilis*, a HP1287 F47Y mutant was created to investigate on the importance of Y47 in restoring the activity, but the HP1287 activity remained low. Moreover, the position of the H86 in the HP1287 active site seems to suggest an additional interaction with the substrate, not possible in TenA from *B. subtilis*. These findings suggested that the enzyme is specific for another slightly different substrate, that probably comes from the thiamin degradation in the stomach.

The third chapter describes the cloning, expression, purification and crystallization of HP1028, a putative adhesin important for stomach colonization. To identify a detergent able to inhibit aggregation, protein samples were incubated with several detergents and analyzed by DLS (Dynamic Light Scattering). Since no homologue structures are known, to get initial phases to solve the structure, a selenomethionine derivative was produced. To improve the anomalous signal of selenium, a double mutant was created, adding two methionines in the amino acid sequence. Preliminary data of native and mutant HP1028 were collected at 2.4 Å.

HP0175 is a Peptidyl Prolyl *cis,trans*-Isomerase that induces apoptosis of gastric epithelial cells through TLR4 (described in chapter IV). After cloning, the protein was expressed and purified in very high yield and its secondary structure was analyzed by CD. To reduce the excessively high solubility of the protein, a lysines methylation approach was adopted. Some crystallization trials were performed.

Finally, the last chapter (chapter V) describes HP0421, a cholesterol-α-glucosyltransferase that promotes immune evasion. *H. pylori* cells containing high levels of cholesterol are easily sequestered by macrophages. Intrinsic α-glucosylation of cholesterol abrogates phagocytosis of *H. pylori* and T cell activation. This part of the PhD project was carried out at the Max Planck Institute for Infection Biology in Berlin, under the supervision of Prof. Thomas F. Meyer. The recombinant protein, expressed in *E. coli*, did not bind strongly to the affinity IMAC-Ni^{2+} resin, probably because of the degradation of the C-terminal His-tag. To solve the purification problems, two approaches were adopted: the cloning of a new HP0421 construct with an N-terminal His-tag and the purification and the refolding of the HP0421 protein from the pellet, verified by activity tests. Moreover, DLS analysis of HP0421, purified in small yield, showed that the protein was aggregated. Aggregation inhibition tests with detergents were attempted.

Sommario

Helicobacter pylori è un batterio microaerofilo Gram-negativo in grado di instaurare un'infezione cronica nello stomaco umano in più della metà della popolazione mondiale. L'infezione è generalmente asintomatica ma, in alcuni casi, *H. pylori* causa patologie gastroduodenali, quali ulcere gastriche e duodenali, adenocarcinomi e linfomi gastrici. *H. pylori* è caratterizzato da elevata variabilità genetica, non solo all'interno delle sequenze geniche ma anche nel contenuto genico. Una delle differenze più importanti nei ceppi *di H. pylori* è la presenza o l'assenza di una sequenza di DNA di 40 kb chiamata Isola di Patogenicità *cag*, che codifica per un Sistema di Secrezione di tipo IV, che provoca la traslocazione della tossina CagA nelle cellule epiteliali gastriche. Dopo l'entrata nella cellula ospite, CagA induce modifiche cellulari come l'alterazione della struttura cellulare e della motilità, l'alterazione della proliferazione cellulare e delle tight junctions. La maggior parte dei ceppi di *H. pylori* secernono VacA, una tossina che provoca vacuolizzazione delle cellule epiteliali dello stomaco. Tutti i ceppi di *H. pylori* contengono il gene *vacA*, ma la sequenza genica è caratterizzata da un'elevata variabilità, causando modifiche nella funzionalità della tossina. Altre caratteristiche condizionate dalla variabilità genetica sono l'espressione e la forza di legame delle adesine. Tutte queste caratteristiche sono coinvolte nella diversità del grado di virulenza di *H. pylori*, assieme a HP-NAP, una tossina in grado di attivare la risposta immunitaria (Capitolo I).

Sebbene la terapia antibiotica sia in grado di eradicare *H. pylori* nella maggior parte dei pazienti, la resistenza ai chemioterapici da parte dei batteri rimane un problema sempre crescente. Per questo motivo, sono necessarie nuove terapie per l'eradicazione del batterio e, in questo contesto, è importante identificare nuovi bersagli farmaceutici per lo sviluppo di nuovi antibiotici specifici.

Recentemente uno studio di mutagenesi sistematica del genoma di *H. pylori* ha permesso d'identificare nuovi geni coinvolti nella colonizzazione dello stomaco. Dei geni analizzati, 29% hanno evidenziato un effetto nella colonizzazione gastrica. Tra questi sono stati identificati geni di virulenza già noti, geni importanti per la virulenza in altri patogeni e geni codificanti per proteine ignote.

Scopo di questo progetto di dottorato è determinare la struttura tridimensionale di proteine di *H. pylori* importanti per la colonizzazione dello stomaco e la patogenesi e ottenere informazioni sulla loro putativa funzione. In particolare, sono state studiate le proteine

HP1287, omologo di TenA, la putativa adesina HP1028, HP0175, una tossina con attività PPIasica e HP0421, una colesterol-α-glucosiltrasferasi che inibisce la risposta immunitaria.
La maggior parte del progetto è stata effettuata presso il laboratorio di Biologia Strutturale del Prof. Zanotti al Dipartimento di Chimica Biologica dell'Università di Padova e presso l'Istituto Veneto di Medicina Molecolare (VIMM) di Padova.
La strategia di lavoro adottata comprendeva analisi bioinformatiche, amplificazione dei geni selezionati mediante PCR a partire dal DNA genomico di *H. pylori* (ceppo CCUG 17874), clonaggio in un vettore di espressione in fusione con un His-tag N-terminale e l'espressione della proteina in cellule batteriche di *Escherichia coli*. Le proteine ricombinanti solubili sono state poi purificate usando due passaggi cromatografici e concentrate per le prove di cristallizzazione. Le proteine HP1028 e HP1287 sono state successivamente cristallizzate e, nel caso di HP1287, è stata risolta la struttura mediante diffrattometria ai raggi X. Inoltre, sono state utilizzate anche tecniche per verificare la qualità del campione per la cristallizzazione, come il DLS (Dynamic Light Scattering) e per analizzare la struttura secondaria, come il Dicroismo circolare (CD). Tecniche di mutagenesi sono state impiegate per studi funzionali e per fini cristallografici.
Dopo un'introduzione generale, nel capitolo II è descritta l'analisi strutturale e funzionale di HP1287 di *H. pylori*. HP1287 è un omologo di TenA di *B. subtilis*, un enzima coinvolto nel recupero di precursori della tiamina. La proteina è stata clonata, espressa, purificata e cristallizzata. La struttura è stata risolta mediante Sostituzione Molecolare ad una risoluzione di 2.7 Å. La struttura evidenzia un elevato grado di similarità con TenA di *B. subtilus*, con la sola differenza della sostituzione di un amminoacido nel sito attivo (una tirosina sostituita con una fenilalanina in HP1287). Poiché HP1287 sembrava non essere attiva nello stesso substrato di TenA di *B. subtilis*, per verificare l'importanza della tirosina in posizione 47 nel ripristino dell'attività catalitica, è stato creato un mutante F47Y, che però non ha evidenziato una maggior capacità di catalisi rispetto alla HP1287 nativa. Inoltre, la collocazione dell'istidina 86 nel sito attivo di HP1287, sembrerebbe suggerire un'interazione dell'amminoacido con il substrato non possibile nel caso di TenA di *B. subtilis*. Tutte queste informazioni suggeriscono come questa proteina potrebbe metabolizzare un composto leggermente diverso da quello di TenA di *B. subtilis*, che deriva dalla degradazione della tiamina nello stomaco.
Il terzo capitolo descrive il clonaggio, l'espressione, la purificazione e la cristallizzazione di HP1028, una putativa adesina importante per la colonizzazione dello stomaco. Per identificare un detergente in grado d'inibire l'aggregazione, campioni proteici sono stati incubati con

numerosi detergenti e analizzati mediante DLS (Dynamic Light Scattering). Non essendo nota alcuna struttura di alcuna proteina omologa, per ottenere fasi approssimate iniziali per la risoluzione della struttura, la proteina HP1028 è stata espressa sostituendo le metionine con selenometionine e purificata. Per amplificare il segnale anomalo del selenio, è stato creato un doppio mutante contenente due ulteriori metionine nella sequenza amminoacidica. Dati preliminari della proteina HP1028 nativa e mutante sono stati raccolti ad una risoluzione massima di 2.4 Å.

HP0175 è una Peptidil-Prolil-*cis,trans*-Isomerasi che induce l'apoptosi di cellule gastriche attraverso il TLR4 (descritta nel capitolo IV). Dopo il clonaggio del gene nel plasmide d'espressione, la proteina è stata espressa e purificata in elevata quantità e la struttura secondaria è stata analizzata mediante CD. Per ridurre l'eccessivamente alta solubilità della proteina, è stata adottata una metodica di metilazione delle lisine e sono state effettuate numerose prove di cristallizzazione della proteina derivatizzata.

Infine, nell'ultimo capitolo (Capitolo V) è descritta la proteina HP0421, una colesterol-α-glucosiltransferasi che promuove l'immuno-evasione. Cellule batteriche di *H. pylori* contenenti alti livelli di colesterolo sono facilmente fagocitate dai macrofagi. La glicosilazione del colesterolo inibisce l'uptake di *H. pylori* da parte dei macrofagi e l'attivazione dei linfociti T. Questa parte del progetto di dottorato è stata svolta presso l'istituto Max Planck di Berlino, sotto la supervisione del Prof. Thomas F. Meyer. La proteina ricombinante, espressa in *E. coli*, sembrava non essere in grado di legare la resina d'affinità IMAC-Ni^{2+}, probabilmente a causa della degradazione dell'His-tag C-terminale. Per risolvere questi problemi durante la fase di purificazione, sono stati adottati due approcci: il clonaggio d'un nuovo costrutto di HP0421 con un His-tag N-terminale e la purificazione e il refolding della proteina HP0421 dal pellet batterico, validato da test di attività. Inoltre, studi di DLS su HP0421, purificata in modeste quantità, hanno evidenziato lo stato di aggregazione della proteina. Per inibire l'interazione aspecifica tra le molecole proteiche, sono state effettuate prove di disaggregazione con detergenti, verificate mediante analisi DLS.

Chapter I

General introduction

1.1 *Helicobacter pylori*

Helicobacter pylori is a Gram-negative bacterium that colonizes the human stomach, an ecological niche characterized by very acidic pH, a lethal condition for most microbes. *H. pylori* is so well adapted to this unfriendly environment that it is able to establish a life-long chronic infection (Montecucco, Rappuoli 2001). The presence of microbes was already known since 1893 (Bizzozero 1893), but *Helicobacter pylori* was identified only in 1982 by Marshall and Warren (Marshall, Warren 1984), who won the Nobel Prize for this discovery in 2005. In the last 27 years, this bacterium became object of intense studies because of its correlation with several gastric diseases. *H. pylori* is an S-shaped microaerophilic bacterium, 0.5 × 5 μm in length with a tuft of 5 to 7 polar sheathed flagella (Mobley, Mendz & Hazell 2001). *H. pylori* is colonizing the stomach of more than half of the human population; the infection is often acquired early in life (almost always before the age of 10 years) by fecal-oral or oral-oral route (Cover, Blaser 2009). Most infected people are asymptomatic, with moderate inflammation detectable only by biopsy and histology. However, an important minority of them (15–20%) develop during their life severe gastroduodenal pathologies, including stomach and duodenal ulcers, adenocarcinomas and stomach lymphomas (Montecucco, Rappuoli 2001). *H. pylori* is the only formally recognized definitive bacterial carcinogen for humans and is estimated to be responsible for 5.5% of all human cancer cases, or approximately 592,000 gastric cancer cases per year (Suerbaum, Josenhans 2007).

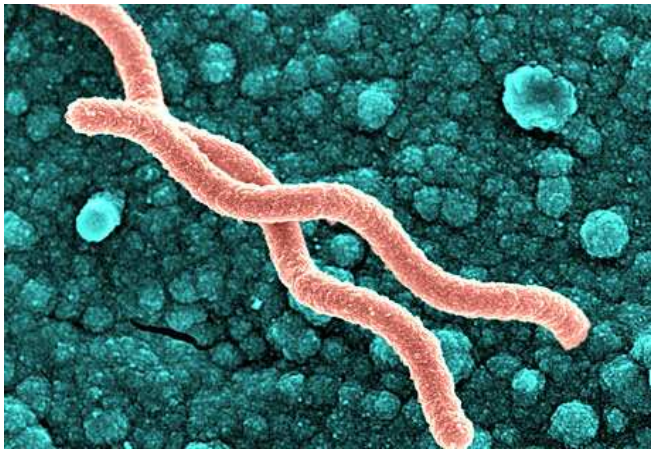

Fig. 1.1. Scanning electron microscopy image of *H. pylori.* Adapted from Encarta web site.

1.2. Genetic variability of *H. pylori*

The first genome of *H. pylori*, strain 26695, was sequenced in 1997: it consists of a circular chromosome with a size of 1,667,867 base pairs (bp), an average G + C content of 39% and 1,590 predicted coding sequences (Tomb et al. 1997). Following the discovery of *H. pylori*, it was noted that this pathogen has an extraordinary genetic heterogeneity, and that almost every isolate from unrelated patients appears to have a unique 'fingerprint'. *H. pylori* may have evolved in a host different from humans and possibility he has made a species jump about 10,000 years ago (Dailidiene et al. 2004; Suerbaum, Achtman 2004). *H. pylori* has

accompanied ancient human migrations and traces of those migrations still remains in modern populations (Suerbaum, Achtman 2004). Individuals can be colonized with multiple strains and strains have been shown to change during chronic colonization (Suerbaum, Josenhans 2007). The polymorphism of *H. pylori* are the result of point mutations, substitution, insertions and deletions of one or more genes and of chromosome arrangements. A particular group of polymorphisms are single or double bases short-tandem repeats close to 5' of contingency genes, that are hypermutable owing to slippage within the DNA repeats. This slippage results in frequent shifting into and out of frame, leading to on-off switching of the associated gene products, and thereby phase or antigenic variation (Saunders et al. 1998). Because of the types in genes in which these repeats are found, this phenomenon could improve the adaptation to changes of the host environment. Depending on the number (*n*) of these contingency genes, an *H. pylori* population colonizing a host's stomach may consist of as many as 2^n competing bacterial lineages. Given the divergence among *H. pylori* strains, the population diversity and overall fitness of the species can be further enhanced by recombination among different bacterial lineages (which may often occur by DNA transformation, conjugation and R/M restriction/modification systems), because it generates new potentially advantageous genotypes much more efficiently than by simple recurrent mutation (Blaser, Berg 2001; Kang, Blaser 2006). Moreover, the anatomical and physiological variation of macroniches in the stomach provides one basis for the diversification and levels of complexity found in subpopulations of *H. pylori* (Kang, Blaser 2006).

1.3. Survival systems in *H. pylori*

H. pylori colonizes the human stomach, which, due to the highly acidic pH, is a hostile environment for most microorganisms and is considered the first line of defense against most gastrointestinal pathogens. It has been estimated that exposure to gastric acids kills more than 99.9 percent of ingested Salmonella and Vibrio (Gorden, Small 1993). Although *H. pylori* is a neutralophile, it is able not only to survive in but also to colonize the human stomach: it has evolved specialized processes that allow maintenance of the pH of its cytoplasm and of the surrounding liquid at a level that enables the organisms to survive and grow (Sachs et al. 2005). *H. pylori* survives in this acidic environment by producing abundant quantities of urease, which hydrolyses urea into ammonia and CO_2, resulting in the generation of a pH neutral micro-environment (Dunn, Phadnis 1998). The amount of urease produced by the bacterium varies with culture conditions and may reach as much as 10% of the total bacterial

proteins (Montecucco, Rappuoli 2001). 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 or free 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).

Urea is taken up by *H. pylori* through a proton-gated channel, and its hydrolysis by urease generates ammonia that buffers the cytosol and the periplasm, and creates a neutral layer around the bacterial surface (Montecucco, Rappuoli 2001; Weeks et al. 2000). *H. pylori* urease is a dodecamer and consists of two different subunits of 61.7 kDa (α) and 26.5 kDa (β), forming a huge complex whose molecular mass has been estimated to be about 600 kDa. TEM (*Transmission Electron Microscopy*) studies reveal an 1.1 MDa double ring dodecameric assembly of 13 nm diameter (Ha et al. 2001). The biosynthesis of urease is governed by a seven-gene cluster, including the subunits of the urease and the accessory proteins that are responsible for Ni^{2+} uptake and insertion into the active site of the apo-enzyme (Ha et al. 2001); gene *ureI* encodes a membrane protein with homology to putative amide transporters and its absence impairs acidic survival. UreI is an acid-activated urea transporter crucial for acid resistance of *H. pylori* (Weeks et al. 2000). It shows characteristics of a pH-gated urea channel (nonsaturable, voltage independent), having low open probability at neutral pH and high open probability at pH 5.0 and below, with half-maximal opening at pH 5.9 (Weeks et al. 2000).

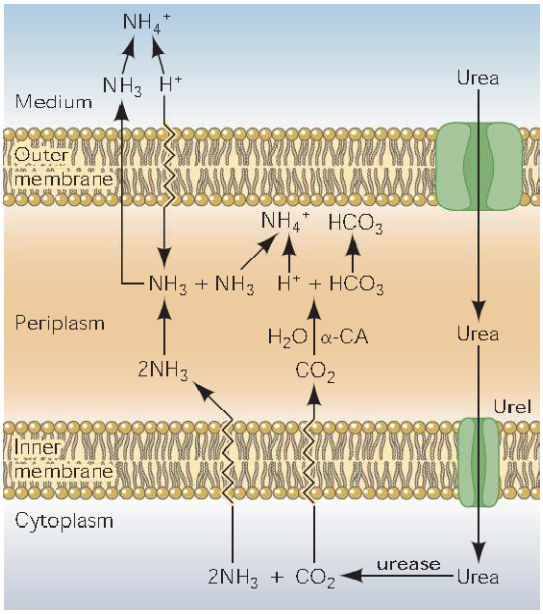

Fig. 1.2. *H. pylori* pH-buffering mechanism. With acidification, UreI opens and urea moves into the cytoplasm, increasing intrabacterial urease activity. Urease produces $2NH_3$ and CO_2, gases that readily exit the cell into the periplasm. Protons entering the cytoplasm are neutralized by NH_3, forming NH_4^+, whereas cytoplasmic carbonic anhydrase generates HCO_3^-, which is a stronger buffer at neutral pH than NH_3. Similarly, the NH_3 that effluxes into the periplasm can neutralize entering acidity and protons coming from the periplasmic carbonic anhydrase activity, that produces HCO_3^- and H^+ from CO_2. A second NH_4^+ is formed along with HCO_3^-, the latter providing buffering in the range of pH 6.1 (Adapted from Sachs et al. 2005).

Although the NH_3 produced by urease activity is able to neutralize entering protons, the pKa of the NH_4^{+}/NH_3 pair is 9.2 and would not buffer the periplasm effectively at a relatively neutral pH, allowing potentially lethal transient acidification of the periplasm. Therefore, NH_3 alone would not be able to maintain the periplasmic pH relatively constant at 6.1. On the other hand, pH 6.1 is the effective pKa of

HCO_3^-, generated by carbonic anhydrase. Because of this, both NH_3 and CO_2 production from intrabacterial urease activity enable acid acclimation by *H. pylori*, NH_3 to neutralize entering protons and HCO_3^- to buffer the periplasm (Sachs et al. 2005).

Other important genes involved in acid regulation are HP0165 and HP0166, which form the ArsSR system. Recent studies have shown that the two-component system ArsRS is one of the mechanisms responsible for regulation of many genes in *H. pylori* in response to an acidic environment, which regulates most of the acid acclimation genes identified so far (Pflock et al. 2006a). In *H. pylori*, the environmental pH is probably sensed directly by the histidines in the periplasmic domain of the histidine kinase sensor protein ArsS (HP0165), which autophosphorylates itself in response to pH changes and transfers the phosphoryl group to its cognate response regulator, ArsR (HP0166). The phosphorylated ArsR may function both as an activator and repressor of pH-responsive target genes by interacting with the promoter regions (Pflock et al. 2006b). In particular, the phosphorylated ArsR induced by low pH acts as an activator in interacting with one of the promoters of the HP1186 carbonic anhydrase gene (Wen, Moss 2009).

1.4. Stomach colonization

H. pylori is a good swimmer and reaches the thick mucus layer that covers and protects the epithelial lining of the stomach mucosa. Here, propelled by its flagella, the helicoidal-shaped bacterium travels across the viscous mucus film like a screw into a cork; non-motile mutants cannot colonize the stomach (Montecucco, Rappuoli 2001).

However, most *H. pylori* avoid the acidic lumen of the stomach by swimming toward the mucosal cell surface, using their polar flagella and a chemotaxis mechanism. Swimming and chemotaxis are important to avoid being swept into hostile microenvironments and cleared with the mucous flow (Amieva, El-Omar 2008). The motility of *H. pylori* is thus considered a colonization factor, since there is evidence from several studies that less motile strains are less able to colonize or survive in the host than fully motile strains (O'Toole 2000). The structure of the *H. pylori* flagellum is similar to those of the bacteria *E. coli* and *S. thyphimurium* (Mobley, Mendz & Hazell 2001): it is composed of a basal body, a hook and a flagellar filament. The flagella consist mainly of the flagellins FlaA and FlaB. Both genes coding for these flagellins are necessary for the full motility of *H. pylori* (Andersen 2007). The flagellins have a similar molecular mass of 53 kDa and share a quite high sequence homology (58% identity) (Suerbaum, Josenhans & Labigne 1993). The 70 x 16 nm *H. pylori* hook is composed of FlgE subunits of 78kDa, whereas *fliD* gene encodes a hook-associated that helps

the flagellin monomers to be incorporated into the growing flagellum (O'Toole, Kostrzynska & Trust 1994). Other genes are involved in flagella building and regulation but their roles are still unclear (O'Toole 2000).

1.5. Adhesion to the gastric cells

H. pylori comes into contact with the mucin layer that covers the epithelial cells either by an active or a passive process, as *H. pylori* moves actively toward areas with the highest concentrations of urea and bicarbonate (chemo-attraction) which are found in the mucosa (Yoshiyama et al. 1999).

Cell adhesion is an essential step in *H. pylori* colonization, but one that is very difficult to dissect at the molecular level. Because several copies of putative binding molecules are present on the *H. pylori* cell surface, and electron microscopy reveals extensive areas of adhesion to the host cells, it is difficult to distinguish the relative role of each type of adhesion. In fact, when many ligands are present on the same particle, even single weak binding interactions become relevant to the establishment of a very strong overall interaction, owing to their combined action (Montecucco, Rappuoli 2001). Not all *H. pylori* strains contain all of the adhesins which contribute to the strain variation in *H. pylori*. In addition, individual types of host's mucin genetically determined might facilitate the colonization of *H. pylori* compared to other types of mucin. The primary colonization may even take place in the oral cavity, as *H. pylori* has been shown to adhere to MG2 in the human salivary mucin (Andersen 2007). Adherence to sialic acid in the mucin seems to be a common feature in most *H. pylori* strains. *H. pylori* has at least six adhesins to sialic acid, of which three genes (*hpa*A, *nap*, *sap*A) have been identified (Wadstrom, Hirmo & Boren 1996) and one of them (*HpaA*) has been demonstrated to be essential for *H. pylori* colonization in mice (Carlsohn et al. 2006).

A well-known *H. pylori* adhesin is BabA, which binds specifically to the Lewis B (LeB) antigen in mucin MUC5AC (Van de Bovenkamp et al. 2003). In a study by Linden et al. (Linden et al. 2004), all of the *H. pylori* strains tested adhered to MUC5AC at acid pH, whereas only BabA-positive strains adhered to MUC5AC and MUC1 at neutral pH.

Another well-characterized adhesin is SabA, that has been shown to bind sialylated glycoconjugates, in particular the salivary mucin MUC7 (Walz et al. 2009): this adhesin is frequently switched "on" or "off", suggesting a response to changing conditions in the stomach (Dossumbekova et al. 2006).

Other adhesins include AlpA/AlpB, whose exact role in *H. pylori* infection remains unclear (Maeda, Mentis 2007). Anyway, *alpA*/*alpB*-deleted mutants were poor colonizers of the stomachs of C57BL/6 mice, and were associated with lower mucosal levels of proinflammatory effectors KC and IL-6 (Lu et al. 2007). Dossumbekova et al. (Dossumbekova et al. 2006) reported that mutagenesis of the recently discovered adhesin *oip*A resulted in lower adherence to gastric epithelia in vitro, but did not alter the epithelial IL-8 secretion.

1.6. Virulence and colonization factors

H. pylori colonization is typically followed by infiltration of the gastric mucosa by polymorphonuclear leukocytes, macrophages and lymphocytes. A strong correlation exists between gastric infiltration by neutrophils, mucosal damage and development of duodenal ulcer disease in *H. pylori* infections (D'Elios, Montecucco & de Bernard 2007). Some virulence factors present in *H. pylori* are strongly linked to a pathologic phenotype.

1.6.1. VacA

One of the main *H. pylori* virulence factors is the VacA secreted toxin. Because of its pleiotropic cellular effects, the *H. pylori* vacuolating cytotoxin VacA represents an important paradigm for understanding the actions of multifunctional bacterial toxins (Cover, Blanke 2005). VacA is produced as a 140 kDa protoxin and is secreted *via* the type Va or autotransporter pathway. Proteolytic processing of the protoxin during secretion yields the mature toxin (88 kDa) that spontaneously forms a flower-shaped dodecameric aggregate of 30 nm in diameter (Sewald, Fischer & Haas 2008). Mature 88-kDa VacA toxin molecules are secreted as soluble proteins into the extracellular space, but can also remain localized on the surface of the bacterium (Ilver et al. 2004). The secreted toxin can assemble into water-soluble oligomeric structures (Cover, Hanson & Heuser 1997), and can insert into planar lipid bilayers to form anion-selective membrane channels. The single polypeptide forming the mature VacA tends to be nicked through the generation of two fragments, an N-terminal (34 kDa, p37) and a C-terminal (54 kDa, p58), which remain non-covalently associated (Lupetti et al. 1996).

There is considerable, either quantitative or qualitative, variation in VacA released by various *H. pylori* strains. Signal sequences can be grouped into at least three different types (s1a, s1b and s2); another highly divergent segment, referred to as the m-region, is present within p58 (Atherton et al. 1995). Cell type-specific binding has been attributed to differences in the m1

and m2 alleles, with m1 VacA binding and inducing vacuoles in HeLa cells, whereas m2 VacA vacuolates rabbit kidney (RK13) and primary epithelial cells, but not HeLa cells (Lupetti et al. 1996). The s2 type signal sequence appears to be poorly efficient, since most strains containing it fail to release the toxin (Atherton et al. 1995). Strains which express an s1/m2 toxin produce large quantities of toxin, which assembles into the correct structure (Pagliaccia et al. 1998).

The first documented effect of VacA was its ability to induce cell vacuolization by channel formation. The more N-terminal region of p55, composed of strands b3, b6 and b9, is thought to function in VacA oligomerization. This is supported by experimental data indicating that expression of the p33 domain and the N-terminal region (100 aa) of p55 are sufficient for vacuolation of target cells (Ye, Willhite & Blanke 1999). Further studies showed that amino acids 351–360 are required for VacA protein–protein interactions (Torres, McClain & Cover 2006).

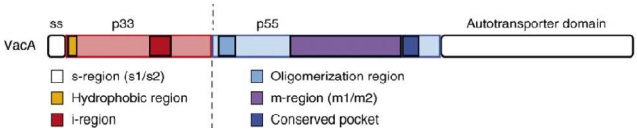

Fig. 1.3. VacA functional domains. Mature secreted VacA comprises the p33 and p55 domains. The signal sequence (ss) and autotransporter domains of the protoxin are processed during transport over the inner and outer membranes of H. pylori, respectively. An N-terminal hydrophobic region contains three GXXXG-motifs for transmembrane dimerization and pore formation (Sewald et al. 2008).

Gangwer et al. (Gangwer et al. 2007) docked the crystal structure of p55 into a cryo-electron microscopy (cryo-EM) map of the VacA oligomer and were able to predict the structure and orientation of the p33 domain, explaining its function in pore formation and oligomerization. The p33 domain seems to extend the right-handed β-helix fold of the p55 domain, which is in agreement with observations of VacA monomers in cryo-EM pictures. In addition to a function of the N-terminal part of p55 and the p33 domain in oligomerization to hexamers, Gangwer et al. hypothesize that three loops mediate the p55–p55 interaction between two layers in the dodecameric form of VacA, that could explain the oligomerization of purified VacA in aqueous solution (Gangwer et al. 2007).

The crucial step for channel formation is insertion into the membrane. The hydrophobic N-terminal region (about 30 aa) of p33 comprises three GXXXG-motifs that are thought to span the membrane and are essential for channel activity (McClain et al. 2003). The membrane insertion process of the hydrophobic region could be initiated by a change of pH. Treatment of VacA oligomers at pH < 5 results in disassembly of the hexamer into monomers and induces conformational changes that increase membrane insertion into liposomes and plasma (Sewald, Fischer & Haas 2008).

The acid-activated VacA toxin is internalized and it has been localized to 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 insert into membranes and form anion-selective channel, including leakage of ions (like Fe^{3+} and Ni^{2+}) and other small molecules from gastric epithelial cells (Cover, Blanke 2005). One of the function of VacA might be to increase the supply of essential nutrients, necessary for growth of *H. pylori*, from the underlying tissue (Montecucco, Rappuoli 2001).

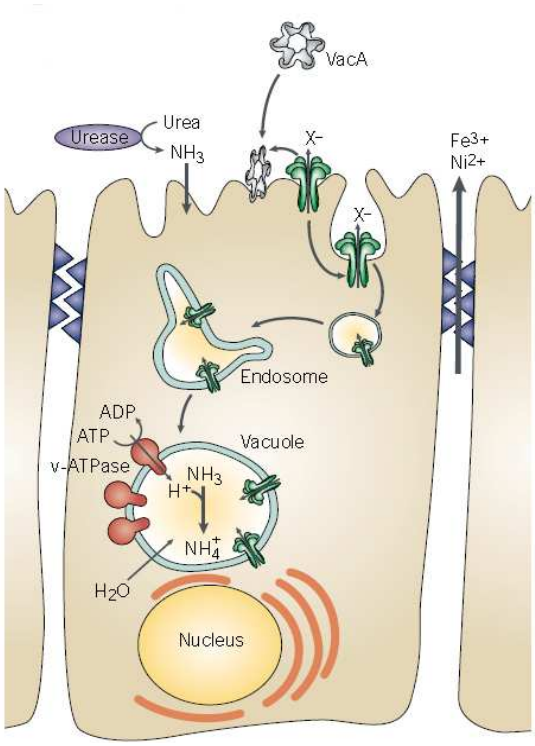

Fig. 1.4. VacA vacuole formation mechanism. The toxin binds to the apical portion of epithelial cells and inserts into the plasma membrane, forming a hexameric anionselective channel. These channels release bicarbonate and organic anions from the cell cytosol to support bacterial growth. The toxin channels are slowly endocytosed and reaching late endosomal compartments, they increase their permeability to anions with 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).

VacA-induced vacuoles are acidic because their limiting membrane contains the vacuolar-type ATPase proton pump (v-ATPase), the operation of which 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). In late endosomes/lysosomes, the action of the v-ATPase proton pump builds up an electrochemical proton gradient that progressively depresses its further activity. The necessary counter ion is normally provided by a Cl^- channel present on the same membrane, which is essential for acidification (Gunther et al. 1998). The anion-selective channel activity of VacA strongly promotes the proton pumping activity of the v-ATPase, which compensates the increasing anion concentration, leading to an enhanced accumulation of protons. This, in turn, will lead to an accumulation of ammonia and any other weak bases present in the medium. Even a limited uptake of osmotic species is expected to be sufficient to cause a significant increase in osmotic pressure given the limited membrane-free space of late endosomal compartments (Montecucco, Rappuoli 2001).

1.6.2. HP-NAP

One of the major pro-inflammatory factors produced by *H. pylori* is the neutrophil-activating protein (HP-NAP). It is a dodecameric protein of 150 kDa with a structure similar to bacterioferritins, including a central cavity for iron accumulation (Zanotti et al. 2002). HP-NAP is a virulence factor that attracts and activates neutrophils, and promotes their endothelial adhesion and the production of oxygen radicals and chemokines (D'Elios et al. 2007). HP-NAP crosses the epithelia and becomes in close contact with inflammatory cells resident in the tissue following inflammation (Montemurro et al. 2002), promoting leukocyte adhesion to the endothelium *in vivo* (D'Elios, Montecucco & de Bernard 2007) and activating the mast cells to release tumor necrosis factor-α (TNF-α) (Montemurro et al. 2002). This toxin acts via toll-like receptors (TLRs), being able to activate NF-kB after TLR2 activation (Amedei et al. 2006). The signal transduction involves the increase of cytosolic Ca^{2+} and the phosphorylation of proteins, leading to the assembly of functional NADPH oxidase on the neutrophil plasma membrane (Montecucco, Rappuoli 2001). HP-NAP, by acting on both neutrophils and monocytes, significantly contributes to the induction of IL-12- and IL-23, which has the potential to drive the differentiation of antigen-stimulated T cells towards a polarized Th1 phenotype (Amedei et al. 2006). Moreover, this toxin has been shown to increase the synthesis of tissue factor (TF) and plasminogen activator inhibitor-2 in mononuclear cells. Macrophage TF production and procoagulant activity are cross-regulated by Th1 and Th2 cells: in particular, IFN-γ and other Th1 cytokines are required for optimal TF synthesis, whereas Th2 cytokines such as IL-4, IL-10 and IL-13 are inhibitory (Del Prete et al. 1995).

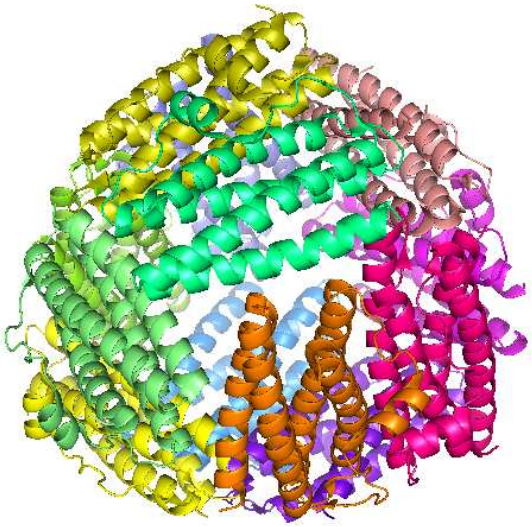

Fig. 1.5. Ribbon representation of Neutrophil activating protein (HP-NAP). *Helicobacter pylori* Neutrophil activating protein is a dodecamer (Adapted from Zanotti et al. 2002).

1.6.3. The *cag* Pathogenicity Island

The *cag* (*cytotoxin associated genes*) Pathogenicity Island from *H. pylori* is a 40 kb DNA sequence, containing about 30 genes, that encode a Type IV protein secretion system (T4SS) for toxin CagA translocation (Montecucco, Rappuoli 2001). CagA was recognized as a marker for the *cag* PAI region which is present in virulent strains but missing in avirulent *H.*

pylori isolates (Censini et al. 1996). A major breakthrough in the study of CagA came when five groups independently reported that the *cag* PAI encodes a functional type IV secretion system (T4SS) which injects CagA into host cells (Covacci, Rappuoli 2000). Pathogenicity islands seem to increase the fitness of bacteria in a given environment by providing them with environment-specific functions. The *cag* PAI 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).

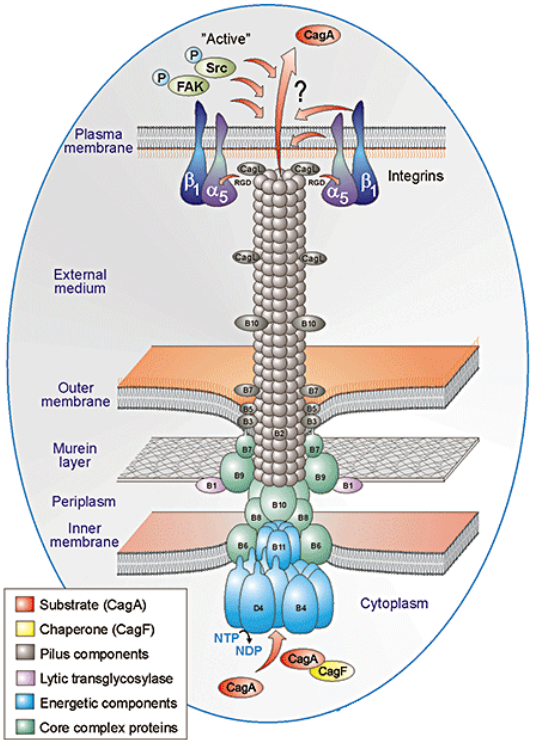

Fig. 1.6. A model of the T4SS encoded from the *cag* Pathogenicity Island. The *cag* Pathogenicity Island form *Helicobacter pylori* traslocates CagA toxin into the gastric cells (Adapted from Backert et al. 2008).

It is possible to find some homologies between some genes of *H. pylori cag* PAI and *Agrobacterium tumefaciens* T4SS (Llosa, O'Callaghan 2004). The T4SSs are a large group of transporter machines in many Gram-negative bacteria which are ancestrally related to conjugation systems (Cascales, Christie 2003). 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). The T4SS of *A. tumefaciens* is a secretion system used by this promiscuous plant pathogen to deliver DNA and proteins into plant cells. This process introduces bacterial DNA into the plant genome and causes tumour formation. The T4SS is composed of at least 12 proteins termed VirB1-11 and VirD4 and can be categorized into three groups: cytoplasmic or inner membrane proteins, a core complex located in the periplasm or membrane and a pilus or surface structure that projects beyond the outer membrane (Zhong et al. 2007). One protein, VirD4, is dedicated to the binding of the DNA substrate. Two ATPase (VirB4 and VirB11) are classified as cytoplasmic or inner membrane proteins. VirB11 and VirD4 proteins assemble into hexameric structures, whereas VirB4 is a homodimer. VirB7, VirB8, VirB9 and VirB10 are important components of the core complex. The pilus of the T4SS is composed of VirB2 and VirB5, which are considered as major and minor subunits, respectively (Zhong et al. 2007). Recently, the structure of a T4SS core,

encoded by pKM101 from *E. coli*, has been determined at low resolution using Cryo-Electron Microscopy (Fronzes et al. 2009). The core, comprising VirB7, VirB8 and VirB10 homologues, has an overall dimension of 185 Å in both height and diameter. It is made of two main layers, labeled inner (I) and outer (O). Both layers are either in part (the I layer) or entirely (the O layer) composed of two walls. The O layer is made by two regions: the cap, which is 110 Å in diameter and 40 Å high, and the main body, which is 185 Å in diameter and 60 Å high. In the cap, the internal wall forms a 20 Å hole that decreases to 10 Å in the proximity of the main body. In the main body, the internal wall forms a chamber that is 110 Å wide and 30 Å high. The I layer resembles a cup, that is linked to the O layer by thin stretches of density. The two walls of the cup merge into a single wall near the base. The internal wall defines a chamber that is as wide as the chamber in the O layer, but double the height at 60 Å. The O layer is mainly composed of the C-terminal domain of TraF/VirB10, the C terminus of TraO/VirB9, and TraN/VirB7, whereas the I layer is composed of the N-terminal regions of TraF/VirB10 and TraO/VirB9.

