UNIVERSITÀ
DEGLI STUDI
DI PADOVA

UNIVERSITA' DEGLI STUDI DI PADOVA

Department of Biological Chemistry

School of Bioscience and Biotechnology

Curriculum: Biotechnology

XXIV cycle

Pathogenesis-related Proteins From *Helicobacter pylori*: Structural and Functional Studies

Director of the PhD School : **Prof. Giuseppe Zanotti**

Coordinator of the Curriculum : **Prof. Giorgio Valle**

Supervisor : **Prof. Giuseppe Zanotti**

PhD Candidate: Md. Munan Shaik

December 2011

UNIVERSITÀ
DEGLI STUDI
DI PADOVA

UNIVERSITA' DEGLI STUDI DI PADOVA

Department of Biological Chemistry

School of Bioscience and Biotechnology
Curriculum: Biotechnology
XXIV cycle

Pathogenesis-related Proteins From *Helicobacter pylori*: Structural and Functional Studies

Md. Munan Shaik

To My Family

Contents

List of Publications

1. **Shaik MM**, Cendron L, Percudani R, Zanotti G. (2011) The Structure of *Helicobacter pylori* HP0310 Reveals an Atypical Peptidoglycan Deacetylase. PLoS ONE. 6(4): e19207. doi:10.1371/journal.pone.0019207.

2. **Shaik MM**, Cendron L, Zanotti G. (2011) The crystal structure of ADP-L-glycero-D-mannoheptose-6-epimerase (HP0859) from *Helicobacter pylori*. Biochemica et Biophysica Acta. 1814:1641–1647. doi:10.1016/j.bbapap.2011.09.006.

3. **Shaik MM**, Cendron L, Zanotti G. (2011) The crystal structure of *Helicobacter pylori* ceuE (HP1561) reveals an ABC transporter substrate binding protein (To be submitted).

4. **Shaik MM**, Cendron L, Zanotti G. (2011) The Crystal Structure of Glutamine Synthetase (*Hp*GS) from *Helicobacter pylori* (Manuscript in preparation).

List of oral or poster presentations

1. **Md Munan shaik**, Laura Cendron and Giuseppe Zanotti – Structural and Functional characterization of protein related to pathogenesis and colonization from *Helicobacter pylori*. Venetian Institute of Molecular Medicine 109th Annual Retreat, Marostica, October 21-22, 2011, Italy. (Oral Presentation).

2. **Md Munan shaik**, Laura Cendron, and Giuseppe Zanotti - Structural and Functional Studies of Proteins Involved in Pathogenesis and Colonization of *Helicobacter pylori*. Centro Biologico Interdipartimentale "A. Vallisneri**"** BioPhD Day, March 31 - April 01, 2011, Italy. (Poster and Oral presentation).

3. **Md Munan shaik**, Laura Cendron, Sara Giannetti, Marco Salamina and Giuseppe Zanotti - Structural and Functional Studies of Proteins Involved in Pathogenesis and Colonization of *Helicobacter pylori***.** Presented on Venetian Institute of Molecular Medicine, SAB II meeting, February 20-21, 2011, Italy. (Poster Presentation).

4. **Md Munan shaik**, Laura Cendron, Ilenia Bertipaglia, Lorenza Sisinni, Sara Giannetti and Giuseppe Zanotti – Structural and Functional characterization of protein related to pathogenesis and colonization from Helicobacter pylori. Venetian Institute of Molecular Medicine 9th Annual Retreat, Marostica, November 12-13, 2010, Italy. (Poster presentation).

5. **Md Munan shaik**, Laura Cendron and Giuseppe Zanotti – Structural and Functional characterization of protein from *Helicobacter pylori*. XVII BCA/CCP4 Summer School in Macromolecular Crystallography, Oxford University, September 5-11, 2010, Oxford, UK. (Poster presentation).

6. **Md Munan Shaik**, L. Cendron, R. Berni, G. Zanotti- Structural and functional studies of conjugational regulatory proteins from *Enterococcus faecalis*. SILS 2010: XVIII Meeting of the Italian Society for Synchrotron Radiation Padua, 24-26, June 2010, Italy. (Poster presentation).

Summary

Summary

The bacterium *Helicobacter pylori* is recognized as one of the most successful bacterial pathogens. It colonizes the stomach of more than half of the world's population and it presents a very high infection rate in developing countries. Most infected people are asymptomatic, however an important minority of them (15–20%) develops during life severe gastroduodenal pathologies, including stomach and duodenal ulcers, adenocarcinomas and stomach lymphomas. Until now, nine *H. pylori* strains have been completely sequenced and this pathogen presents high genetic variability, not only in the gene sequences but also in genes content. Many of the genes functions are annotated based on sequence similarity and there are around 45% genes with unknown function. The most remarkable differences in *H. pylori* virulent strains compared to non-virulent ones is the presence or absence of the so-called *cag*-PAI (a 40-kb DNA sequence named *cag* Pathogenicity Island), that encodes a Type IV Secretion System, responsible for the translocation of the CagA toxin into host epithelial cells. Although *H. pylori* can be successfully eradicated by antibiotics in many patients, rising antibiotic resistance poses a serious problem. New therapies are required to eradicate *H. pylori* infection and the search for new targets could allow the development of new treatment strategies. Recently, new factors important for colonization and establishment of infection have been proposed, using multiple approaches. A pull of these *H. pylori* proteins have been cloned, expressed in *E. coli* and purified for structural and functional studies.

The aim of this thesis was to determine the three dimensional structural and characterize the function of *H. pylori* proteins important for stomach colonization and pathogenesis. In particular, we have concentrated our efforts on cell wall modification enzymes (peptidoglycan deacetylase), LPS biosynthesis (ADP-L-*glycero*-D-*manno*-heptose-6-epimerase, rfaD), periplasmic substrate binding protein of ABC transporter (ceuE), key enzymes in nitrogen assimilation pathway (Glutamine synthetase), secreted immunogenic disulfide isomerase (DsbG).

The strategy employed includes bioinformatic analyses, molecular cloning of the gene from PCR amplification, construction of vectors for cloning, expression of the protein in *E. coli*. After expression analysis and optimization of conditions, the solubility of the recombinant proteins was checked. The soluble recombinant protein was then purified using different chromatography techniques, and eventually characterized by analytical gel filtration, mass

spectrometry, UV spectroscopy. Techniques to make the protein samples more suitable for crystallization, as DLS (**D**ynamic **L**ight **S**cattering), and to investigate the secondary structure, as CD (**C**ircular **D**ichroism), were used. The proteins were concentrated for crystallization trials. X-ray diffraction data of the crystals were measured at the ESRF synchrotron (Grenoble, France). Functional characterizations of the proteins were accomplished using fluorescence spectroscopy, CD, ITC, UV spectroscopy and ELISA.