Focusing on the structure of *H. pylori cag* Pathogenicity Island, only the function of few proteins has been determined and the structure of just four proteins have been solved: Cagα (Yeo et al. 2000), CagS (Cendron et al. 2007), CagZ (Cendron et al. 2004) and CagD (Cendron et al. 2009). A systematic mutagenesis study of the *cag* PAI revealed that 18 of 27 genes are essential for CagA translocation into host cells and 14 for a *cag* PAI-dependent induction of the chemokine IL-8 (Fischer et al. 2001). If we compare the *H. pylori cag* Pathogenicity Island to the *A. tumefaciens* T4SS on the basis of functional and structural homologies of the components, it is possible to find many parallelisms. In *H. pylori* the conserved T4SS components include ATPases VirB4 (CagE/HP0544) and VirB11 (Cagα/HP0525) and the coupling protein VirD4 (Cagβ/HP0524) (Zhong et al. 2007). The proteins forming the core complex, CagT/HP0532, CagX/HP0528, and CagY/HP0527 in *H. pylori* are homologue of VirB7, VirB9, VirB10 respectively. The T4SS pilus base is associated with VirB7 (CagT/HP0532) and VirB9 (CagW/HP0529) proteins (Rohde et al. 2003). The pilin has been demonstrated to be the VirB2-homologue CagC/HP0546 (Andrzejewska et al. 2006), but the pilus is covered by CagY, that contains a variable number of repeats probably useful to avoid the immune-response (Rohde et al. 2003), and by the adhesin CagL, that not only anchors the type-IV secretion apparatus to the host surface through binding to integrin, but also promotes signal transduction (Kwok et al. 2007). Cagγ/HP0523 is thought to be a transglycosylase that lyses the murein cell wall to facilitate assembly of the T4SS across the bacterial cell wall (Zhong et al. 2007) and finally,

CagF/HP0543 is a chaperone-like protein which binds close to the C-terminal secretion signal of the CagA effector protein and is crucial for the translocation of CagA (Pattis et al. 2007). After injection of the CagA toxin inside the cell, CagA is tyrosine-phosphorylated by host cell kinases (Stein, Rappuoli & Covacci 2000). The phosphorylation sites are Glu-Pro-Ile-Tyr-Ala (EPIYA) motifs repeated up to five times in the C-terminal half of CagA and the number of repeats could be a reason of different pathogenicities (Backert, Selbach 2008). The kinases responsible for phosphorylation of CagA are known oncogenes: Src family kinases (SFKs) which control cytoskeletal processes, cell proliferation and differentiation in normal cells, but are also key players in carcinogenesis (Selbach et al. 2002) and Abl kinases (Tammer et al. 2007). After phosphorylation, CagA acquires the ability to bind specifically to SHP-2, a cytoplasmic protein tyrosine phosphatase having two tandem-repeated Src homology-2 (SH2) domains, named N-SH2 and C-SH2, on the N-terminal half and a protein tyrosine phosphatase (PTP) domain on the C-terminal half. Binding of tyrosine-phosphorylated CagA to the SH2 domains induces a conformational change in SHP-2, resulting in the activation of SHP-2 phosphatase activity (Higashi et al. 2002b). SHP-2 is required for full activation of the Ras-MAP kinase cascade in response to growth factor-receptor interaction and also plays an important role in cell morphogenesis as well as cell motility (Neel, Gu & Pao 2003). As a result, gastric epithelial cells containing CagA elicit cell-morphological transformation termed the hummingbird phenotype, which is characterized by elongated cell shape with dramatic cytoskeletal rearrangements and elevated cell motility (Hatakeyama 2009). Recently, focal adhesion kinase (FAK) has been identified as a substrate for CagA-deregulated SHP-2. Upon dephosphorylation at the activating tyrosine phosphorylation sites of FAK, SHP-2 downregulates FAK kinase activity, resulting in decreased cell-extracellular matrix interaction. This in turn causes elevated cell motility that underlies the hummingbird phenotype (Tsutsumi et al. 2006). Expression of CagA in gastric epithelial cells evokes sustained Erk MAP kinase activation. Since prolonged Erk activation promotes G1-to-S phase progression, CagA stimulates cell proliferation sustaining Erk activation (Tsutsumi et al. 2006). CagA indirectly affects the activities of transcription factors through multiple distinct mechanisms. CagA activates serum responsive element (SRE)-dependent transcription in a phosphorylation-independent manner (Hirata et al. 2002). CagA also activates NF-κB, which induces pro-inflammatory cytokines, such as interleukin IL-8. CagA has been shown to activate the nuclear factor of activated T cells (NFAT) by eliciting nuclear translocation of the cytoplasmic NFAT in gastric epithelial cells, in a way independent from CagA phosphorylation (Yokoyama et al. 2005). Moreover, CagA disrupts the tight junctions and

causes loss of epithelial apical-basolateral polarity. This CagA activity is mediated by the specific interaction of CagA with partitioning-defective-1b (PAR1b)/microtubule affinity regulating kinase-2 (MARK2) (Saadat et al. 2007). PAR1b, a serine/threonine kinase, has been shown to act as a master regulator of cell polarity. Indeed, cell polarization is established and maintained by the mutually exclusive distribution of PAR1 and atypical protein kinase C (aPKC)/PAR3/PAR6 complex (aPKC complex). In polarized epithelial cells, PAR1b localizes to the basolateral membrane, whereas the aPKC complex specifically localizes to the apical membrane. CagA directly binds to the kinase domain of PAR1b in a manner independent of CagA tyrosine phosphorylation and inhibits the PAR1b kinase activity. This in turn causes junctional and polarity defects and subsequent disorganization of the epithelial monolayer (Saadat et al. 2007). Since PAR1b exists as a homodimer in cells, two CagA proteins bind to a PAR1b dimer via the CM sequence and thereby passively form a CagA dimer, which is crucial for stable CagA–SHP-2 interaction. Thus, the CagA–PAR1b–SHP-2 complex coordinates cell polarity defects with an oncogenic signal to promote epithelial cell transformation (Hatakeyama 2009).

1.7. *H. pylori* and gastroduodenal diseases

Although *H pylori* typically colonizes the human stomach for many decades without adverse consequences, the presence of *H pylori* is associated with an increased risk of several diseases, including peptic ulcers, noncardia gastric adenocarcinoma, and gastric mucosa associated lymphoid tissue (MALT) lymphoma (Cover, Blaser 2009).

The risks of peptic ulcer disease and noncardia gastric adenocarcinoma are determined in part by characteristics of the *H pylori* strain with which an individual is colonized. Most of the *H. pylori* polymorphisms associated with various disease risks are found in genes that encode bacterial products that interact with host tissue. Several studies have shown that *cag* PAI-positive *H. pylori* strains, particularly in Western countries, are associated with a higher risk of peptic ulcer disease, premalignant gastric lesions, and gastric cancer than the strains that lack the *cag* PAI (Basso et al. 2008). Moreover, the number of tyrosine phosphorylation (EPIYA) motifs in CagA proteins correlates with gastric cancer risk. Strains that express forms of VacA that are active *in vitro* are associated with a higher risk of disease than the strains that express inactive forms of VacA (Higashi et al. 2002a). Similarly, strains that express BabA and OipA (HopH) OMPs are also associated with a higher risk of disease than the strains that lack these factors (Cover, Blaser 2009).

One of the pathologies promoted by *H. pylori* is peptic ulceration, a common disease of elderly that is increasing in developed countries with non-steroidal anti-flammatory drugs implicated as a major cause (Blaser 1998): the *H. pylori* eradication should give short term and long term benefits.

Another important pathology is the gastric cancer, the most common cancer in several areas of the world, most notably Japan, Korea and China. Most cases of stomach cancer are diagnosed between the ages of 50 and 70 years, but younger cases are more frequently seen in families with a hereditary risk of stomach cancer (Hatakeyama 2009). Histologically, there are two major types of gastric carcinoma, the intestinal type, which is associated more commonly with environmental perturbations, and the diffuse type, which is ascribed etiologically more often to host genetic factors. Recent epidemiological studies have indicated that *H. pylori* plays a key role in the development of both intestinal-type and diffuse-type gastric carcinomas (Uemura et al. 2001). In particular, the dominant form, the intestinal-type, seems to be related to *cag*$^{+}$ strains of *H. pylori* (Blaser 1998).

Moreover, *H pylori* infection significantly increased the risk for gastric MALT lymphoma (Wotherspoon et al. 1991). *H pylori* infection through its diverse virulence factors triggers inflammatory responses by attracting and activating neutrophils, which release ROS. The organism also modulates immunologic responses that not only perpetuate the infection, but also stimulate the growth of malignant B cells. The oxygen radicals and other inflammatory products may cause a wide range of genetic damage, and therefore might have a role in the acquisition of genetic abnormalities in gastric MALT lymphoma (Farinha, Gascoyne 2005). Early stage gastric MALT lymphomas can be cured in about two thirds of cases by antibiotic eradication of the bacterium (Farinha, Gascoyne 2005).

Studies in healthy volunteers have shown that 80–100% of gastroesophageal reflux episodes are associated with abnormal or transient lower esophageal sphincter relaxation (Penagini, Carmagnola & Cantu 2002). However, motor alterations alone are not sufficient to cause gastroesophageal reflux and must be accompanied by reflux content that is aggressive to esophageal mucosal cells. Gastric juice secretion (mainly hydrochloric acid and pepsin) was identified as the most important factor in the induction of reflux esophagitis. *H. pylori* could be involved because of the production of gastric mucosal interleukins IL-8 and IL-1b, which stimulate gastrin secretion by endocrine G cells, increasing the proliferation and the secretion of acid by parietal cells (Souza, Lima 2009).

Finally, *H. pylori* infection is associated to dyspepsia. This disease is defined as a chronic or recurrent pain centered in the upper abdomen (Selgrad, Kandulski & Malfertheiner 2008).

Dyspeptic symptoms have a high prevalence in the population and may overlap with other domains of functional gastrointestinal disorders, such as with functional esophageal disorders or irritable bowel syndrome (Selgrad, Kandulski & Malfertheiner 2008). Large population studies have shown that *H. pylori* is more frequently detected in dyspeptic patients than in controls (O'Morain 2006) and it causes disease because of disturbed acid secretion (el-Omar et al. 1995).

1.8. Therapy of *H. pylori* infection

As already recommended in the original Maastricht Consensus Report, treatment regimens should be simple, well tolerated and cost-effective. The first-line therapy is a proton pump inhibitor (PPI) triple therapy, consisting of PPI, clarithromycin and amoxicillin/or metronidazole in populations with less than 15–20% clarithromycin resistance rate (Malfertheiner et al. 2007). Although *H. pylori* can be successfully eradicated by antibiotics and proton pump inhibitors in most patients, increasing antibiotic resistance in the bacterium remains a serious problem (D'Elios, Andersen 2007). Therefore, there still exists a strong rationale for development of effective vaccines against *H. pylori*. Over the last year, several approaches were 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, but they require adjuvants to elicit effective protection. They have been studied extensively in animal models but minimally in humans. About human trials, the most successful studies exploited the bacterium *Salmonella typhi* carrying *H. pylori* antigens or inactivated *H. pylori* cells.

During a preliminary study, a *S. typhi* strain, deleted of virulence genes and modified to constitutively express *H. pylori* urease, was orally administered to 8 uninfected adult volunteers, but it was ineffective in producing any detectable mucosal or humoral immune responses to the urease antigen (DiPetrillo et al. 1999). The clinical trial was repeated using *S. typhimurium* instead of *S. typhi*: in some volunteers antibody response against *H. pylori* urease was detected (Angelakopoulos, Hohmann 2000). Another prophylactic *Salmonella*-based vaccine named Ty21a(pDB1) was prepared by expressing the *H. pylori* urease in the common live typhoid vaccine. The Ty21a vaccine was tested orally in uninfected volunteers with only moderate efficacy but with no major adverse effects (Bumann et al. 2001). Oral inactivated *H. pylori* whole cell vaccine showed specific B-cell responses in uninfected volunteers but could not eradicate the bacterium from already infected individuals, showing

efficacy of prophylactic but not therapeutic use of *H. pylori* whole cell vaccine (Losonsky, Kotloff & Walker 2003).

New vaccine types are going to be explored in the near future: live vector vaccines, DNA vaccines, microsphere vaccines, ghost vaccines. Briefly, DNA vaccines carry DNA sequences encoding *H. pylori* antigens and they are known because of their safety and efficacy. Microspheres effectively induce humoral and mucosal immunity as well as cell mediated immunity, whereas bacterial ghosts are empty cell envelopes without cytoplasmic contents that retain their cellular morphology with native antigenic structures (Agarwal, Agarwal 2008).

In conclusion, new therapies are required to eradicate *H. pylori* infection and looking for new pharmaceutical targets could be useful for the development of new antibiotics in the future.

Chapter II

Crystal structure and enzymatic characterization of HP1287 from *Helicobacter pylori*

This chapter has been adapted from:

Nicola Barison[1,2], Laura Cendron[1,2], Alberto Trento[2], Alessandro Angelini[2,*] and Giuseppe Zanotti[1,2,3]. Structural and mutational analysis of TenA protein (HP1287) from the *Helicobacter pylori* thiamin salvage pathway – evidence of a different substrate specificity. FEBS Journal 2009; 276(21):6227-35.

[1] Department of Biological Chemistry, University of Padua, Italy
[2] Venetian Institute of Molecular Medicine (VIMM), Padua, Italy
[3] Institute of Biomolecular Chemistry of CNR, Padua, Italy
*Present address: Laboratory of Therapeutic Proteins and Peptides–LPPT, Institute of Chemical Sciences and Engineering, Ecole Polytechnique Federal de Lausanne (EPFL), Lausanne, Switzerland

2.1. Introduction

2.1.1. Thiamin metabolism in bacteria

Thiamin pyrophosphate (vitamin B_1) is an essential cofactor for several important enzymes of the carbohydrate metabolism. Many microorganisms, but also plants and fungi, synthesize thiamin, but the latter is not produced by vertebrates (Rodionov et al. 2002). The thiamin metabolism process is a complex process and its biosynthesis is not fully understood (Jurgenson, Begley & Ealick 2009). Moreover the regulation of thiamin is controlled transcriptionally through the use of thiamin riboswitches (Serganov et al. 2006; Thore, Leibundgut & Ban 2006). Thiamine monophosphate (ThMP) is formed through the coupling of two independently synthesized moieties, the hydroxymethylpyrimidine pyrophosphate derivative (HMP-PP) and hydroxyethylthiazole phosphate derivative (HET-P) (Rodionov et al. 2002). The best-studied thiamin biosynthetic pathways are those of *Escherichia coli* and *Bacillus subtilis*, which utilize very similar pathways. In all organisms, the thiazole and pyrimidine moieties of thiamin monophosphate are generated in separate branches of the pathway and then joined by a coupling enzyme. ThMP is converted to the active form of the cofactor thiamin diphosphate (ThDP) by a specific kinase (Jurgenson, Begley & Ealick 2009). Structural and mechanistic studies on thiamin biosynthetic enzymes have played a key role in increasing our understanding of thiamin pyrophosphate biosynthesis (Settembre, Begley & Ealick 2003).

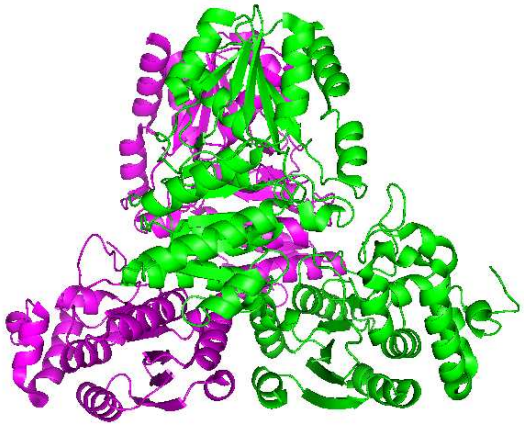

Fig. 2.1. Ribbon representation of 1-deoxy-D-xylulose-5-phosphate synthase from *E. coli.* PDB accession code: 2O1S.

The thiazole moiety (4-methyl-5-β-hydroxyethylthiazole or THZ) is made through three steps. First, glyceraldehyde 3-phosphate and pyruvate are coupled together by 1-deoxy-D-xylulose-5-phosphate synthase (Dxs) to give 1-deoxy-D-xylulose-5-phosphate (DXP). Surprisingly, Dxs requires thiamine diphosphate to catalyze this reaction. The structure of Dxs from *E. coli* and *Deinococcus radiodurans* shows that the enzyme is a dimer composed by three domains, each including 5 or 6-stranded β-sheets. The active site is comprised between the domains I and II of the same monomer (Xiang et al. 2007).

In the second step, several enzymes are involved. First of all, the sulfur carrier protein ThiS undergoes an adenylylation by ThiF. The 3D structure of ThiS has been determined by NMR: it consists of a five-stranded mixed β-sheet with an α-helix, which crosses over between strands β2 and β3, and a 3_{10}-helix between strands β4 and β5 (Wang et al. 2001). The adenylyltransferase ThiF was bound to ThiS tightly enough to be copurified. Crystal structures of ThiF-ThiS complex exists as a dimer of ThiF with one ThiS molecule bound to

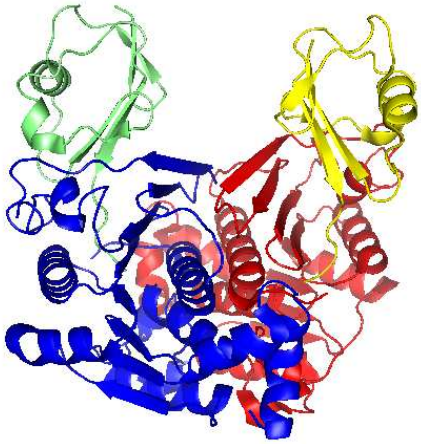

Fig. 2.2.A. Ribbon representation of ThiF-ThiS from *E. coli*. PDB accession code: 1ZUD.

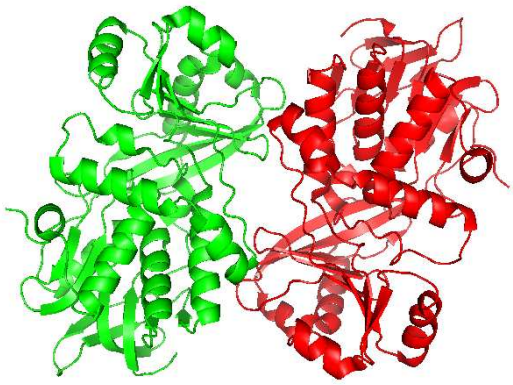

Fig. 2.2.B. Ribbon representation of ThiO from *Bacillus subtilis*. PDB accession code: 1NG3.

each ThiF protomer (Lehmann, Begley & Ealick 2006). The adenylation of ThiS is required for the activation of the enzyme C-terminal group. The sulfur is transferred by ThiI (*E. coli*) or NifS (in *B. subtilis*) to ThiS, displacing the C-terminal oxygen atom. This generates ThiS-COS^{-}, which is the source of the sulfur atom in the thiazole ring (Jurgenson, Begley & Ealick 2009). IscS is also required for the biosynthesis of iron-sulfur clusters and may be responsible for sulfur incorporation in molybdopterin (Kessler 2006). Finally, glycine (by ThiO in *B. subtilis*) or tyrosine (by ThiH in *E. coli*) is converted to dehydroglycine. The structure of ThiO from *B. subtilis* is known (Settembre et al. 2003). It is a tetramer with 222-point symmetry. Each ThiO protomer interacts with each of the other protomers in the tetramer. The ThiO monomer presents two separate domains. One domain belongs to the glutathione reductase type 2 family and is responsible for binding flavin adenine dinucleotide (FAD), and the other domain binds substrate. The thiocarboxy C-terminus of ThiS, along with DXP and dehydroglycine, are all coupled together by thiazole synthase, ThiG, to give thiazole phosphate carboxylate tautomer (Jurgenson, Begley & Ealick 2009). Thiazole synthase (ThiG) is a tetramer with 222 symmetry and it has a rectangular shape of size 85 Å x 75 Å x 35 Å. The monomer is a $(\alpha\beta)_8$ barrel with similarities to the aldolase class 1 and flavin mononucleotide dependent oxidoreductase and phosphate binding superfamilies (Settembre et al. 2004). Interactions between ThiG and ThiS involve two main areas. The first area, called

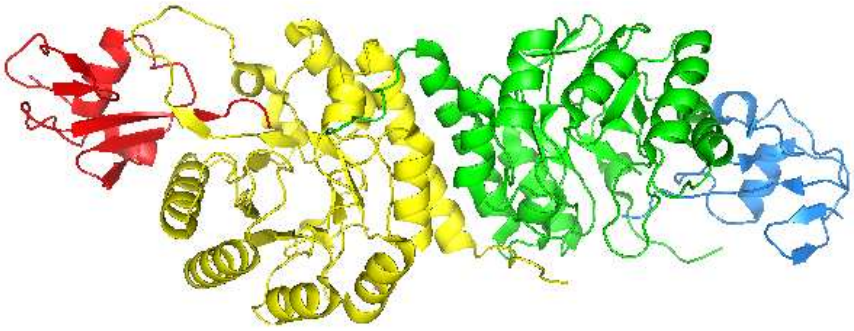

Fig. 2.3.A. Ribbon representation of ThiG-ThiS complex from *Bacillus subtilis*. PDB accession code: 1TYG.

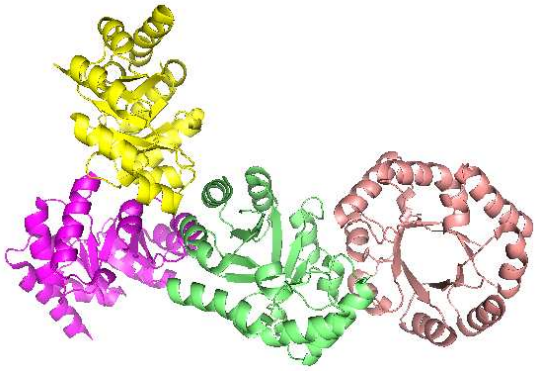

Fig. 2.3.B. Ribbon representation of TenI from *Bacillus subtilis.* PDB accession code: 1YAD.

the "clamp" loop, is mainly hydrophobic and involves the extended β2-α2 loop of thiazole synthase and the C-terminal tail of ThiS, that is located inside the ThiG loop. The second area is dominated by hydrophobic contacts as well and involves conserved surface hydrophobic residues (Settembre et al. 2004). The final product is not thiazole phosphate (THZ-P), but a carboxy thiazole phosphate tautomer, which must aromatize to form the final product. The enzyme TenI (*B. subtilis*) then aromatizes the thiazole tautomer to the thiazole phosphate carboxylate (Jurgenson, Begley & Ealick 2009).

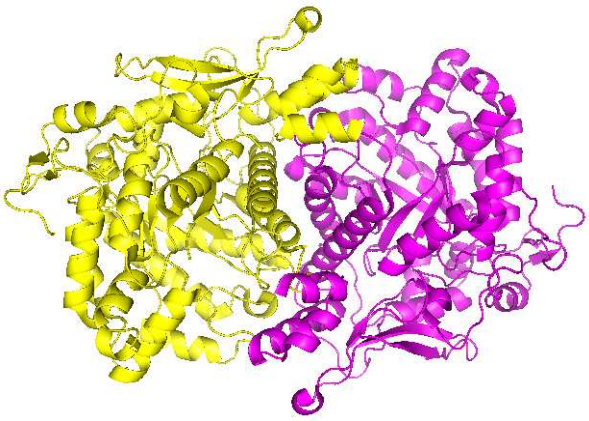

Fig.2.4. Ribbon representation of ThiC from *Caulobacter crescentus.* PDB accession code: 3EPM.

The biosynthesis of hydroxymethylpyrimidine (HMP) moiety involves two genes. In the first step, the 4-amino-5-hydroxymethyl-2-methylpyrimidine phosphate (HMP-P) ring is generated through a complicated rearrangement reaction catalyzed by ThiC, from 5-aminoimidazole ribotide (AIR) (Jurgenson, Begley & Ealick 2009). The crystal structure of ThiC from *Caulobacter crescentus* has been determined (Chatterjee et al. 2008). The enzyme is a homodimer. Each protomer contains a bound metal ion, Zn(II). Each protomer consists of three domains. The first domain contains a novel fold consisting of

six α-helices and five β-strands. The second domain comprises a $(\beta\alpha)_8$ barrel fold. Domain 3 forms an antiparallel three-helix bundle followed by a loop that extends into the top of the $(\beta\alpha)_8$ barrel of the adjacent protomer. The active site is located within a cavity at the C-terminal end of the $(\beta\alpha)_8$ barrel. Most of the residues in the active site cavity are highly conserved among the HMP-P synthase orthologues. A large peak assigned as a Zn(II) with tetrahedral geometry is located within the active site and is coordinated by the Nε2 of His417, the Nε2 of His481 and two water molecules with unusually long distances (Chatterjee et al. 2008). The cofactor S-adenosylmethionine (SAM) and the [4Fe-4S] cluster were modeled into the active site. Hydrogen bonds to the phosphate group of the substrate analog IMR (imidazole ribotide) come from the side chains of Tyr277, His313, and Arg377 as well as the amide nitrogen atom of Gly335. The iron atoms of the iron sulfur cluster bind to the thiol side chains of Cys561, Cys564, Cys569 and the carboxylate moiety of SAM.

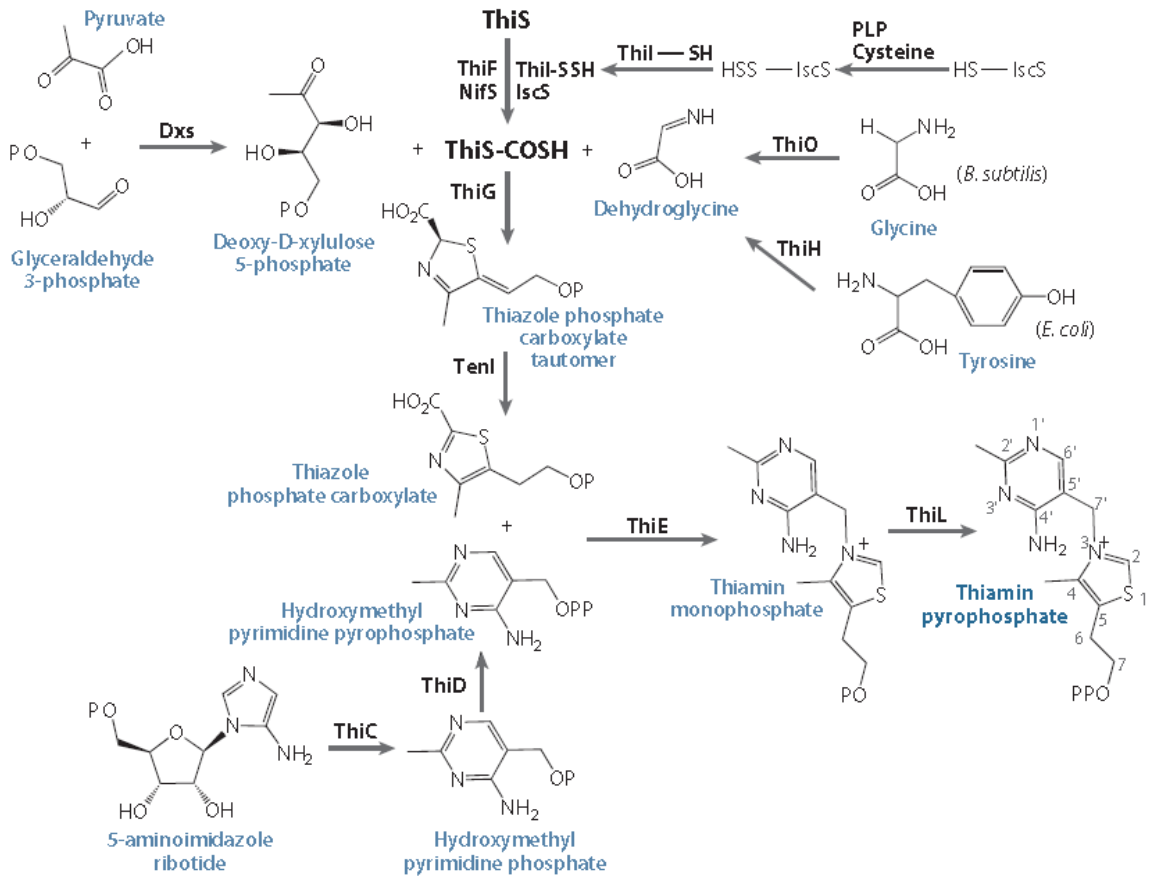

Fig. 2.5. *De novo* synthesis of thiamin in bacteria. Adapted from Jurgenson et al. 2009.

The last step for the synthesis of hydroxymethylpyrimidine pyrophosphate derivative is the 4-amino-5-hydroxymethyl-2-methylpyrimidine phosphate kinase (ThiD). ThiD from *Salmonella thyphimurium* is a dimer and each monomer consists of eight α-helices, two 3_{10}-

helices and ten β-strands, arranged in an αβα sandwich fold. The dimers interaction domains involve the C-terminus. The ThiD fold is similar to a family of ribokinase and is able to catalyze the phosphorylation of the HMP to HMP-PP. Owing to the possibility of using both the hydroxymethylpyrimidine derivative HMP or the hydroxymethylpyrimidine phosphate derivative HMP-P as substrate, the enzyme shows flexibility in the active site: the phosphate group of HMP-P is able to move into different phosphate-binding pockets through rotation about the C5-C7 bond, thus allowing it to be both a product and a substrate (Cheng et al. 2002).

The key step for the thiamin biosynthesis is the coupling of the two main moieties, the hydroxymethylpyrimidine pyrophosphate derivative and the hydroxyethylthiazole phosphate derivative. This reaction is catalyzed by ThiE, a dimer whose structure is known. The structure of ThiE reveals a $(\beta\alpha)_8$-barrel fold, where helices α2–8 and α10 surround the central β-barrel. ThiE dimerisation involves helices α3 from each protomer, that align parallel to each other along the dimer interface (Peapus et al. 2001). The proposed mechanism of the reaction starts with the loss of the PP group from C5 of HMP-PP to generate a carbocation that is stabilized through the delocalized π system of the pyrimidine ring. Following a proton transfer from the N4 nitrogen atom of the pyrimidine ring to the PP leaving group, a pyrimidine imine

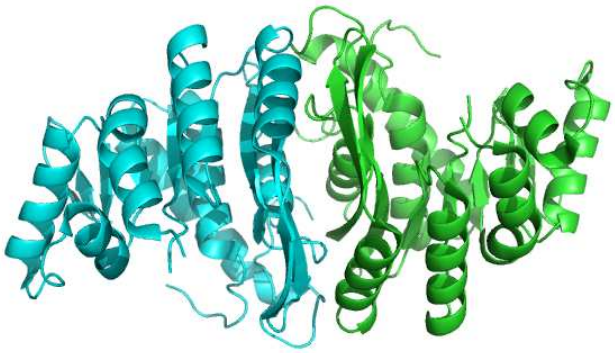

Fig. 2.6.A Ribbon representation of ThiD *from S. typhimurium*. PDB accession code: 1JXH.

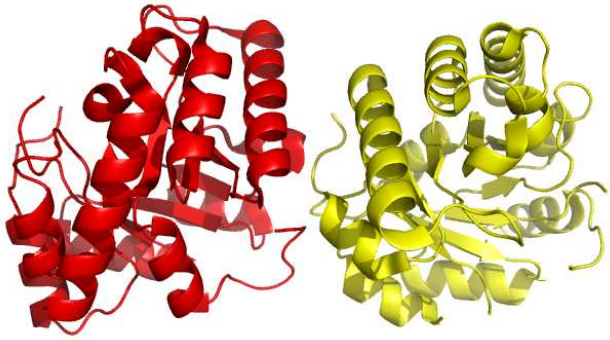

Fig. 2.6.B. Ribbon representation of ThiE from *B. subtilis.* PDB accession code: 1G4E.

methide intermediate is formed, which then undergoes nucleophilic attack from the nitrogen atom of the THZ-P to form thiamin phosphate (Peapus et al. 2001). The final step for the thiamin synthesis is the phosphorylation of thiamin, carried out by ThiL. ThiL contains two domains: the first domain is a half-barrel with four very long β-strands that are involved in the dimerisation. The second domain is an α/β domain formed by six-stranded antiparallel β-sheet and six α-helices and is linked to the first domain by a short connective loop between β5 and β6 (McCulloch et al. 2008). Interestingly, five Mg^{2+} ions are bound to the protein and are observed to coordinate the phosphate moieties of thiamine monophosphate and the AMP

analog AMP-PCP. Three of them are bound to the two phosphate groups of AMP-PCP and are coordinated to aspartate residues. The fourth magnesium ion is coordinated to the oxygen atom of the β-phosphate group and two aspartate and one threonine residues. A fifth magnesium ion is found coordinated to the α-phosphate and β-phosphate groups of thiamine pyrophosphate and to an aspartic residue. The enzymatic mechanism requires some movement of the magnesium ions (McCulloch et al. 2008).

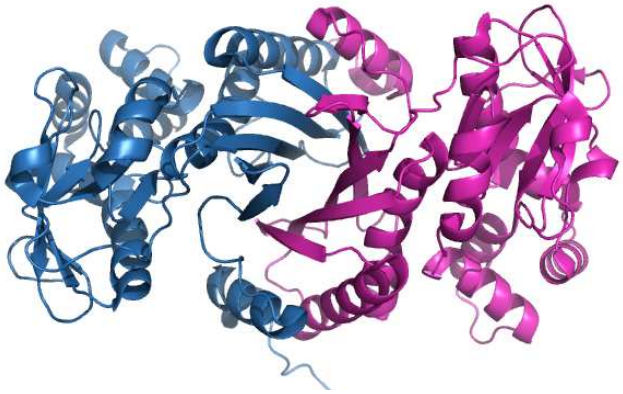

Fig. 2.7. Ribbon representation of ThiL from *Aquifex aeolicus*. PDB accession code: 3C9R.

Thiamine catabolism exists as well and is associated to vitamin deficiency. Thiamine is degraded into the thiazole and pyrimidine moieties by an enzyme called thiaminase. Two classes of thiaminase have been identified: thiaminase I and II. Thiaminase I has been associated with early mortality syndrome in predatory fish of the Laurentian Great Lakes and the New York Finger Lakes and in Atlantic salmon from the Baltic Sea. Fish eggs affected by this disease contain low levels of thiamine between the time the fish are hatched and their first feeding (Honeyfield et al. 2005). The ill fishes are believed to be affected by *Alosa pseudoharengus* that expresses high level of thiaminase I. The difference between thiaminase I and II is the nucleophile used to separate the two moieties of thiamin. Thiaminase I can use aniline, cysteine, dithiothreitol, pyridine, quinoline, and veratrylamine as substrates, thiaminase II is able to use only water (Costello et al. 1996). Moreover, thiaminase I and II have also different sequences and structures. The structure of thiaminase I from *Bacillus thiaminolyticus* is ellipsoidial with dimensions 70 Å × 42 Å × 35 Å. Thiaminase I is monomeric and has two distinct globular domains (identified as N- and C- domains), linked by a deep groove. Each of these domains has an α/β fold and they share a similar structure, with a central *β*-pleated sheet, flanked on both sides by α-helices. A cleft is located in between the N and C-domains and six tyrosine residues and four acidic residues are present in the cleft. The active site was identified using the mechanism-based inhibitor 4-amino-6-chloro-2,5-dimethylpyrimidine, which bonds irreversibly to the active-site cysteine residue (Cys113). Thiamin is positioned in the active site by two hydrogen bonds between the pyrimidine moiety and Asp272, one of which occurs through an intervening water molecule. Glu241 then activates Cys113 by deprotonation for attack at C6 of the pyrimidine to form a zwitterionic intermediate.

Nucleophilic attack and protonation by Glu241 result in cleavage of the bond between the thiazole and pyrimidine and release of products (Campobasso et al. 1998).

The only thiaminase II whose structure has been determined, is TenA. It was firstly classified as enzyme involved in thiamin metabolism because of this residual activity (Toms et al. 2005), but the function and the structure of this enzyme are discussed later.

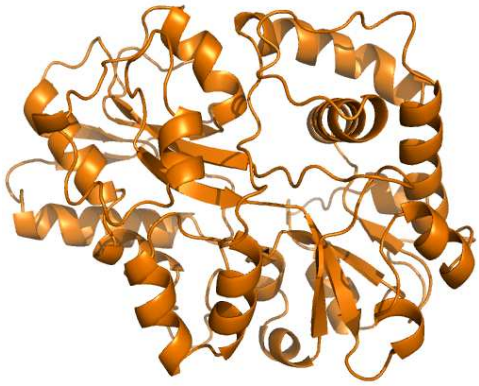

Fig. 2.8.A. Ribbon representation of thiaminase I from *B. thiaminolyticus.* PDB accession code: 2THI.

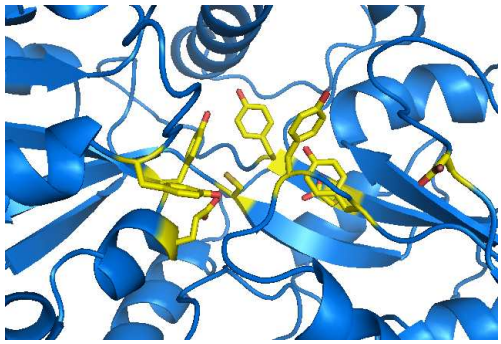

Fig. 2.8.B. Active site of thiaminase I from *Bacillus thiaminolyticus*.

2.1.2. Thiamin salvage pathway

Instead of *de novo* synthesis, the main portions of thiamin can be salvaged. In bacteria, thiazole alcohol derivative (THZ) can be converted to thiazole phosphate derivative (THZ-P) by thiazole kinase (ThiM). The structure of ThiM from *Bacillus subtilis* shows an homology to ribokinase and other kinases (Sigrell et al. 1998). ThiM is a trimer and every monomer contains nine β-strands and 12 α-helices, but the additional β-sheet, that acts like a flap in

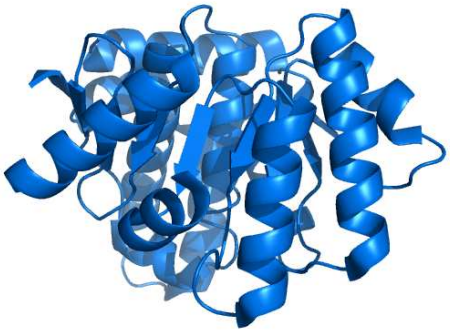

Fig. 2.9.A. Ribbon representation of the ThiM monomer from *B. subtilis*. PDB accession code: 1EKQ.

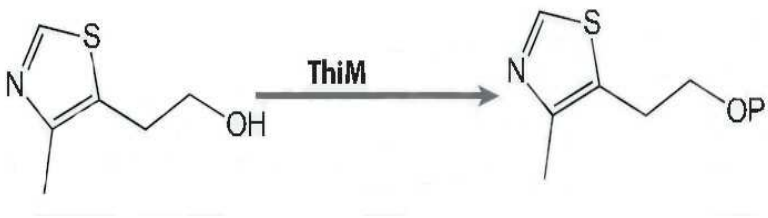

Fig. 2.9.B. Reaction catalyzed by ThiM. Adapted from Jurgenson et al. 2009.

ribokinases, is missing, replaced by some residues belonging to the following protomer (Campobasso et al. 2000). The active site contains THZ-P and the substrate ATP. When the substrate is located inside the active site, the enzyme shows a conformation that facilitates the transfer of the phosphate moiety to the hydroxyl group of the thiazole. The thiazole ring is

bound to the enzyme through a hydrogen bond to the amide nitrogen atom of Met45. The phosphate moiety of THZ-P interacts with the side chain of Ser198. Arg121 binds to the β-phosphate of ATP and Thr168 binds to the α-phosphate through the hydroxyl moiety of its side chain. The ribose ring interacts through a hydrogen bond between the 2' position of the ribose ring and the carboxylate moiety of Asp172.

Another enzyme involved in salvage of thiamin is ThiN, which catalyzes the pyrophosphorylation of the thiazole hydroxyl group from thiamin. Iterative BLAST searches suggest that the bacterial thiamin pyrophosphate kinase (ThiN) shares at least some structural homology with THI80 from yeast and mouse, whose structure is known (Baker et al. 2001; Timm et al. 2001). Each monomer consists of an α/β domain and a β-sandwich domain. Both domains contribute to dimer formation. Each of the two active sites is located in a cleft between the N-terminal domain of one protomer and the C-terminal domain of the other.

The key enzyme for the salvage of the thiamin pyrimidine moiety is TenA, whose function and structure are discussed later.

2.1.3. Function of TenA from *Bacillus subtilis*

TenA is an enzyme present in several eubacteria and archaea, usually found in a cluster with other genes involved in thiamin metabolism. In *Bacillis subtilis*, TenA is part of the thiazole biosynthetic operon (TenA-TenI-ThiO-ThiS-ThiG-ThiF-ThiD) and is known to be strongly repressed by thiamin (Lee et al. 2001). The first identified function of the *tenA* gene from *Bacillus subtilis* was the transcriptional control of extracellular enzymes (Pang, Nathoo & Wong 1991). Overexpression of *tenA* can stimulate eight- to ninefold the production of several extracellular enzymes like neutral proteases, alkaline proteases and levansucrase at the transcriptional level and, functionally speaking, *tenA* appears similar to *degQ*, *degR*, *senS*, and *senN*. Dependence on a functional *degS* product suggested that *tenA* acts indirectly to enhance the production of extracellular enzymes. It may stimulate the protein kinase activity of the *degS* product and result in an increase in the level of the activated *degU* transcription activator.

An enzymatic activity of TenA as thiaminase II has been identified: TenA enzyme was incubated with thiamin for 1 minute, the degradation products were purified by RP-HPLC and identified by NMR (Toms et al. 2005).

Recently, the physiological activity of TenA was identified. In *Bacillus halodurans*, the gene *tenA* is clustered with the transporter ThiXYZ, involved in HMP analogs transporters (Rodionov et al. 2002) and YlmB. In basic soil, where *B. halodurans* lives, but even in milder

conditions, where *B. subtilis* grows, the thiamin is degraded in several compounds. Thiamin decomposition reaction is presumably catalyzed by metal-ion-bound hydroxide on the clay surface (Xu et al. 2001). One of these thiamin-degradation compounds is N-formyl-4-amino-5-aminomethyl-2-methylpyrimidine (FAMP), that could be produced by the mechanism shown in Fig. 2.11: water addition to the thiazolium heterocycle of the compound 1 gives 2, which then undergoes ring opening to give 3. Tautomerization to 4 followed by hydrolysis gives FAMP (5) (Jenkins et al. 2007). FAMP is transported inside the bacterial cell by the specific ABC transporter ThiXYZ, whose subunit ThiY is the putative substrate-binding component and has been demonstrated to bind the FAMP (Jenkins et al. 2007). FAMP is metabolized by YlmB to 4-amino-5-aminomethyl-2-methylpyrimidine (AMP). AMP is subsequently converted by TenA to the thiamin precursor HMP. The previous identified reaction catalized by TenA, thiamin hydrolysis, does not follow Michaelis-Menten kinetics and competition assays show that TenA catalyzes the hydrolysis of AMP 100 times faster than it catalyzes the hydrolysis of thiamin: these findings suggest that the salvage of the pyrimidine moiety of thiamin is the main physiologic activity of TenA (Jenkins et al. 2007).

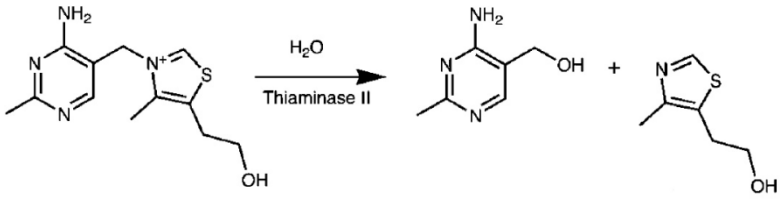

Fig. 2.10. The firstly identified reaction of TenA: the thiamin is degradated in the two main moieties, the pyrimidine derivative and the thiazole derivative. Adapted from Toms et al. 2005.

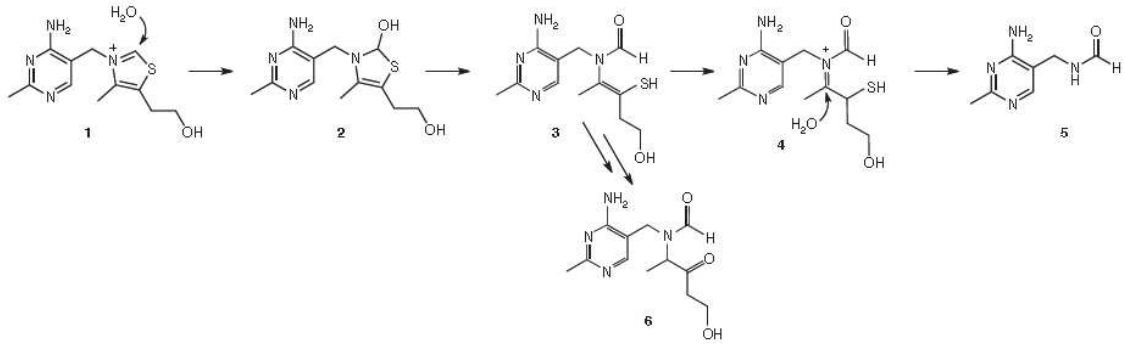

Fig. 2.11. The proposed mechanism of thiamin degradation to FAMP. Adapted from Jenkins et al. 2007.

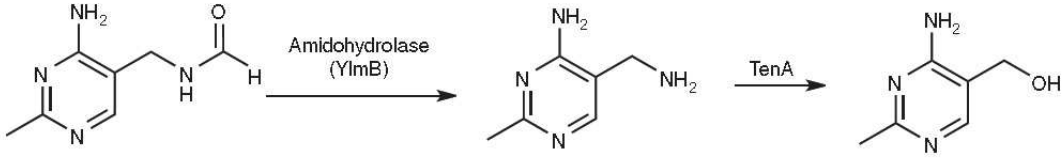

Fig. 2.12. YlmB removes the formyl group of FAMP, producing AMP, that is metabolized by TenA to HMP. Adapted from Jenkins et al. 2007.

2.1.4. Structure of TenA from *B. subtilis*

The structure of TenA from *Bacillus subtilis* has been determined as well as that from other species (Toms et al. 2005), like *Pyrobaculum aerophilum* (2GM7) , *Pyrococcus horikoshii* (1UDD) (Itou et al. 2004), *Pyrococcus furiosus* (1RTW) (Benach et al. 2005), *Sulfolobus solfataricus* (2QZC). TenA is a tetramer with 222 point symmetry and has a rectangular shape with approximate dimensions of 70 Å x 65 Å x 45 Å. The TenA monomer is roughly rectangular shaped as well, with dimensions of 46 Å x 30 Å x 25 Å. The TenA fold consists of 11 α-helices and a deep acidic pocket is located in the central region of the protein surrounded by helices α4, α5, α9, and α10. The pocket is flanked by several aromatic and acidic residues and has a volume of 700 $Å^3$. TenA from *B. subtilis* has been crystallized in complex with HMP as well. The electron density for crystals grown in the presence of HMP showed one HMP molecule per monomer inside the active site. The pyrimidine ring of the HMP molecule is sandwiched between two tyrosine residues (Tyr47 and Tyr139), with a ring stacking distance of 3.56 Å. The binding of HMP is also stabilized by several electrostatic and hydrogen-bonding interactions: the 4-amino group and the N3 nitrogen atom of the pyrimidine ring both form hydrogen bonds with Asp44. The N1 atom of HMP is close to Glu205, whereas the sulfur atom of Cys135 is located about 3 Å from the C6 position of the pyrimidine ring. The structures of TenA from other bacteria are very similar to *B. subtilis* one, with just slight differences in the number of α-helices. Interestingly, in TenA structures from *Pyrococcus horikoshii* (Itou et al. 2004) and *Pyrococcus furiosus* (Benach et al. 2005) a residual electron density, corresponding probably to a co-purified compound, has been identified in between the aromatic rings of the conserved tyrosine residues Tyr47 and Tyr139.

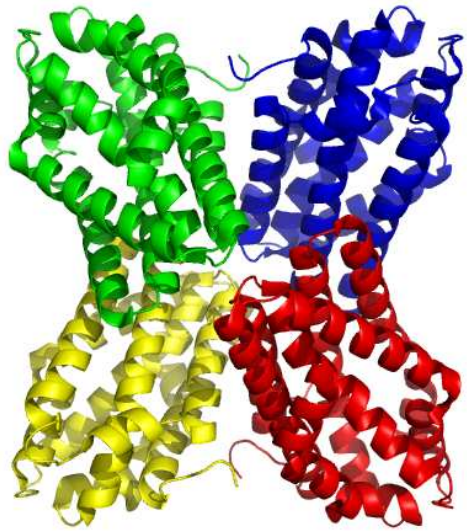

Fig. 2.13. Ribbon representation of TenA from *B. subtilis*. PDB accession codes: 1YAF, 1YAK.

The mechanism of the reaction has been clarified by mutagenesis studies, whose kinetic parameters relative to the mutants are shown in table 2.1.

Protein	K_m(μM)	K_{cat}(min^{-1})	K_{cat}/K_m($\mu M^{-1}min^{-1}$)	Relative K_{cat}/K_m
Native	11.8 ± 1.6	22 ± 0.5	1.9 ± 0.3	1
E205A	251 ± 31.5	0.27 ± 0.006	0.001 ± 0.0001	0.0005
D44A	337 ± 94	0.09 ± 0.004	0.0003 ± 0.00008	0.0003
Y112F	<25	0.31 ± 0.02	>0.012 ± 0.0007	>0.01
Y47F	<25	0.75 ± 0.03	>0.029 ± 0.0012	>0.03
C135A	Inactive	Inactive	Inactive	Inactive

Table 2.1. Kinetic parameters of TenA *wild-type* and mutants

The catalytic residue is Cys135, that adds to C6 of the pyrimidine ring generating anion 2 (Fig. 2.14.). The leaving group ammonia is expelled, giving 3. The addition of a water molecule and the expulsion of the catalytic cysteine give the final product. A network of other residues promotes the enzymatic catalysis: the catalytic residue Cys135 is deprotonated by Tyr47, Tyr112 and the glutamic acid in position 205, that protonates the pyrimidine, stabilizing the resulting anion. The aspartic acid in position 44 seems to be important for the substrate binding.

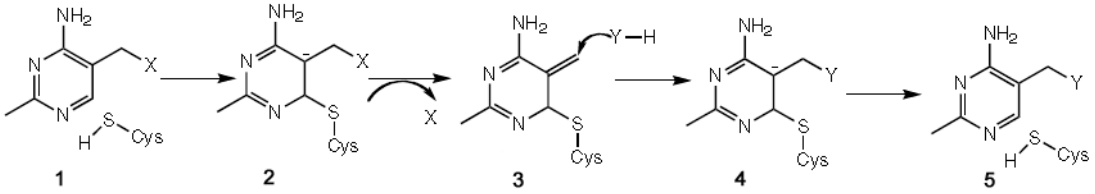

Fig. 2.14. Steps of the reaction catalyzed by TenA from *B. subtilis*. Adapted from Jenkins et al. 2008.

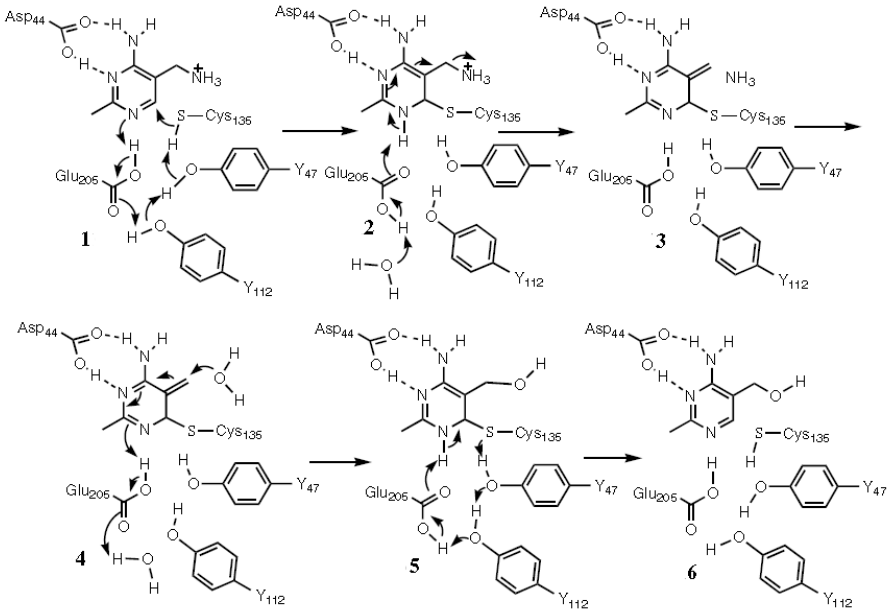

Fig. 2.15. Proposed mechanism of the reaction catalyzed by TenA from *B. subtilis.* Adapted from Jenkins et al. 2008.

2.1.5. HP1287 from *Helicobacter pylori*

TenA homologue in *Helicobacter pylori* is HP1287. The alignment of the amino acid sequence of HP1287 shows 33% identity and 51% similarity to the TenA protein from *B. subtilis*. HP1287 sequence alignment with TenA from *B. subtilis* and other bacteria reveals that the residues important for the catalysis are conserved, with the only exception of Tyr47, substituted from Phe47 in *H. pylori*: Cys135, Tyr112, D44 are maintained, whereas the glutamic acid in position 205 is substituted by Glu207 in HP1287. Recently, a transposon mutagenesis method in a mouse model of infection has identified HP1287 within a pool of candidates that might contribute to stomach colonization and persistence (Baldwin et al. 2007), raising intriguing questions about the putative roles of the corresponding protein product.

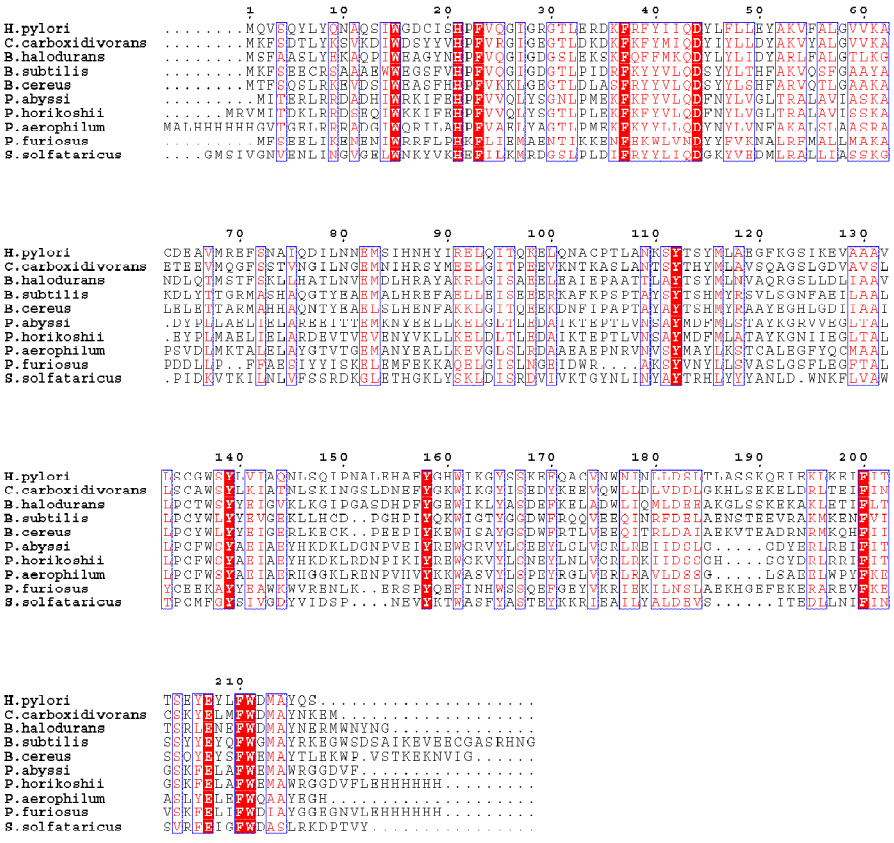

Fig.2.16. Sequence alignment of HP1287 from *H.pylori* and TenA from other bacteria

The *HP1287* gene cloned in our lab belongs to the *H. pylori* CCUG17874 strain. The sequence below corresponds to the full-length protein added 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.

```
       -24        -14          -4           7         17         27
MHHHHHHGKP IPNPLLGLDS TENLYFQ‖GID PFTMQVSQYL YQNAQSIWGD CISHPFVQGI

        37         47         57         67         77         87
GRGTLERDKF RFYIIQDYLF LLEYAKVFAL GVVKACDEAV MREFSNAIQD ILNNEMSIHN

        97        107        117        127        137        147
HYIRELQITQ KELQNACPTL ANKSYTSYML AEGFKGSIKE VAAAVLSCGW SYLVIAQNLS

       157        167        177        187        197        207
QIPNALEHAF YGHWIKGYSS KEFQACVNWN INLLDSLTLA SSKQEIEKLK EIFITTSEYE

       217
YLFWDMAYQS
```

‖ corresponds to the cleavage site for TEV protease.

HP1287	Properties
Number of amino acids	223
Molecular Weight	25643.2
Theoretical pI	5.23
Abs (1 g/l)	1.886

Table 2.2. Some properties of HP1287 protein, referred to the construct after cleavage with TEV protease.

2.2. Materials and Methods

2.2.1. Cloning, expression and purification of *H. pylori* HP1287

The *HP1287* gene was amplified by PCR from genomic *H. pylori* CCUG17874, using the primers 5'-CACCATGCAAGTTTCACAATATCTGTA-3' (forward, Topoisomerase recognition site underlined) and 5'-TTATCAACTTTGATACGCCATATCC-3' (reverse). It was then cloned into the pET151 vector (pET151; Invitrogen) in frame with an N-terminal His-tag flanked by a TEV proteolysis site, using a TOPO® Cloning kit by Invitrogen.

E. coli BL21(DE3) cells, harboring the pET151-*HP1287* plasmid, were grown in LB medium supplemented with 100 μg/mL ampicillin and the protein expression induced by 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG). The bacterial pellet was resuspended in 50 mM phosphate pH 7.4, 300 mM NaCl; cells lysis was performed by a two-step method, via incubation with lysozyme (1 mg/ml, 1 h, 4 °C) and sonication. The lysate was centrifuged to remove cell debris and loaded into a column containing 4 ml of Ni^{2+} charged Chelating Sepharose™ (GE Healthcare). After extensive washing using the lysis buffer, supplemented with 20mM imidazole, the resin was incubated overnight, at 4°C and under mild shaking,

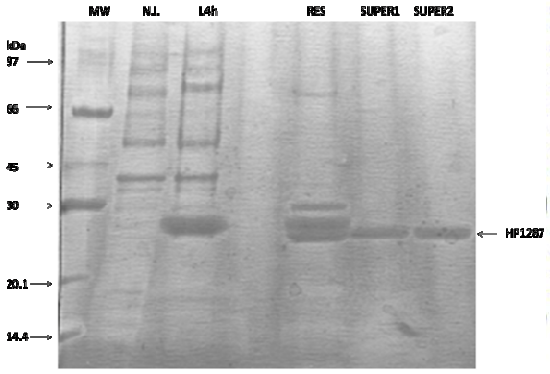

Fig. 2.17.A. SDS-PAGE of samples coming from different steps of HP1287 purification.

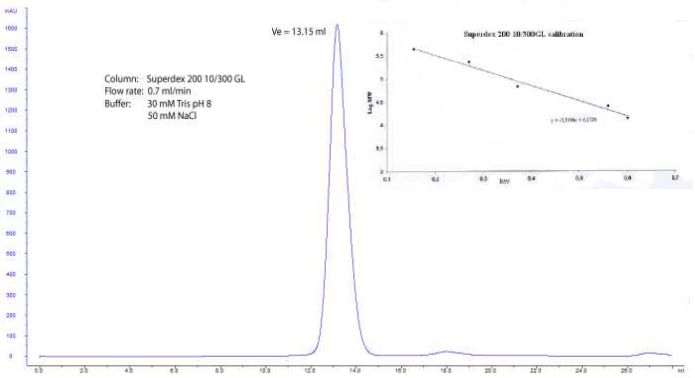

Fig. 2.17.B. Elution profile of HP1287 gel filtration chromatography.

with recombinant His_6-TEV protease. The supernatant was recovered by centrifugation, filtered and supplemented with 2 mM octyl-β-D-glucopyranoside to prevent HP1287 aggregation. The proteolytic product was further purified by Superdex 200™ 10/300 GL (GE Healthcare), equilibrated with 30 mM Tris pH 8, 50 mM NaCl. The protein was eluted as a single peak, roughly corresponding to a tetramer and migrated as a single 25 kDa species on SDS-PAGE (theoretical mass: 25,643.2 Da, confirmed by Mass Spectrometry). HP1287 was concentrated to 10 mg/ml for crystallization purposes.

Expression and purification of HP1287 F47Y were performed in the same conditions used for the native enzyme.

2.2.2. Mutagenesis of HP1287

The F47Y mutation was performed with QuikChange Site-Directed Mutagenesis Kit (Stratagene). The primers used were 5'-TATATCATTCAGGATTATTTGT**AT**CTTTTAGA ATACGCTAAGGTG-3' (forward, the mutagenesis codon is underlined) and 5'-TTAGCGTA TTCTAAAAG**AT**ACAAATAATCCTGAATGATATAAAAAC -3' (reverse). The pET151 HP1287 plasmid was amplified using *PfuTurbo* DNA polymerase and incubated with *Dpn I* to digest the template plasmid. Mutated plasmids were afterwards transformed into *E. coli* Top10 competent cells and selected on LB agar plates containing ampicillin (100 μg/ml).

2.2.3. Crystallization and structure determination

The best crystals of HP1287 were obtained at 20°C by vapour diffusion technique using a 4 mg/ml protein stock solution and 0.1 M Tris pH 8.5, 1.1 M Lithium sulphate, as precipitant. In particular, the highest quality crystals were obtained by the seeding technique with the help of the Oryx8 drop maker (Douglas Instruments). Crystallization of HP1287 F47Y was performed at 4°C.

Fig. 2.18.A. Oryx8 Protein Crystallization System by Douglas Instruments.

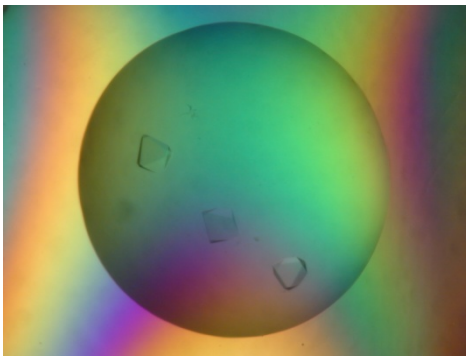

Fig. 2.18.B. HP1287 Crystals.

A preliminary diffraction data set at 3 Å resolution was measured at the XRD1 beamline of ELETTRA synchrotron (Trieste, Italy), whilst the best resolution data set for the wild-type enzyme (2.7 Å resolution) was collected at the ESRF beamline ID23-2 (Grenoble, France). An entire set of data was measured at 100°K from one crystal, using the precipitant solution including 20% glycerol as cryoprotectant. Crystals belong to space group $I4_122$, with cell parameters a = b = 148.42 Å, c = 233.52 Å. A dataset of the F47Y variant was measured at the ID14-4 beamline at a maximum resolution of 2.4 Å. The datasets were processed and scaled with programs MOSFLM and SCALA (Collaborative Computational Project 1994), respectively. As confirmed by the structure determination, the asymmetric unit contains two monomers, corresponding to a V_M of 6.27 $Å^3$/Da and a solvent content of about 80% of the

crystal volume. The structure was solved by molecular replacement with the Molrep software, using structure 1TO9 as the starting model. A twofold non-crystallographic axis relates the monomers A and B, whilst the other two monomers are generated by a crystallographic two-fold axis. The refinement was performed using the package CNS (Brunger et al. 1998) and, in the final steps, with Refmac (Murshudov et al. 1999). Several cycles of automatic refinement and manual model building reduced the crystallographic R_{factor} for the wild-type enzyme to 0.236 (R_{free} 0.257) for all the data from 125 Å to 2.7 Å resolution. All residues are clearly visible in the electron density. In monomer A, four additional residues at the N-terminus, deriving from the cloning construct, are also visible. The F47Y variant was refined starting from the molecular model of the wild-type enzyme, after substituting the mutated residue. The TLS refinement procedure (Painter, Merritt 2006) was introduced in the last cycles of refinement. Since the mutant diffracts to a higher resolution, the quality of its model presents slightly better statistics: R_{factor} = 21.8 and R_{free} = 23.0. The quality of both models, assessed using software Procheck (Laskowski et al. 1993), is as expected or better for a structure at this resolution.

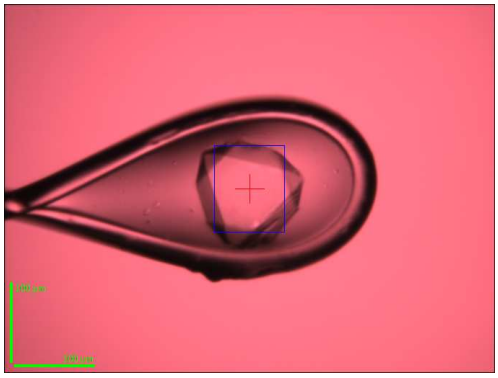

Fig.2.19. HP1287 crystal in a loop.

	Wild-Type	**Mutant F47Y**
Space Group	$I4_122$	$I4_122$
Cell parameters (Å)	a=b=148.42, **c**=233.52	a=b=148.73, **c**=233.57
Resolution (Å)	125-2.7 (2.85-2.70)	78 – 2.4 (2.53-2.40)
Independent reflections	36057 (5136)	51200 (7412)
Multiplicity	9.4 (8.8)	9.9 (10.2)
Completeness (%)	99.8 (99.0)	99.7 (100)
<I/σ(I)>	9.4 (3.8)	8.6 (3.7)
R_{merge}	0.081 (0.526)	0.153 (0.460)
Refinement		
Total number of atoms, including solvent	3623	3637
Mean B value (Å^2)	53.2	26.8
R_{cryst}	23.6(36.1)	21.8 (27.5)
R_{free} (%)	25.7(34.4)	23.0 (30.0)
Ramachandran plot [%]		
Most favored	90.2	94.6
Additionally allowed	8.1	5.4
Generously allowed	1.7	0.0
Overall G-factor	0	0.1
R.m.s. on bonds length [Å], angles (°)	0.016, 1.6	0.010, 1.2

Table 2.3. Statistics on data collection and refinement. A wavelength of 0.9794 Å was used. A CCD detector was positioned at a distance of 150 mm from the sample. Rotations of 1° were performed.

2.2.4. Enzymatic activity test

Hydrolytic activity towards the substrate 4-amino-5-aminomethyl-2-methylpyrimidine (Interchim, Montluçon, France) was determined, as described previously (Jenkins et al. 2008), by monitoring the release of ammonia through the glutamate dehydrogenase assay (Day, Keillor 1999). Recombinant HP1287 with a concentration of 2.4 μM, was added to a mixture of 5 units of glutamate dehydrogenase, 5 mM α-ketoglutarate, 0.1 mM EDTA, 0.250 mM NADPH and 20–480 μM 4-amino-5-aminomethyl-2-methylpyrimidine in two different buffers (20 mM sodium phosphate at pH 8 and 50 mM MES at pH 6). The reaction was followed by monitoring the decrease in A_{340} as a result of the enzymatic consumption of NADPH. The HP1287 enzyme concentration was calculated by measuring A_{280} and applying the theoretical extinction coefficient 48360 $M^{-1}cm^{-1}$, as estimated by ProtParam (Wilkins et al. 1999). The collected data were fitted to the Michaelis-Menten equation using GraphPad prism, version 5 (GraphPad Software Inc., San Diego, CA, USA), evaluating the initial rates by using the absorbance values at a fixed time in the linear segment of the registered curves.

To determine thiaminase II activity, 5 μM HP1287 was incubated overnight at 20 ºC with a mixture containing 2.5 mM thiamin, 30 mM Tris, 50 mM NaCl (pH 8.0). An aliquot of 100 μL from the reaction mixture was heated to 95 ºC for 5 min and centrifuged at 35 000 g to remove denatured protein. The reaction products were purified by RP-HPLC on a C_{18} column (Grace Vydac, WR Grace & Co-Conn, Columbia, MD, USA) in 20 mM phosphate buffer (pH 6.6). The elution of HMP, thiamin and thiazole was obtained using a gradient of methanol to a final concentration of 50% and was monitored by measuring A_{254}. Reaction products were identified by NMR and MS.

To evaluate thiaminase I activity, 1 μM HP1287 was incubated at room temperature with 100 μM 4-nitrothiophenolate, 800 μM thiamin in 50 mM phosphate buffer (pH 7.2), 100 mM NaCl, 2 mM Tris(2-carboxyethyl)phosphine (Hanes, Kraft & Begley 2007). The enzymatic activity was monitored at 411 nm for 15 min using a Shimadzu UV-2501PC spectrophotometer (Shimadzu Corp., Kyoto, Japan).

2.3. Results

2.3.1. Structure of wild-type HP1287

H. pylori HP1287 was produced starting from the *H. pylori* CCUG17874 genomic DNA. The protein was expressed in *Escherichia coli* with an N-terminal His-tag, cleaved by TEV protease after affinity chromatography and purified by gel filtration. The crystals obtained, despite their relatively large size, present a modest diffracting power, even when using a very brilliant synchrotron source. This may be ascribed to the very loose packing of the protein tetramers in the crystal cell, which leaves a large amount of empty space, ~ 80% of the volume, filled with solvent. The alignment of the amino acid sequence of HP1287 shows 33% identity and 51% similarity to the TenA protein from *B. subtilis*.

The 3D structure of the monomer is quite similar to that of the other members of the TenA family of known structure: twelve α-helices, labeled A–L, are arranged in a complex topology, as previously described (Toms et al. 2005). The assignment of secondary structure elements, made according to the software Procheck (Laskowski et al. 1993), is illustrated in Fig. 2.20. The slightly different number of a-helices, compared to other members of the same family, is a result of some pairs of helices, such D–E, H–I and J–K, comprising long helices interrupted by kinks, which break each long a-helix in two shorter ones. The superposition of

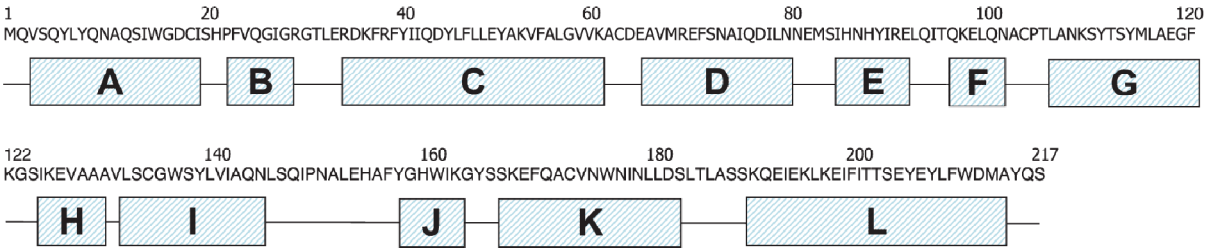

Fig. 2.20. Representation of Secondary Structure of HP1287.

the Cα atoms of one monomer with that of the other members of the family gives a rmsd of 1.7, 1.3 and 1.4 Å for models 1RTW, 1UDD and 1TYH, respectively. The major differences are observed in two regions: in the long stretch comprising residues 94–106 that connects helices E to G and includes the short helix F, and in residues 148–157 that connect helix I to helix J. Helix E is also slightly shifted with respect to the other models. The quaternary organization of the enzyme is that of a tetramer presenting 222 symmetry. One of the two-fold axes coincides with a crystallographic one, so that a dimer is present in the asymmetric unit. The only contacts in the dimer are made through α-helix G and its symmetry mate, which accounts for the burying of 1630 $Å^2$ of the solvent-accessible surface of each monomer. This dimer interacts with a second one burying a much large surface (7700 $Å^2$), which involves

parts of a-helices C, D, G and L and connections between helices C–D, G–H, K–L and F–G. The molecular weight from the gel filtration experiment indicates that the tetramer corresponds to the physiological unit.

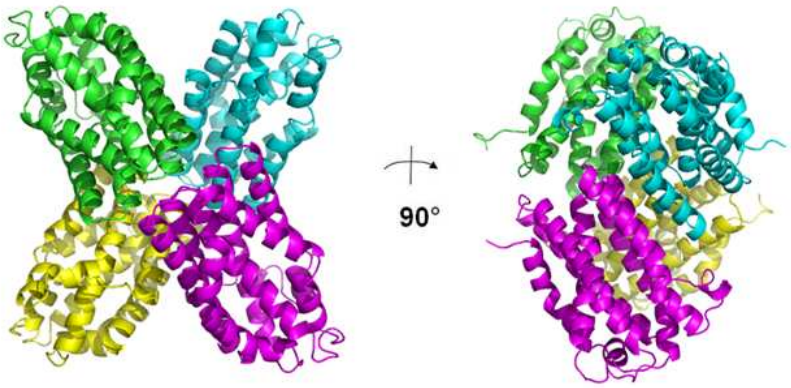

Fig.2.21.A. On the left, ribbon representation of HP1287 tetramer, on the right, ribbon diagram of the same tetramer equivalent to 90° rotation from the left view.

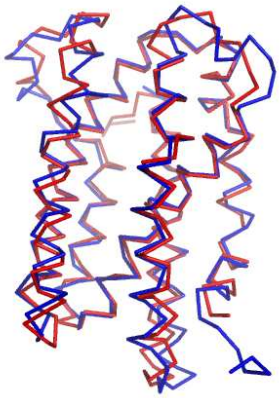

Fig.2.21.B. Superposition of main chains. In blue HP1287 from *H. pylori*, In red TenA from *B. subtilis*.

2.3.2. Structure of HP1287 F47Y

Because of the very low catalytic activity of the wild-type enzyme (see next below and Discussion), a mutant where Phe47 was substituted by a tyrosine was prepared, as described in the experimental procedure. Crystals of the F47Y variant are isomorphous with those of the wild-type protein and the two crystal structures are practically superimposable: the rmsd between equivalent Cα atoms is 0.73 *Å*. In particular, Tyr47 keeps the position previously held by phenylalanine, whereas the only significant difference between the two structures involves residues 79–84 of chain A. These connect α-helices D and E, but although, in the wild-type protein, they present only some irregularities, such that helices D and E can be considered as two parts of a long α-helix, in the F47Y variant, these become completely unfolded, breaking the continuity between the two helices. The same situation does not occur in the other monomer defining the asymmetric unit, where the electron density in this region is not very clearly defined.

2.3.3. Enzymatic activity and the putative catalytic site

A small cavity is present in each monomer, located among helices C, G, I and L. This cavity, which has been demonstrated to host the substrate in the *B. subtilis* enzyme (Toms et al. 2005), is quite a long tunnel that extends inside each protein monomer from the protein surface. The inner part of the cavity, which is connected to the solvent through a long tunnel,

is lined by residues Phe210 and 47, Trp211, Tyr51 and 139, Asp44 and Glu207 (Fig. 2.22.A). In one of the two monomers of the asymmetric unit (monomer D in our labeling system), a residual electron density is visible, whereas, in monomer A, the cavity is empty. Noticeably, an unknown ligand was also found in TenA from *P. horikoshii* (Itou et al. 2004) and from *P. furiosus* (Benach et al. 2005). The flat electron density likely corresponds to an endogenous compound of the *E. coli* where the protein was produced, or to a reagent used during

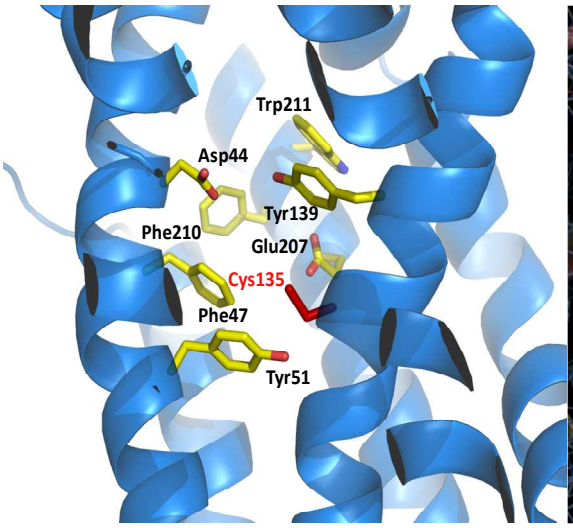

Fig. 2.22.A. HP1287 active site. In red the catalytic residue.

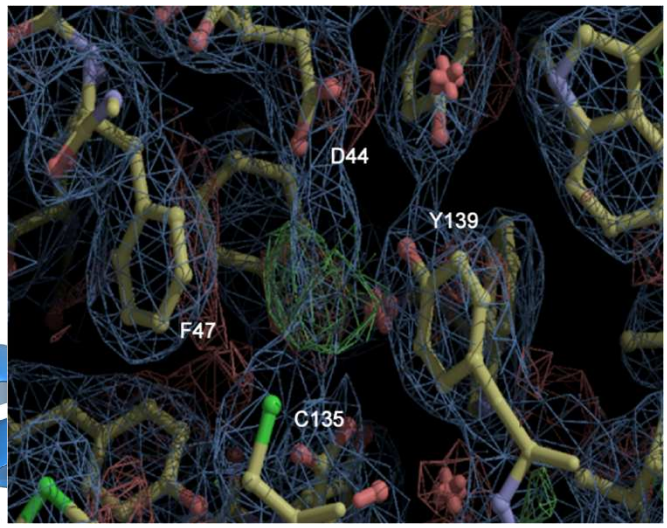

Fig. 2.22.B. Electron density map of HP1287 active site. A residual electron density is located between F47 and Y139 (Fourier map σ=1.8, Fourier difference map σ=3.8)

purification, possibly imidazole. It approximately mimics the HMP bound to the *B. subtilis* enzyme (PDB code: 1YAK). In our model, the pyrimidine ring is stacked with the aromatic rings of Tyr139 and Phe47 (where the latter replaces Tyr47 present in the *B. subtilis* enzyme) and lies coplanar with side chains of Cys135 and Asp44, as shown in Fig. 2.21.B.