After a general introduction about the bacterium (Chapter 1), in chapter 2 the three dimensional structure and the enzymatic activity of a putative peptidoglycan deacetylase is described. Peptidoglycan deacetlyase (HP0310, *Hp*PgdA) from *H. pylori* has been indicated as the enzyme responsible for a modification of peptidoglycan that counteracts the host immune response. The enzyme, which belongs to the polysaccharide deacetylases protein family, is a homo-tetramer. The four polypeptide chains, each folded into a single domain characterized by a non-canonical TIM-barrel fold, are arranged around a four-fold symmetry axis. The active site, one per monomer, contains a heavy ion coordinated in a way similar to other deacetylases. However, the enzyme showed no in vitro activity on the typical polysaccharide substrates of peptidoglycan deacetylases. In striking contrast with the known peptidoglycan deacetylases, *Hp*PgdA does not exhibit a solvent-accessible polysaccharide binding groove, suggesting that the enzyme binds a small substrate at the active site.

The crystal structure of the last enzyme of the biosynthesic pathway of core oligosaccharides (L, D-heptose) of LPS is discussed in details in chapter 3. *H. pylori* owes much of the integrity of its outer membrane on lipopolysaccharides (LPSs). Together with their essential structural role, LPSs contribute to the bacterial adherence properties, as well as they are well characterized for the capability to modulate the immune response. In *H. pylori* the core oligosaccharide, one of the three main domains of LPSs, shows a peculiar structure in the branching organization of the repeating units, which displayed further variability when different strains have been compared. We present here the crystal structure of ADP-L-*glycero*-D-*manno*-heptose-6-epimerase (HP0859, rfaD), the last enzyme in the pathway that produces L-*glycero*-D-*manno*-heptose starting from sedoheptulose-7-phosphate, a crucial compound in the synthesis of the core oligosaccharide. In a recent study, a HP0859 knockout mutant has been characterized, demonstrating a severe loss of lipopolysaccharide structure and a significant reduction of adhesion levels in an infection model to AGS cells, if compared with the wild type strain, in good agreement with its enzymatic role. The crystal structure

reveals that the enzyme is a homo-pentamer, and NAD is bound as a cofactor in a highly conserved pocket. The substrate-binding site of the enzyme is very similar to that of its orthologue in *E. coli*, suggesting also a similar catalytic mechanism. The other enzymes of the pathway are also discussed in terms of their three-dimensional structure.

The crystal structure of the periplasmic ABC transporter substrate-binding protein is discussed in chapter 4. Host-derived iron sources, including heme compounds released from damaged tissues, utilized by *H. pylori* have been identified. *H. pylori* can use heme as a sole iron source. Heme transport across the periplasmic space and into the cytoplasm is affected by an active transport system comprising a soluble periplasmic binding protein, a cytoplasmic permease, and an ATPase (ABC transporter). A periplasmic heme-binding protein have been identified and characterized in several different pathogenic bacteria, but until now no protein was recognized as heme-binding in *H. pylori*. HP15161 (ceuE) was, in fact, annotated as an ABC transporter periplasmic iron-binding protein, but the homologous gene from *H. mustalae* was reported to be involved in nickel transport. To elucidate the structural features of this putative periplasmic ABC transporter binding protein, the protein was cloned, expressed and purified in good yield in *E. coli*, crystallized, its structure determined in apo form and in complex with nickel. The structure was solved by means of single anomalous dispersion experiments on crystals of a selenomethionine variant of the protein. The overall structure is folded in two main domains connected by a long α-helix. It shares the common features of other bacterial periplasmic ABC transporter heme- or vitamin B12-binding protein. The substrate binding site is generated in between these two domain. The crystal structure revels that nickel is not the physiological substrate for this protein and that the substrate-binding cavity is more similar to a heme-binding groove. *In vitro* binding activity with heme was performed with different techniques (Fluorescence and ITC), which confirms that ceuE in *H. pylori* is a periplasmic heme binding protein responsible for heme uptake in conjunction with other members of the ABC transporters family.

The crystal structure of the key enzyme in the unique nitrogen assimilation pathway glutamine synthetase is discussed in chapter 5. Glutamine synthetase (GS) catalyzes the synthesis of glutamine, a central intermediate in nitrogen metabolism, from ATP, glutamate, and ammonia in a divalent metal ion dependent reaction. Ammonia, which is also a preferred nitrogen source for *H. pylori*, is available in plentiful quantity owing to urease activity. It is assimilated into proteins and other nitrogenous compounds through a single nitrogen

incorporation pathway, mediated by GS, an enzyme encoded by gene Hp0512. The absence of an allosteric regulation site (adenylation site) and of other key enzymes in the single nitrogen assimilation pathways makes *Hp*GS a potential target for structural studies. In order to elucidate the structural features of the synthetase and to establish a possible regulatory mechanism of the enzyme, glutamine synthetase (*Hp*GS) from *H. pylori* was cloned, expressed and purified in good yield in *E. coli*, crystallized and its structure determined. The enzyme is a dodecamer, organized in two hexamers. The latter are held together mainly by hydrophobic and hydrogen bonding interactions. The N-terminal helix assembles above the hexameric ring and is exposed to solvent. The C-terminal helix, called the `helical thong,' is inserted into a hydrophobic hole in the eclipsed subunit on the opposite hexameric ring. In addition, the central channel of the dodecamer is lined by six four-stranded L-sheets, each built from an antiparallel loop contributed by subunits of opposite rings. The structure of a monomer consists of a smaller N-terminal domain and a larger C-terminal domain. The dodecameric enzyme contains 12 active sites, which can be described as a 'bifunnel', in which ATP and glutamate bind at opposite ends. The ATP binding site is located at the top of the bifunnel, because it opens to the external 6-fold surface of GS. At the junction of the bifunnel there are two divalent cations-binding sites. The adenylation site that in all the most similar homologs of HpGS contains the consensus sequence NLYDLP, is replaced in *H. pylori* by sequence NLFKLT (residues 405 to 410). Since the Tyr407 residue is the well-conserved target of adenylation and *H. pylori* glutamine synthetase apparently lacks that residue, no adenylation occurs for this enzyme within this motif and no significant structural changes are observed in the adenylation loop.

In chapter 6 the cloning, expression and characterization of secreted immunogenic protein DsbG is discussed. Proteins that are involved in disulfide isomerization in bacteria have been identified. They are located in the membranes or in the periplasm and are called Dsb's for disulfide bond formation. These proteins catalyze the introduction of disulfide bridges, isomerization (shuffling) of incorrectly introduced disulfide bonds and reduction (removal) of inappropriate disulfide bonds. Secreting proteins is a way many pathogenic bacteria use to interact with hosts. Several secreted proteins, either residing in or transiting through the periplasmic space, form disulfide bonds after translocation. HP0231 has a sequence similarity with *E. coli* DsbG and contains the CXXC motif. HP0231 have been already identified as an immunogenic protein, recognized by patient sera. 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 N-terminal His-tag. To solve the purification problems, two approaches were adopted: the cloning of a new HP0231 construct with a C-terminal His-tag and the purification and the refolding of the protein from the inclusion bodies. Finally, refolding procedure was optimized and the protein was purified to high homogeneity. HP0231 was characterized employing immunological techniques and blot with patient blood sera, and by bioinformatics tools.

Riassunto

Il batterio *Helicobacter pylori* è riconosciuto essere uno dei più diffusi patogeni umani: esso colonizza circa la metà della popolazione umana, con una elevata velocità di propagazione nei paesi in via di sviluppo. Benché molti soggetti infetti siano asintomatici, una significativa minoranza di questi (15-20%) sviluppano durante la loro vita patologie duodenali gravi, che includono ulcera gastrica e duodenale, adeno-carcinoma e linfoma dello stomaco. Fino ad oggi sono stati completamente sequenziati nove diversi ceppi di *H. pylori*, in quanto esso presenta una alta variabilità genetica, non solo nelle sequenze geniche, ma anche nel contenuto di geni. Molti dei geni del batterio sono annotati solo sulla base dell'omologia di sequenza e per circa il 45% di essi la funzione è incerta o addirittura ignota. La differenza più significativa tra i ceppi virulenti del batterio rispetto ai ceppi non virulenti è la presenza o assenza della cosiddetta *cag*-PAI (una sequenza del DNA di 40-kb definita "isola di patogenicità *cag*"), un inserto genico che codifica per un sistema di secrezione di tipo IV, responsabile della trasloscazione della tossina CagA nelle cellule epiteliali. Benché *H. pylori* in molti pazienti possa essere sradicato mediante antibiotici, l'aumento della resistenza in alcuni ceppi rappresenta un problema emergente. Nuove terapie sono richieste per combattere il batterio e l'individuazione di nuovi bersagli farmacologici può essere utile per sviluppare nuove strategie di trattamento. Recentemente, nuovi fattori importanti per la colonizzazione e per lo stabilirsi dell'infezione sono stati identificati. In questo lavoro di tesi, un gruppo di queste proteine sono state clonate, espresse in *E. coli* e purificate allo scopo di effettuare studi strutturali.

Obiettivo della tesi era quello di determinare la struttura tridimensionale e caratterizzare la funzione di proteine importanti per la colonizzazione dello stomaco e per la patogenesi. In particolare, gli sforzi sono stati concentrati sugli enzimi coinvolti nella modifica della parete cellulare (peptidoglicano deacetilasi), nella biosintesi di LPS (ADP-L-glycero-D-manno-eptoso-6-epimerasi, rfaD), su una periplasmic-substrate binding protein di un trasportatore ABC (ceuE), su enzimi chiave nel ciclo di assimilazione dell'azoto (glutamina sintasi) e sulla disolfuro ìsomerasi (DsbG), una proteina immunogenica secreta.

La strategia applicata è consistita in una analisi bioinformatica preliminare,

nell'ottenimento del gene a partire da amplificazione mediante PCR, nella costruzione di un vettore per la clonazione e nell'espressione della proteina in *E. coli*. Dopo analisi dell'espressione e ottimizzazione delle condizioni, è stata analizzata la solubilità della proteina ricombinante. Quest'ultima è stata quindi purificata mediante diverse tecniche cromatografiche, ed eventualmente caratterizzata per gel-filtrazione analitica, spettrometria di massa, spettroscopia UV. Sono state usate tecniche per rendere i campioni di proteina più adatti per la cristallizzazione, quali DLS (dynamic light scattering) e per investigarne la struttura secondaria, quali CD (dicroismo circolare). La proteina è stata quindi concentrata prima di essere sottoposta ai test di cristallizzazione. I dati di diffrazione sono stati misurati ai sincrotroni ESRF (Grenoble, Francia). La caratterizzazione funzionale delle proteine è stata eseguita usando spettroscopia di fluorescenza, CD, ITC, UV ed ELISA.

Dopo una introduzione generale sul batterio (Capitolo 1), nel capitolo 2 viene descritta la struttura tridimensionale e l'attività enzimatica di una putativa peptidoglicano deacetilasi. HP0310 (HpPdgA) da *H. pylori* è stato indicato come l'enzima responsabile della modifica del peptidoglicano che serve a minimizzare la risposta immunitaria da parte dell'ospite. L'enzima, che appartiene alla famiglia delle polisaccaride deacetilasi, è un omo-tetramero. Le quattro catene polipeptidiche, ciascuna avvolta in un dominio singolo caratterizzato da un TIM-barrel non canonico, sono arrangiate attorno ad un asse di rotazione quaternario. Il sito attivo, uno per monomero, contiene uno ione coordinato in modo simile ad altre deacetilasi. L'enzima non presenta però *in vitro* attività sui tipici substrati delle peptidoglicano deaetilasi. In netto contrasto con altre peptidoglicano deacetilasi conosciute, HpPdgA non ha un sito di legame accessibile ad una molecola ingombrante quale un polisaccaride, suggerendo che l'enzima leghi nel proprio sito attivo un substrato di piccole dimensioni.

Nel capitolo 3 viene discussa in dettaglio la struttura cristallina dell'ultimo degli enzimi del ciclo della biosintesi dell'oligosaccaride (L,D-eptoso) del core di LPS. *H. pylori* deve molta dell'integrità della sua membrana esterna ai lipopolisaccaridi (LPS). Insieme al loro esenziale ruolo strutturale, gli LPS contribuiscono alle proprietà di aderenze del batterio, come anche alla modulazione della risposta immunitaria. L'oligosaccaride del *core* del batterio, uno dei tre principali domini dell'LPS, presenta una struttura peculiare

nell'organizzazione della ramificazione delle unità che si ripetono. Queste mostrano ulteriore variabilità quando si confrontano ceppi diversi. In questo capitolo viene presentata la struttura cristallina della ADP-L-glicero-D-manno-eptoso-6-epimerasi (HP0859, rfaD), l'ultimo enzima del ciclo che produce L-glicero-D-manno-eptoso partendo da sedoeptuloso-7-fosfato, un composto cruciale nella sintesi dell'oligosaccaride del *core*. In uno studio recente è stato caratterizzato un mutante *knok-out* di HP0859 che mostra, in un modello di infezione in cellule AGS, una seria perdita di struttura del lipopolisaccaride e una significativa riduzione dei livelli di adesione, se paragonato ai ceppi *wild-type*. La struttura cristallina rivela che l'enzima è un omo-pentamero e che NAD è legato come cofattore in una cavità altamente conservata. Il sito di legame del substrato è molto simile a quello del suo ortologo in *E. coli*, suggerendo anche un simile meccanismo catalitico. Altri enzimi del ciclo sono discussi nei termini della loro struttura tridimensionale.