Activity data at pH 8 indicate that the wild-type enzyme is poorly active on 4-amino-5 aminoethyl-2-methylpyrimidine: with a k_{cat} and K_M of 1.7 ± 0.2 min^{-1} and 58 ± 22 µM, respectively. The F47Y variant appears to be poorly active as well: with a k_{cat} and K_M of 0.06 ± 0.006 min)1 and 68 ± 16 µM, respectively. At pH 6, the activity is absent. Moreover, the enzyme does not present any activity on thiamin degradation.

Protein	K_m(µM)	K_{cat}(min^{-1})	K_{cat}/K_m($\mu M^{-1}min^{-1}$	Relative K_{cat}/K_m
HP1287 wild-type	58 ± 22	1.7 ± 0.2	0.029 ± 0.009	0.0015
HP1287 F47Y	68 ± 16	0.06 ± 0.006	0.00088 ± 0.00024	0.0005

Table 2.4. Activity of HP1287 wild-type and F47Y.

2.3.4. Thiamin metabolism in *H. pylori*

A comparative analysis of the thiamin biosynthetic pathway of more than 80 bacterial genomes was performed (Rodionov et al. 2002). The *H. pylori* genome includes two genes that code for enzymes possibly involved in the phosphorylation of HMP and hydroxyethylthiazole (HET), ThiD (HP0844) and ThiM (HP0845), respectively, and one responsible for the coupling of the HMP and HET moieties, corresponding to ThiE (HP0843) (Rodionov et al. 2002). By contrast, the bacterium apparently lacks the genes devoted to the biosynthesis of the thiamin precursor HMP and HET moieties. Moreover, the two genes HP1290 and HP1291 could define a divergon with the gene coding for the TenA enzyme, located far away from genes ThiD, ThiM and ThiE, which are likely involved in the thiamin biosynthesis pathway (Rodionov et al. 2002). Indeed, HP1290 shares a significant sequence similarity with PnuT, a component of the PnuC family of nonphosphorylated N-ribosylnicotinamide transporters (Rodionov et al. 2002). HP1291 is similar (34% amino acid sequence identity) to the thiamin pyrophosphokinase from *Bacteroides thetaiotamicron* (PDB code: 2OMK) and shares 24% identity with the mouse enzyme (PDB code: 2F17) (Liu, Timm & Hurley 2006).

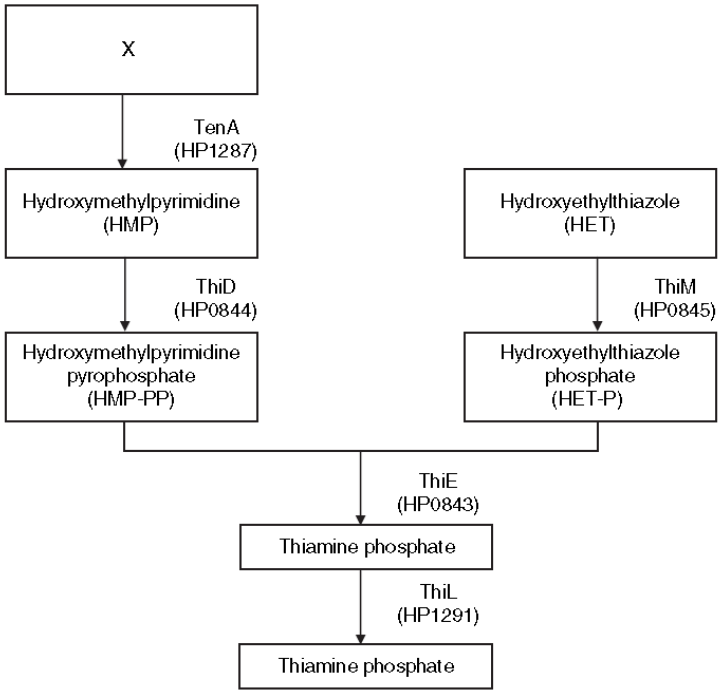

Fig. 2.23. Thiamin metabolism in *H. pylori*.

A homology model of all these proteins, with the exception of the putative transporter HP1290, was built using the swiss-model server (Arnold et al. 2006). The analysis of these structures (Fig. 2.23.) indicates that their active sites are structurally well preserved and that HP0843, HP0844, HP0845 and HP1291 can be considered as orthologues of ThiE, ThiD, ThiM and ThiL, respectively.

2.4. Discussion

The structure of HP1287 is very similar to that of *B. subtilis* TenA, with the few differences involving mainly the regions between helices E and G together with I and J, thus confirming that, from the structural point of view, it belongs to the thiaminase II enzymes family. The structure of the active site of the *B. subtilis* TenA enzyme is well characterized and, upon comparison with *H. pylori* TenA, a high degree of structural similarity is observed, with the exception of mutations in position 47, from Tyr to Phe, and position 51, from Phe to Tyr. The hypothesized mechanism for the reaction of *B. subtilis* TenA (Jenkins et al. 2008) assumes that the thiol group of Cys135 adds to C6 of the pyrimidine ring, favoring the exit of the aminic group. The subsequent addition of a water molecule and the expulsion of the active cysteine complete the reaction. Asp44 is positioned to stabilize and orient the binding of the substrate, and Tyr112, Glu205 (207 in HP-TenA) and Tyr47 assist the reaction. All these residues, with the exception of Tyr47, are present in our structure and their positions in the active site are conserved. Because the activity of our enzyme towards 4-amino-5-aminomethyl-2-methylpyrimidine is very modest, this suggests that a tyrosine at position 47 could play a crucial role in catalytic efficiency. Furthermore, our activity data are in good agreement with those obtained for the mutant Y47F of the *B. subtilis* enzyme (Jenkins et al. 2008): k_{cat} and K_M in the latter are reduced to values comparable to those found for the *H. pylori* enzyme. Tyr51, which replaces the phenylalanine present in other enzymes of this family, despite its close proximity to the substrate, is unable to compensate for the absence of Tyr47 because its orientation is incorrect with respect to the substrate. Mutation Y47F appears to be peculiarly conserved in *H. pylori* because it is present in all the strains sequenced to date, whereas, in most of the other bacteria, a tyrosine is present in this position. To test the role of Tyr47, the mutant F47Y was prepared. This mutation does not perturb the active site, which becomes even more similar to that of the *B. subtilis* enzyme. Nevertheless, the catalytic activity remains very low. A careful comparison of the active sites of the enzymes from the two species shows that, despite a complete conservation of the residues known until now to be involved in the catalytic mechanism, another significant difference is present in the *H. pylori* enzyme. In the latter enzyme, His86, which belongs to α-helix E, points towards the center of the active site cavity, making it smaller. Moreover, His86 is at a distance allowing possible interaction with the substrate. His86 is also present in the amino acid sequence of *B. subtilis* enzyme, although this part of α-helix E is distorted and the histidine points to the exterior of the proteins, towards the solvent. All these previous observations suggest that the active site of TenA has been slightly modified to act towards a different substrate: the hydroxyl group of

Tyr51 and His86 could be correctly positioned in the active site with respect to a different, unknown pyrimidine derivative. The presence of a limited number of enzymes involved in the thiamin biosynthesis in *H. pylori*, and the peculiar environment in which it thrives in, leads to the hypothesis of the existence of a reduced thiamin biosynthetic pathway. Indeed, degradation products of thiamin (Hartman et al. 1984) can be present in the stomach during digestion as a result of the processing and storage of foods (Richardson, Finley 1985). At the same time, the very acidic environment of the stomach makes the accumulation of FAMP very unlikely because it is mainly a base-degraded derivative of thiamin. We tentatively suggest the presence of an as yet unidentified peculiar precursor, deriving from the human stomach food assumption or processing, which is internalized through an unknown receptor in

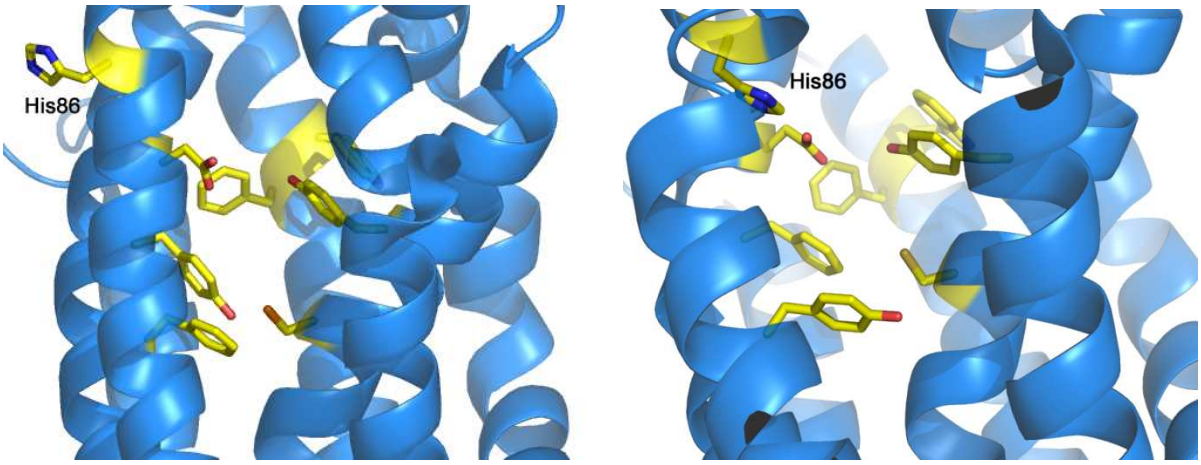

Fig. 2.24. The position of His86 in *Bacillus subtilis* TenA (on the left) and *Helicobacter pylori* HP1287 (on the right). His86 in H. pylori could be close enough to interact with the unknown substrate.

cooperation with the PnuC analog HP1290 transporter. It is converted by TenA to HMP, which is subsequently phosphorylated by ThiD (HP0844) to the activated compound HMP-PP. Phosphorylation of HET is catalyzed by ThiM (HP0845). The final synthetic reaction that combines the two, giving rise to thiamin phosphate, is promoted by ThiE (HP0843) and the conversion to thiamin pyrophosphate by HP1291, which consequently has been labeled ThiL. It must be considered that FAMP, which has been identified as the starting point of the thiamin salvage pathway in *B. halodurans* (Jenkins et al. 2007), apparently cannot play the same role in *H. pylori* because the amidohydrolase enzyme YlmB is also absent. In the earliest studies concerning TenA, the protein from *B. subtilis* was found to play an indirect role in the control of gene expression of degradative enzymes, mainly alkaline protease arpE (Pang, Nathoo & Wong 1991); however, on the basis of all subsequent findings with respect to this class of proteins, this role appears to be unlikely, at least in *H. pylori*. We cannot exclude the possibility that TenA, besides being an enzyme involved in thiamin biosynthesis, plays another relevant (despite still not being characterized) role in *H. pylori* and other

bacteria. Finally, the pivotal role of TenA in the thiamin biosynthetic route as the first enzyme of the pathway is in agreement with the relevance of this protein in the stomach colonization process, where the *tenA* gene has been found among the approximately 350 genes that could play a relevant role in its colonization and persistence (Baldwin et al. 2007).

Chapter III

Cloning, Expression, Purification and Crystallization of HP1028 from *Helicobacter pylori*

3.1. Introduction

3.1.1. Identification of new *H. pylori* genes involved in stomach colonization

Recently, new genes involved in *H. pylori* stomach colonization were identified by mutagenesis (Baldwin et al. 2007). A C57BL/6 mouse infection model was used to query a collection of 2,400 transposon mutants in two different bacterial strain backgrounds (NSH79 and NSH57) for *H. pylori* genetic loci contributing to colonization of the stomach. Behavior of transposon insertions in 758 different gene loci was monitored by microarray-based tracking of transposon mutants. Of the loci measured, 223 (29%) had a predicted colonization defect. These included previously described *H. pylori* virulence genes, genes implicated in virulence in other pathogenic bacteria and 81 hypothetical proteins. In particular, *HP1028* seemed to be important for stomach colonization in NSH79 background, but it looked like not to be a candidate mutant in the NSH57 strain background. The location of the transposon for the clones detected in the NSH57 screen was examined and the insertion in gene *HP1028* at nucleotide 150 (out of 450) was found. For the NSH57 screen, the insertion in *HP1028* appeared attenuated.

The uncharacterized colonization candidate *HP1028* gene was better analyzed by making independent null alleles. Null alleles were made in both the NSH79 and NSH57 mouse adapted strain backgrounds to determine if the phenotype was strain specific or not. These mutants were tested in a 1:1 competition experiment with wild-type bacteria. After 1 week of infection, the stomach was harvested and the competitive index was determined (CFU mutant bacteria/CFU wild-type bacteria). *HP1028* gave a phenotype in both strain backgrounds (NSH79 and NSH57), confirming that the leaky mutation was the reason of the different phenotypes in the two analyzed bacterial strains.

3.1.2. Function of HP1028

The specific function of *HP1028* in *H. pylori* is unknown, but some hypothesis on the putative function could be performed. The bioinformatic tools Gaggle identifies *HP1028* as a *cag* Pathogenicity Island-functionally associated gene (Shannon et al. 2006).

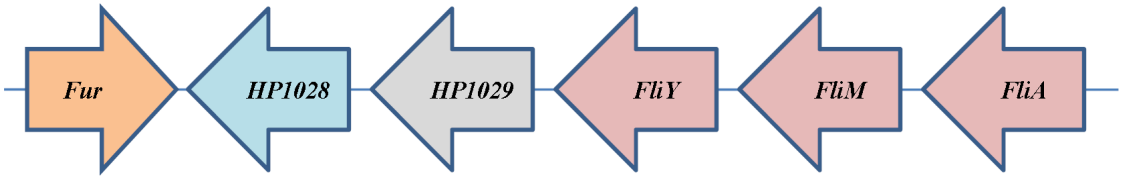

Fig. 3.1. Graphic representation of *HP1028* gene localization into *H. pylori* genome

HP1028 and *HP1029* are connected via gene cluster operon edges to *FliY* (*HP1030*) and *FliM* (*HP1031*) that encode key flagellar switch proteins (Fig. 3.1.). In this operon is also present an alternate sigma factor (σ28) gene *fliA (HP1032*), which has been implicated in mediating transcription of FlaA, the major flagellar subunit required for both motility (Josenhans, Labigne & Suerbaum 1995) as well as gastric colonization (Eaton et al. 1996). Close to the *HP1028* gene is also located *HP1027 (Ferric Uptake Regulator protein, fur),* encoding a protein involved in the modulation of urease expression in response to nickel (van Vliet et al. 2001). Since there is no known structure of any homologue protein, Homology modeling methods failed to predict a putative HP1028 fold. Moreover, the HP1028 protein sequence was manually submitted to Robetta, a structure prediction server (Chivian et al. 2003). Among the various proteins analyzed the most striking was the match of predicted three dimensional structure of HP1028 to β-chain of ebulin (PDB 1HWM), a ricin-like toxin (Fig. 3.2) (Pascal et al. 2001). Proteins with the conserved ricin domain are ribosome inactivating proteins widely distributed in plants, fungi, algae and bacteria. The β-chain is responsible for attaching the toxin to cell surface galactosides and facilitating delivery to the cytosol. The ebulin β-chain contains 31% β-sheet and 6% α-helix and has two distinct structural domains. Each of these domains contains three homologous subdomains designated α, β, and γ. Two of these subdomains, 1α and 2γ, have shown galactoside-binding features by crystallographic methods (Pascal et al. 2001).

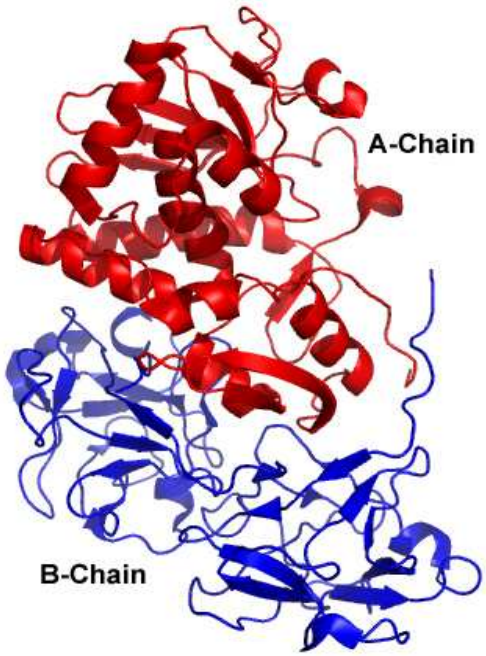

Fig. 3.2. Ribbon representation of Ebulin from *Sambucus ebulus* (PDB accession code 1HWN). The β-chain is colored in blue.

A simple similarity analysis using BLASTp (www.blast.ncbi.nlm.nih.gov/Blast.cgi) does not match any sequence of proteins whose function or structure is known, but it identifies a

conserved domain, named COG4731, present in several bacteria, like Burkholderiales, Campylobacterales, Caulobacterales, Rhizobiales and Spirochaetales. The function and the structure of COG4731 domain, belonging to DUF2147 superfamily, are unknown. These hypothesized functions of HP1028, coupled to its importance in *H. pylori* survival in stomach environment, implicate it in a likely role in *H. pylori* pathogenesis.

3.1.3. HP1028 gene and protein features

The *HP1028* gene was cloned from the *H. pylori* genome of CCUG17874 strain. Analysis with SignalP (www.cbs.dtu.dk/services/SignalP/) bioinformatics tool identifies a signal peptide at the N-terminus and a predicted proteolytic site between 16th and the 17th amino acid of HP1028 sequence. The sequence below corresponds to recombinant protein, without signal peptide, added of a N-terminal His-tag flanked by a cleavage sequence for TEV protease, coming from the construct codified by pET151 expression vector.

-24	-14	-4	23	33	43
MHHHHHHGKP	**IPNPLLGLDS**	**TENLYFQ‖GID**	**PFT**VELPGIY	QTQEFLYMKS	SFVEFFEHNG
53	63	73	83	93	103
KFYAYGISDV	DGSKAKKDKL	NPNPKLRNRS	DKGVVFLSDL	IKVGKRSYKG	GKAYNFYDGK
113	123	133	143	153	163
TYYVRVAQNS	NGDLEFTSSY	DKWGYLGKTF	TWKRLSDEEI	KNLKLKRFNL	DEVLKTIKDS

PI

‖ corresponds to the cleavage site for TEV protease.

HP1028	Properties
Number of amino acids	155
Molecular Weight (Da)	17946.3
Theoretical pI	9.36
Abs (1 g/l)	1.526

Table 3.1. Some properties of HP1028 protein, referred to the construct after cleavage with TEV protease.

3.2. Materials and Methods

3.2.1. Cloning of *HP1028*

HP1028 open reading frame was PCR amplified from the genomic DNA of *H. pylori* strain CCUG17874, using the primers 5'-CACCGTAGAGTTGCCTGGAATTTATC-3' (forward, Topoisomerase recognition site underlined) and 5'-CTATTAGATAGGGCTATCTTTAATG -3' (reverse). The PCR product was inserted into the expression pET151/D-TOPO directional vector, using TOPO® Cloning kit by Invitrogen.

3.2.2. Mutagenesis

Since the amino acid sequence of HP1028 contains only one internal methionine residue, because of the removal of the signal peptide containing the initial N-terminal methionine, two mutants with one and two additional methionine residues were produced in order to solve the phase problem, using QuikChange® Site-directed Mutagenesis Kit (Stratagene). The mutated amino acids, L129M and A110M, were selected as the most variable amino acids in a multiple alignment of all the HP1028 sequences available from different *H. pylori* strains. A single mutant containing the mutation L129M and a double mutant L129M/A110M were created. For the L129M mutation the following primers were used: FW 5'-CTATGATAAATGGGGGTATATGGGCAAGACTTTCACTTGG-3' and RV 5'-CCAAGTGAAAGTCTTGCCCATATACCCCCATTTATCATAG-3', whereas for A110M mutation FW 5'-GACCTACTACGTGAGAGTCATGCAAAATTCAAACGGGG-3' and RV 5'-CCCCGTTTGAATTTTGCATGACTCTCACGTAGTAGGTC-3' were used. HP1028 double and single mutants were obtained by consecutive PCR reactions using a high fidelity thermostable DNA polymerase (*Pfu*Turbo DNA polymerase, Stratagene). Single clones were then sequenced to confirm the occurrence of the desired mutations. *pET151 HP1028 L129M* and *pET151 HP1028 L129M A110M* were obtained for the transformation in B834(DE3) *E. coli* competent cells.

3.2.3. Dynamic Light Scattering (DLS) analysis

A sample of HP1028, in buffer 30 mM Tris pH 7.5, 250 mM NaCl, 10 mM galactose, coming from preliminary purification trials was analysed at DLS Zetasizer Nano ZS (Malvern Instruments). Since the protein was aggregated, aggregation-inhibition tests were performed using detergents at concentration lower than the CMC value (Gall, Ruff & Moras 2003). 20 µl of samples of HP1028 protein, at 5 mg/ml concentration, were incubated with different

detergents at 0.5 CMC concentration. The detergents used included C12E9, C12E8, n-Dodecyl-β-D-maltoside, CTAB, LDAO, n-Octyl-β-D-thioglucoside, n-Octanoylsucrose, MEGA®-8, n-Heptyl-β-D-thioglucopyranoside, CHAPS, FOS-Choline®-9. C12E8 seemed to inhibit HP1028 aggregation at 55 µM (0.5 CMC) concentration. Further DLS analysis for crystals optimization purposes revealed that adding C12E8 at 0.25 CMC (27.5 µM) concentration to the buffer was sufficient to inhibit aggregation.

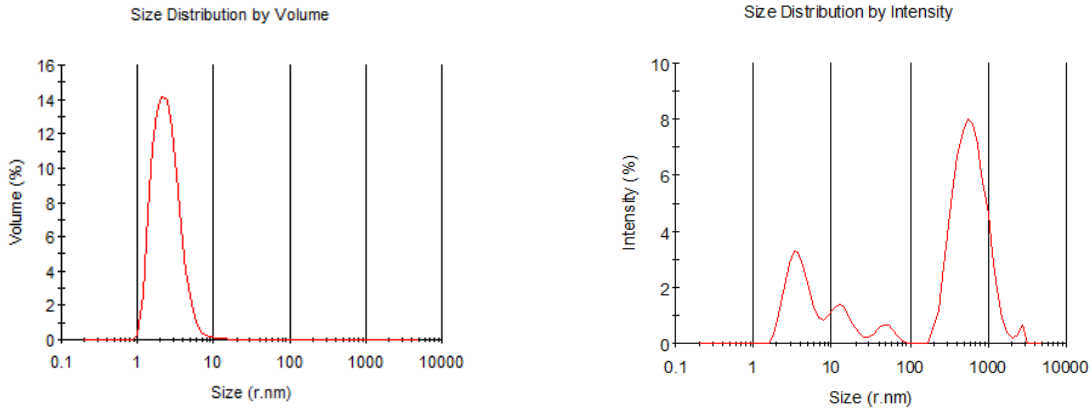

Fig. 3.3. Distribution diagrams of HP1028 without detergent: Volume (%) vs Size radius (nm) (left) and Intensity (%) vs Size radius (nm) (right).

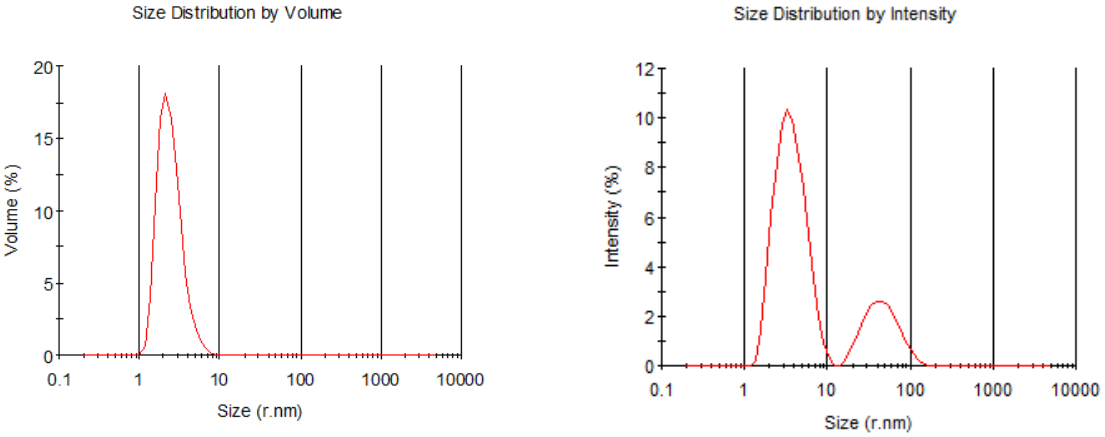

Fig. 3.4. Distribution diagrams of HP1028 incubated with 0.5 CMC C12E8: Volume (%) vs Size radius (nm) (left) and Intensity (%) vs Size radius (nm) (right).

3.2.4. Expression and Purification

The plasmid *pET151-HP1028* codifying for native HP1028 was used to transform *E. coli* BL21(DE3) strain. Cells harboring *pET151-HP1028* were grown at 37°C in 3 liters LB medium containing 100 µg/ml ampicillin until 0.6 OD_{600}. The protein expression was induced by adding 1 mM isopropyl-β-thiogalactopyranoside (IPTG) and incubating the bacterial cells in mild shaking for 5 h at 30 °C. Cells were harvested by centrifugation at 11 000 g for 8 min and suspended in 30 mM Tris-HCl, pH 7.5, 250 mM NaCl, 10 mM galactose, 0.05% NaN_3,

55 µM C12E8 (0.5 CMC). Re-suspended cells were lysated, by incubating with lysozyme 1 mg/ml for 1 h at 4°C and by sonication in presence of protein inhibitors (Protease Inhibitor Cocktail, Sigma-Aldrich). After centrifugation (25 min at 40 000 g), the soluble fraction was loaded onto a 1 ml His-Trap column (GE Healthcare) equilibrated with the buffer. After two washing steps with the buffer supplemented with 20 mM imidazole and 40 mM imidazole, the protein was eluted in imidazole gradient to 500 mM. The His-tag was cleaved by enzymatic digestion with 1.5 mg Tobacco Etch Virus (TEV) protease, carried out adding 0.5 mM EDTA and 1 mM DTT to protein solution and incubating O.N. at 4 °C. The protein was purified

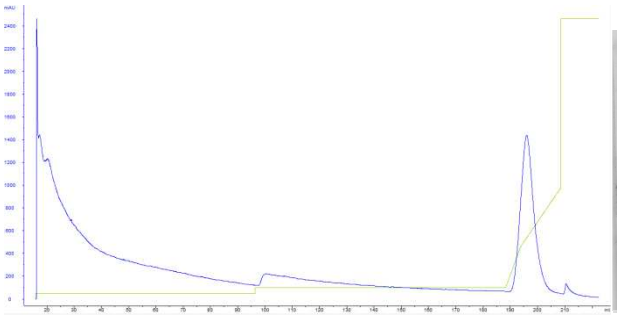

Fig. 3.5.A Chromatogram of HP1028 Ni^{2+}-affinity.

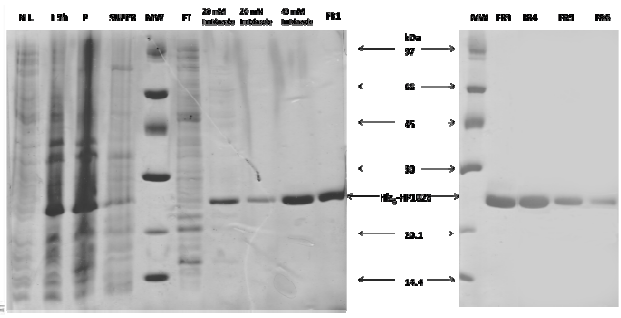

Fig. 3.5.B. SDS-PAGE of samples coming from different steps of affinity chromatography. Strangely, His_6-HP1028 seemed to have an higher apparent MW.

from the cleavage solution by diluting the imidazole concentration to 20 mM with buffer, loading the protein solution onto a 1 ml His-Trap column and collecting the flow through.

The protein was concentrated by ultrafiltration (Vivaspin 15R 5,000 MW, Sartorius) for loading on XK 16/20 size exclusion column (GE Healthcare) loaded with about 24 ml Sephacryl S-100 HR (GE Healthcare). The monodisperse protein was separated from the aggregated fraction and concentrated to 50 mg/ml for crystallization purposes. The proteins concentration was determined by UV/VIS spectroscopy (280 nm, Cary 50 Bio UV-Visible spectrophotometer, Varian Inc.) using the theoretical absorption coefficient.

Protein purity was checked by SDS-PAGE at the end of the purification step.

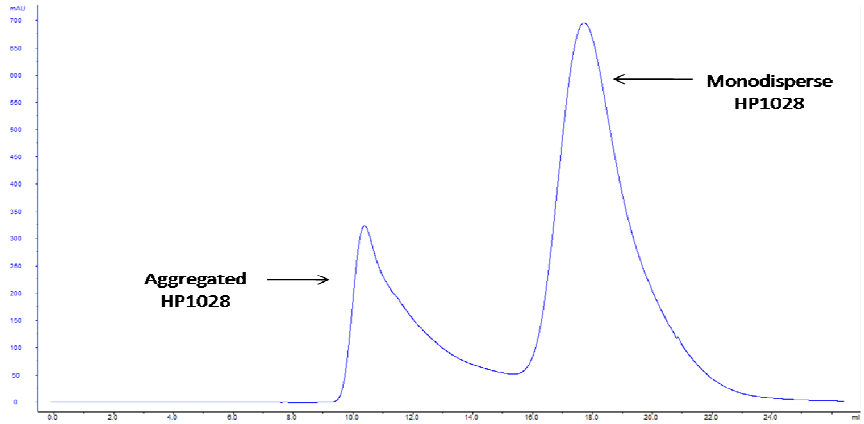

Fig. 3.6.A. Elution profile of HP1028 gel filtration chromatography.

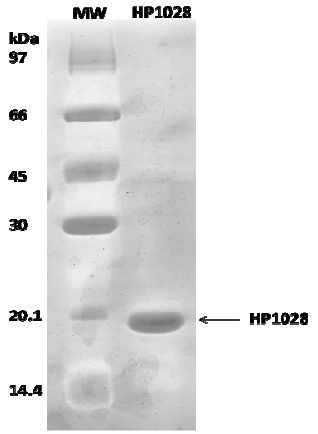

Fig. 3.6.B. SDS-PAGE of HP1028 at the end of the purification.

Expression and purification of HP1028 single (L129M) and double (L129M/A110M) mutants were performed using the same method, but the expression of the mutant proteins was performed using a seleno-methionine containing medium. For the expression of seleno-methionine derivated proteins, a *Met* auxotroph *E. coli* strain (B834) was used. The protein expressions were performed in 2 liters of a seleno-methionine-containing medium, composed by minimal medium M9 added of glucose 2%, vitamin B_1, small quantity of metals, amino acids and seleno-methionines. In some purifications, for crystal optimization a lower concentration of C12E8 was used (27.5 μM, 0.25 CMC) in purification buffers.

3.2.5. Western Blotting

HP1028 samples were re-suspended in loading buffer and separated by SDS-PAGE. After the SDS-PAGE, the resolved proteins were transferred electrophoretically onto a nitrocellulose membrane (HybondTM, GE Healthcare) and incubated with primary antibody against His-tag in blocking reagent O.N. at 4 °C. The membrane was washed three times in PBS-Tween (0.15% Tween 20), 3% BSA and then incubated for 1 h at room temperature with secondary antibody (anti-mouse IgG, linked to alkaline phosphatase) diluted in blocking reagent. The membrane was washed as described above and developed using 5-bromo-4-chloro-3-indolyl-1-phosphate (BCIP) and nitro blue tetrazolium (NBT) (Western Blue®, Promega), following the manufacturer procedure.

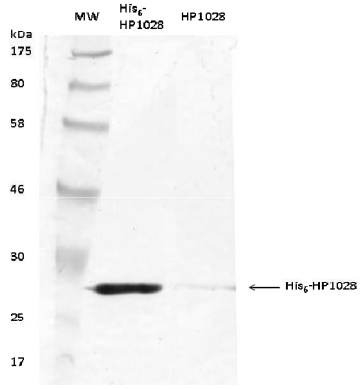

Fig.3.7. Western Blotting of His_6-HP1028.

3.2.6. Circular dichroism analysis

To determine the secondary structure of native HP1028, a Circular Dichroism analysis was performed. CD measurements of the protein diluited to 1 mg/ml final concentration were performed using a JASCO-J720 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. Ten spectra were accumulated and averaged followed by baseline correction by subtraction of the buffer. Mean residue weight ellipticities were 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 showed 48% β-sheet, 6% α-helix, 30% random coil and 16% β-turn. Therefore, the protein was predicted to be mainly a β-protein.

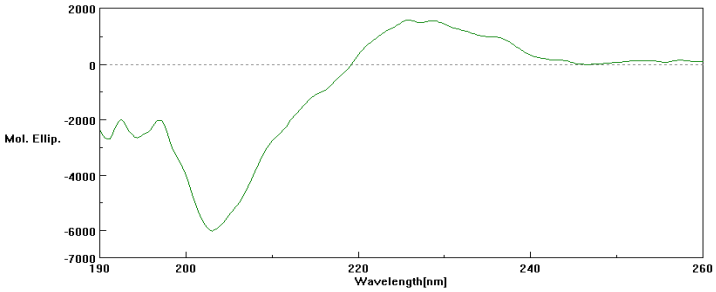

Fig. 3.8. Far-UV CD spectum of native HP1028.

3.2.7. Crystallization and preliminary X-ray diffraction trials

The purified native HP1028 protein has been diluted to 35 mg/ml for crystallization trials. Crystallization conditions were identified using sitting drop methods and Oryx8 crystallization robot (Douglas Instruments). Several crystallization precipitants were used from different commercial kits: Structure Screen I and II from Molecular Dimensions, The PACT Suite, MB Class I and II from Qiagen. General crystallization trials were performed mixing 0.35 µl protein solution with 0.35 µl precipitant on 96-well plates (Douglas Instruments), containing 80 µl precipitant in every well. Three crystallization temperatures were tested: 4°C, 10°C and 20°C. Preliminary crystals of native HP1028 were obtained in two different conditions: with the screening solution 9 and with the screening solution 35 from Structure Screen I. The first crystal typology, obtained at 30 mg/ml protein concentration with screening solution I 35, composed by 0.2 M Lithium sulphate monohydrate, 0.1 M Tris pH 8.5, 30% PEG 4000, had a monolithic cubic shape (Fig. 3.9.). Preliminary diffraction trials reveal that the crystals belonged to P23 space group with cubic unitary cell and diffracted to 4 Å maximum resolution, even after incubation of the crystals with different cryoprotectant solutions (precipitant added of 15% glycerol or 5% Ethylenglycole or 5% PEG 400). The crystals obtained at 35 mg/ml with screening solution I 9, composed by 20% isopropanol, 0.1 M Sodium Citrate pH 5.6, 20% PEG 4000, had a long-shape and they appeared twinned. Crystals belong to space group $P2_12_12_1$ and preliminary diffraction data were collected at 3.3 Å maximum resolution, using a cryoprotectant solution containing the precipitant added of 10% Glycerol. Crystals

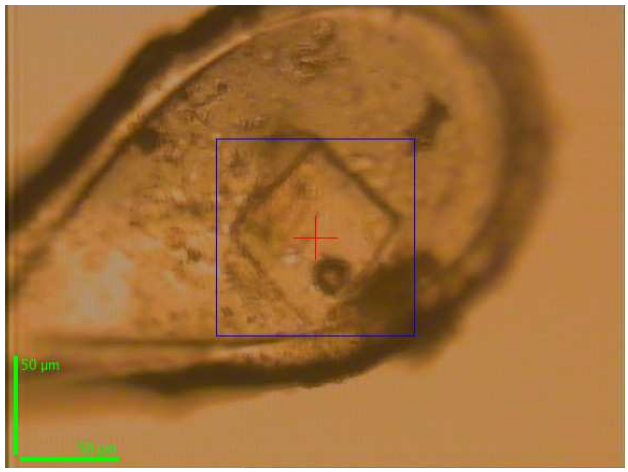

Fig. 3.9. Native HP1028 crystal grown in Structure screen I 35 mounted in a loop.

optimization trials of both the crystal types were performed: changing the components of the precipitant, the dimension of the drops, the temperature, the concentration of the protein, addition of adjutants, small organic molecules and detergents, and using *seeding* techniques. No higher resolution data were collected for the first crystal type, obtained with Structure Screen I solution 35, even if crystals were big enough for data collection, since they poorly diffract, probably because of the high solvent content. All the methods used for crystals optimization did not permit to improve the quality of the native HP1028 crystals obtained with both the precipitants. On the other hand, HP1028 L129M and HP1028 129M/A110M crystals, obtained with screening solution I 9, diffracted to higher resolution (Fig. 3.10.). HP1028 L129M/A110M crystals were optimized reducing the concentration of C12E8 detergent to 0.25 CMC (27.5 μM) in the purification buffer.

Some other trials to improve the quality of the crystals were attempted. The vapour-diffusion crystallization drops were added of detergents (Detergent screen, Hampton research) at 0.25 CMC and 0.125 final concentrations. ZWITTERGENT® 3-14, FOS-Choline® -12 and CHAPSO seemed to slightly improve the crystals quality. Diffraction data were collected from HP1028 L129M and HP1028 L129M/A110M crystals, respectively, to 2.4 Å and 2.5 Å maximum resolution (see paragraph 3.2.8.).

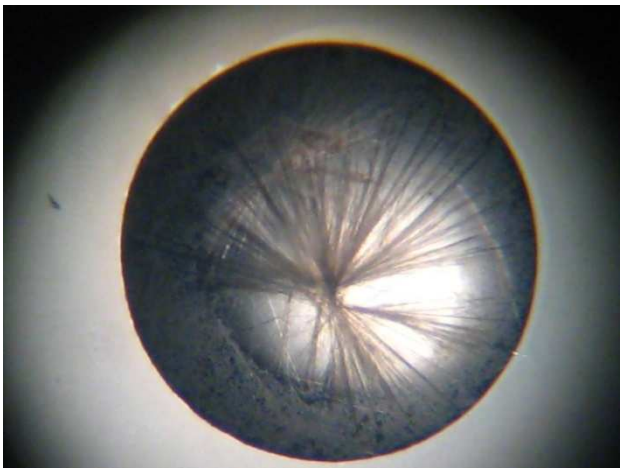

Fig. 3.10. HP1028 L129M A110M crystals grown at 35 mg/ml protein concentration in 20% isopropanol, 0.1 M Sodium Citrate pH 5.6, 20% PEG 4000 at 20°C.