Nel capitolo 4 viene descritta la struttura tridimensionale di una *binding-protein* al trasportatore periplasmico ABC. E' noto dalla letteratura che il batterio può utilizzare l'eme come sola sorgente di ferro, e sono state identificate sorgenti di ferro, derivanti dall'ospite, utilizzate da *H. pylori*, inclusi composti provenienti da gruppi eme da tessuti danneggiati. Il trasporto di eme entro il citoplasma è effettuato da un sistema di trasporto attivo che comprende una proteina periplasmica solubile, una permease citoplasmatica e una ATPase (ABC *transporter*). La proteina periplasmica che lega l'eme è stata identificata e caratterizzata in vari batteri patogeni, ma fino ad ora non era ancora stata identificata una *heme-binding protein* in *H. pylori*. Hp1561 (ceuE) era annotata come una "ABC transporter periplasmic binding protein", ma un gene omologo from *H. mustalae* era stato riportato essere coinvolto nel trasporto del nichel. Per chiarire le caratteristiche strutturali di questa putativa proteina di trasporto, essa è stata clonata, espressa e purificata con buona resa in *E. coli*, cristallizzata e la sua struttura determinata nella forma *apo*- and in complesso con il Ni(II). La struttura è stata risolta per mezzo di esperimenti di dispersione anomala singola (SAD) su cristalli di Se-metionina. Il modello molecolare è costituito da due domini, collegati da una lunga α-elica e presenta le caratteristiche generali di altre "heme-binding periplasmic ABC transporters" o di "B12 binding-proteins". Il sito di legame del substrato è localizzato tra i due domini. La struttura cristallina suggerisce che il Ni(II) non è il legante naturale

della proteina e che la cavità di legame assomiglia di più a quella delle heme-binding proteins. Sono stati effettuati anche studi di legame *in vitro* con tecniche diverse (fluorescenza e ITC), che hanno confermato che ceuE in *H. pylori* è una "periplasmic heme-binding protein", responsabile per l'assunzione dell'eme.

La struttura cristallina di un enzima chiave nell'unico ciclo di assimilazione dell'azoto in *H. pylori* è discusso nel capitolo 5. La glutamina sintetasi (GS) catalizza la sintesi di glutamina, un intermedio centrale nel metabolismo dell'azoto, da ATP, glutammato e ammoniaca, in una reazione dipendente da un catione bivalente. L'ammoniaca, che è anche una sorgente preferita di azoto per *H. pylori*, è disponibile in grande quantità, grazie all'attività ureasica del batterio. E' assimilato in proteine e altri composti contenenti azoto attraverso un singolo ciclo di incorporazione dell'azoto, mediato da GS, un enzima codificato dal gene hp0512. L'assenza di un sito si regolazione allosterico (sito di adenilazione) e di altri enzimi chiave nel ciclo di assimilazione dell'azoto rende HpGS un interessante soggetto per gli studi strutturali, per chiarirne le caratteristiche strutturali e il meccanismo regolatorio. La glutamina sintetasi (HpGS) di *H. pylori* è stata clonata, espressa, purificata e cristallizzata e la sua struttura determinata. L'enzima è un dodecamero, i cui monomeri sono tenuti assieme soprattutto da interazioni idrofobiche e legami ad idrogeno tra i due anelli esamerici. L'elica N-terminale, chiamata "elica stringa", è inserita in una cavità idrofobia nella subunità eclissata sull'anello esamerico opposto. In aggiunta, il canale centrale del dodecamero è formato da sei fogli L a quattro fogli beta, ciascuno costituito da un *loop* antiparallelo cui contribuiscono subunità in anelli opposti. La struttura di un monomero consiste di un piccolo dominio N-terminale e di un dominio C-terminale più grande. L'enzima dodecamerico contiene 12 siti attivi, ciascuno dei quali può essere descritto come un "*bifunnel*", in cui ATP e glutammato legano da lati opposti. Il sito di legame dell'ATP è localizzato nella parte alta del *bifunnel*, alla giunzione del quale ci sono due siti per il legame di cationi bivalenti. Il sito di adenilazione che in tutti gli enzimi omologhi contiene la sequenza consenso NLYDLP è sostituita in *H. pylori* da NLFKLT (residui da 405 a 410). Poiché la tirosina 407 è il bersaglio conservato dell'adenilazione e *H. pylori* apparentemente manca di questo residuo, non c'è adenilazione nell'enzima e non si osservano modifiche strutturali significative in questo *loop* a confronto con altre strutture di GS.

Nel capitolo 6 vengono discussi la clonazione, l'espressione e la caratterizzazione della proteina immunogenica secreta DsbG. Le proteine coinvolte nell'isomerizzazione dei ponti disolfuro sono ben note. Esse sono localizzate nelle membrane o nel periplasma e sono chiamate Dsb (*Disulfide bond formation*). Questi enzimi catalizzano l'introduzione di ponti disolfuro, la loro isomerizzazione o rimozione. Secernere alcune proteine è un modo attraverso il quale un batterio può interagire con l'ospite e in *H. pylori* questo avviene attraverso uno dei tre sistemi di secrezione di tipo IV presenti. Molte tra le proteine secrete, alcune delle quali risiedono nello spazio periplasmico, formano ponti a disolfuro dopo la traslocazione. HP0231 ha una sequenza simile a Dsbg di *E. coli* e contiene un motivo CXXC. Essa è stata già identificata come una proteina immunogenica riconosciuta da sieri di pazienti. La proteina ricombinante, espressa in *E. coli*, non lega fortemente alla resina di affinità IMAC-Ni^{2+}, probabilmente a causa di degradazione dell'His-tag presente all'N-terminale. Per risolvere il problema della purificazione sono perciò stati adottati due approcci: è stato preparato un nuovo costrutto di HP0231 con His-tag al C-terminale, e in parallelo la proteina è stata purificata dai corpi di inclusione e poi rifoldata. La proteina è infine stata purificata ad alta omogeneità. HP0231 è stata caratterizzata usando tecniche immunologiche e *blot* con sieri di pazienti.

Chapter One

General Introduction

1. Introduction

1.1. *Helicobacter pylori*

The bacterium *Helicobacter pylori is* recognized as one of the most successful bacterial pathogens. It colonizes the stomach of more than half of the world's population and presents very high infection rates in developing countries (Rothenbacher and Brenner, 2003). The presence of some microbes in the human stomach was already known since 1893, but *H. pylori* was identified only in 1982 by Marshall and Warren (Marshall and Warren 1984; Nobel Prize 2005). In the last 30 years, this bacterium has become object of intense research because of its correlation with several gastric diseases. *H. pylori* is an S-shaped microaerophilic, slow growing, non-spore forming gram-negative rod bacterium, 0.5 - 5 µm in length with a clump of 5 to 7 polar-sheathed flagella (O'Rourke and Bode, 2001) (Figure 1).

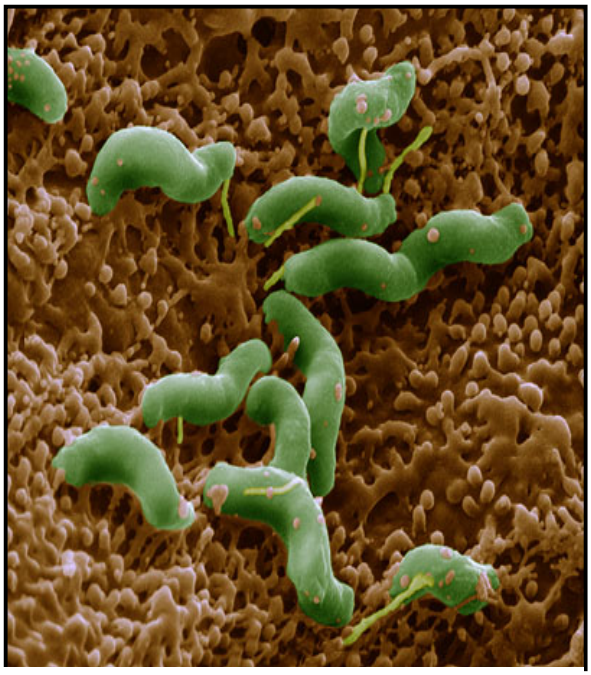


Figure 1. Scanning electron microscopy image of *H. pylori.* Adapted from "thehealthage" web site.

H. pylori has probably accompanied humans for tens of thousands of years (Ghose *et al.,* 2002; Falush *et al.*, 2003; Linz *et al.,* 2007) and it has been hypothesized that *H. pylori* colonization could have provided benefits to its human carriers and hence provided a selective advantage during long periods of human history (Blaser, 2006). The infection is often acquired early in life (almost always before the age of 10) by fecal-oral or oral-oral route (Blaser, 1998). Epidemiological studies suggested that *H. pylori* is mainly transmitted vertically within families and does not spread epidemically (Kivi *et al.,* 2003). Most infected people are asymptomatic, with moderate inflammation detectable only by biopsy and histology. The clinical upshot of *H. pylori* infection is determined by multiple factors: host genetic predisposition (cytokine polymorphisms), (Amieva and El-Omar, 2008), heterogeneity presented by different bacterial strain and environmental factors (Forman and Burley, 2006). However, an important minority of affected people (15–20%) develops during life severe gastroduodenal pathologies, including stomach and duodenal ulcers, adenocarcinomas and stomach lymphomas (Montecucco and 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 and Josenhans, 2007), a connection that prompted the World Health Organization to affirm *H. pylori* as the first bacterial class 1 carcinogen in 1994 (Blaser, 1998; Peek, 2002; Suerbaum and Michetti, 2002).

1.2. *Genetic variability*

Sampling the bacterium in human populations from Asia, Africa, and South America has demonstrated that *H. pylori*-human coevolution started about 58,000 years ago (Linz *et al.,* 2007; Yamaoka *et al.,* 2002; Falush *et al.,* 2003; Devi *et al.,* 2007). 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' After the initial infection and the subsequent colonization of the mucosal layer, in order to persist in close contact with epithelial cells the bacterium is continuously facing harsh physiological conditions and a vigorous immune response. The need for adaptation to the extremely changing microenvironment of individual hosts (van Vliet *et al.,* 2001a) is probably the cause of a high degree of interstrain genetic variation observed in the *H. pylori* population worldwide (Kang and Blaser, 2006). To successfully adapt to the environment, *H. pylori* accumulates different kinds of mutations: point mutations and changes in genomic structure, like inversion, transposition, translocation and duplication, gene gain and gene loss, gene fusion and gene split, gene fragmentation (pseudogene), insertion and deletion (Tillier and Collins, 2000; Blaser, 1994). These changes, along with intergenic regions (Bereswill *et al.,* 2000) direct repeats (Aras *et al.,* 2003a) present in all isolates, are thought to provide specific mechanisms to subvert host immunity (Cooke *et al.,* 2005). In addition, the order of genes can vary among strains (Jiang *et al.,* 1996). Direct repeats had also important role in *H. pylori* DNA diversification (Aras *et al.,* 2003b). *H. pylori* has a high rate of mutation and recombination that allow a quasi-panmictic population for rapid adaptation to a new environment (Suerbaum *et al.,* 1998; Schreiber *et al.,* 2005)]. Furthermore, the inter-strains diversity of *H. pylori* is extended by plasmids (Hofreuter and Haas, 2002) and there is evidence that active import of DNA via specific uptake machineries (Karnholz *et al.,* 2006) promotes genetic variability in the gastric pathogen (de Reuse and Bereswill, 2007).

H. pylori possesses a reduced genome, only from 1.5 to 1.7 Mb. Until now, nine *H. pylori* genome sequences are available from public databases: 26695 (Tomb *et al.,* 1997), J99 (Alm *et al.,* 1999), P12 (Fischer *et al.,* 2010), HPAG1 (Oh *et al.,* 2006), or G27 (Baltrus *et al.,* 2009), Shi470 (Devi *et al.,* 2006), B38 (Thiberge *et al.,* 2010), 51(CP000012) and 52 (CP001680). They represent the genetic information of isolates from patients with various diseases, from gastritis to cancer, and from different geographic regions (Lara-Ramı'rez *et al.,* 2011). Comparison of the genome of the nine strains showed that the core genome of *H. pylori* consists of 1186 genes, of which 22 genes represent a particular adaptation to the human stomach niche; about 450 genes are found to be present in at least two strains, whilst each individual strain presents a unique number of genes (Figure 2). *H. pylori* contains a high proportion of pseudogenes whose genesis was principally caused by homopolynucleotide (HPN) mutations. Such mutations are reversible and facilitate the control of gene expression through the change of the DNA structure. The reversible mutations and a quasi-panmictic feature could allow such genes or gene fragments to be frequently transferred within or between populations (Lara-Ramı´rez *et al.,* 2011).

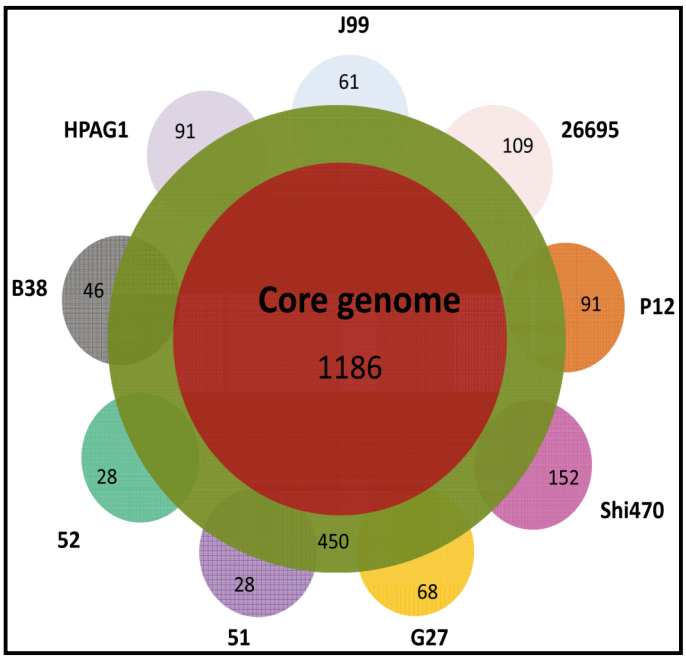


Figure 2. Genetic Variability of *H. pylori*. The central cycle represents core genome. The secondary cycle indicates the genes that are shared by at least two genomes. Other small partial cycles show the strain-specific genes. The number within the cycles means the number of genes in this category (Lara-Ramı´rez *et al.,* 2011).

1.3. *Colonization*

The successful persistence of a life long colonization of the human stomach by *H. pylori* is achieved through a combination of factors, which address the different challenges presented by the harsh environment. In general, the human stomach is well protected against bacterial infections, but *H. pylori* can survive in it thanks to a bundle of polar flagella, a potent urease activity and some other additional features, including the ability to adhere to the gastric epithelial cells.