To increase further on the anomalous signal of the selenium at some wavelength, oxidation of the selenium in the seleno-methionine derivatized protein was attempted. The selenium oxidation was performed following a protocol present in literature (Sharff et al. 2000). A HP1028 L129M/A110M sample was oxidated with hydrogen peroxide 0.1% for 90 seconds and subsequently diluted in buffer and concentrated for crystallization trials. No crystals of oxidated HP1028 L129M/A110M were obtained.

To obtain initial approssimate phases to solve the structure by MAD, as well as the selenomethionine derivated protein, an Iodine-derivatized protein was produced (Miyatake, Hasegawa & Yamano 2006). A 0.5 μl drop of 0.67 M KI, 0.47 M I_2 was located next to the crystallization drop, containing HP1028 crystals, and the vapour-diffusion chamber was

sealed. The two droplets were incubated for 1 hour at room temperature. The Iodine-derivatized crystals diffracted to a very low resolution.

To obtain initial phases by MIR method, HP1028 L129M crystals were soaked with precipitant added of 10 mM $PtCl_6$, $PtCl_2$, $AuCl_3$, $TICl_3 \cdot H_2O$, $AgNO_3$, $IrCl_3 \cdot H_2O$, $C_8H_8HgO_2$, $Eu(NO_3)_3 \cdot 6H_2O$, $GdCl_3 \cdot H_2O$ or 0.3 µl of 50 mM solution of each heavy metal was added to 1.5 µl crystallization drops before crystals formation. Derivatized crystals obtained in both the methods diffracted poorly.

3.2.8. Data collection

Cryoprotectants used for the crystal type obtained with precipitant Structure Screen 9 were 20% isopropanol, 0.1 M Sodium Citrate pH 5.6, 20% PEG 4000 added of 5% Glycerol or 10% PEG400.

The soaked crystals were then mounted on a cryo-loop and immediately frozen under a nitrogen stream at 100 K. Diffraction data of HP1028 L129M, crystallized in precipitant Structure Screen I 9 were collected on beamline ID14-4 using a Q315r ADSC X-ray detector at the European Synchrotron Radiation Facility (ESRF) in Grenoble, with 1° oscillation per image. Instead, HP1028 L129M/A110M data were collected at ID23-1 (ESRF) with 1° oscillation per image using an ADSC Q315R detector.

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). Both the crystals of HP1028 L129M and of HP1028 L129M/A110M belong to $P2_12_12_1$ space group with slightly different cell parameters: a=34.1644 b=78.9431, c=130.8406 Å and a=34.1503, b=80.9116, c=131.3469 Å, respectively. Data at different wavelengths, corresponding to remote and peak, for the MAD experiment were collected from several crystals. Some of the data-collection statistics are shown in Tables 3.2 and 3.3. MAD approaches to determine the positions of Selenium atoms and to calculate the phases were performed using the *SHARP/autoSHARP* suite (Bricogne et al. 2003).

3.3. Results and Discussion

The native HP1028 protein was expressed in *E. coli* in a soluble form, without the putative signal peptide identified by SignalP. After purification trials by Ni^{2+}-affinity chromatography, the HP1028 protein resulted to be aggregated from DLS analysis. To inhibit the hydrophobic unspecific interactions between protein molecules, protein samples were incubated with different detergents at concentration lower than CMC and one detergent (C12E8) was identified as able to disaggregate the protein (Fig. 3.3., Fig. 3.4.). A detergent concentration lower than CMC was used to cover aggregation-prone hydrophobic surfaces of HP1028 and to avoid micelles-formation in protein solution. Monodisperse HP1028 protein was purified in high yield for crystallization purposes in a C12E8-containing buffer using Ni^{2+}-affinity chromatography (Fig.3.5.A) and size exclusion chromatography (Fig. 3.6.A.), that suggested a monomeric state of HP1028 protein. High purity of the protein solution at the end of the purification step was confirmed by SDS–PAGE (Fig. 3.6.B). Interestingly, the apparent molecular weight of His_6-HP1028 (Fig. 3.5.B.) seemed to be higher than expected, about 27 kDa, whereas the HP1028 protein after cleavage of the His-tag had a normal SDS-PAGE migration pattern (Fig. 3.6.B.). Western Blotting analysis (Fig. 3.7.) was performed on samples after affinity chromatography using an antibody against His-tag to validate the presence of His_6-HP1028.

On the other hand, further DLS analysis indicated a dimeric state of HP1028 protein: these discrepancies in Molecular weight estimation form different techniques could be caused by the low resolution of the manually-packed size exclusion column used for the purification, the binding of detergent molecules on the protein surface and the hydrodynamic volume of the protein that could be different in presence of His_6-tag or not.

Interestingly, HP1028 was purified with a sephacryl-containing gel filtration column, because HP1028 seemed to bind the Sephadex resin, contained in Superdex columns. Since Sephadex resins are made by dextran, a cross-linked polysaccharide, it is possible to hypothesize an interaction between the HP1028 protein and carbohydrates, as suggested by the bioinformatic tool Robetta. Circular dichroism analysis (Fig. 3.8) indicated a prevalence of β-sheet secondary structure (48%), like ebulin β-chain, enhancing the hypothesis of a similarity between the two structures. Following this hypothesis, the protein buffer was added of 10 mM galactose, with the purpose of stabilizing the protein during the crystallization trials and, if the protein-galactose interaction exists and is strong enough, of identifying the binding site by co-crystallization. If this hypothesis was confirmed, HP1028 could be identified as an adhesin or a carrier for carbohydrates.

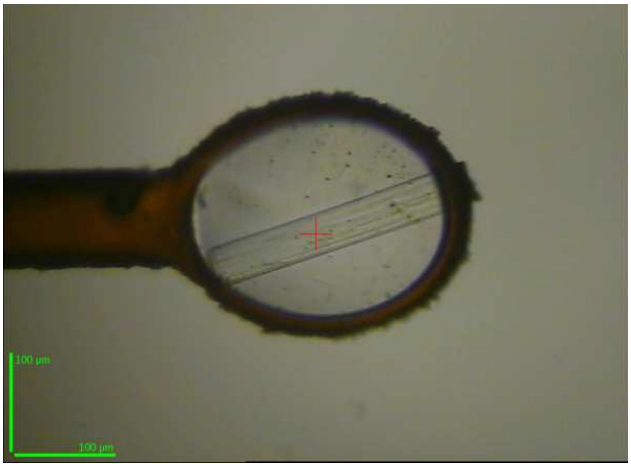

Fig. 3.11. A crystal of HP1028 L129M A110M with seleno-methionines in a loop.

To obtain approximate phases by MAD method, a single mutant (L129M) and a double mutant (L129M/A110M) HP1028 were produced. Single and double mutant HP1028 proteins were expressed, marked with seleno-methionines and purified by affinity chromatography and gel filtration, similarly to the native HP1028 protein. Optimized crystals of L129M and L129M/A110M HP1028 with seleno-methionines, obtained with Structure screen solution I 9, were used for data collection.

After extensive screening of these crystals, data sets were collected at 2.4 Å and 2.5 Å, respectively, for single mutant and double mutant protein (Tables 3.2. and 3.3). The crystals belong to orthorhombic $P2_12_12_1$ space group.

Data set	**Remote**	**Peak**
Beamline	ID14-4	ID14-4
Wavelength	0.8726	0.9795
Space group	$P2_12_12_1$	$P2_12_12_1$
Unit cell parameters (Å)	a=34.2096 b=79.2777, c=131.7389	a=34.1644 b=78.9431, c=130.8406
Resolution (Å)	79.31-2.85 (2.85-2.70)	78.94-2.53 (2.53-2.40)
Total No. of reflections	70998 (10588)	188390 (26824)
No. of unique reflections	10461 (1497)	14580 (2078)
Multiplicity	6.8 (7.1)	12.9 (12.9)
Anomalous multiplicity	3.6 (3.7)	6.9 (6.7)
Completeness (%)	99.9 (100.0)	100.0
Anomalous completeness	100.0 (100.0)	100.0
R_{merge}	0.131 (0.395)	0.159 (0.430)
I/σ(I)	10.8 (5.0)	15.9 (5.3)

Table 3.2. Data-collection and processing statistics relative to HP1028 L129M. Values in parentheses are for the highest resolution shell.

Data set	Remote	Peak
Beamline	ID23-1	ID23-1
Wavelength	0.97206	0.97887
Space group	$P2_12_12_1$	$P2_12_12_1$
Unit cell parameters (Å)	a=34.3751 b=82.3462 c=131.1900	a=34.1503 b=80.9116 c=131.3469
Resolution (Å)	82.48-2.95 (2.95-2.80)	80.85-2.64 (2.64-2.50)
Total No. of reflections	56527 (8355)	73968 (10450)
No. of unique reflections	9628 (1383)	13201 (1880)
Multiplicity	5.9 (6.0)	5.6 (5.6)
Anomalous multiplicity	3.3 (3.3)	3.1 (3.0)
Completeness (%)	99.2 (100.0)	99.4 (98.7)
Anomalous completeness	98.0 (99.2)	97.8 (95.8)
R_{merge}	0.165 (0.498)	0.112 (0.499)
I/σ(I)	12.4 (4.7)	14.4 (5.2)

Table 3.3. Data-collection and processing statistics relative to HP1028 L129M/A110M. Values in parentheses are for the highest resolution shell.

Based on the molecular weight of the protein (17946.3Da) and the volume of the asymmetric unit, the Matthews parameters (Matthews 1968) for one, two, three and four molecules in the asymmetric unit are 5.06, 2.53, 1.69 and 1.26 $Å^3$/Da respectively. This suggests the presence of two molecules in the asymmetric unit, which corresponds to 51.4 % solvent content. Several trials were performed to phase HP1028 L129M and HP1028 L129M/A110M using MAD method and different combinations of data sets but as yet, we have not been able to obtain a solution. This is probably attributable to the low quality of the anomalous data and to the low multiplicity of some data sets. Further data are going to be collected to attempt to solve the three-dimensional structure.

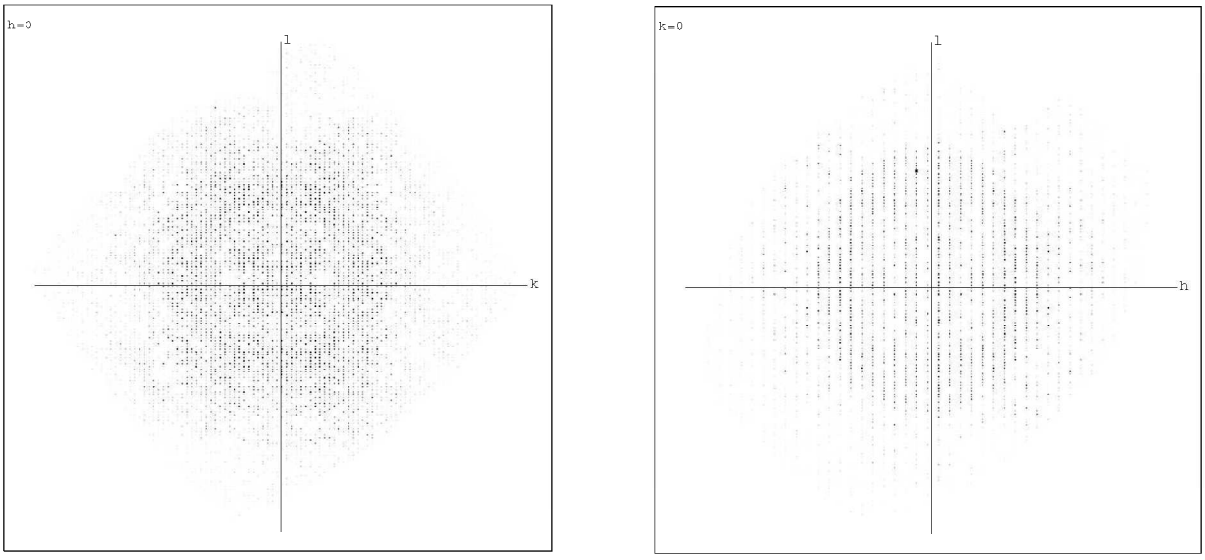

Fig. 3.12.A. Reciprocal space planes of HP1028: 0kl (left) and h0l (right)

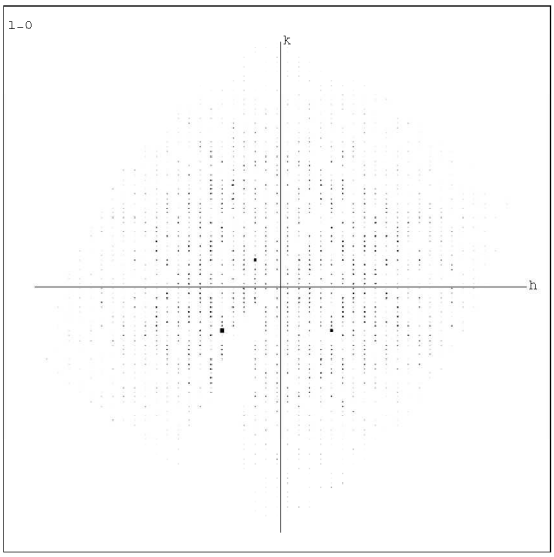

Fig. 3.12.B. Reciprocal space plane hk0 of HP1028

Chapter IV

Cloning, Expression and Purification of HP0175 from *Helicobacter pylori*

4.1. Introduction

4.1.1. Identification of HP0175 in *Helicobacter pylori* secretome

HP0175 from *Helicobacter pylori* was previously identified in 1998 by proteomic studies (McAtee et al. 1998). After analysis of the protein antigens of the gastric pathogen *Helicobacter pylori* recognized by human sera, HP0175 was identified as one of the highly and consistently reactive antigens. Further clinical studies (Atanassov et al. 2002) identified HP0175 as one of five antigens of *H. pylori* preferentially recognized by antibodies in patients affected by gastroduodenal ulcers and non-ulcer dyspepsia, suggesting a possible use of this protein in *H. pylori* vaccine development. Studies (Kim et al. 2002; Bumann et al. 2002) of *H. pylori* secretome revealed that HP0175 was secreted in the medium and, since this protein was too small to contain the multiple domains necessary for autotransport (Schmitt, Haas 1994), a type II-like secretion mechanism was hypothesized, *via* the general secretory pathway of Gram-negative bacteria.

4.1.2. HP0175 is a toxin with Peptidyl-prolyl *cis,trans*-Isomerase activity

PPIases catalyse the *cis-trans* isomerization of proline imidic peptide bonds in oligopeptides and include distinct classes of proteins, characterized by different drug specificity and by different structures: cyclosporin A-binding cyclophilins, FK506-binding proteins and parvulin-like PPIases that do not bind immunosuppressants (Shaw 2002).

Cyclophilins could play important roles on protein trafficking, assembly, immune-system modulation and cell signaling (Obchoei et al. 2009) and are composed by an eight-stranded β-barrel that forms a hydrophobic binding site for Cyclosporin A (Ke et al. 1991). FK506-binding proteins (FKBPs) are chaperons most highly expressed in the nervous system (Shim et al. 2009), they play an important role in the T-cell activation (Kang et al. 2008) and are characterized by an amphipathic five stranded β-sheet located around a single short α-helix (Michnick et al. 1991). Parvulin-like PPiases consist of a half β-barrel, four antiparallel strands surrounded by four α-helices, and have been demonstrated to be involved in cell regulatory mechanisms (Ranganathan et al. 1997). Recently, Parvulin-like PPiases have been shown to be involved in cell signaling. Their deregulation can lead to cancerogenesis. In particular, the parvulin-like PPIase Pin1 controls the conversion of peptidyl-proline bond conversion from *cis* to *trans*, only when the preceding serine or threonine is phosphorylated (Finn, Lu 2008). Pin1 overexpression is found in several types of cancer, suggesting a key role in regulation of proteins that facilitate persistent proliferative capacity.

Some clues have been reported that confirm the role of the three main classes of PPIases in bacterial infection and virulence. The streptococcal cyclophilin SlrA is involved in pneumococcal colonization (Hermans et al. 2006). The Mip protein of *Legionella pneumophila*, a member of the FK506-binding protein family, is involved in its entry into host cells and intracellular replication (Kohler et al. 2003).

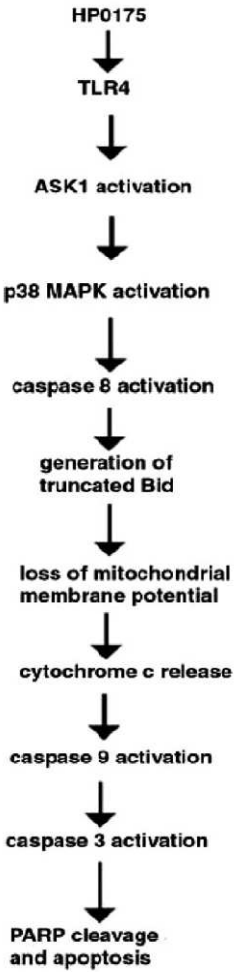

Fig. 4.1. Apoptosis pathway induced by HP0175. Adapted by Basak et al. 2005.

In a more recent study, Basak et al. identified the Peptidyl-prolyl *cis,trans*-Isomerase (PPIase) activity of Parvulin-like HP0175 and its ability to induce apoptosis (Basak et al. 2005). The HP0175 apoptosis ability was confirmed through a KO *H. pylori* strain, where interaction of the protein with Toll-Like Receptor 4 (TLR4) was demonstrated. Through TLR4-binding HP0175 activates several mechanisms. After TLR4 interaction, HP0175 is able to activate apoptosis signal-regulating kinase 1 (ASK1), p38 MAPK and caspase 8. Subsequently, caspase 8 mediates cleavage of Bid factor, leading to deregulation of the mitochondrial membrane potential, release of cytochrome *c* and activation of caspase 9 and 3. Finally, the last step of HP0175-induced cell death is characterized by PARP cleavage and consequent apoptosis.

Moreover, HP0175 was demonstrated to induce IL-6 release from THP-1 (Human acute monocytic leukemia) cell line and from PBMCs (Peripheral Blood Mononuclear Cells) (Pathak et al. 2006). HP0175-promoted TLR4-dependent signaling pathways were demonstrated to activate the MAPKs, ERK and p38 MAPK through PI3K and Ras/Rac1. This, subsequently, activated MSK1, a nuclear kinase central leading to HP0175-driven NF-κB activation. NF-κB activation was necessary for driving HP0175-induced *IL-6* gene expression. Since MSK1 is able to phosphorylate histone H3 as well as the p65 subunit of NF-κB, MSK1 could most likely influence HP0175-promoted chromatin modifications at *IL-6* promoter level, by inducing *IL-6* gene expression. HP0175-stimulated histone H3 phosphorylation on Ser10 and p65 phosphorylation were dependent on MSK1 and chromatin immunoprecipitation assays confirmed that interaction of p65 and phospho-Ser10 H3 with the *IL-6* promoter was correlated on MSK1, which seems to play a central role in HP0175-mediated induction of IL-6.

Finally, it was demonstrated that HP0175 trans-activates EGFR and stimulates EGFR-dependent VEGF production in the gastric cancer cell line AGS (Basu et al. 2008). VEGF plays an important role in *H. pylori*-associated gastroduodenal diseases (Uemura et al. 2001) and is overexpressed in human gastric adenocarcinomas (Brown et al. 1993). VEGF function is pleiotropic: it is involved in the reconstruction of normal mucosal architecture required during healing, but at the same time, VEGF promotes gastric carcinomas by supporting tumor-associated angiogenesis (Basu et al. 2008).

Exogenous HP0175 was able to stimulate EGFR activation with involvement of cyclooxygenase-2. Additionally, TLR4-mediated EGFR transactivation seemed to be independent from *cag* Pathogenicity Island. The TLR4 was co-localized with the lipid raft marker flotillin, suggesting that the binding of HP0175 to TLR4 triggers the translocation of TLR4 to lipid rafts, a prerequisite for EGFR activation. AGS cells have been reported to express c-Src, Lyn, and c-Fgr (Stein et al. 2002): the translocation of TLR4 to lipid rafts have been demonstrated to facilitate its interaction with Lyn, followed by Lyn-dependent phosphorylation of TLR4 on tyrosine residues. Some observations suggested that VEGF activation could be stimulated via HIF-1α (Basu et al. 2008).

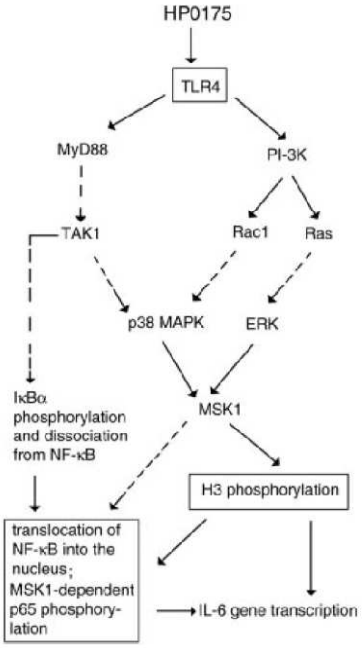

Fig. 4.2. HP0175-mediated IL-6 release mechanism. Adapted from Pathak et al. 2006.

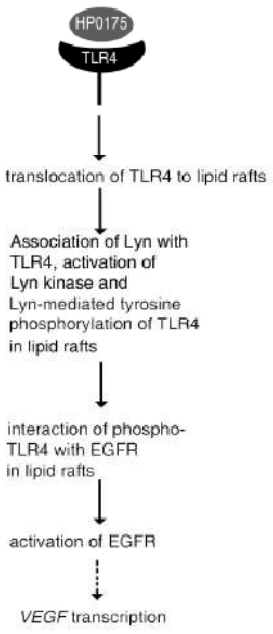

Fig. 4.3. VEGF activation pathway promoted by HP0175. Adapted from Basu et al. 2008.

4.1.3. HP0175 similarity analysis

HP0175 amino acidic sequence analysis identifies this protein as belonging to the Parvulin-like family of PPIases. Analysis with BLASTp detects 31% identity and 45% similarity with the C-terminal catalytic domain of chain A of SurA from *Escherichia coli*, whose structure is known. SurA is a chaperone that facilitates correct folding of outer membrane proteins in Gram-negative bacteria. SurA is packed in the crystal with a tetramer in the asymmetric unit and every protomer is constituted by four domains (N, P1, P2, and C), grouped in two parts. The larger part, called “core domain”, is composed of N, P1 and C segments, whereas the

smaller one includes the P2 domain, which is connected to the core domain by two 25-30 Å-long segments. Because of the domain organization, SurA is an asymmetric dumbbell. The

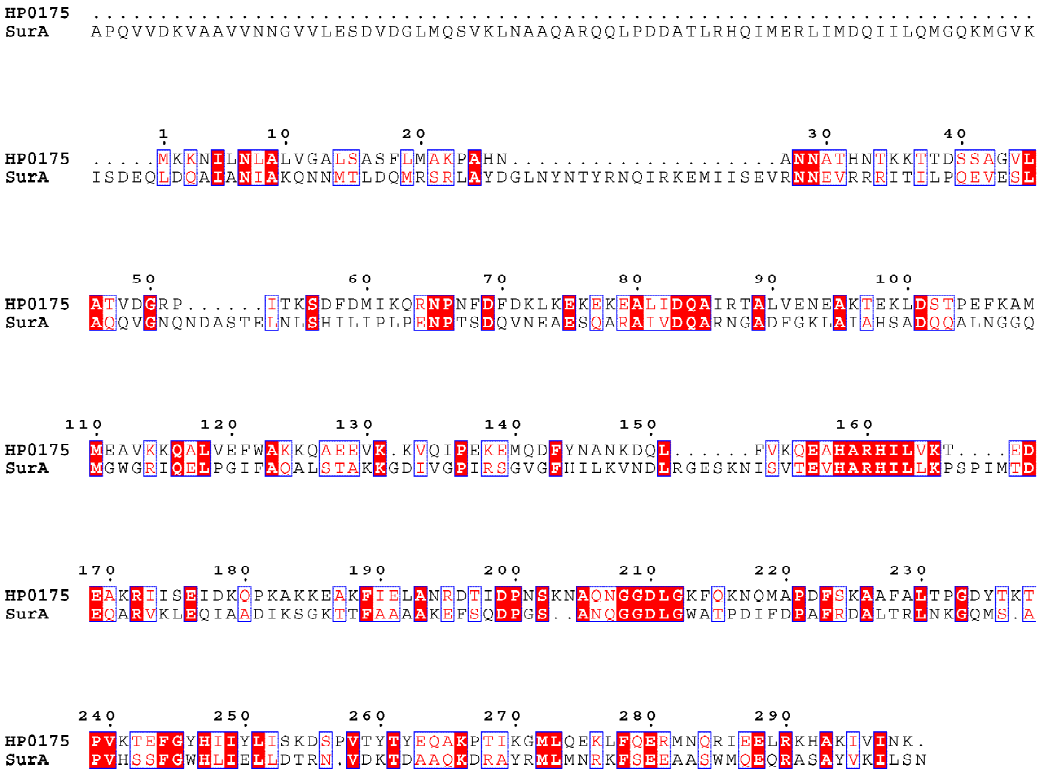

Fig. 4.4. Sequence alignment of HP0175 from *H. pylori* and SurA from *E. coli*.

core unit begins at the N-terminus with a pair of short antiparallel β strands followed by α-helices. P1 domain, connected with the N-terminal domain through a short linker, shares the topology of the Parvulin-like family of PPIases, similar to human Pin1 one (Ranganathan et al. 1997). The scaffold of the PPIase domain is a four-stranded β-sheet surrounded by four α-helices. The C-terminal domain is composed by a long α-helix located between the N-terminal and P1 domain and by a short β-strand antiparallel to the initial β strand of the N-terminus. The C-terminal helix partially occludes the catalytic cleft of P1. The P2 domain has a Parvulin-like fold as well and shares sequence similarity with the P1 domain.

The homology regions between SurA and HP0175 involve the N-domain, P1 and part of P2 domains. A putative structure of HP0175 was predicted using the bioinformatic Homology Modelling tool ESyPred3D (www.fundp.ac.be/sciences/biologie/urbm/bioinfo/esypred/) using as template SurA from *E. coli*.

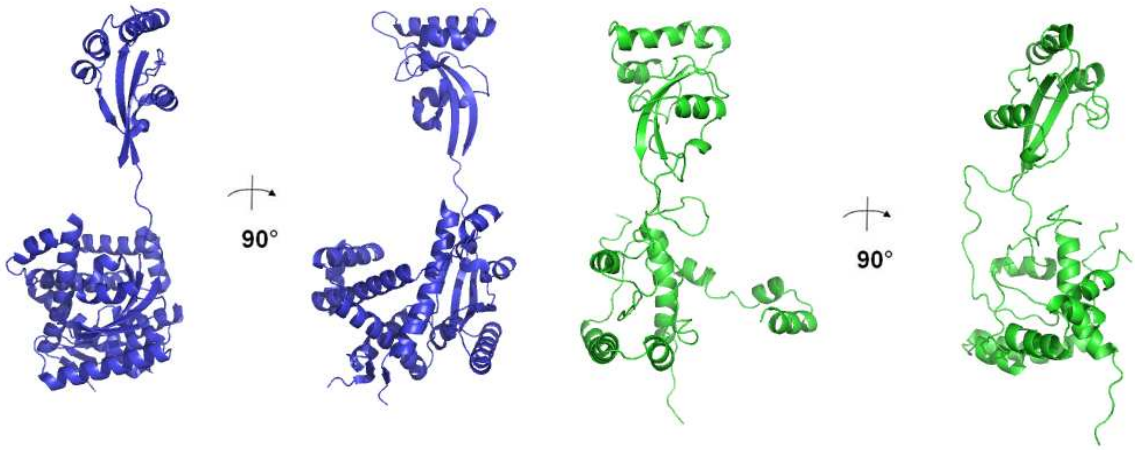

Fig. 4.5. Crystal structure of SurA from *E. coli* (left) and Homology Modelling structure of HP0175 from *H. pylori* (right).

4.1.4. HP0175 gene and protein features

The *HP0175* gene was cloned from the *H. pylori* genome of CCUG17874 strain. Analysis with SignalP (www.cbs.dtu.dk/services/SignalP/) bioinformatic tool identifies a signal peptide at the N-terminus and a predicted proteolytic site between 26th and the 27th amino acid of HP0175 sequence. The sequence below corresponds to the recombinant protein, without signal peptide, added of a N-terminal His-tag flanked by a cleavage sequence for TEV protease, coming from the construct codified by pET151 expression vector.

-24	-14	-4	33	43	53
MHHHHHHGKP	**IPNPLLGLDS**	**TENLYFQ‖GID**	**PFT**ANNATHN	TKKTTDSSAG	VLATVDGRPI
63	73	83	93	103	113
TKSDFDMIKQ	RNPNFDFDKL	KEKEKEALID	QAIRTALVEN	EAKTEKLDST	PEFKAMMEAV
123	133	143	153	163	173
KKQALVEFWA	KKQAEEVKKV	QIPEKEMQDF	YNANKDQLFV	KQEAHARHIL	VKTEDEAKRI
183	193	203	213	223	233
ISEIDKQPKA	KKEAKFIELA	NRDTIDPNSK	NAQNGGDLGK	FQKNQMAPDF	SKAAFALTPG
243	253	263	273	283	293
DYTKTPVKTE	FGYHIIYLIS	KDSPVTYTYE	QAKPTIKGML	QEKLFQERMN	QRIEELRKHA

KIVINK

‖ corresponds to the cleavage site for TEV protease.

HP0175	Properties
Number of amino acids	276
Molecular Weight (Da)	31611.0
Theoretical pI	9.10
Abs (1 g/l)	0.457

Table 4.1. Some properties of HP0175 protein, referred to the construct after cleavage with TEV protease

4.2. Materials and methods

4.2.1. Cloning of HP0175

HP0175 gene was PCR amplified from chromosomal DNA of *H. pylori* strain CCUG17874, using the primers 5'-CACCGCAAATAACGCTACGCATAAC-3' (forward, Topoisomerase recognition site underlined) and 5'-CTATTACTTGTTGATAACAATTTTAGC -3' (reverse) The PCR amplification product was inserted into the expression pET151/D-TOPO directional plasmid, using TOPO® Cloning kit by Invitrogen and following the manufacturer procedure.

4.2.2. Expression and Purification

The expression vector resulting from the cloning step was used to transform *E. coli* BL21 (DE3) (Invitrogen) for protein expression. Cells harboring *pET151-HP0175* were grown at 37 °C in 2 liters LB medium containing 100 μg/ml ampicillin. 1 mM isopropyl-β-thiogalactopyranoside (IPTG) was added to the medium to induce protein expression and the culture was incubated for 4 h at 30 °C under mild shaking (200 rpm). The bacteria were harvested by centrifugation and stored at -80 °C. The bacterial pellet was re-suspended in 30 mM MOPS pH 7.2, 200 mM NaCl and cell lysis was performed by incubation with lysozyme (1 mg/ml, 1 h, 4 °C) and by sonication after addition of Protease Inhibitor Cocktail (Sigma-Aldrich). The lysed cell suspension was cleared of debris by centrifugation (40000 g, 25 min, 4 °C) and the supernatant was loaded onto a 5 ml HisTrap FF column (GE Healthcare), equilibrated with the buffer used for bacterial re-suspension.

After three extensive washes with 2 %, 5% and 10 % solutions of Elution buffer (30 mM MOPS pH 7.2, 200 mM NaCl, 500 mM Imidazole), the protein was eluted by applying a gradient from 10 % to 100 % Elution buffer. Because of the huge amount of protein, probably greater than the column binding capacity, His_6-HP0175 was present in all the fractions collected from the chromatography.

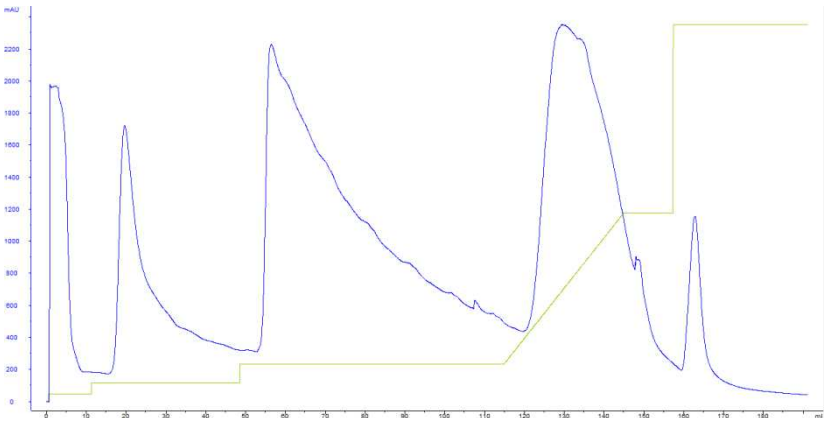

Fig. 4.6. Affinity chromatography of HP0175. HP0175 was present in all the peaks of the chromatogram.

The fractions containing His_6-HP0175 were pooled and incubated O.N. at 4 °C with 6 mg TEV, concentrated by ultrafiltration (10,000 MWCO, Millipore), diluted with 30 mM MOPS

pH 7.2, 200 mM NaCl to 8 mM final concentration imidazole and loaded again into the affinity column. The protein fraction not bound the column, i.e. without His-tag, was concentrated for the second step of purification, a size exclusion chromatography. Because of the very high yield of purified protein (300 mg), several chromatographic runs using an HiLoad 26/60 Superdex 200™ (GE Healthcare) column were performed. The elution volume showed that HP0175 is a dimeric protein. The protein, that revealed a very high solubility, was concentrated to 350 mg/ml for crystallization trials.

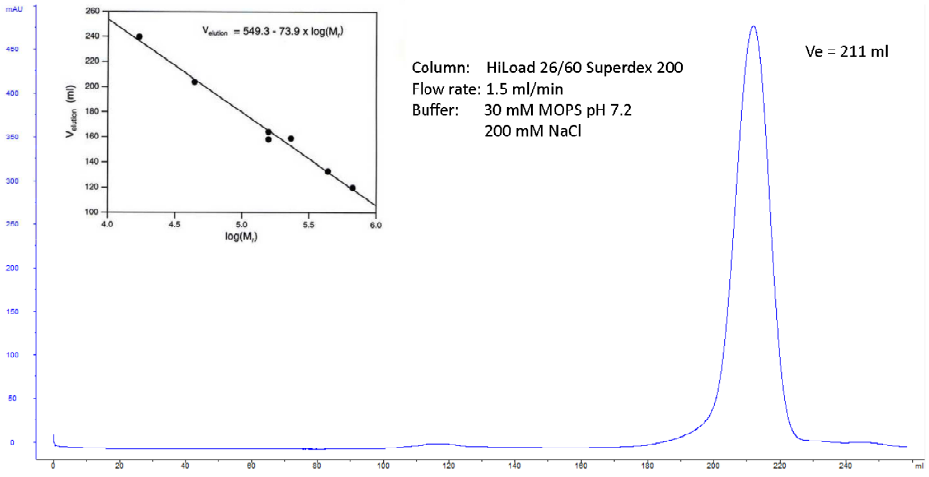

Fig. 4.7.A. Size exclusion chromatography of HP0175. The elution volume suggests a dimeric form of HP0175.

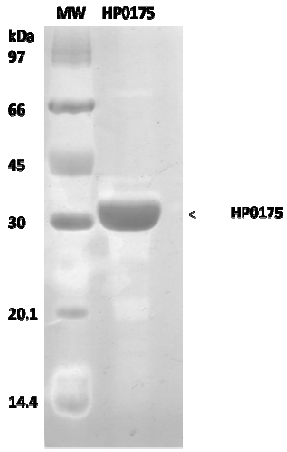

Fig. 4.7.B. SDS-PAGE of HP0175 at the end of the purification process.

4.2.3. Western Blotting

Samples from different steps of HP0175 purification were investigated for the presence of the His-tag. After loading of the samples into a SDS-PAGE, the proteins were transferred on a nitrocellulose membrane by electrophoresis (HybondTM, GE Healthcare). After blocking of the membrane using BSA (3%) solution in PBS for 1 h and incubation of the primary antibody O.N. at 4 °C, three washing steps using 3% BSA in PBS were performed. The secondary antibody was incubated for 1 h at room temperature and washed for three times with PBS. Finally, the membrane was developed by using 5-bromo-4-chloro-3-indolyl-1-phosphate (BCIP) and nitro blue tetrazolium (NBT) (Western Blue®, Promega).

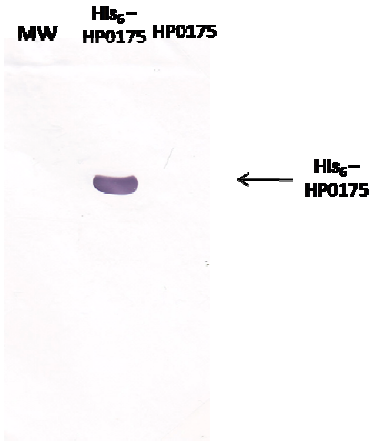

Fig. 4.8. Western Blotting of HP0175 before and after cleavage step with TEV protease.

4.2.4. Limited Proteoysis

To increase the possibilities of HP0175 crystallization, trying to remove the flexible loops of the protein and to obtain a protein core suitable for crystallization, a limited proteolysis of HP0175 was attempted (Forsgren, Lamont & Persson 2009; Danley et al. 2000). Different proteases and different incubation times were tested. The proteases tested were thrombin, trypsin, chimotrypsin, thermolysin at 1:1000 protease/HP0175 ratio. 1 µg of protease with 1 mg of HP0175 were incubated at 37°C in 200 µl buffer Tris pH 8, 50 mM NaCl. Incubation times attempted were 15, 30 and 60 minutes for all the proteases. All the samples coming from the proteolysis tests were loaded into SDS-PAGE, which shows a 25-27 kDa degradation product in a proteolysis mix incubated for 1 h with trypsin. To scale up the process, 18 mg of HP0175 were incubated with 18 µg trypsin for 1 h at 37°C in 10 ml buffer 30 mM Tris pH 8, 50 mM NaCl. The proteolysis reaction was stopped by adding PMSF (2 mM), the proteolysis mix was concentrated by ultrafiltration (10,000 MWCO, Millipore) and loaded into a Superdex 200 10/300GL column (GE Healthcare). The HP0175 proteolysis product was concentrated (10,000 MWCO, Millipore) for crystallization purposes.

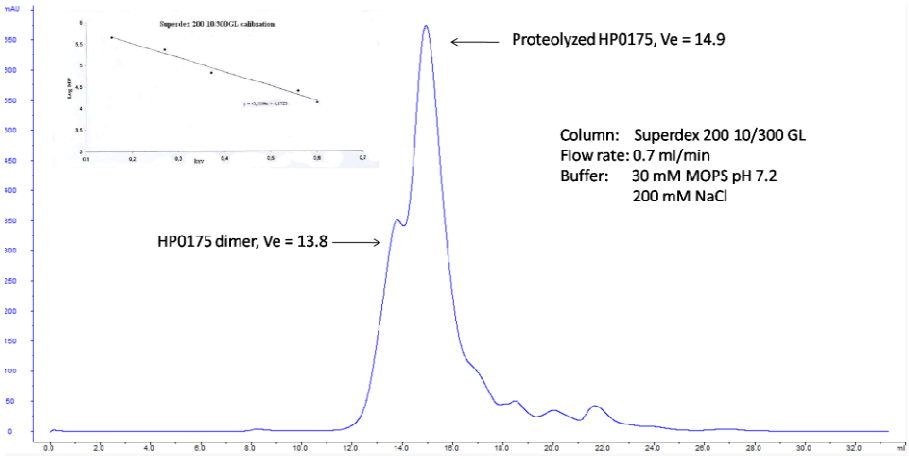

Fig. 4.9. Size exclusion chromatogram of proteolytic product of HP0175.

4.2.5. HP0175 Lysines methylation

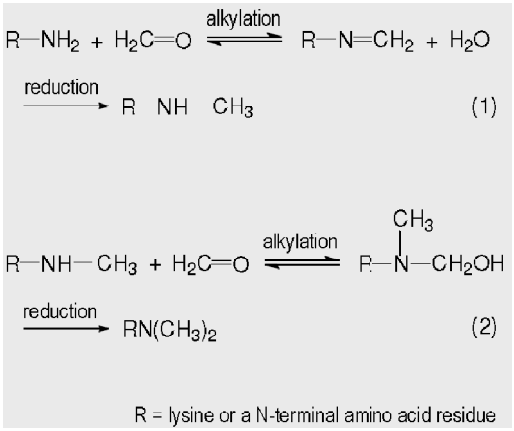

Fig. 4.10. Lysine methylation method (adapted from JBS Methylation Kit manual).

It is generally accepted that chemical modification of proteins sometimes could facilitate crystallization. Protein flexibility can come from large-scale motions of domains, surface loops and N- or C- terminus. These flexible parts could be reduced or eliminated by redesigning the protein construct or removing disordered terminal extensions with limited proteolysis (Walter et al. 2006; Kobayashi, Kubota & Matsuura 1999). Methylation of lysines is one of the

most common methods to engineering the protein surface to reduce its entropy (Walter et al. 2006). To perform this chemical modification for HP0175, that contains 39 lysines per monomer, JBS Methylation Kit by Jena Bioscience was used. It employs formaldehyde as alkylating reagent and dimethylamine borane complex as reducing agent in mild reaction conditions (Walter et al. 2006). In particular, 10 mg of HP0175 were used for every methylation step, following the manufacturer's procedure. The methylated protein was subsequently purified by size exclusion chromatography.

4.2.6. Crystallization trials

HP0175 samples at different protein concentrations, until 350 mg/ml, were used for crystallization tests. Crystallization trials were carried out using Oryx8 crystallization device (Douglas Instruments). Both sitting drops and microbatch drops were performed, screening different precipitant solutions from several kits, Structure Screen I and II by Molecular Dimensions, PACT Suite, JCSG Suite, Anions and Cations Suites, PEGs and PEGs II, $AmSO_4$, MPD Suite, MbClass I and II Suites by Qiagen. Only few crystallization conditions showed a microcrystalline precipitate at high protein concentration and some optimization trials were attempted, modifying protein concentration and precipitant composition. No crystals were obtained from these trials.

Similar crystallization trials were performed with proteolysed and methylated HP0175, but no crystals were obtained.

To reduce the very high HP0175 solubility, the HP0175 purification buffer was changed. Since HP0175 is a basic protein, positively-charged at physiological pH, KSCN, a salt able to decrease the solubility of positively-charged proteins (Hofmeister 1888), was added to the protein buffer instead of sodium chloride. A 350 mg/ml-concentrated protein sample was diluted 1:100 in 30 mM MES 6.2, 200 mM KSCN and concentrated again by ultrafiltration (10,000 MWCO, Millipore). Moreover, since the theoretical HP0175 Isoelectric point (pI) is about 9.1, the protein buffer was substituted with 30 mM Tris pH 8.5, 50 mM, to exploit the inferior protein solubility in proximity of pH values of the buffer close to protein pI. Several crystallization conditions were attempted, but no crystals were obtained.

4.3. Results and Discussion

The *HP0175* gene, amplified without the DNA sequence codifying for the putative secretion signal, was cloned in frame with an N-terminal His-tag in pET151 expression vector. HP0175 was expressed in BL21(DE3) *E. coli* cells. The protein was subsequently purified by affinity chromatography (Fig. 4.6.) and size exclusion chromatography (Fig. 4.7.A) in very high yield. The high purity of HP0175 protein at the end of the purification step was checked by SDS-PAGE (Fig. 4.7.B) and His-tag cleavage was tested by Western-Blotting (Fig. 4.8). Since no crystallization conditions were identified even using a very high concentration of protein (350 mg/ml), probably because of the extremely high solubility of HP0175, some methods to decrease the solubility of the protein were attempted. Two protein buffers were used. The first buffer contained a salt able to decrease the solubility (KSCN) of positively-charged proteins. The second buffer had a pH closer to HP0175 pI and a different salt concentration. Both the buffers did not change significantly the protein solubility.

To remove flexible parts of the protein and to reduce protein surface entropy, limited proteolysis and lysines methylation were attempted. Limited proteolysis with trypsin gave a 25-27 kDa proteolytic product, purified by size exclusion chromatography (Fig. 4.9). Both proteolysed and methylated HP0175 were used to attempt to find crystallization conditions, that have been not identified yet.

Chapter V

Cloning, Expression and Purification trials of HP0421 from *Helicobacter pylori*

5.1. Introduction

5.1.1. Growth-inhibition effect of α1,4-GlcNAc-capped *O*-glycans mucins against *H. pylori*

H. pylori infection is asymptomatic in most of the infected individuals, suggesting that natural defense mechanisms occur in the human stomach (Kawakubo et al. 2004).

Gastric mucins are classified into two types, based on their histochemical properties (Ota et al. 1991): the first type is a mucin derived by surface mucous cells and displayed on MUC5AC core protein (Reis et al. 1999), whereas the second type is a mucin displayed on MUC6 core protein secreted by the Brunner's glands in the duodenum, glands of the gallbladder and in a few by acinar cells of the pancreas, glands of the antrum and by mucopeptic cells of the neck zone of the body region of the stomach (Reis et al. 2000). These two types of mucins form the surface mucous gel layer (SMGL), which reveals an alternating laminated array (Kobayashi et al. 2009).

H. pylori associates only with surface mucous cell-type mucin and two carbohydrate molecules, Lewis *b* and sialyl dimeric Lewis X displayed in surface mucous cells, serve as specific ligands for *H. pylori* adhesins, BabA and SabA, respectively (Ilver et al. 1998; Mahdavi et al. 2002). *H. pylori* seldom colonizes deeper portions of gastric mucosa, where gland mucous cells produce mucins having terminal α1,4-linked *N*-acetylglucosamine (α1,4-linked-GlcNAc) residues bound to core 2-branched *O*-glycans [GlcNAcα1→4Galβ1→4GlcNAcβ1→6(GlcNAcα1→4Galβ1→3)GalNAcα→Ser/Thr], abbreviated as α1,4-GlcNAc-capped *O*-glycans (Hidaka et al. 2001). The inhibition of *H. pylori* growth in this part of the human stomach is caused specifically by α1,4-GlcNAc-capped *O*-glycans, which have protective properties against the bacterium (Kawakubo et al. 2004). Recombinant soluble CD43, the preferential core protein of α1,4-GlcNAc-capped *O*-glycans, was produced in CHO cells (Chinese Hamster Ovary cells) with and without the *O*-glycans. Effect on *H. pylori* growth was tested *in vitro* incubating the two CD43 isoforms in presence of bacterial cells. *H. pylori* cells incubated with α1,4-GlcNAc-capped *O*-glycans-containing CD43 grew slower than the cells incubated with only CD43. The *H. pylori* growth reduction rate was dependent on the dose of α1,4-GlcNAc-capped *O*-glycans-CD43 and was demonstrated to be effective even using CD34 as core protein instead of CD43. Additionally, *p*-nitrophenyl-α-*N*-acetylglucosamine and natural gastric mucins, containing the glycan portion, were able to suppress *H. pylori* growth. The α1,4-GlcNAc-capped *O*-glycans

antibiotic activity was confirmed by incubating AGS cells transfected with α1,4-*N*-acetylglucosaminyltransferase with *H. pylori* cells (Kawakubo et al. 2004).

5.1.2. Morphological effects in *H. pylori* cells and the role of cholesteryl-α-D-glucopyranoside

In addition to bacterial growth inhibition, other phenotypes were observed in *H. pylori* cells incubated with α1,4-GlcNAc-capped *O*-glycans mucins, like reduction of motility, cell elongation, segmental narrowing and folding (Kawakubo et al. 2004). These *O*-glycans induced morphological abnormalities similar to those produced by β-lactamase inhibitors, responsible of disruption of peptidoglycan wall biosynthesis. This finding suggested an inhibition of cell wall biosynthesis caused by these mucins (Kobayashi et al. 2009).

The *H. pylori* wall contains α-cholesteryl glucosides, that represent 25% of total lipids, including cholesteryl-α-D-glucopyranoside (α-CG) (Haque et al. 1995). Mass spectrometry analysis of *H. pylori* wall components reveals a reduction of cholesteryl-α-D-glucopyranoside in bacterial cells incubated with α1,4-GlcNAc-capped *O*-glycans-CD43. Enzymatic studies showed that *H. pylori* lysate was able to catalyze the glycosylation of cholesterol, suggesting the presence of a cholesterol α-glucosyltransferase expressed by the bacterium (Kawakubo et al. 2004).

5.1.3. Identification of HP0421 as a Cholesterol-α-glucosyltransferase

To identify the gene encoding the cholesterol-α-glucosyltransferase in *H. pylori* genome, a BLAST similarity research was performed, starting from the amino acid sequence of the sterol-β-glucosyltransferase from *Arabidopsis thaliana*, but no similar enzymes were found (Lebrun et al. 2006). Based on previously identified correlations between cholesterol and diacylglycerol-using enzymes (Cases et al. 1998), a BLAST analysis using the amino acids sequence of diacylglycerol-α-glycosyltransferase from *A. laidlawii* identified a significant similarity (42% similarity and 19% identity) with HP0421 from *H. pylori* (Lebrun et al.

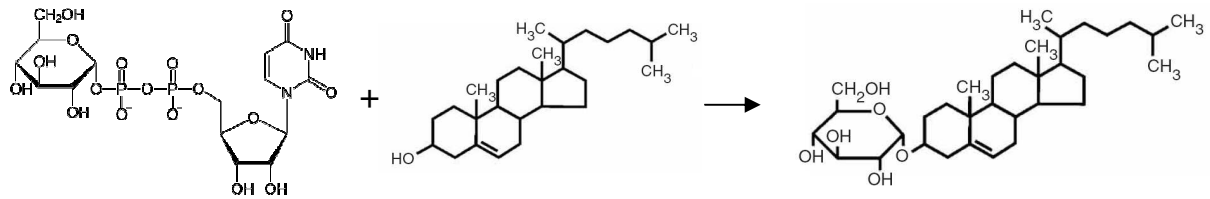

Fig. 5.1. Chemical reaction catalyzed by HP0421 protein. This enzyme catalyzes the glucosylation of cholesterol using UDP-glucose.

2006). HP0421 enzymatic activity was validated by creating a knock-out strain of *H. pylori*. Deletion of *HP0421* gene caused the loss of cholesterol-α-glucosyltransferase activity determined by *in vitro* activity assay on *H. pylori* lysates. Further analysis detects the absence of three glycolipids in K.O. bacterial cells, including cholesteryl-α-D-glucopyranoside (α-CG), demonstrating that *HP0421* gene was required for α-CG synthesis (Lebrun et al. 2006). *E. coli* cells, expressing HP0421, were incubated with several sterols and UDP-[^{14}C]: this assay reveals the enzyme selectivity for cholesterol as substrate. Localization assays in *H. pylori* fractions showed as HP0421 enzymatic activity was present in membrane fraction, demonstrating the protein association to the membrane (Lebrun et al. 2006). Inhibition

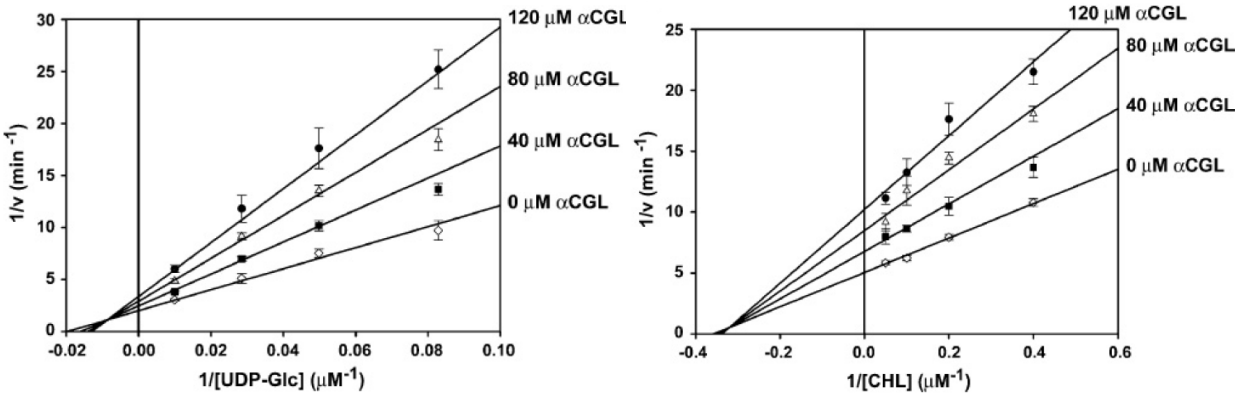

Fig. 5.2. Lineweaver-Burk plot corresponding to HP0421 enzymatic assays in the presence of α-GC as inhibitor. HP0421 was incubated with 5 μM cholesterol and different concentrations of UDP-glucose (left). Several concentrations of cholesterol was added to HP0421 and 20 μM UDP-glucose (right). Adapted from Lee et al., 2008.

kinetic studies, adding α-CG as inhibitor, suggested that HP0421 works in an ordered Bi-Bi manner, and that cholesterol is added to the enzyme/substrate complex after the binding of UDP-glucose to the enzyme. Based on this mechanism, the enzyme-UDP-glucose complex presumably induces a conformational change that permits the binding to an acceptor substrate (Lee et al. 2008; Kobayashi et al. 2009). Finally, α1,4-GlcNAc-capped *O*-glycan was demonstrated to inhibit the cholesterol-α-glucosyltransferase HP0421 (Lee et al. 2008). Further studies, in particular structural studies, could be important to develop α1,4-GlcNAc-capped *O*-glycan-based drug and to identify a low molecular weight inhibitor specific for the enzyme to eradicate *H. pylori* infection.

5.1.4. Cholesterol glycosylation causes immune inhibition

Since it is not able to synthesize *de novo* sterols, *H. pylori* requires exogenous cholesterol for its growth. Recently, *H. pylori* cells have been demonstrated to follow the cholesterol gradient, to associate to cholesterol-rich membranes and to extract cholesterol from membranes of gastric epithelial cells by direct contact (Wunder et al. 2006). Cholesterol

extraction was demonstrated to be connected to destruction of lipid rafts. CD55, a protein bound to glycosylphosphatidylinositol and associated to lipid rafts (Legler et al. 2005), was expressed in AGS cells in fusion with reporter protein GFP. Infection of the AGS cells with *H. pylori* was linked to a decrease of fluorescence in infected cells, suggesting a reduction of lipid raft integrity (Wunder et al. 2006).

Additionally, *H. pylori* cells were incubated with different concentrations of cholesterol and used to infect macrophages (Wunder et al. 2006). *H. pylori* cells-phagocytosis was demonstrated to be increased against cholesterol-rich bacterial cells. Further studies in T cells activation *in vitro* demonstrated that cholesterol stimulated antigen-specific T cell activation on *H. pylori* infection of APC. These findings were confirmed by studies *in vivo* in mice. Anyway, to escape phagocytosis, *H. pylori* is able to glycosylate cholesterol. Cholesterol glycosylation was demonstrated to be time dependent and to provide protection from phagocytosis (Wunder et al. 2006). Additionally, studies on HP0421-deleted *H. pylori* cells showed that the mutant strain was not able to synthesize cholesterol glycosides and was easily internalized by macrophages, indicating as cholesterol-α-glucosyltransferase HP0421 is essential to avoid phagocytosis. Finally, immune evasion from T cells by *H. pylori* expressing HP0421 was demonstrated *in vitro* and *in vivo*, comparing *H. pylori* strains deleted of *HP0421* gene with reconstituted strains (Wunder et al. 2006).

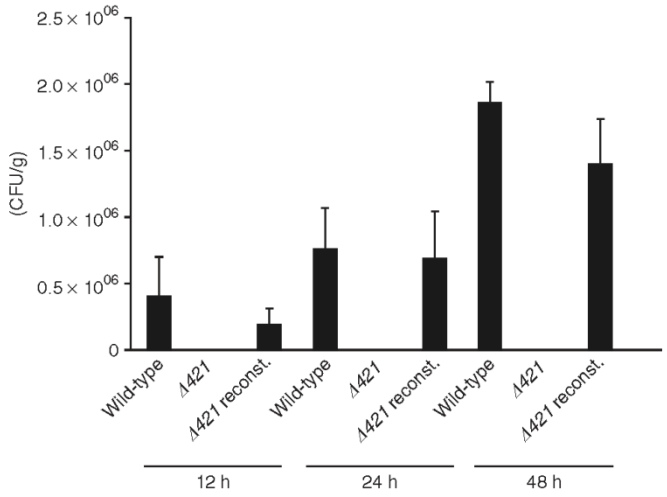

Fig.5.3. Cholesterol α-glucosylation is essential for *H.pylori* infection. C57BL/6 mice were infected with wild-type, Δ421 mutant and reconstituted *H. pylori*. Bacterial charge was evaluated harvesting the stomach. Adapted from Wunder et al. 2006.

5.1.5. HP0421 Homology modeling and structure prediction

HP0421 is a member of the glycosyltransferase family 4 (GT4) and shares some sequence similarity with several members of the family with diacylglycerol glycosyltransferase and monoglycosyldiacylglycerol glycosyltransferase activity (Lebrun et al. 2006). Some of them are *A. laidlawii* diacylglycerol glucosyltransferase and diacylglycerol-3-glucose-(1→2)-glucosyltransferase, *A. thaliana* sulfolipid synthase and diacylglycerol-3-galactose-(1→6)-galactosyltransferase, *S. pneumonia* diacylglycerol glucosyltransferase and diacylglycerol-3-

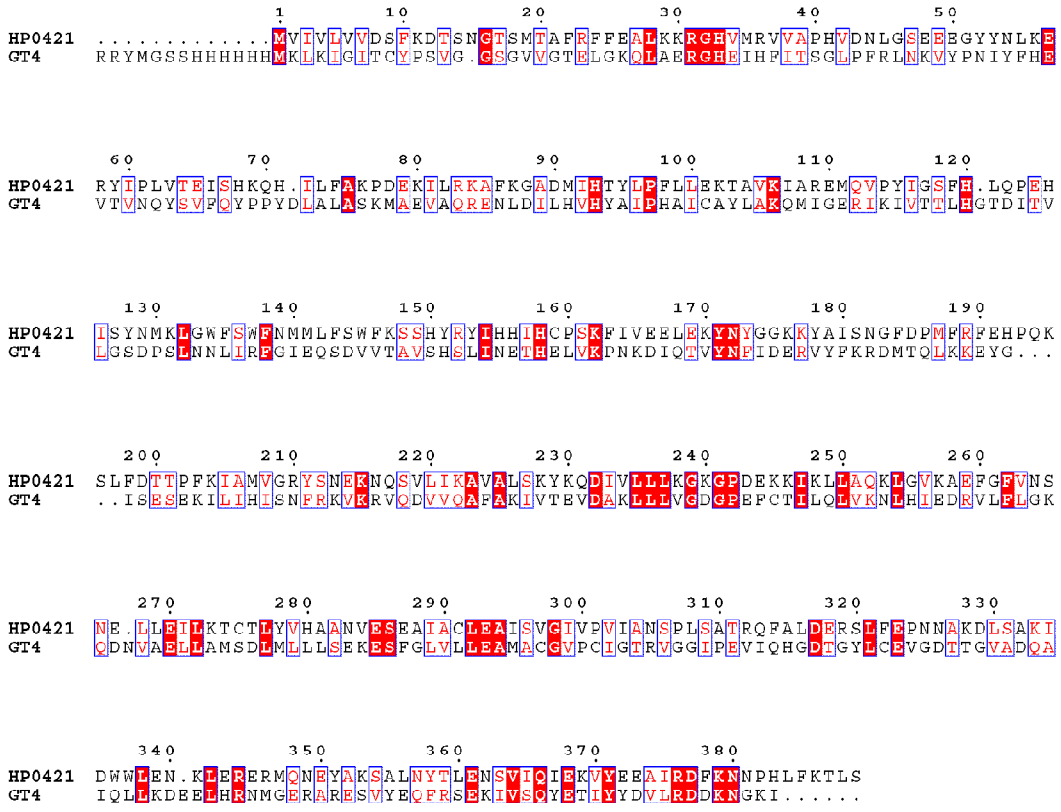

Fig. 5.4. Sequence alignment of HP0421 from *H. pylori* and GT4 Glycosyltransferase from *Bacillus anthracis*

glucose-(1→2)-galactosyltransferase, but also proteins from *Clostridium thermocellum*, *Deinococcus radiodurans*, *Helicobacter hepaticus*, *Lactobacillus johnsonii* and *T. maritima* (Lebrun et al. 2006).

An Homology Modelling with the ESyPred3D server, based on GT4 Glycosyltransferase from *Bacillus anthracis* (Ruane, Davies & Martinez-Fleites 2008) as a template which shares 24% identity and 43% similarity with a 172-amino acids domain with HP0421, was performed successfully.

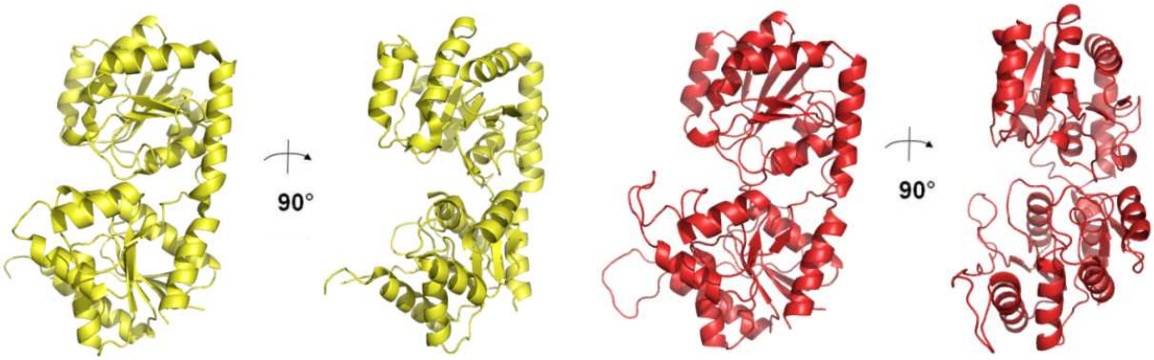

Fig. 5.5. Crystal structure of chain A of GT4 Glycosyltransferase from *Bacillus anthracis* template (left) and Homology Modeling structure of HP0421 from *H. pylori* (right).

GT4 Glycosyltransferase from *Bacillus anthracis* is a tetramer and every monomer folds in two ‘‘Rossmann-like’’ β/α/β domains divided by a deep crevice in the interdomain region and a kinked C-terminal α-helix, which intersects from the C-terminal domain to contact the N-terminal domain. The N-terminal domain is composed of a six-stranded twisted β-sheet flanked, on both sides of the β-sheet, by six α-helices, whereas the C-terminal domain comprises five β-strands surrounded by four α-helices (Ruane, Davies & Martinez-Fleites 2008).

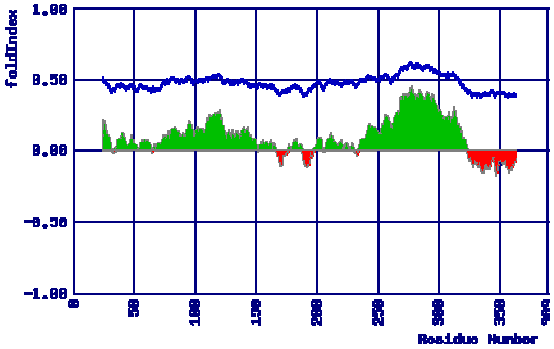

Fig. 5.6. Folding and hydrophobicity of HP0421 predicted using Foldindex bioinformatic program.

HP0421 analysis to detect the presence of a transmembrane domain, using several prediction methods at the ExPASy Molecular Biology server, reveals contradictory results (Lebrun et al. 2006). A folding prediction program (Foldindex, www.bip.weizmann.ac.il/fldbin/findex/) suggests the presence of an unfolded domain at the C-terminal HP0421 domain.

5.1.6. HP0421 gene and protein feature

HP0421 gene, from 26695 *H. pylori* strain, was previously cloned in the expression vector pET24d in frame with a C-terminal His-tag. The amino acidic sequence of the recombinant protein (C-His_6-tag HP0421) is represented below.

```
        10         20         30         40         50         60
MVIVLVVDSF KDTSNGTSMT AFRFFEALKK RGHVMRVVAP HVDNLGSEEE GYYNLKERYI

        70         80         90        100        110        120
PLVTEISHKQ HILFAKPDEK ILRKAFKGAD MIHTYLPFLL EKTAVKIARE MQVPYIGSFH

       130        140        150        160        170        180
LQPEHISYNM KLGWFSWFNM MLFSWFKSSH YRYIHHIHCP SKFIVEELEK YNYGGKKYAI

       190        200        210        220        230        240
SNGFDPMFRF EHPQKSLFDT TPFKIAMVGR YSNEKNQSVL IKAVALSKYK QDIVLLLKGK

       250        260        270        280        290        300
GPDEKKIKLL AQKLGVKAEF GFVNSNELLE ILKTCTLYVH AANVESEAIA CLEAISVGIV

       310        320        330        340        350        360
PVIANSPLSA TRQFALDERS LFEPNNAKDL SAKIDWWLEN KLERERMQNE YAKSALNYTL

       370        380        390        400        410
ENSVIQIEKV YEEAIRDFKN NPHLFKTLSR DPNSSSVDKL AAALEHHHHH H
```

A new construct encoding recombinant HP0421 in frame with a N-terminal His-tag was cloned in pET28a expression plasmid. The amino acidic sequence of the second recombinant protein construct (N-His_6-tag HP0421) is represented below.

```
       -11          -1          10          20          30          40
MGSSHHHHHH SSGLVPR‖GSH MVIVLVVDSF KDTSNGTSMT AFRFFEALKK RGHVMRVVAP

        50          60          70          80          90         100
HVDNLGSEEE GYYNLKERYI  PLVTEISHKQ HILFAKPDEK ILRKAFKGAD MIHTYLPFLL

       110         120         130         140         150         160
EKTAVKIARE MQVPYIGSFH  LQPEHISYNM KLGWFSWFNM MLFSWFKSSH YRYIHHIHCP

       170         180         190         200         210         220
SKFIVEELEK YNYGGKKYAI  SNGFDPMFRF EHPQKSLFDT TPFKIAMVGR YSNEKNQSVL

       230         240         250         260         270         280
IKAVALSKYK QDIVLLLKGK  GPDEKKIKLL AQKLGVKAEF GFVNSNELLE ILKTCTLYVH

       290         300         310         320         330         340
AANVESEAIA CLEAISVGIV  PVIANSPLSA TRQFALDERS LFEPNNAKDL SAKIDWWLEN

       350         360         370         380
KLERERMQNE YAKSALNYTL  ENSVIQIEKV YEEAIRDFKN NPHLFKTLS
```

‖ corresponds to the cleavage site for thrombin protease.

HP0421	Properties
Number of amino acids	389
Molecular Weight (Da)	44965.1
Theoretical pI	8.96
Abs (1 g/l)	1.175

Table 5.1. Some properties of HP0421 protein, referred to the native protein without amino acids from the constructs

5.2. Materials and methods

5.2.1. Cloning of N-terminal His-tag HP0421

HP0421 gene was amplified by PCR from chromosomal DNA of *H. pylori* strain 26695, using the DNA polymerase AmpliTaq Gold® by Applied Biosystems and the primers 5'-GCACATATGGTTATTGTTTTAGTCGTGGATAG-3' (forward) and 5'-AATAGGATCCTTATTATGATAAGGTTTTAAAGAGATGGG-3' (reverse), composed by a flanking part and two restriction site (underlined), for NdeI and BamHI respectively. The PCR amplification product was cloned in pGEM®-T cloning plasmid by Promega, exploiting the overhangs between the vector and PCR products, generated by a DNA polymerase that adds a single deoxyadenosine to the 3'-ends of the amplified fragments. The reaction is performed incubating the sample at 70 ºC with AmpliTaq Gold® DNA polymerase into a 0.2 mM dATP-containing buffer for 30 minutes. After transformation in *E. coli* DH5α and selection of positive clones, the *pGEM-HP0421* plasmid was double digested with NdeI and BamHI, the *HP0421* gene was purified and inserted into a pET28a plasmid (Novagen) in frame with a sequence encoding an His_6-tag followed by a Thrombin cleavage site at the N-terminus of the recombinant construct (N-His_6-HP0421).

5.2.2. Expression and solubility estimation of C-terminal His-tag HP0421

C-terminal His-tag HP0421 solubility (C-His_6-HP0421) was tested in two buffers 30 mM HEPES pH 7, 300 mM NaCl and 30 mM Tris pH 8, 500 mM NaCl with different expression conditions. HP0421 expression in *E. coli* BL21(DE3) harbouring *pET24d-HP0421* was induced with 0.5 mM IPTG and performed in 250 ml total LB medium added with 50 μg/ml kanamycin at 30 ºC and at 18 ºC for 4 hours and O.N, respectively. After bacteria re-suspension in 10 ml buffer, lysis was performed by incubation with 1 mg/ml lysozyme at 4 ºC for 1 hour and by sonication. After centrifugation, pellet and supernatant samples have been loaded into a SDS-PAGE to evaluate the solubility of the expressed protein. The protein solubility, estimated by SDS-PAGE, was about 40% in all the expression trials.

To express C-terminal His-tag HP0421 in high yield, *E. coli* BL21(DE3) harbouring pET24d-HP0421 were grown in 3 liters LB supplemented with kanamycin (50 μg/ml) with vigorous shaking; at an OD_{600} of 0.6, expression of HP0421 was induced by the addition of 0.5 mM IPTG and the culture incubated for a further 4 h at 30 °C under mild shaking. The bacteria were harvested by centrifugation and stored at -80 °C. Expression of N-terminal His-tag HP0421 was performed in the same way of C-terminal His-tag HP0421.

5.2.3. Purification trials of C-terminal His-tag HP0421

To purify C-terminal His-tag HP0421 (C-His_6-tag HP0421), a first step by affinity IMAC-Ni^{2+} chromatography was performed. After bacteria re-suspension in 35 ml buffer 30 mM HEPES pH 7, 300 mM NaCl, lysis was performed by incubation with 1 mg/ml lysozyme at 4 ºC for 1 hour and by sonication, after adding of protease inhibitor cocktail (Roche, Complete-mini EDTA-free). After centrifugation, the supernatant was loaded into a His-Trap column (GE Healthcare) equilibrated with the buffer. After a washing step with 20 mM imidazole, the protein was eluted in imidazole gradient to 500 mM. The presence of C-His_6-tag HP0421 was detected by SDS-PAGE. The SDS-PAGE, loaded with fraction samples from elution, showed

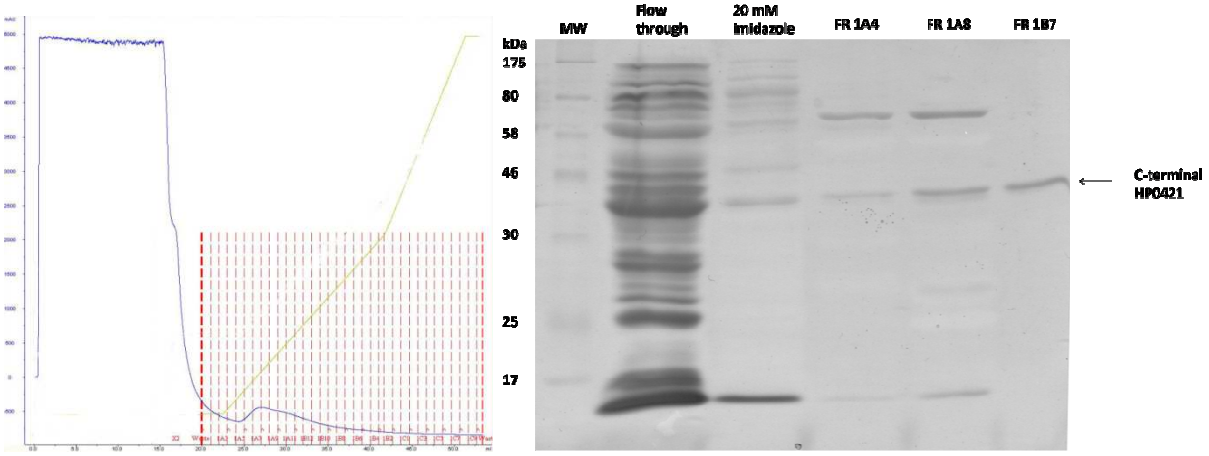

Fig. 5.7. IMAC-Ni^{2+}-affinity chromatography of C-terminal His-tag HP0421 (left). SDS-PAGE corresponding to the affinity chromatography (right).

a protein weakly binding to the nickel, because most of the HP0421 protein eluted with a low imidazole concentration (till 40 mM) together with other *E. coli* proteins.

To improve protein binding, trials were performed using different buffers (Tris, HEPES, PBS), testing different pH values (7-8.5) and modifying the ionic strength (50-500 mM NaCl), but no improvements of protein binding were obtained.

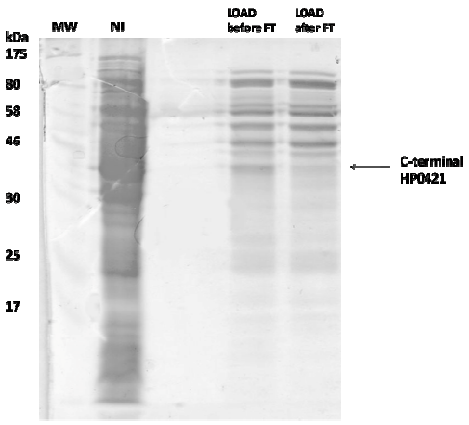

Fig. 5.8. SDS-PAGE of HP0421 samples before ionic exchange chromatography. The protein tends to flocculate at low salt concentration.

Other chromatographic steps were attempted to purify the protein.

The first method tested was ionic exchange chromatography. The affinity chromatography fractions containing C-His_6-tag HP0421 were collected and diluted in 30 mM MES pH 6 till a final NaCl concentration of about 6 mM. The sample, that showed some flocculation, was filtrated with 0.22 μm filter and

loaded into a Cationic Exchange Chromatography column (Mono S HR 5/5, GE Healthcare). The elution step was performed with NaCl gradient to 1 M. SDS-PAGE revealed that HP0421 flocculated before the loading into the column. To inhibit HP0421 flocculation, different buffers with different pH and slightly higher final ionic strength (25 mM NaCl) were used and even an Anionic Exchange Chromatography was attempted.

Samples, coming from IMAC-Ni^{2+} affinity chromatography and containing C-His_6-tag HP0421, were used to perform a size exclusion chromatography. For this purification step, a HiLoad 26/60 Superdex 200™ (GE Healthcare) column was used. HP0421 revealed an unusual elution volume, almost corresponding to the total column volume. HP0421 was not successfully purified using this method, owing to the presence of *E. coli* proteins with a molecular weight very similar to that of HP0421.

Hydrophobic Interaction Chromatography was also attempted. Preliminary protein solubility trials failed to identify the ammonium sulphate concentration in which HP0421 was soluble and sufficient to allow it to bind to the hydrophobic matrix.

Finally, affinity chromatography and ionic exchange chromatography were attempted on denaturing conditions (8 M urea), but the yield of purified denatured HP0421 was not high enough for refolding purposes.

5.2.4. HP0421 Refolding

Purification of the insoluble fraction of HP0421 from inclusion bodies is briefly described. After HP0421 expression in *E. coli* and lysis of the bacterial cells by lysozyme (1 mg/ml, 1 h, 4°C) and sonication (in buffer 50 mM Tris pH 8, 1% Triton X-100, 0.1% sodium deoxicolate, 100 mM NaCl, 0.1% NaN_3, 10 mM DTT), the pellet, obtained by centrifugation at 20,000 rpm for 25 min, was re-suspended in a Triton and EDTA-containing buffer (50 mM Tris pH 8, 0.5% Triton X-100, 100 mM NaCl, 1 mM EDTA, 1 mM DTT, 0.1% NaN_3) and submitted to sonication.

The Triton, contained into the buffer, is requested to solubilize proteins present in the bacterial pellet. Almost all the *E. coli* proteins and a small amount of recombinant protein were removed in this way: at the end of the process, HP0421 was still present in the pellet because of its large amount. Several washing steps were performed, until the bacterial pellet contained the only pure HP0421 protein, whereas the last washing step was performed in a buffer without Triton X-100. The purity of HP0421 was checked by SDS-PAGE.

The pellet, after centrifugation, has been re-suspended in 7 M guanidine pH 5 to solubilize the HP0421 protein in a denatured form. The sample has been used for the HP0421 refolding

process. Small aliquots of the sample have been diluted to 1:100 into several refolding buffers, containing different concentrations of gluthatione reduced/oxidized system, of EDTA or of $MgCl_2$ and $CaCl_2$ and incubated O.N. to identify the best conditions to refold the protein. After dialysis 1:10 against a Tris buffer, the quality of the refolding process has been evaluated by activity test. The best refolding conditions, after optimization of component concentrations, contained: 100 mM Tris pH 8, 300 mM NaCl, 0.5 M guanidine, 0.5 M

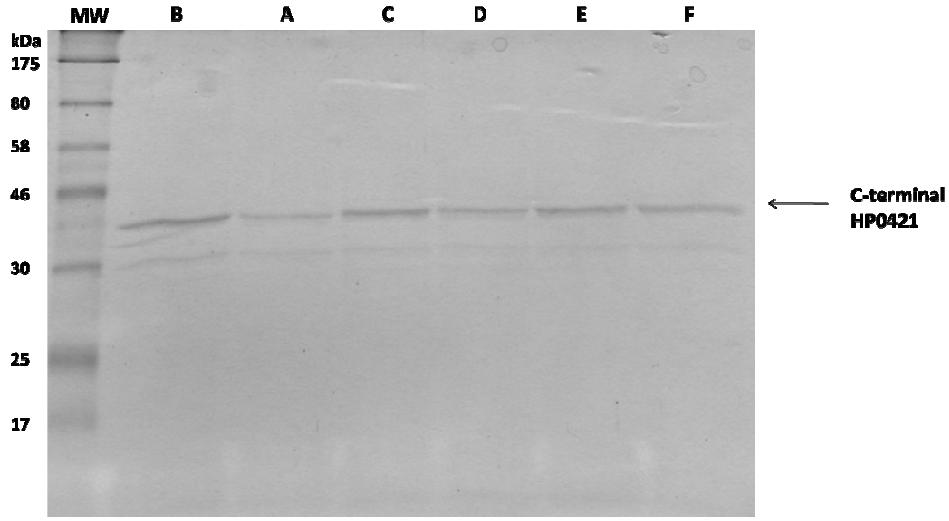

Fig. 5.9. SDS-PAGE relative to HP0421 refolding trials. The lanes correspond to different refolding conditions.

sucrose, 1.4 M arginine, 10 mM EDTA, 10 mM reduced gluthatione, 1 mM oxidated gluthatione, 30 µM cholesterol.