1.3.1. *Survival in stomach*

Acid tolerance

The human stomach is a unique ecological niche characterised by a very acid pH. It is a harsh 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 and Small, 1993). *H. pylori* is not an acidophilus bacterium and is able to survive in the acid lumen only for a short period, sufficient to enter the highly viscous mucosa, reach the gastric epithelium, find nutrient and multiply (Suerbaum and Josenhans, 1999). It has evolved specialized processes by dedicating a gene cluster (*ure*AB and *ure*EFGHI) to the biosynthesis of Urea Amidohydrolase (EC 3.5.1.5, also named Urease), that allows 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). This is accomplished mainly by producing abundant quantities of urease, which hydrolyses urea into ammonia and CO_2, resulting in the generation of a pH buffered micro-environment (Dunn and Phadnis, 1998). In fact, urease-defective *H. pylori* mutants cannot colonise the stomach (Eaton *et al.,* 1991). Urease is one of the main antigens recognised by the human immune response to *H. pylori* (Del Giudice *et al.,* 2001). It is produced in a large amount (10–15% of total proteins by weight). Urease is a cytoplasmic enzyme, but it has been hypothesized that it becomes associated with the surface of the organism in a process called 'altruistic autolysis', in which some of the bacteria lyses spontaneously, followed by adsorption of the enzyme onto the surface of the remaining intact bacteria (Phadnis *et al.,* 1996; Marcus and Scott, 2001). Up to 30% of the total urease is present on the surface (Dunn *et al.,* 1997). Surface or free urease has a pH optimum activity 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). *H. pylori* urease is a dodecamer and consists of two different subunits of 61.7 kDa (α) and 26.5

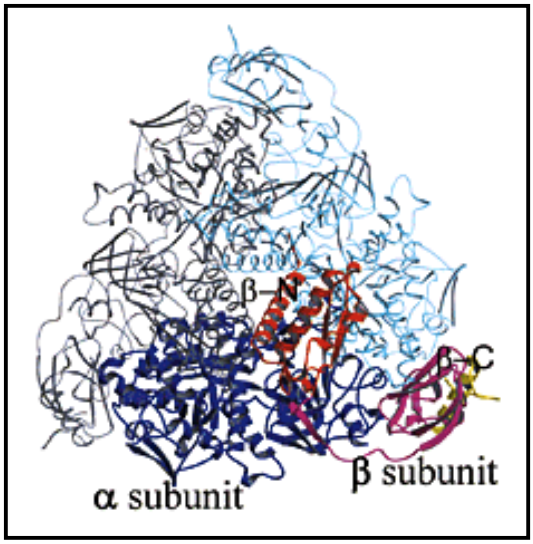


Figure 3. Structure of *H. pylori* urease. Ribbon diagram of the trimeric unit. Each unit consists of the α-subunit (blue,) and two physically distinct domains of the β-subunit (red, magenta), Ha *et al.,* 2001.

kDa (β), forming a massive complex whose molecular mass has been estimated to be about 600 kDa. TEM (Transmission Electron Microscopy) studies revealed a 1.1 MDa double ring dodecameric assembly of approximately 13 nm in diameter (Ha *et al.,* 2001) (Figure 3). A proton-gated urea channel, UreI (HP0071) regulates the intake of urea in the bacterial cytoplasm (Weeks *et al.,* 2000). The urea channel UreI is a six helices transmembrane protein regulated positively by protons, which opens at acidic pH values to permit more urea to enter into the cytoplasm for a maximal production of ammonia and carbamate, and closes at neutral pH to avoid over-alkalinization, which is lethal to *H. pylori* (Clyne *et al.,* 1995). UreE (HP0070), UreF (HP0069), UreG (HP0068) and UreH (HP0067) are instead accessory components, which probably participate as chaperones during the Urease enzyme assembly and/or in the activation/inhibition of its catalysis.

Another important system employ by *H. pylori* for acid regulation is the ArsSR system encoded by the genes HP0165 and HP0166, 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) It 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 and Moss, 2009).

The metal-dependent regulators NikR (HP1338) and Fur (HP1027) were also shown to play an important role in the transcriptional regulation of acid resistance. Fur controls intracellular iron homeostasis *via* concerted expression of iron-uptake and iron-storage genes (van Vliet *et al.,* 2002a; Delany *et al.,* 2001a and 2001b). The transcription of Fur is repressed at low pH (Bijlsma *et al.,* 2002; Bury-Mone *et al.,* 2004), and this was shown to be mediated by NikR (van Vliet *et al.,* 2004).

Flagellar motility

Flagellar motility, including chemotaxis, is one of the most predominant colonization and virulence factors of *H. pylori* (Kavermann *et al.,* 2003). The helicoidal-shaped bacterium propelled by its flagella travels across the viscous mucus film like a screw into a cork; non-

motile mutants cannot colonize the stomach (Montecucco and Rappuoli, 2001). In *H. pylori* around fifty putative proteins are predicted to be involved in expression, secretion, and assembly of this complex flagellar machinery (Tomb *et al.,* 1997; Alm *et al.,* 1999), formed by at least twenty proteins. The structure of the *H. pylori* flagellum is similar to those of the bacteria *E. coli* and *S. thyphimurium* (Spohn and Scarlato, 2001): it is composed of a basal body, a hook and a flagellar filament. The flagellum consists 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 70 x 16 nm *H. pylori* hook is composed of FlgE subunits of 78kDa, whereas *fliD* gene encodes a hook-associated protein that helps the flagellin monomers to be incorporated into the growing flagellum (O'Toole *et al.,* 1994). Other genes are involved in flagella building and regulation, but their roles are still unclear (O'Toole *et al.,* 2000). In addition, *H. pylori* was shown to possess the enzymatic ability to disrupt the oligomeric structure of mucin (Windle *et al.,* 2000), which may assist the pathogen to move freely in the mucus layer. The movement of *H. pylori* towards the stomach mucosa is guided by chemotactic factors, which include urea and bicarbonate ions (Yoshiyama *et al.,* 1999).

1.3.2. *Adhesion to the gastric cells*

Approximately 20% of *H. pylori* in the stomach are found adherent to the surface of the mucus layer that covers the epithelial cells. The bacterium comes into contact with mucins both by flagellar movement and chemo-attraction (abundant urea and bicarbonate are present in the mucosa) (Yoshiyama *et al.,* 1999). Cell adhesion is an essential step in *H. pylori* colonization but not all *H. pylori* strains contain all of the adhesins, which contribute to the strain variation in *H. pylori*. Along with individual types of host's, mucins genetically determined might facilitate the colonization of *H. pylori* compared to other types of mucins. *H. pylori* adheres strongly to human gastric epithelial cells using fucosylated glycoproteins and sialylated glycolipids as cellular receptors (Mahdavi *et al.,* 2002; Boren *et al.,* 1993) (Figure 4). Adherence to sialic acid in the mucus 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 *et al.,* 1996) and one of them (*HpaA*) has been demonstrated to be essential for *H. pylori* colonization in mice (Carlsohn *et al.,* 2006). The best characterized *H. pylori* adhesin are the two outer membrane proteins BabA and SabA. BabA, which is 75 kDa surface exposed protein binds specifically to the fucosylated Lewis B (LeB) antigen in mucin MUC5AC (Van de Bovenkamp *et al.,* 2003). SabA, that has been shown to bind sialylated glycoconjugates, in particular to 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).