Optimized refolding protocols are as follow. 1 ml of 40 mg HP0421 (in 7 M guanidine pH 5 solution, 0.1 mM DTT) was rapidly injected via a small gauge needle into 200 ml of vigorously stirred refolding buffer. Buffer was saturated with nitrogen, sealed with parafilm and incubated 3 days at 6°C without stirring. Refolding solution was subsequently dialysed over night against 2 liters 20 mM Tris pH 8, 100 mM NaCl, 1 mM EDTA.

5.2.5. HP0421 Activity Test

HP0421 ability to glycosylate cholesterol starting from UDP-glucose was investigated by a specific activity test. 5 µl of refolded protein (or sample coming from different purification steps) was mixed with 0.1 mM cholesterol, 80 µM UDP-^{14}C-glucose (300 mCi/mmol, 10 µCi/400 µl) in PBS pH 7.5 to 50 µl total volume and incubated at 37 ºC for one hour. After the addition of 0.01% Bromophenol blue, all the lipids have been extracted by adding 50 µl hexane/isopropanol 3:2 and purified by TLC (Thin Layer Chromatography). The silica TLC plate has been developed in chlorophorm/methanol 85:15 and exposed O.N. to a Phosphor-imaging film to detect the radioactivity of the product.

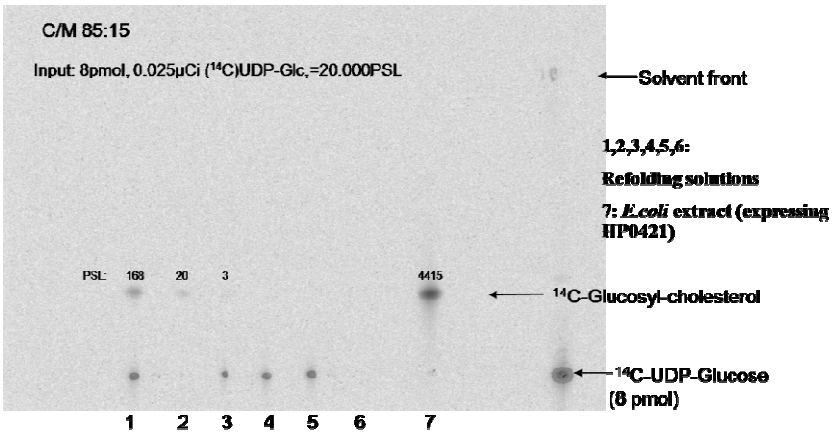

Fig. 5.10. TLC from *HP0421* activity test. 1,2,3,4,5,6: HP0421 protein obtained from different refolding buffers, 7: *E.coli* cytosolic extract expressing HP0421

5.2.6. Western Blotting

The western blotting method was used to validate the presence of HP0421 in purified fractions, using polyclonal antibodies against the protein, and to detect the His-tag of N-terminal and C-terminal His-tag HP0421 after expression and during purification steps.

Protein samples were added with loading buffer and separated by SDS-PAGE. Following SDS-PAGE, the resolved proteins were transferred electrophoretically onto a nitrocellulose membrane (HybondTM, GE Healthcare) and incubated with primary antibody in blocking (3% BSA in PBS) reagent O.N. at 4 °C. The membrane was washed three times in PBS-Tween (0.15 % Tween 20), and then incubated for 1 h with secondary antibody (anti-mouse IgG, coupled to horseradish peroxidase) in presence of blocking reagent. The membrane was washed in PBS-Tween 0.15% three times and developed using the SuperSignal® Chemiluminescent substrate (Pierce).

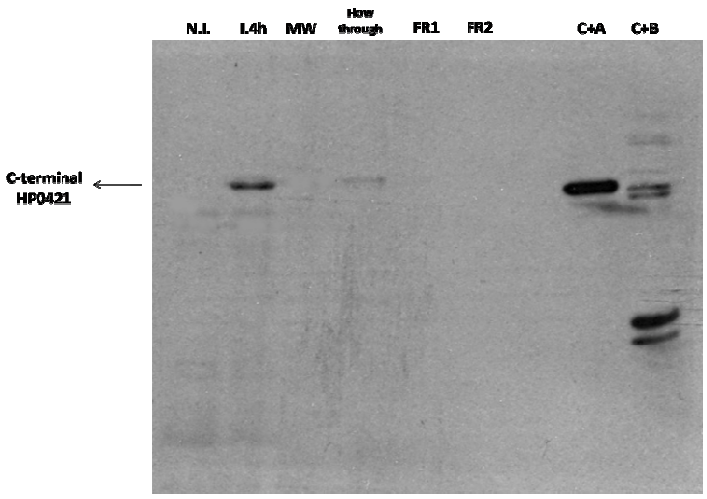

Fig. 5.11. Western blotting of C-His$_6$-tag HP0421 using an antibody against His-tag.

Because HP0421 migrates in SDS-PAGE with lower apparent mass than expected, samples from affinity chromatography was checked for the presence of the protein, using polyclonal antibodies specific for the protein.

Additionally, the presence of His-tag in C-His$_6$-tag HP0421 and N-His$_6$-tag HP0421 samples was validated by Western blotting using a specific

antibody against His-tag. To confirm that the protein detected by the antibody against His-tag was HP0421, other two Western blottings were performed using polyclonal antibodies specific for HP0421.

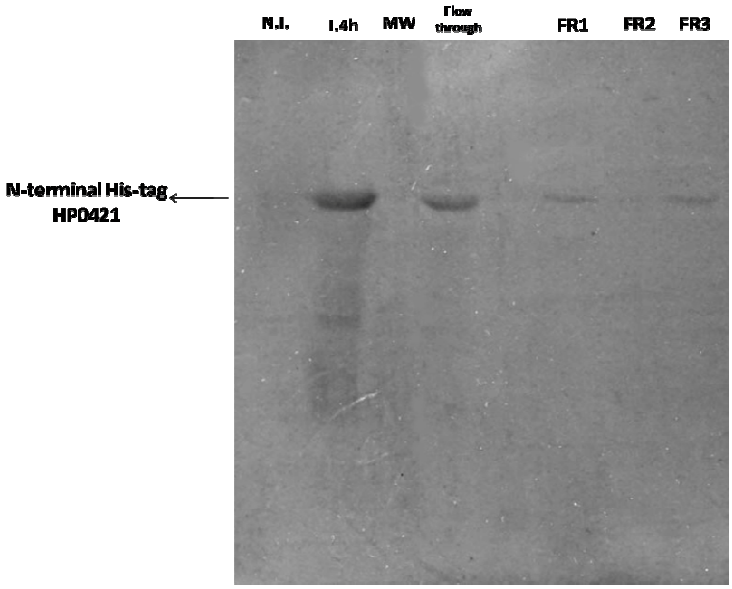

Fig. 5.12. Western blotting of N-His$_6$-tag HP0421 of fractions from affinity chromatography, using an antibody against His-tag.

5.2.7. Dynamic Light Scattering (DLS) analysis

Samples of C-His$_6$-tag HP0421 and of N-His$_6$-tag HP0421 protein purified by refolding and by chromatographic methods were analyzed at DLS Zetasizer Nano ZS (Malvern Instruments). Since the HP0421 protein seemed to be aggregated at DLS analysis, HP0421 aliquots at 0.6 mg/ml were incubated with different detergents at 0.5 CMC concentration. The detergents used were: C12E8, n-Dodecyl-β-D-maltoside, n-Octanoylsucrose, MEGA®-8, CHAPS, FOS-Choline® -9. Only the n-Dodecyl-β-D-maltoside seemed to inhibit partially the protein aggregation.

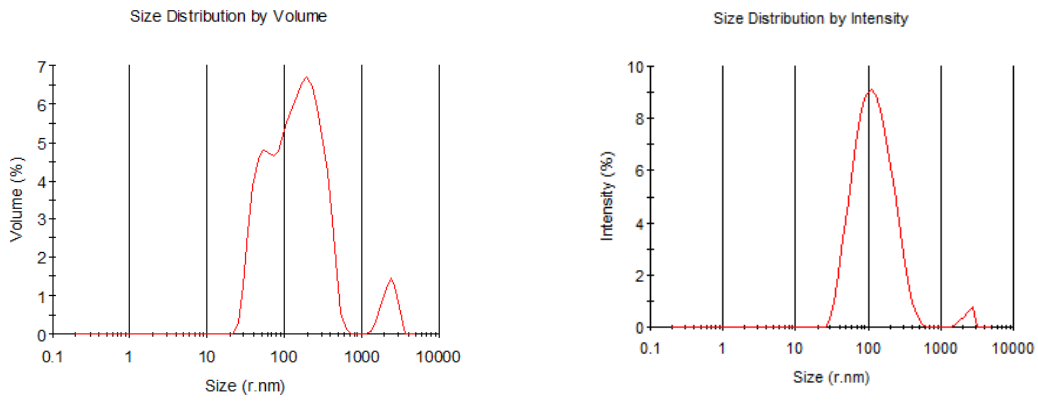

Fig. 5.13. Distribution diagrams of HP0421 without detergent: Volume (%) vs Size radius (nm) (left) and Intensity (%) vs Size radius (nm) (right).

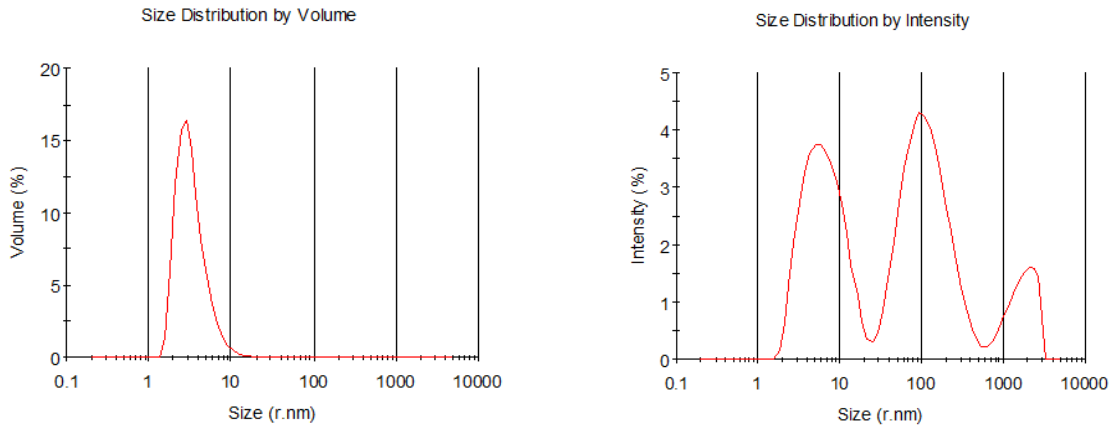

Fig. 5.14. Distribution diagrams of HP0421 with 0,5 CMC n-Dodecyl-β-D-maltoside: Volume (%) vs Size radius (nm) (left) and Intensity (%) vs Size radius (nm) (right).

Affinity chromatography trials of N-His_6-tag HP0421 in a n-Dodecyl-β-D-maltoside(0.5 CMC)-containing buffer did not show a stronger binding of HP0421 with the IMAC-Ni^{2+} and no higher yields of protein were obtained.

5.3. Results and Discussion

HP0421 gene had been previously cloned in frame with a C-terminal His-tag into *pET24d* expression vector. *E. coli* BL21(DE3) cells harbouring *pET24d HP0421* were used for expression and solubility trials. C-His_6-tag HP0421 solubility was tested in two different buffers with different expression conditions and estimated about 40%. To purify C-His_6-tag HP0421 in high yield, a 3 liters culture was set up and an affinity IMAC-Ni^{2+} chromatography was performed. The SDS-PAGE, loaded with fraction samples from chromatographic steps, exhibited weak binding to the nickel matrix, because most of the HP0421 protein eluted with a low imidazole concentration together with others *E. coli* proteins and just a small amount of pure HP0421, less than 200 μg, could be obtained at the end of the process (Fig. 5.7). To increase C-His_6-tag HP0421 binding strength, binding trials were performed using different buffers, but the protein binding was not improved.

Therefore, to remove the protein contaminants, a second purification step was planned. Three purification methods were attempted. The first purification trial tested was Ionic Exchange Chromatography: but in low saline content buffer the protein was almost insoluble (Fig. 5.8.). Subsequently, a gel filtration chromatography was performed but the purification was unsatisfactory, due to the presence of protein contaminants with a molecular weight similar to that of the C-His_6-tag HP0421, that the low resolution of the size exclusion column could not separate successfully. The third attempted chromatography was HIC (Hydrophobic Interaction Chromatography). Preliminary protein solubility trials were performed after addition of ammonium sulphate, but in presence of a salt concentration sufficiently high to allow the binding of the protein to the hydrophobic resin, C-His_6-tag HP0421 was not soluble.

Moreover, the possible causes of the C-His_6-tag HP0421 weak binding to the affinity IMAC-Ni^{2+} resin were investigated. Western blotting studies revealed that C-His_6-tag HP0421 lacked the C-terminal His-tag sequence during the purification steps, but not after expression in *E. coli*, probably because of gradual C-terminal protein degradation (Fig. 5.11.). C-terminal degradation could be caused by flexibility of C-terminal domain, as predicted by the Foldindex program (Fig. 5.6).

Since the C-His_6-tag HP0421 could not be purified in sufficient amount by affinity IMAC-Ni^{2+} chromatography or other standard chromatography methods, to solve the purification problem, two new strategies were planned: the cloning of a N-terminal His-tag construct and the purification of insoluble fraction of HP0421 from inclusion bodies in *E. coli*, followed by protein refolding. A C-His_6-tag HP0421 refolding method was set up (Fig. 5.9) and the success of the process was evaluated by activity tests (Fig. 5.10.). The refolding protocol was

optimized, but, unfortunately, this did not increase the yield of HP0421 protein obtained at the end of the process (lower than 2% of maximum recovery) in comparison with the amount of C-His_6-tag HP0421 protein purified by affinity chromatography.

To circumvent the C-terminal degradation problems of C-His_6-tag HP0421, which occurred during the purification step, a new N-terminal His-tag protein construct was prepared (N-His_6-tag HP0421) and expressed. Affinity chromatography IMAC-Ni^{2+} revealed that the new N-terminal His-tag construct seemed to slightly improve the binding to the affinity resin compared to the C-terminal His-tag construct, although the average yield was still low, 0.4 mg from 2 liters of *E. coli* culture. Anyway, Western blotting studies showed that most of the N-His_6-tag HP0421 protein, even if it still contained an His-tag, was present in the flow through of the affinity column (Fig. 5.12.).

The possible causes of this phenomenon were investigated. DLS analysis revealed that the weak binding of N-His_6-tag HP0421 protein construct to the affinity resin was probably caused by protein aggregation (Fig. 5.13.). Samples of aggregated N-His_6-tag HP0421 protein were incubated with different detergents to identify one able to inhibit protein aggregation. The only detergent that could slightly inhibit unspecific protein molecules interaction was n-Dodecyl-β-D-maltoside (Fig. 5.14.).

Unfortunately, further affinity chromatography trials using a n-Dodecyl-β-D-maltoside-containing buffer did not show any improvement of N-His_6-tag HP0421 binding to the affinity IMAC-Ni^{2+} column.

Appendix

Appendix A: Abbreviations and Symbols

α-CG	Cholesteryl-α-D-glucopyranoside
Å	Angstrom
aa	amino acid
Abs	Absorption
ADP	Adenosine Di-Phosphate
AIR	5-aminoimidazole ribotide
AMP	4-amino-5-aminomethyl-2-methylpyrimidine
ATP	Adenosine Tri-Phosphate
B. subtilis	*Bacillus subtilis*
cag	cytotoxin associated gene (gene)
Cag	Cytotoxin associated gene (associated protein)
CCD	Charge-Coupled Device
CHAPS	3-[(3-Cholamidopropyl)-dimethylammonio]-1-propane sulfonate / N,NDimethyl-3-sulfo-N-[3-[[3α,5β,7α,12α)-3,7,12 trihydroxy-24-oxocholan-24-yl]amino]propyl]-1-propanaminium hydroxide, inner salt
CMC	Critical micelle concentration
$C_{12}E_8$	Octaethyleneglycol Mono-n-dodecyl Ether
$C_{12}E_9$	Polyoxyethylene(9)dodecyl ether / Thesit ® / α-Dodecyl-ω-hydroxy-poly(oxy-1,2-ethanediyl)
CTAB	Hexadecyltrimethylammonium bromide / Cetyltrimethylammonium bromide/ Cetrimonium bromide / Palmityltrimethylammonium bromide
Da	Dalton
DTT	DiThioThreitol
DXP	1-deoxy-D-xylulose-5-phosphate
E. coli	*Escherichia coli*
EDTA	Ethylene Diamino Tetracetic Acid
ESRF	European Synchrotron Radiation Facility
FAD	flavin adenine dinucleotide
FAMP	N-formyl-4-amino-5-aminomethyl-2-methylpyrimidine
F(hkl)	Structure factor amplitude
Fobs,Fcalc	Observed and calculated structure factor amplitudes
FOS-Choline® -9	n-Nonylphosphocholine
FPLC	Fast Protein Liquid Chromatography
Fur	Ferric Uptake Regulator protein
Hepes	N-[2-Hydroxyethyl] piperazine-N'-[2-ethanesolfonic] acid
HET	*see* THZ
HET-P	*see* THZ-P
HMP	4-amino-5-hydroxymethyl-2-methylpyrimidine

HMP-P	4-amino-5-hydroxymethyl-2-methylpyrimidine phosphate
HMP-PP	4-amino-5-hydroxymethyl-2-methylpyrimidine pyrophosphate
H. pylori	*Helicobacter pylori*
I	Measured Intensity of the diffraction spots
IMR	Imidazole ribotide
IPTG	IsoPropyl-β-D-ThioGalactopyranoside
LB	Luria Bertani liquid medium
LDAO	LaurylDimethylAmine Oxide (detergent)
MAD	Multiple Anomalous Dispersion
mAu	milli Absorption unit
MEGA®-8	Octanoyl-N-methylglucamide
MES	2-(N-Morpholin) ethansulfonate
MIR	Multiple Isomorphous Replacement
MS	Mass Spectrometry
MW	Molecular Weight
n-Octanoylsucrose	Sucrose monocaproylate / n-Octanoyl-β-D-fructofuranosyl-α-D-glucopyranoside
n-Octyl-β-D-thioglucoside	n-Octyl-β-D-thioglucopyranoside / 1-s-Octyl-β-D-thioglucoside
O.N.	Over Night
OD	Optical Dispersion
ORF	Open Reading Frames
PAI	PAthogenicity Island
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
PEG	PolyEthylene Glycol
pI	Isoelectric point
r.m.s.d.	Root-mean-square deviation
RP-HPLC	Reversed Phase-High Performance Liquid Chromatography
SAD	Single Anomalous Dispersion
SDS	Sodium Dodecyl Sulfate
SDS-PAGE	SDS-PolyAcrylamide Gel Electrophoresis
SAM	S-adenosylmethionine
SEM	Scanning Electron Microscopy
ThDP	Thiamine diphosphate
ThMP	Thiamine monophosphate
THZ	4-methyl-5-β-hydroxyethylthiazole
THZ-P	4-methyl-5-β-hydroxyethylthiazole phosphate
TLC	Thin Layer Chromatography
TLR4	Toll-Like Receptor 4
Tris	2-Amino-2-(hydroxymethyl)-1,3-propanediol
Triton	Octylphenoxypolyethoxyethanol Polyethylene Glycol-p-isooctylphenyl Ether
σ(I)	Standard deviation of the measured Intensities (I)

Amino acids

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

Appendix B: Crystallographic formulas

$$completeness = \frac{number\ of\ unique\ reflections\ measured}{total\ number\ of\ unique\ reflections}$$

$$multiplicity = \frac{total\ number\ of\ intensity\ measurements}{total\ number\ of\ unique\ reflections\ measured}$$

$$R_{factor} = \frac{\sum_{hkl}\left||F_o(hkl)|-|F_c(hkl)|\right|}{\sum_{hkl}|F_o(hkl)|}$$

$$R_{free} = \frac{\sum_{test\ set}\left||F_o|-|F_c|\right|}{\sum_{test\ set}|F_o|}$$

$$R_{merge} = \frac{\sum_{hkl}|I_{hkl}-\bar{I}_{hkl}|}{\sum_{hkl} I_{hkl}}$$

$$V_M = \frac{Volume\ of\ unit\ cell}{Total\ molecular\ mass\ in\ unit\ cell}$$

References

References

Agarwal, K. & Agarwal, S. 2008, "Helicobacter pylori vaccine: from past to future ", *Mayo Clinic proceedings.Mayo Clinic,* vol. 83, no. 2, pp. 169-175.

Amedei, A., Cappon, A., Codolo, G., Cabrelle, A., Polenghi, A., Benagiano, M., Tasca, E., Azzurri, A., D'Elios, M.M., Del Prete, G. & de Bernard, M. 2006, "The neutrophil-activating protein of Helicobacter pylori promotes Th1 immune responses ", *The Journal of clinical investigation,* vol. 116, no. 4, pp. 1092-1101.

Amieva, M.R. & El-Omar, E.M. 2008, "Host-Bacterial Interactions in Helicobacter pylori Infection ", *Gastroenterology,* vol. 134, no. 1, pp. 306-323.

Andersen, L.P. 2007, "Colonization and infection by Helicobacter pylori in humans ", *Helicobacter,* vol. 12 Suppl 2, pp. 12-15.

Andrzejewska, J., Lee, S.K., Olbermann, P., Lotzing, N., Katzowitsch, E., Linz, B., Achtman, M., Kado, C.I., Suerbaum, S. & Josenhans, C. 2006, "Characterization of the pilin ortholog of the Helicobacter pylori type IV cag pathogenicity apparatus, a surface-associated protein expressed during infection ", *Journal of Bacteriology,* vol. 188, no. 16, pp. 5865-5877.

Angelakopoulos, H. & Hohmann, E.L. 2000, "Pilot study of phoP/phoQ-deleted Salmonella enterica serovar typhimurium expressing Helicobacter pylori urease in adult volunteers ", *Infection and immunity,* vol. 68, no. 4, pp. 2135-2141.

Arnold, K., Bordoli, L., Kopp, J. & Schwede, T. 2006, "The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling ", *Bioinformatics (Oxford, England),* vol. 22, no. 2, pp. 195-201.

Atanassov, C., Pezennec, L., d'Alayer, J., Grollier, G., Picard, B. & Fauchere, J.L. 2002, "Novel antigens of Helicobacter pylori correspond to ulcer-related antibody pattern of sera from infected patients ", *Journal of clinical microbiology,* vol. 40, no. 2, pp. 547-552.

Atherton, J.C., Cao, P., Peek, R.M.,Jr, Tummuru, M.K., Blaser, M.J. & Cover, T.L. 1995, "Mosaicism in vacuolating cytotoxin alleles of Helicobacter pylori. Association of specific vacA types with cytotoxin production and peptic ulceration ", *The Journal of biological chemistry,* vol. 270, no. 30, pp. 17771-17777.

Backert, S. & Selbach, M. 2008, "Role of type IV secretion in Helicobacter pylori pathogenesis ", *Cellular microbiology,* vol. 10, no. 8, pp. 1573-1581.

Baldwin, D.N., Shepherd, B., Kraemer, P., Hall, M.K., Sycuro, L.K., Pinto-Santini, D.M. & Salama, N.R. 2007, "Identification of Helicobacter pylori genes that contribute to stomach colonization ", *Infection and immunity,* vol. 75, no. 2, pp. 1005-1016.

Baker, L.J., Dorocke, J.A., Harris, R.A. & Timm, D.E. 2001, "The crystal structure of yeast thiamin pyrophosphokinase ", *Structure (London, England : 1993),* vol. 9, no. 6, pp. 539-546.

Basak, C., Pathak, S.K., Bhattacharyya, A., Pathak, S., Basu, J. & Kundu, M. 2005, "The secreted peptidyl prolyl cis,trans-isomerase HP0175 of Helicobacter pylori induces apoptosis of gastric epithelial cells in a TLR4- and apoptosis signal-regulating kinase 1-dependent manner ", *Journal of immunology (Baltimore, Md.: 1950),* vol. 174, no. 9, pp. 5672-5680.

Basso, D., Zambon, C.F., Letley, D.P., Stranges, A., Marchet, A., Rhead, J.L., Schiavon, S., Guariso, G., Ceroti, M., Nitti, D., Rugge, M., Plebani, M. & Atherton, J.C. 2008, "Clinical relevance of Helicobacter pylori cagA and vacA gene polymorphisms ", *Gastroenterology,* vol. 135, no. 1, pp. 91-99.

Basu, S., Pathak, S.K., Chatterjee, G., Pathak, S., Basu, J. & Kundu, M. 2008, "Helicobacter pylori protein HP0175 transactivates epidermal growth factor receptor through TLR4 in gastric epithelial cells ", *The Journal of biological chemistry*, vol. 283, no. 47, pp. 32369-32376.

Benach, J., Edstrom, W.C., Lee, I., Das, K., Cooper, B., Xiao, R., Liu, J., Rost, B., Acton, T.B., Montelione, G.T. & Hunt, J.F. 2005, "The 2.35 A structure of the TenA homolog from Pyrococcus furiosus supports an enzymatic function in thiamine metabolism ", *Acta crystallographica.Section D, Biological crystallography*, vol. 61, no. Pt 5, pp. 589-598.

Bizzozero, G. 1893, "
Uber die schlauchformingen Drusen des Magendarmkanals und die Beziehung ihres Epithels zu dem Oberflachenepithel de Schleimhaut", *Arch. Mikr. Anat.*, vol. 42, no. 82-152.

Blaser, M.J. 1998, "Helicobacter pylori and gastric diseases ", *BMJ (Clinical research ed.)*, vol. 316, no. 7143, pp. 1507-1510.

Blaser, M.J. & Berg, D.E. 2001, "Helicobacter pylori genetic diversity and risk of human disease ", *The Journal of clinical investigation*, vol. 107, no. 7, pp. 767-773.

Bohm, G., Muhr, R. & Jaenicke, R. 1992, "Quantitative analysis of protein far UV circular dichroism spectra by neural networks ", *Protein engineering*, vol. 5, no. 3, pp. 191-195.

Bricogne, G., Vonrhein, C., Flensburg, C., Schiltz, M. & Paciorek, W. 2003, "Generation, representation and flow of phase information in structure determination: recent developments in and around SHARP 2.0 ", *Acta crystallographica.Section D, Biological crystallography*, vol. 59, no. Pt 11, pp. 2023-2030.

Brown, L.F., Berse, B., Jackman, R.W., Tognazzi, K., Manseau, E.J., Senger, D.R. & Dvorak, H.F. 1993, "Expression of vascular permeability factor (vascular endothelial growth factor) and its receptors in adenocarcinomas of the gastrointestinal tract ", *Cancer research*, vol. 53, no. 19, pp. 4727-4735.

Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., Read, R.J., Rice, L.M., Simonson, T. & Warren, G.L. 1998, "Crystallography & NMR system: A new software suite for macromolecular structure determination ", *Acta crystallographica.Section D, Biological crystallography*, vol. 54, no. Pt 5, pp. 905-921.

Bumann, D., Aksu, S., Wendland, M., Janek, K., Zimny-Arndt, U., Sabarth, N., Meyer, T.F. & Jungblut, P.R. 2002, "Proteome analysis of secreted proteins of the gastric pathogen Helicobacter pylori ", *Infection and immunity*, vol. 70, no. 7, pp. 3396-3403.

Bumann, D., Metzger, W.G., Mansouri, E., Palme, O., Wendland, M., Hurwitz, R., Haas, G., Aebischer, T., von Specht, B.U. & Meyer, T.F. 2001, "Safety and immunogenicity of live recombinant Salmonella enterica serovar Typhi Ty21a expressing urease A and B from Helicobacter pylori in human volunteers ", *Vaccine*, vol. 20, no. 5-6, pp. 845-852.

Campobasso, N., Costello, C.A., Kinsland, C., Begley, T.P. & Ealick, S.E. 1998, "Crystal structure of thiaminase-I from Bacillus thiaminolyticus at 2.0 A resolution ", *Biochemistry*, vol. 37, no. 45, pp. 15981-15989.

Campobasso, N., Mathews, I.I., Begley, T.P. & Ealick, S.E. 2000, "Crystal structure of 4-methyl-5-beta-hydroxyethylthiazole kinase from Bacillus subtilis at 1.5 A resolution ", *Biochemistry*, vol. 39, no. 27, pp. 7868-7877.

Carlsohn, E., Nystrom, J., Bolin, I., Nilsson, C.L. & Svennerholm, A.M. 2006, "HpaA is essential for Helicobacter pylori colonization in mice ", *Infection and immunity*, vol. 74, no. 2, pp. 920-926.

Cascales, E. & Christie, P.J. 2003, "The versatile bacterial type IV secretion systems ", *Nature reviews.Microbiology*, vol. 1, no. 2, pp. 137-149.

Cases, S., Smith, S.J., Zheng, Y.W., Myers, H.M., Lear, S.R., Sande, E., Novak, S., Collins, C., Welch, C.B., Lusis, A.J., Erickson, S.K. & Farese, R.V.,Jr 1998, "Identification of a gene encoding an acyl CoA:diacylglycerol acyltransferase, a key enzyme in triacylglycerol synthesis ", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 95, no. 22, pp. 13018-13023.

Cendron, L., Couturier, M., Angelini, A., Barison, N., Stein, M. & Zanotti, G. 2009, "The Helicobacter pylori CagD (HP0545, Cag24) protein is essential for CagA translocation and maximal induction of interleukin-8 secretion ", *Journal of Molecular Biology*, vol. 386, no. 1, pp. 204-217.

Cendron, L., Seydel, A., Angelini, A., Battistutta, R. & Zanotti, G. 2004, "Crystal structure of CagZ, a protein from the Helicobacter pylori pathogenicity island that encodes for a type IV secretion system ", *Journal of Molecular Biology*, vol. 340, no. 4, pp. 881-889.

Cendron, L., Tasca, E., Seraglio, T., Seydel, A., Angelini, A., Battistutta, R., Montecucco, C. & Zanotti, G. 2007, "The crystal structure of CagS from the Helicobacter pylori pathogenicity island ", *Proteins*, vol. 69, no. 2, pp. 440-443.

Censini, S., Lange, C., Xiang, Z., Crabtree, J.E., Ghiara, P., Borodovsky, M., Rappuoli, R. & Covacci, A. 1996, "cag, a pathogenicity island of Helicobacter pylori, encodes type I-specific and disease-associated virulence factors ", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 93, no. 25, pp. 14648-14653.

Chatterjee, A., Li, Y., Zhang, Y., Grove, T.L., Lee, M., Krebs, C., Booker, S.J., Begley, T.P. & Ealick, S.E. 2008, "Reconstitution of ThiC in thiamine pyrimidine biosynthesis expands the radical SAM superfamily ", *Nature chemical biology*, vol. 4, no. 12, pp. 758-765.

Cheng, G., Bennett, E.M., Begley, T.P. & Ealick, S.E. 2002, "Crystal structure of 4-amino-5-hydroxymethyl-2-methylpyrimidine phosphate kinase from Salmonella typhimurium at 2.3 A resolution ", *Structure (London, England : 1993)*, vol. 10, no. 2, pp. 225-235.

Chivian, D., Kim, D.E., Malmstrom, L., Bradley, P., Robertson, T., Murphy, P., Strauss, C.E., Bonneau, R., Rohl, C.A. & Baker, D. 2003, "Automated prediction of CASP-5 structures using the Robetta server ", *Proteins*, vol. 53 Suppl 6, pp. 524-533.

Collaborative Computational Project, N.4. 1994, "The CCP4 suite: programs for protein crystallography ", *Acta crystallographica.Section D, Biological crystallography*, vol. 50, no. Pt 5, pp. 760-763.

Costello, C.A., Kelleher, N.L., Abe, M., McLafferty, F.W. & Begley, T.P. 1996, "Mechanistic studies on thiaminase I. Overexpression and identification of the active site nucleophile ", *The Journal of biological chemistry*, vol. 271, no. 7, pp. 3445-3452.

Covacci, A. & Rappuoli, R. 2000, "Tyrosine-phosphorylated bacterial proteins: Trojan horses for the host cell ", *The Journal of experimental medicine*, vol. 191, no. 4, pp. 587-592.

Cover, T.L. & Blanke, S.R. 2005, "Helicobacter pylori VacA, a paradigm for toxin multifunctionality ", *Nature reviews.Microbiology*, vol. 3, no. 4, pp. 320-332.

Cover, T.L. & Blaser, M.J. 2009, "Helicobacter pylori in health and disease ", *Gastroenterology*, vol. 136, no. 6, pp. 1863-1873.

Cover, T.L., Hanson, P.I. & Heuser, J.E. 1997, "Acid-induced dissociation of VacA, the Helicobacter pylori vacuolating cytotoxin, reveals its pattern of assembly ", *The Journal of cell biology*, vol. 138, no. 4, pp. 759-769.

Dailidiene, D., Dailide, G., Ogura, K., Zhang, M., Mukhopadhyay, A.K., Eaton, K.A., Cattoli, G., Kusters, J.G. & Berg, D.E. 2004, "Helicobacter acinonychis: genetic and rodent infection studies of a Helicobacter pylori-like gastric pathogen of cheetahs and other big cats ", *Journal of Bacteriology*, vol. 186, no. 2, pp. 356-365.

Danley, D.E., Haggan, M.E., Cunningham, D., Fennell, K.F., Pauly, T.A. & LeMotte, P.K. 2000, "A crystallizable form of RIIbeta regulatory domain obtained by limited proteolysis ", *Acta crystallographica.Section D, Biological crystallography*, vol. 56, no. Pt 8, pp. 1038-1041.

Day, N. & Keillor, J.W. 1999, "A continuous spectrophotometric linked enzyme assay for transglutaminase activity", *Analytical Biochemistry*, vol. 274, no. 1, pp. 141-144.

Del Prete, G., De Carli, M., Lammel, R.M., D'Elios, M.M., Daniel, K.C., Giusti, B., Abbate, R. & Romagnani, S. 1995, "Th1 and Th2 T-helper cells exert opposite regulatory effects on procoagulant activity and tissue factor production by human monocytes ", *Blood*, vol. 86, no. 1, pp. 250-257.

D'Elios, M.M., Amedei, A., Cappon, A., Del Prete, G. & de Bernard, M. 2007, "The neutrophil-activating protein of Helicobacter pylori (HP-NAP) as an immune modulating agent ", *FEMS immunology and medical microbiology*, vol. 50, no. 2, pp. 157-164.

D'Elios, M.M. & Andersen, L.P. 2007, "Helicobacter pylori inflammation, immunity, and vaccines ", *Helicobacter*, vol. 12 Suppl 1, pp. 15-19.

D'Elios, M.M., Montecucco, C. & de Bernard, M. 2007, "VacA and HP-NAP, Ying and Yang of Helicobacter pylori-associated gastric inflammation ", *Clinica chimica acta; international journal of clinical chemistry*, vol. 381, no. 1, pp. 32-38.

DiPetrillo, M.D., Tibbetts, T., Kleanthous, H., Killeen, K.P. & Hohmann, E.L. 1999, "Safety and immunogenicity of phoP/phoQ-deleted Salmonella typhi expressing Helicobacter pylori urease in adult volunteers ", *Vaccine*, vol. 18, no. 5-6, pp. 449-459.

Dossumbekova, A., Prinz, C., Mages, J., Lang, R., Kusters, J.G., Van Vliet, A.H., Reindl, W., Backert, S., Saur, D., Schmid, R.M. & Rad, R. 2006, "Helicobacter pylori HopH (OipA) and bacterial pathogenicity: genetic and functional genomic analysis of hopH gene polymorphisms ", *The Journal of infectious diseases*, vol. 194, no. 10, pp. 1346-1355.

Dunn, B.E. & Phadnis, S.H. 1998, "Structure, function and localization of Helicobacter pylori urease ", *The Yale journal of biology and medicine*, vol. 71, no. 2, pp. 63-73.

Eaton, K.A., Suerbaum, S., Josenhans, C. & Krakowka, S. 1996, "Colonization of gnotobiotic piglets by Helicobacter pylori deficient in two flagellin genes ", *Infection and immunity*, vol. 64, no. 7, pp. 2445-2448.

el-Omar, E., Penman, I., Ardill, J.E. & McColl, K.E. 1995, "A substantial proportion of non-ulcer dyspepsia patients have the same abnormality of acid secretion as duodenal ulcer patients ", *Gut*, vol. 36, no. 4, pp. 534-538.

Farinha, P. & Gascoyne, R.D. 2005, "Helicobacter pylori and MALT lymphoma ", *Gastroenterology*, vol. 128, no. 6, pp. 1579-1605.

Finn, G. & Lu, K.P. 2008, "Phosphorylation-specific prolyl isomerase Pin1 as a new diagnostic and therapeutic target for cancer ", *Current cancer drug targets*, vol. 8, no. 3, pp. 223-229.

Fischer, W., Puls, J., Buhrdorf, R., Gebert, B., Odenbreit, S. & Haas, R. 2001, "Systematic mutagenesis of the Helicobacter pylori cag pathogenicity island: essential genes for CagA translocation in host cells and induction of interleukin-8 ", *Molecular microbiology*, vol. 42, no. 5, pp. 1337-1348.

Forsgren, N., Lamont, R.J. & Persson, K. 2009, "A crystallizable form of the Streptococcus gordonii surface antigen SspB C-domain obtained by limited proteolysis ", *Acta crystallographica.Section F, Structural biology and crystallization communications*, vol. 65, no. Pt 7, pp. 712-714.

Fronzes, R., Schafer, E., Wang, L., Saibil, H.R., Orlova, E.V. & Waksman, G. 2009, "Structure of a type IV secretion system core complex ", *Science (New York, N.Y.)*, vol. 323, no. 5911, pp. 266-268.

Gall, A.L., Ruff, M. & Moras, D. 2003, "The dual role of CHAPS in the crystallization of stromelysin-3 catalytic domain ", *Acta crystallographica.Section D, Biological crystallography*, vol. 59, no. Pt 3, pp. 603-606.

Gangwer, K.A., Mushrush, D.J., Stauff, D.L., Spiller, B., McClain, M.S., Cover, T.L. & Lacy, D.B. 2007, "Crystal structure of the Helicobacter pylori vacuolating toxin p55 domain ", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 41, pp. 16293-16298.

Gorden, J. & Small, P.L. 1993, "Acid resistance in enteric bacteria ", *Infection and immunity*, vol. 61, no. 1, pp. 364-367.

Gunther, W., Luchow, A., Cluzeaud, F., Vandewalle, A. & Jentsch, T.J. 1998, "ClC-5, the chloride channel mutated in Dent's disease, colocalizes with the proton pump in endocytotically active kidney cells ", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 95, no. 14, pp. 8075-8080.

Ha, N.C., Oh, S.T., Sung, J.Y., Cha, K.A., Lee, M.H. & Oh, B.H. 2001, "Supramolecular assembly and acid resistance of Helicobacter pylori urease ", *Nature structural biology*, vol. 8, no. 6, pp. 505-509.