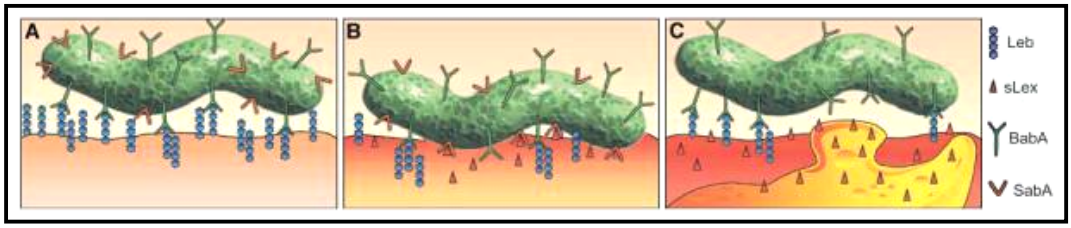


Figure 4. *H. pylori* adherence in health and disease. This figure illustrates the proficiency of *H. pylori* for adaptive multistep mediated attachment (Mahadavi et al., 2002).

Other adhesins identified to date include the Outer membrane proteins AlpA/AlpB (HP0912/HP0913) (Odenbreit *et al.,* 1999) and HopZ (HP0009) (Peck *et al.,* 1999), which are known to mediate the attachment to epithelial cells by defining a macromolecular complex on the bacterial surface with other outer membrane proteins. Recently, OipA (HP0638), an outer membrane protein associated with more virulent strains of *H. pylori*, was identified as a potential colonization factor (Yamaoka *et al.,* 2000; Dossumbekova *et al.,* 2006), associated with the presence of duodenal ulceration and gastric cancer, high *H. pylori* density and severe neutrophil infiltration (Yamaoka *et al.,* 2006). OipA mutagenesis resulted in reduced bacterial adherence to gastric epithelial cells, but did not alter IL-8 secretion *in vitro* (Dossumbekova *et al.,* 2006).

1.3.3. *Persistence*

H. pylori has evolved mechanisms to survive the severe environment of the stomach: it dynamically moves through the mucosal layer, sticks to the epithelium, escapes immune responses, and accomplishes persistent colonization (Suarez *et al.,* 2006). The humoral and cellular immune responses elicited by *H. pylori* colonization in most cases does not outcome in a complete clearance of the bacteria (Algood and Cover, 2006), and often significant biological costs are generated by such long-term host pathogen relationships (Polk and Peek, 2010). The failure of the host response to clear infections of *H. pylori* could reflect down-regulatory mechanisms that constraint the resulting immune responses to prevent risky inflammation as a means to protect the host. Consequently, the chronic immune response induced may be insufficient or misdirected, and could thus offer a colonization benefit for the bacteria by providing enhanced accessibility of nutrients or of adhesion places. An example of this is the consequential increase in class II major histocompatibility complex (MHC) and

CD74, induced by IFN-γ and IL-8, that are used as receptors by H pylori (Beswick *et al.,* 2005a, b; Fan *et al.,* 1998).

H. pylori has been shown to employ numerous mechanisms to antagonize, impair, or subvert host responses. Among them we can include an increasing genetic diversity (Blaser and Atherton, 2004), the inhibition of the production of nitric oxide by macrophages, the inhibition of phagocytosis (Gobert *et al.,* 2002), the interference with both uptake and processing of antigens by antigen-presenting cells (APCs) through its VacA virulence factor (Molinary *et al.,* 1998), the suppression of T cell proliferation and activation through interference of the calcineurin associated IL-2 signaling pathway (Gobert *et al.,* 2003), the induction of selective T cell apoptosis (Gebert *et al.,* 2003), the mimicry of the gastric epithelial fucosylated (Lewis) antigens, which helps to evades host adaptive responses (Wirth *et al.,* 1997) and the presention of an antigenic variation of the surface proteins including a critical pilus molecule, CagY (Aras *et al.,* 2003), which tightly adheres to the epithelial cells by an array of adhesin (Mahadavi *et al.,* 2002). These and other multiple interpretations on the scenery of the immune response during *H. pylori* infection have led to models that help to elucidate how *H. pylori* can persist in the gastric environment by generating an incompetent immune response. The ineffective immune response, together with some others host factors, determines the severity and onset of the disease.

Toll-like receptors (TLRs) are single component, non-catalytic cell-surface receptors responsible for the recognition of structurally conserved molecules derived from microorganism. They are employed by the innate immune system to discriminate pathogen-associated molecular patterns (Takeda *et al.,* 2003). *H. pylori* flagella are not recognized by TLR5, which are able to recognize bacterial flagella such as those of *Salmonella typhimurium* (Gewirtz *et al.,* 2003). *H. pylori* DNA are highly methylated and likely minimize the recognition by TLR9, which recognizes the largely unmethylated DNA of most bacteria (Takata *et al.,* 2002). *H. pylori's* LPS present some composition variation compared with that of other enteric bacteria. The latter are due to lipid A core modifications (Muotiala *et al.,* 1992) and/or variation in sugar composition. The latter is able to stimulate macrophage's TLR4 but not gastric epithelial TLR4 (Backhed *et al.,* 2003). Although *H. pylori* is relatively masked from innate immune sensors present on cell surfaces of the host, cag positive strains do stimulate NF-κB activation in epithelial cells (Foryst-Ludwig and Naumann, 2000), evidently through recognition by Nod1, intracellular pathogen-

recognition molecule responsible for the recognition of pathogen-derived soluble components (peptidoglycan) (Giardin *et al.,* 2003). *H. pylori* has also evolved a mechanism to modify, through the enzyme peptidoglycan deacetylase, their cell wall murein layer to avoid immune recognition by the host (Wang *et al.,* 2010). The relative contributions to *H. pylori* persistence of the different host management and evasion strategies are not established, perhaps differing in individual hosts, but the subsistence of these diverse mechanisms implies that immune supervision of the gastric lumen is prevailing, and that bacterial continued existence requires its subversion (Blaser and Atherton, 2004).

1.4. *Virulence factors of H. pylori*

H. pylori persists in the stomach in spite of activation of both the innate and adaptive immune systems (Pinto-Santini and Salama, 2005). Bacterial colonization is typically followed by infiltration of the gastric mucosa by polymorphonuclear leukocytes, macrophages and lymphocytes. To colonize the host, *H. pylori* employs an array of machineries. Virulence factors have been studied for this bacterium in an effort to show a relationship between bacterial phenotype with specific manifestations of disease, and to elucidate mechanisms of pathogenesis. Virulence factors with potential predictive for specific pathologies include the presence of the cag-pathogenicity island, specific vacuolating cytotoxin A (*vacA*), proteins induced by contact with epithelium (*iceA*) alleles, and blood group antigen-binding adhesion (*babA2*) genes (D'Elios *et al.,* 2007).