Hanes, J.W., Kraft, C.E. & Begley, T.P. 2007, "An assay for thiaminase I in complex biological samples ", *Analytical Biochemistry*, vol. 368, no. 1, pp. 33-38.

Haque, M., Hirai, Y., Yokota, K. & Oguma, K. 1995, "Lipid profiles of Helicobacter pylori and Helicobacter mustelae grown in serum-supplemented and serum-free media ", *Acta Medica Okayama*, vol. 49, no. 4, pp. 205-211.

Hartman, J.C., Carlin, J.T., Scheide, J.D. & Ho, C.T. 1984, "
Volatile products formed from the thermal degradation of thiamin at high and low moisture levels", *Journal of Agricultural and Food Chemistry*, vol. 32, pp. 1015-1018.

Hatakeyama, M. 2009, "Helicobacter pylori and gastric carcinogenesis ", *Journal of gastroenterology*, vol. 44, no. 4, pp. 239-248.

Hermans, P.W., Adrian, P.V., Albert, C., Estevao, S., Hoogenboezem, T., Luijendijk, I.H., Kamphausen, T. & Hammerschmidt, S. 2006, "The streptococcal lipoprotein rotamase A (SlrA) is a functional peptidyl-prolyl isomerase involved in pneumococcal colonization ", *The Journal of biological chemistry*, vol. 281, no. 2, pp. 968-976.

Hidaka, E., Ota, H., Hidaka, H., Hayama, M., Matsuzawa, K., Akamatsu, T., Nakayama, J. & Katsuyama, T. 2001, "Helicobacter pylori and two ultrastructurally distinct layers of gastric mucous cell mucins in the surface mucous gel layer ", *Gut*, vol. 49, no. 4, pp. 474-480.

Higashi, H., Tsutsumi, R., Fujita, A., Yamazaki, S., Asaka, M., Azuma, T. & Hatakeyama, M. 2002a, "Biological activity of the Helicobacter pylori virulence factor CagA is determined by variation in the tyrosine phosphorylation sites ", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, no. 22, pp. 14428-14433.

Higashi, H., Tsutsumi, R., Muto, S., Sugiyama, T., Azuma, T., Asaka, M. & Hatakeyama, M. 2002b, "SHP-2 tyrosine phosphatase as an intracellular target of Helicobacter pylori CagA protein ", *Science (New York, N.Y.)*, vol. 295, no. 5555, pp. 683-686.

Hirata, Y., Maeda, S., Mitsuno, Y., Tateishi, K., Yanai, A., Akanuma, M., Yoshida, H., Kawabe, T., Shiratori, Y. & Omata, M. 2002, "Helicobacter pylori CagA protein activates serum response element-driven transcription independently of tyrosine phosphorylation ", *Gastroenterology*, vol. 123, no. 6, pp. 1962-1971.

Hofmeister, F. 1888, "Zur lehre der wirkung der salze. Zweite mittheilung", *Arch Exp Pathol Pharmakol*, vol. 24, pp. 247-260.

Honeyfield, D.C., Brown, S.B., Fitzsimons, J.D. & Tillitt, D.E. 2005, "Early Mortality Syndrome in Great Lakes Salmonines ", *Journal of Aquatic Animal Health*, vol. 17, no. 1, pp. 1 <last_page> 3.

Ilver, D., Arnqvist, A., Ogren, J., Frick, I.M., Kersulyte, D., Incecik, E.T., Berg, D.E., Covacci, A., Engstrand, L. & Boren, T. 1998, "Helicobacter pylori adhesin binding fucosylated histo-blood group antigens revealed by retagging ", *Science (New York, N.Y.)*, vol. 279, no. 5349, pp. 373-377.

Ilver, D., Barone, S., Mercati, D., Lupetti, P. & Telford, J.L. 2004, "Helicobacter pylori toxin VacA is transferred to host cells via a novel contact-dependent mechanism ", *Cellular microbiology*, vol. 6, no. 2, pp. 167-174.

Itou, H., Yao, M., Watanabe, N. & Tanaka, I. 2004, "Structure analysis of PH1161 protein, a transcriptional activator TenA homologue from the hyperthermophilic archaeon Pyrococcus horikoshii ", *Acta crystallographica.Section D, Biological crystallography*, vol. 60, no. Pt 6, pp. 1094-1100.

Jenkins, A.H., Schyns, G., Potot, S., Sun, G. & Begley, T.P. 2007, "A new thiamin salvage pathway ", *Nature chemical biology*, vol. 3, no. 8, pp. 492-497.

Jenkins, A.L., Zhang, Y., Ealick, S.E. & Begley, T.P. 2008, "Mutagenesis studies on TenA: a thiamin salvage enzyme from Bacillus subtilis ", *Bioorganic chemistry*, vol. 36, no. 1, pp. 29-32.

Josenhans, C., Labigne, A. & Suerbaum, S. 1995, "Comparative ultrastructural and functional studies of Helicobacter pylori and Helicobacter mustelae flagellin mutants: both flagellin subunits, FlaA and FlaB, are necessary for full motility in Helicobacter species ", *Journal of Bacteriology*, vol. 177, no. 11, pp. 3010-3020.

Jurgenson, C.T., Begley, T.P. & Ealick, S.E. 2009, "The structural and biochemical foundations of thiamin biosynthesis ", *Annual Review of Biochemistry*, vol. 78, pp. 569-603.

Kang, J. & Blaser, M.J. 2006, "Bacterial populations as perfect gases: genomic integrity and diversification tensions in Helicobacter pylori ", *Nature reviews.Microbiology*, vol. 4, no. 11, pp. 826-836.

Kang, C.B., Hong, Y., Dhe-Paganon, S. & Yoon, H.S. 2008, "FKBP family proteins: immunophilins with versatile biological functions ", *Neuro-Signals*, vol. 16, no. 4, pp. 318-325.

Kawakubo, M., Ito, Y., Okimura, Y., Kobayashi, M., Sakura, K., Kasama, S., Fukuda, M.N., Fukuda, M., Katsuyama, T. & Nakayama, J. 2004, "Natural antibiotic function of a human gastric mucin against Helicobacter pylori infection ", *Science (New York, N.Y.)*, vol. 305, no. 5686, pp. 1003-1006.

Ke, H.M., Zydowsky, L.D., Liu, J. & Walsh, C.T. 1991, "Crystal structure of recombinant human T-cell cyclophilin A at 2.5 A resolution ", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 88, no. 21, pp. 9483-9487.

Kessler, D. 2006, "Enzymatic activation of sulfur for incorporation into biomolecules in prokaryotes ", *FEMS microbiology reviews*, vol. 30, no. 6, pp. 825-840.

Kim, N., Weeks, D.L., Shin, J.M., Scott, D.R., Young, M.K. & Sachs, G. 2002, "Proteins released by Helicobacter pylori in vitro ", *Journal of Bacteriology*, vol. 184, no. 22, pp. 6155-6162.

Kobayashi, M., Kubota, M. & Matsuura, Y. 1999, "Crystallization and improvement of crystal quality for x-ray diffraction of maltooligosyl trehalose synthase by reductive methylation of lysine residues ", *Acta crystallographica.Section D, Biological crystallography*, vol. 55, no. Pt 4, pp. 931-933.

Kobayashi, M., Lee, H., Nakayama, J. & Fukuda, M. 2009, "Carbohydrate-dependent defense mechanisms against Helicobacter pylori infection ", *Current Drug Metabolism*, vol. 10, no. 1, pp. 29-40.

Kohler, R., Fanghanel, J., Konig, B., Luneberg, E., Frosch, M., Rahfeld, J.U., Hilgenfeld, R., Fischer, G., Hacker, J. & Steinert, M. 2003, "Biochemical and functional analyses of the Mip protein: influence of the N-terminal half and of peptidylprolyl isomerase activity on the virulence of Legionella pneumophila ", *Infection and immunity*, vol. 71, no. 8, pp. 4389-4397.

Kwok, T., Zabler, D., Urman, S., Rohde, M., Hartig, R., Wessler, S., Misselwitz, R., Berger, J., Sewald, N., Konig, W. & Backert, S. 2007, "Helicobacter exploits integrin for type IV secretion and kinase activation ", *Nature*, vol. 449, no. 7164, pp. 862-866.

Laskowski, R.A., MacArthur, M.W., Moss, D.S. & Thornton, J.M. 1993, "*PROCHECK*: a program to check the stereochemical quality of protein structures", *Journal of Applied Crystallography*, vol. 26, no. 2, pp. 283-291.

Lebrun, A.H., Wunder, C., Hildebrand, J., Churin, Y., Zahringer, U., Lindner, B., Meyer, T.F., Heinz, E. & Warnecke, D. 2006, "Cloning of a cholesterol-alpha-glucosyltransferase from Helicobacter pylori ", *The Journal of biological chemistry*, vol. 281, no. 38, pp. 27765-27772.

Lee, H., Wang, P., Hoshino, H., Ito, Y., Kobayashi, M., Nakayama, J., Seeberger, P.H. & Fukuda, M. 2008, "Alpha1,4GlcNAc-capped mucin-type O-glycan inhibits cholesterol alpha-glucosyltransferase from Helicobacter pylori and suppresses H. pylori growth ", *Glycobiology*, vol. 18, no. 7, pp. 549-558.

Lee, J.M., Zhang, S., Saha, S., Santa Anna, S., Jiang, C. & Perkins, J. 2001, "RNA expression analysis using an antisense Bacillus subtilis genome array ", *Journal of Bacteriology*, vol. 183, no. 24, pp. 7371-7380.

Legler, D.F., Doucey, M.A., Schneider, P., Chapatte, L., Bender, F.C. & Bron, C. 2005, "Differential insertion of GPI-anchored GFPs into lipid rafts of live cells ", *The FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, vol. 19, no. 1, pp. 73-75.

Lehmann, C., Begley, T.P. & Ealick, S.E. 2006, "Structure of the Escherichia coli ThiS-ThiF complex, a key component of the sulfur transfer system in thiamin biosynthesis ", *Biochemistry*, vol. 45, no. 1, pp. 11-19.

Leslie, A.G.W. 1992, "Recent changes to the *MOSFLM* package for processing film and image plate data", *Joint CCP4+ESF-EAMCB Newsletter on Protein Crystallography*, vol. 26.

Linden, S., Mahdavi, J., Hedenbro, J., Boren, T. & Carlstedt, I. 2004, "Effects of pH on Helicobacter pylori binding to human gastric mucins: identification of binding to non-MUC5AC mucins ", *The Biochemical journal*, vol. 384, no. Pt 2, pp. 263-270.

Liu, J.Y., Timm, D.E. & Hurley, T.D. 2006, "Pyrithiamine as a substrate for thiamine pyrophosphokinase ", *The Journal of biological chemistry*, vol. 281, no. 10, pp. 6601-6607.

Llosa, M. & O'Callaghan, D. 2004, "Euroconference on the Biology of Type IV Secretion Processes: bacterial gates into the outer world ", *Molecular microbiology*, vol. 53, no. 1, pp. 1-8.

Losonsky, G.A., Kotloff, K.L. & Walker, R.I. 2003, "B cell responses in gastric antrum and duodenum following oral inactivated Helicobacter pylori whole cell (HWC) vaccine and LT(R192G) in H pylori seronegative individuals ", *Vaccine*, vol. 21, no. 5-6, pp. 562-565.

Lu, H., Wu, J.Y., Beswick, E.J., Ohno, T., Odenbreit, S., Haas, R., Reyes, V.E., Kita, M., Graham, D.Y. & Yamaoka, Y. 2007, "Functional and intracellular signaling differences associated with the Helicobacter pylori AlpAB adhesin from Western and East Asian strains ", *The Journal of biological chemistry*, vol. 282, no. 9, pp. 6242-6254.

Lupetti, P., Heuser, J.E., Manetti, R., Massari, P., Lanzavecchia, S., Bellon, P.L., Dallai, R., Rappuoli, R. & Telford, J.L. 1996, "Oligomeric and subunit structure of the Helicobacter pylori vacuolating cytotoxin ", *The Journal of cell biology*, vol. 133, no. 4, pp. 801-807.

Maeda, S. & Mentis, A.F. 2007, "Pathogenesis of Helicobacter pylori infection ", *Helicobacter*, vol. 12 Suppl 1, pp. 10-14.

Mahdavi, J., Sonden, B., Hurtig, M., Olfat, F.O., Forsberg, L., Roche, N., Angstrom, J., Larsson, T., Teneberg, S., Karlsson, K.A., Altraja, S., Wadstrom, T., Kersulyte, D., Berg, D.E., Dubois, A., Petersson, C., Magnusson, K.E., Norberg, T., Lindh, F., Lundskog, B.B., Arnqvist, A., Hammarstrom, L. & Boren, T. 2002, "Helicobacter pylori SabA adhesin in persistent infection and chronic inflammation ", *Science (New York, N.Y.)*, vol. 297, no. 5581, pp. 573-578.

Malfertheiner, P., Megraud, F., O'Morain, C., Bazzoli, F., El-Omar, E., Graham, D., Hunt, R., Rokkas, T., Vakil, N. & Kuipers, E.J. 2007, "Current concepts in the management of Helicobacter pylori infection: the Maastricht III Consensus Report ", *Gut*, vol. 56, no. 6, pp. 772-781.

Marshall, B.J. & Warren, J.R. 1984, "Unidentified curved bacilli in the stomach of patients with gastritis and peptic ulceration ", *Lancet*, vol. 1, no. 8390, pp. 1311-1315.

Matthews, B. 1968, "Solvent contents of protein crystals", *Journal of Molecular Biology*, vol. 33, pp. 491-497.

McAtee, C.P., Lim, M.Y., Fung, K., Velligan, M., Fry, K., Chow, T.P. & Berg, D.E. 1998, "Characterization of a Helicobacter pylori vaccine candidate by proteome techniques ", *Journal of chromatography.B, Biomedical sciences and applications*, vol. 714, no. 2, pp. 325-333.

McClain, M.S., Iwamoto, H., Cao, P., Vinion-Dubiel, A.D., Li, Y., Szabo, G., Shao, Z. & Cover, T.L. 2003, "Essential role of a GXXXG motif for membrane channel formation by Helicobacter pylori vacuolating toxin ", *The Journal of biological chemistry*, vol. 278, no. 14, pp. 12101-12108.

McCulloch, K.M., Kinsland, C., Begley, T.P. & Ealick, S.E. 2008, "Structural studies of thiamin monophosphate kinase in complex with substrates and products ", *Biochemistry*, vol. 47, no. 12, pp. 3810-3821.

Michnick, S.W., Rosen, M.K., Wandless, T.J., Karplus, M. & Schreiber, S.L. 1991, "Solution structure of FKBP, a rotamase enzyme and receptor for FK506 and rapamycin ", *Science (New York, N.Y.)*, vol. 252, no. 5007, pp. 836-839.

Miyatake, H., Hasegawa, T. & Yamano, A. 2006, "New methods to prepare iodinated derivatives by vaporizing iodine labelling (VIL) and hydrogen peroxide VIL (HYPER-VIL) ", *Acta crystallographica.Section D, Biological crystallography*, vol. 62, no. Pt 3, pp. 280-289.

Mobley, H.L.T., Mendz, G.L. & Hazell, S.L. 2001, "*Helicobacter pylori*: physiology and genetics", .

Montecucco, C. & Rappuoli, R. 2001, "Living dangerously: how Helicobacter pylori survives in the human stomach ", *Nature reviews.Molecular cell biology*, vol. 2, no. 6, pp. 457-466.

Montemurro, P., Nishioka, H., Dundon, W.G., de Bernard, M., Del Giudice, G., Rappuoli, R. & Montecucco, C. 2002, "The neutrophil-activating protein (HP-NAP) of Helicobacter pylori is a potent stimulant of mast cells ", *European journal of immunology*, vol. 32, no. 3, pp. 671-676.

Murshudov, G.N., Vagin, A.A., Lebedev, A., Wilson, K.S. & Dodson, E.J. 1999, "Efficient anisotropic refinement of macromolecular structures using FFT ", *Acta crystallographica.Section D, Biological crystallography*, vol. 55, no. Pt 1, pp. 247-255.

Neel, B.G., Gu, H. & Pao, L. 2003, "The 'Shp'ing news: SH2 domain-containing tyrosine phosphatases in cell signaling ", *Trends in biochemical sciences*, vol. 28, no. 6, pp. 284-293.

Obchoei, S., Wongkhan, S., Wongkham, C., Li, M., Yao, Q. & Chen, C. 2009, "Cyclophilin A: potential functions and therapeutic target for human cancer ", *Medical science monitor : international medical journal of experimental and clinical research*, vol. 15, no. 11, pp. RA221-32.

O'Morain, C. 2006, "Role of Helicobacter pylori in functional dyspepsia ", *World journal of gastroenterology : WJG*, vol. 12, no. 17, pp. 2677-2680.

Ota, H., Katsuyama, T., Ishii, K., Nakayama, J., Shiozawa, T. & Tsukahara, Y. 1991, "A dual staining method for identifying mucins of different gastric epithelial mucous cells ", *The Histochemical journal*, vol. 23, no. 1, pp. 22-28.

O'Toole, P. 2000, "Helicobacter pylori motility ", *Microbes and Infection*, vol. 2, no. 10, pp. 1207.

O'Toole, P.W., Kostrzynska, M. & Trust, T.J. 1994, "Non-motile mutants of Helicobacter pylori and Helicobacter mustelae defective in flagellar hook production ", *Molecular microbiology*, vol. 14, no. 4, pp. 691-703.

Pagliaccia, C., de Bernard, M., Lupetti, P., Ji, X., Burroni, D., Cover, T.L., Papini, E., Rappuoli, R., Telford, J.L. & Reyrat, J.M. 1998, "The m2 form of the Helicobacter pylori cytotoxin has cell type-specific vacuolating activity ", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 95, no. 17, pp. 10212-10217.

Painter, J. & Merritt, E.A. 2006, "Optimal description of a protein structure in terms of multiple groups undergoing TLS motion ", *Acta crystallographica.Section D, Biological crystallography*, vol. 62, no. Pt 4, pp. 439-450.

Pang, A.S., Nathoo, S. & Wong, S.L. 1991, "Cloning and characterization of a pair of novel genes that regulate production of extracellular enzymes in Bacillus subtilis ", *Journal of Bacteriology*, vol. 173, no. 1, pp. 46-54.

Papini, E., Gottardi, E., Satin, B., de Bernard, M., Massari, P., Telford, J., Rappuoli, R., Sato, S.B. & Montecucco, C. 1996, "The vacuolar ATPase proton pump is present on intracellular vacuoles induced by Helicobacter pylori ", *Journal of medical microbiology*, vol. 45, no. 2, pp. 84-89.

Pascal, J.M., Day, P.J., Monzingo, A.F., Ernst, S.R., Robertus, J.D., Iglesias, R., Perez, Y., Ferreras, J.M., Citores, L. & Girbes, T. 2001, "2.8-A crystal structure of a nontoxic type-II ribosome-inactivating protein, ebulin l ", *Proteins*, vol. 43, no. 3, pp. 319-326.

Pathak, S.K., Basu, S., Bhattacharyya, A., Pathak, S., Banerjee, A., Basu, J. & Kundu, M. 2006, "TLR4-dependent NF-kappaB activation and mitogen- and stress-activated protein kinase 1-triggered phosphorylation events are central to Helicobacter pylori peptidyl prolyl cis-, trans-isomerase (HP0175)-mediated induction of IL-6 release from macrophages ", *Journal of immunology (Baltimore, Md.: 1950)*, vol. 177, no. 11, pp. 7950-7958.

Pattis, I., Weiss, E., Laugks, R., Haas, R. & Fischer, W. 2007, "The Helicobacter pylori CagF protein is a type IV secretion chaperone-like molecule that binds close to the C-terminal secretion signal of the CagA effector protein ", *Microbiology (Reading, England)*, vol. 153, no. Pt 9, pp. 2896-2909.

Peapus, D.H., Chiu, H.J., Campobasso, N., Reddick, J.J., Begley, T.P. & Ealick, S.E. 2001, "Structural characterization of the enzyme-substrate, enzyme-intermediate, and enzyme-product complexes of thiamin phosphate synthase ", *Biochemistry*, vol. 40, no. 34, pp. 10103-10114.

Penagini, R., Carmagnola, S. & Cantu, P. 2002, "Review article: gastro-oesophageal reflux disease--pathophysiological issues of clinical relevance ", *Alimentary Pharmacology & Therapeutics*, vol. 16 Suppl 4, pp. 65-71.

Pflock, M., Finsterer, N., Joseph, B., Mollenkopf, H., Meyer, T.F. & Beier, D. 2006a, "Characterization of the ArsRS regulon of Helicobacter pylori, involved in acid adaptation ", *Journal of Bacteriology*, vol. 188, no. 10, pp. 3449-3462.

Pflock, M., Kennard, S., Finsterer, N. & Beier, D. 2006b, "Acid-responsive gene regulation in the human pathogen Helicobacter pylori ", *Journal of Biotechnology*, vol. 126, no. 1, pp. 52-60.

Ranganathan, R., Lu, K.P., Hunter, T. & Noel, J.P. 1997, "Structural and functional analysis of the mitotic rotamase Pin1 suggests substrate recognition is phosphorylation dependent ", *Cell*, vol. 89, no. 6, pp. 875-886.

Reis, C.A., David, L., Carvalho, F., Mandel, U., de Bolos, C., Mirgorodskaya, E., Clausen, H. & Sobrinho-Simoes, M. 2000, "Immunohistochemical study of the expression of MUC6 mucin and co-expression of other secreted mucins (MUC5AC and MUC2) in human gastric carcinomas ", *The journal of histochemistry and cytochemistry : official journal of the Histochemistry Society*, vol. 48, no. 3, pp. 377-388.

Reis, C.A., David, L., Correa, P., Carneiro, F., de Bolos, C., Garcia, E., Mandel, U., Clausen, H. & Sobrinho-Simoes, M. 1999, "Intestinal metaplasia of human stomach displays distinct patterns of mucin (MUC1, MUC2, MUC5AC, and MUC6) expression ", *Cancer research*, vol. 59, no. 5, pp. 1003-1007.

Richardson, T. & Finley, J.W. 1985, "Chemical changes in food during processing: papers from a Symposium Held by the Institute of Food Technologists in Anaheim, CA".

Rodionov, D.A., Vitreschak, A.G., Mironov, A.A. & Gelfand, M.S. 2002, "Comparative genomics of thiamin biosynthesis in procaryotes. New genes and regulatory mechanisms ", *The Journal of biological chemistry*, vol. 277, no. 50, pp. 48949-48959.

Rohde, M., Puls, J., Buhrdorf, R., Fischer, W. & Haas, R. 2003, "A novel sheathed surface organelle of the Helicobacter pylori cag type IV secretion system ", *Molecular microbiology,* vol. 49, no. 1, pp. 219-234.

Ruane, K.M., Davies, G.J. & Martinez-Fleites, C. 2008, "Crystal structure of a family GT4 glycosyltransferase from Bacillus anthracis ORF BA1558 ", *Proteins,* vol. 73, no. 3, pp. 784-787.

Saadat, I., Higashi, H., Obuse, C., Umeda, M., Murata-Kamiya, N., Saito, Y., Lu, H., Ohnishi, N., Azuma, T., Suzuki, A., Ohno, S. & Hatakeyama, M. 2007, "Helicobacter pylori CagA targets PAR1/MARK kinase to disrupt epithelial cell polarity ", *Nature,* vol. 447, no. 7142, pp. 330-333.

Sachs, G., Weeks, D.L., Wen, Y., Marcus, E.A., Scott, D.R. & Melchers, K. 2005, "Acid acclimation by Helicobacter pylori ", *Physiology (Bethesda, Md.),* vol. 20, pp. 429-438.

Saunders, N.J., Peden, J.F., Hood, D.W. & Moxon, E.R. 1998, "Simple sequence repeats in the Helicobacter pylori genome ", *Molecular microbiology,* vol. 27, no. 6, pp. 1091-1098.

Schmitt, W. & Haas, R. 1994, "Genetic analysis of the Helicobacter pylori vacuolating cytotoxin: structural similarities with the IgA protease type of exported protein ", *Molecular microbiology,* vol. 12, no. 2, pp. 307-319.

Selbach, M., Moese, S., Hauck, C.R., Meyer, T.F. & Backert, S. 2002, "Src is the kinase of the Helicobacter pylori CagA protein in vitro and in vivo ", *The Journal of biological chemistry,* vol. 277, no. 9, pp. 6775-6778.

Selgrad, M., Kandulski, A. & Malfertheiner, P. 2008, "Dyspepsia and Helicobacter pylori ", *Digestive diseases (Basel, Switzerland),* vol. 26, no. 3, pp. 210-214.

Serganov, A., Polonskaia, A., Phan, A.T., Breaker, R.R. & Patel, D.J. 2006, "Structural basis for gene regulation by a thiamine pyrophosphate-sensing riboswitch ", *Nature,* vol. 441, no. 7097, pp. 1167-1171.

Settembre, E., Begley, T.P. & Ealick, S.E. 2003, "Structural biology of enzymes of the thiamin biosynthesis pathway ", *Current opinion in structural biology,* vol. 13, no. 6, pp. 739-747.

Settembre, E.C., Dorrestein, P.C., Park, J.H., Augustine, A.M., Begley, T.P. & Ealick, S.E. 2003, "Structural and mechanistic studies on ThiO, a glycine oxidase essential for thiamin biosynthesis in Bacillus subtilis ", *Biochemistry,* vol. 42, no. 10, pp. 2971-2981.

Settembre, E.C., Dorrestein, P.C., Zhai, H., Chatterjee, A., McLafferty, F.W., Begley, T.P. & Ealick, S.E. 2004, "Thiamin biosynthesis in Bacillus subtilis: structure of the thiazole synthase/sulfur carrier protein complex ", *Biochemistry,* vol. 43, no. 37, pp. 11647-11657.

Sewald, X., Fischer, W. & Haas, R. 2008, "Sticky socks: Helicobacter pylori VacA takes shape ", *Trends in microbiology,* vol. 16, no. 3, pp. 89-92.

Shannon, P.T., Reiss, D.J., Bonneau, R. & Baliga, N.S. 2006, "The Gaggle: an open-source software system for integrating bioinformatics software and data sources ", *BMC bioinformatics,* vol. 7, pp. 176.

Sharff, A.J., Koronakis, E., Luisi, B. & Koronakis, V. 2000, "Oxidation of selenomethionine: some MADness in the method! ", *Acta crystallographica.Section D, Biological crystallography,* vol. 56, no. Pt 6, pp. 785-788.

Shaw, P.E. 2002, "Peptidyl-prolyl isomerases: a new twist to transcription ", *EMBO reports,* vol. 3, no. 6, pp. 521-526.

Shim, S., Yuan, J.P., Kim, J.Y., Zeng, W., Huang, G., Milshteyn, A., Kern, D., Muallem, S., Ming, G.L. & Worley, P.F. 2009, "Peptidyl-Prolyl Isomerase FKBP52 Controls Chemotropic Guidance of Neuronal Growth Cones via Regulation of TRPC1 Channel Opening ", *Neuron*, vol. 64, no. 4, pp. 471-483.

Sigrell, J.A., Cameron, A.D., Jones, T.A. & Mowbray, S.L. 1998, "Structure of Escherichia coli ribokinase in complex with ribose and dinucleotide determined to 1.8 A resolution: insights into a new family of kinase structures ", *Structure (London, England : 1993)*, vol. 6, no. 2, pp. 183-193.

Souza, R.C. & Lima, J.H. 2009, "Helicobacter pylori and gastroesophageal reflux disease: a review of this intriguing relationship ", *Diseases of the esophagus : official journal of the International Society for Diseases of the Esophagus / I.S.D.E*, vol. 22, no. 3, pp. 256-263.

Stein, M., Bagnoli, F., Halenbeck, R., Rappuoli, R., Fantl, W.J. & Covacci, A. 2002, "c-Src/Lyn kinases activate Helicobacter pylori CagA through tyrosine phosphorylation of the EPIYA motifs ", *Molecular microbiology*, vol. 43, no. 4, pp. 971-980.

Stein, M., Rappuoli, R. & Covacci, A. 2000, "Tyrosine phosphorylation of the Helicobacter pylori CagA antigen after cag-driven host cell translocation ", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 97, no. 3, pp. 1263-1268.

Suerbaum, S. & Achtman, M. 2004, "Helicobacter pylori: recombination, population structure and human migrations ", *International journal of medical microbiology : IJMM*, vol. 294, no. 2-3, pp. 133-139.

Suerbaum, S. & Josenhans, C. 2007, "Helicobacter pylori evolution and phenotypic diversification in a changing host ", *Nature reviews.Microbiology*, vol. 5, no. 6, pp. 441-452.

Suerbaum, S., Josenhans, C. & Labigne, A. 1993, "Cloning and genetic characterization of the Helicobacter pylori and Helicobacter mustelae flaB flagellin genes and construction of H. pylori flaA- and flaB-negative mutants by electroporation-mediated allelic exchange ", *Journal of Bacteriology*, vol. 175, no. 11, pp. 3278-3288.

Tammer, I., Brandt, S., Hartig, R., Konig, W. & Backert, S. 2007, "Activation of Abl by Helicobacter pylori: a novel kinase for CagA and crucial mediator of host cell scattering ", *Gastroenterology*, vol. 132, no. 4, pp. 1309-1319.

Thore, S., Leibundgut, M. & Ban, N. 2006, "Structure of the eukaryotic thiamine pyrophosphate riboswitch with its regulatory ligand ", *Science (New York, N.Y.)*, vol. 312, no. 5777, pp. 1208-1211.

Timm, D.E., Liu, J., Baker, L.J. & Harris, R.A. 2001, "Crystal structure of thiamin pyrophosphokinase ", *Journal of Molecular Biology*, vol. 310, no. 1, pp. 195-204.

Tomb, J.F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., Ketchum, K.A., Klenk, H.P., Gill, S., Dougherty, B.A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E.F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H.G., Glodek, A., McKenney, K., Fitzegerald, L.M., Lee, N., Adams, M.D., Hickey, E.K., Berg, D.E., Gocayne, J.D., Utterback, T.R., Peterson, J.D., Kelley, J.M., Cotton, M.D., Weidman, J.M., Fujii, C., Bowman, C., Watthey, L., Wallin, E., Hayes, W.S., Borodovsky, M., Karp, P.D., Smith, H.O., Fraser, C.M. & Venter, J.C. 1997, "The complete genome sequence of the gastric pathogen Helicobacter pylori ", *Nature*, vol. 388, no. 6642, pp. 539-547.

Toms, A.V., Haas, A.L., Park, J.H., Begley, T.P. & Ealick, S.E. 2005, "Structural characterization of the regulatory proteins TenA and TenI from Bacillus subtilis and identification of TenA as a thiaminase II ", *Biochemistry*, vol. 44, no. 7, pp. 2319-2329.

Torres, V.J., McClain, M.S. & Cover, T.L. 2006, "Mapping of a domain required for protein-protein interactions and inhibitory activity of a Helicobacter pylori dominant-negative VacA mutant protein ", *Infection and immunity*, vol. 74, no. 4, pp. 2093-2101.

Tsutsumi, R., Takahashi, A., Azuma, T., Higashi, H. & Hatakeyama, M. 2006, "Focal adhesion kinase is a substrate and downstream effector of SHP-2 complexed with Helicobacter pylori CagA ", *Molecular and cellular biology*, vol. 26, no. 1, pp. 261-276.

Uemura, N., Okamoto, S., Yamamoto, S., Matsumura, N., Yamaguchi, S., Yamakido, M., Taniyama, K., Sasaki, N. & Schlemper, R.J. 2001, "Helicobacter pylori infection and the development of gastric cancer ", *The New England journal of medicine*, vol. 345, no. 11, pp. 784-789.

Van de Bovenkamp, J.H., Mahdavi, J., Korteland-Van Male, A.M., Buller, H.A., Einerhand, A.W., Boren, T. & Dekker, J. 2003, "The MUC5AC glycoprotein is the primary receptor for Helicobacter pylori in the human stomach ", *Helicobacter*, vol. 8, no. 5, pp. 521-532.

van Vliet, A.H., Kuipers, E.J., Waidner, B., Davies, B.J., de Vries, N., Penn, C.W., Vandenbroucke-Grauls, C.M., Kist, M., Bereswill, S. & Kusters, J.G. 2001, "Nickel-responsive induction of urease expression in Helicobacter pylori is mediated at the transcriptional level ", *Infection and immunity*, vol. 69, no. 8, pp. 4891-4897.

Wadstrom, T., Hirmo, S. & Boren, T. 1996, "Biochemical aspects of Helicobacter pylori colonization of the human gastric mucosa ", *Alimentary Pharmacology & Therapeutics*, vol. 10 Suppl 1, pp. 17-27.

Walter, T.S., Meier, C., Assenberg, R., Au, K.F., Ren, J., Verma, A., Nettleship, J.E., Owens, R.J., Stuart, D.I. & Grimes, J.M. 2006, "Lysine methylation as a routine rescue strategy for protein crystallization ", *Structure (London, England : 1993)*, vol. 14, no. 11, pp. 1617-1622.

Walz, A., Odenbreit, S., Stuhler, K., Wattenberg, A., Meyer, H.E., Mahdavi, J., Boren, T. & Ruhl, S. 2009, "Identification of glycoprotein receptors within the human salivary proteome for the lectin-like BabA and SabA adhesins of Helicobacter pylori by fluorescence-based 2-D bacterial overlay ", *Proteomics*, vol. 9, no. 6, pp. 1582-1592.

Wang, C., Xi, J., Begley, T.P. & Nicholson, L.K. 2001, "Solution structure of ThiS and implications for the evolutionary roots of ubiquitin ", *Nature structural biology*, vol. 8, no. 1, pp. 47-51.

Weeks, D.L., Eskandari, S., Scott, D.R. & Sachs, G. 2000, "A H+-gated urea channel: the link between Helicobacter pylori urease and gastric colonization ", *Science (New York, N.Y.)*, vol. 287, no. 5452, pp. 482-485.

Wen, S. & Moss, S.F. 2009, "Helicobacter pylori virulence factors in gastric carcinogenesis ", *Cancer letters*, vol. 282, no. 1, pp. 1-8.

Wilkins, M.R., Gasteiger, E., Bairoch, A., Sanchez, J.C., Williams, K.L., Appel, R.D. & Hochstrasser, D.F. 1999, "Protein identification and analysis tools in the ExPASy server", *Methods in molecular biology (Clifton, N.J.)*, vol. 112, pp. 531-552.

Wotherspoon, A.C., Ortiz-Hidalgo, C., Falzon, M.R. & Isaacson, P.G. 1991, "Helicobacter pylori-associated gastritis and primary B-cell gastric lymphoma ", *Lancet*, vol. 338, no. 8776, pp. 1175-1176.

Wunder, C., Churin, Y., Winau, F., Warnecke, D., Vieth, M., Lindner, B., Zahringer, U., Mollenkopf, H.J., Heinz, E. & Meyer, T.F. 2006, "Cholesterol glucosylation promotes immune evasion by Helicobacter pylori ", *Nature medicine*, vol. 12, no. 9, pp. 1030-1038.

Xiang, S., Usunow, G., Lange, G., Busch, M. & Tong, L. 2007, "Crystal structure of 1-deoxy-D-xylulose 5-phosphate synthase, a crucial enzyme for isoprenoids biosynthesis ", *The Journal of biological chemistry*, vol. 282, no. 4, pp. 2676-2682.

Xu, J.C., Stucki, J.W., Wu, J., Kostka, J.E. & Sims, G.K. 2001, "Fate of atrazine and alachlor in redox-treated ferruginous smectite", *Environmental toxicology and chemistry / SETAC*, vol. 20, no. 12, pp. 2717-2724.

Ye, D., Willhite, D.C. & Blanke, S.R. 1999, "Identification of the minimal intracellular vacuolating domain of the Helicobacter pylori vacuolating toxin ", *The Journal of biological chemistry*, vol. 274, no. 14, pp. 9277-9282.

Yeo, H.J., Savvides, S.N., Herr, A.B., Lanka, E. & Waksman, G. 2000, "Crystal structure of the hexameric traffic ATPase of the Helicobacter pylori type IV secretion system ", *Molecular cell*, vol. 6, no. 6, pp. 1461-1472.

Yokoyama, K., Higashi, H., Ishikawa, S., Fujii, Y., Kondo, S., Kato, H., Azuma, T., Wada, A., Hirayama, T., Aburatani, H. & Hatakeyama, M. 2005, "Functional antagonism between Helicobacter pylori CagA and vacuolating toxin VacA in control of the NFAT signaling pathway in gastric epithelial cells ", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 27, pp. 9661-9666.

Yoshiyama, H., Nakamura, H., Kimoto, M., Okita, K. & Nakazawa, T. 1999, "Chemotaxis and motility of Helicobacter pylori in a viscous environment ", *Journal of gastroenterology*, vol. 34 Suppl 11, pp. 18-23.

Zanotti, G., Papinutto, E., Dundon, W., Battistutta, R., Seveso, M., Giudice, G., Rappuoli, R. & Montecucco, C. 2002, "Structure of the neutrophil-activating protein from Helicobacter pylori ", *Journal of Molecular Biology*, vol. 323, no. 1, pp. 125-130.

Zhong, Q., Shao, S.H., Cui, L.L., Mu, R.H., Ju, X.L. & Dong, S.R. 2007, "Type IV secretion system in Helicobacter pylori: a new insight into pathogenicity ", *Chinese medical journal*, vol. 120, no. 23, pp. 2138-2142.

Acknowledgements

I would like to express all my sincere gratitude to my supervisor prof. Giuseppe Zanotti for his invaluable assistance and support. His personal understanding, wide knowledge and personal guidance were fundamental for this PhD thesis and for my scientific career: I can never thank him enough.

I would like to thank prof. Thomas F. Meyer that gave me the opportunity to spend 7 months at Max Planck in Berlin and that contributed to enrich my scientific knowledge and experience.

I want to extend my thanks to prof. Roberto Battistutta as well for his helpfulness and kindness and to prof. Paola Spadon.

I must also thank Laura Cendron, who have helped me during these years in assisting my research and in a personal way and for all the nice time spent together. Big thanks to all my lab mates PhD students that helped me in numerous ways, Alessandro, Anton, Lorenza, Sandra, Munan, Ilenia and all the undergraduate students that worked with me: Alberto, Sandro, Luisa, Giulia.

I would like to thank also all the people from VIMM: Elisa, Rosa, Elena, Graziano, Marco, Andrea and Cinzia and all the PhD students from the other labs that I had the pleasure to know during this PhD and during my undergraduate thesis.

From all the friends and nice people that I have met in Berlin, I want to thank first of all Tim and Yuri, Ines, Anita, Angela, Roman, Beatriz, Yujin, Uwe, Christina, Sabrina, Ebbe for their friendship and to make me feel like at home.

Special thanks to all my friends for all their help and support during all these years and for the beautiful time that we spent together, Marco V. & Anna, Marco P., Tommaso, Tigre & Alessia, Marco A. & Lara and to all the other people that supported me: thanks for everything.

I finally want to thank my parents and my brother that have loved me always and my only love Diane, to who I want to say: “We did it! And I will never leave you again!”.