1.4.1. VacA

The Vacuolating cytotoxin A (VacA) is a secreted protein that induces the formation of large cytoplasmic vacuoles in gastric cultured cell lines (Leunk *et al.,* 1988; Cover *et al.,* 1992). VacA is not only an important antigen in the human response to *H. pylori* (Del Giudice *et al.,* 2001), but also an important virulence factor which confers a strong competitive advantage to *H. pylori* wild-type strains with respect to VacA-defective mutants in the colonisation of the stomach (Salama *et al.,* 2001) and is associated with more severe cases of disease (Atherton *et al.,* 1995; van Doorn *et al.,* 1998). 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 30nm in diameter (Sewald *et al.,* 2008). The mature toxin has two domains, p33 and p55. The latter are cleaved after a two-step bacteria secretion process involving a N-terminal signal peptide of 33 amino acids, which directs secretion from the

cytoplasm to the periplasm, and a large C-terminal domain, around 45 kDa, which seems to act like an autotransporter directing export across the outer membrane (Telford *et al.,* 1994; Montecucco and Rappuoli, 2001) (Figure 5).

Figure 5. The vacA gene encodes a 140-kDa protoxin. The mature 88-kDa VacA toxin contains two domains, designated p33 and p55. The mid-region sequence that defines type m1 and m2 forms of VacA is located within p55.

The secreted toxin can assemble into water-soluble oligomeric structures (Cover *et al.,* 1997), and can be inserted into planar lipid bilayers to form anion-selective membrane channels. Wide ranges of cellular effects have been attributed to VacA and it has been shown to act on both epithelial and immune cells (Cover and Blanke, 2005). Such activities include the induction of apoptosis (Galmiche *et al.,* 2000), alteration of the process of antigen presentation (Molinari *et al.,* 1998), inhibition of T cell activation and proliferation (Boncristiano *et al.,* 2003; Gebert *et al.,* 2003; Sundrud *et al.,* 2004).

Cryo-EM structure of VacA revealed that the secreted toxin has a strong tendence to oligomerize into rosettes (Lupetti *et al.,* 1996). The structure resembles “flowers” in which a central ring is surrounded by peripheral “petals”. The crystal structure of p55 VacA domain was determined (PDB code: 2QV3). It is predominantly a right-handed parallel β-helix, a feature that is characteristic of autotransporter passenger domain, but unique among known bacterial protein toxins (Gangwer *et al.,* 2007) (Figure 6).

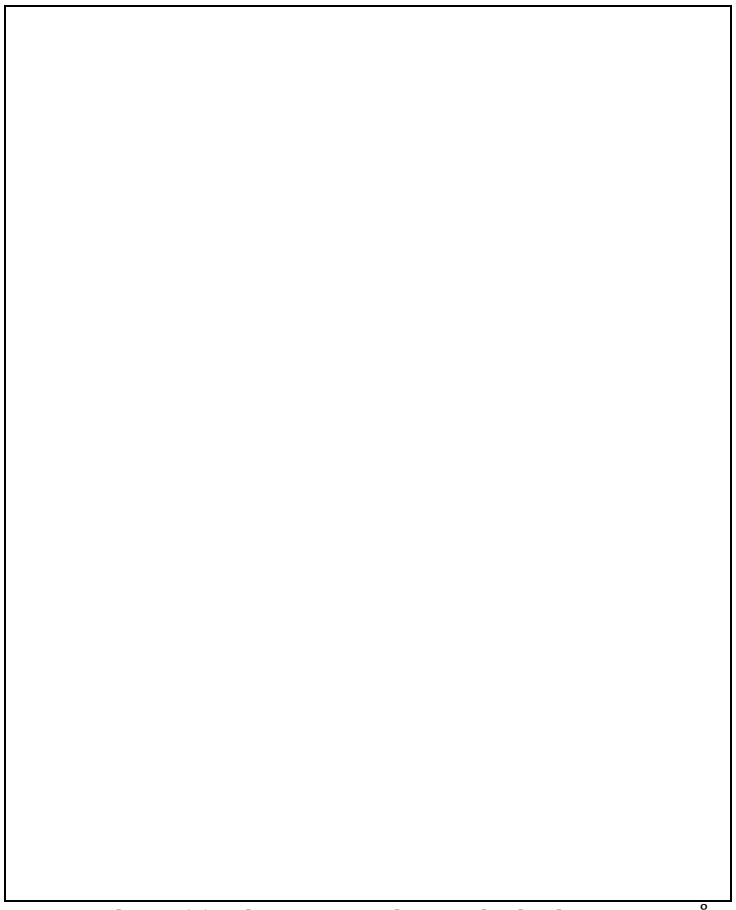


Figure 6. A. Twelve p55 subunits are shown docked into a 19-Å cryo-EM map of a VacA dodecamer (left). In the proposed oligomerization model, p33 interacts with the N-terminal portion of p55 from the neighbouring subunit. B. The VacA p55 fragment adopts a β-helix structure that is composed of three parallel β-sheets (red) connected by loops of varying length and structure (yellow). Adapted from Gangwer *et al.,* 2007.

1.4.2. HP-NAP

H. pylori neutrophil-activating protein (HP-NAP), a highly immunogenic protein, is responsible for the attraction of neutrophils and monocytes by chemotaxis at the site of *H. pylori* infection (Yoshida *et al.,* 1993; Satin *et al.,* 2000). It is a 180 kDa oligomeric protein, composed of 12 identical subunits of about 15 kDa each. HP-NAP promotes the adhesion of human neutrophils to the endothelial cells and stimulates the production of reactive oxygen radicals (ROs) (Yoshida *et al.,* 1993; Evans *et al.,* 1995). 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 *et al.,* 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 and 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 (Del Prete *et al.,* 1995). HP-NAP shows similarities to some bacterial DNA binding proteins (Dps), but it does not bind to DNA at neutral pH like the other Dps proteins. It can store up to 500 atoms of iron (Tonello *et al.,* 1999). The structure of the monomer is a four-helix bundle protein that oligomerizes to form a dodecamer with a central internal iron-containing negative cavity (Zanotti *et al.,* 2002; PDB code: 1JI4). Since iron is not an essential nutrient for *H. pylori* growth and HP-NAP synthesis is not regulated by iron concentration or influenced by the presence of other metals, several researchers have hypothesized that HP-NAP was originally an iron-binding or iron-regulated protein which has evolved to function as a leukocyte activator able to induce mucosal damage, through neutrophils activation, and an easy release of nutrients for the bacterial survival and colonization (Montecucco and Rappuoli, 2001).

1.4.3. CagPAI

The most virulent *H. pylori* strains produce a key toxin, the CagA protein (Odenbrei *et al.,* 2000), which is transported to host cell by a dedicated secretion system belonging to the class of type IV secretion systems (T4SS). The T4SSs are a large group of transporter machines functioning as a molecules transporting device present in many Gram-negative bacteria, that facilitate host–pathogen interaction and⁄or inject toxins into the host cell, ancestrally related to conjugation systems (Dreiseikelmann, 1994; Cascales and Christie, 2003; Backert and Meyer, 2006). T4SS are adapted to transport different substrate (proteins or DNA–protein complexes) to different recipients (bacteria of the same or different species, or to a different organism like fungi, plants or mammalian cells). T4SSs were also reported to be present in *Agrobacterium*, *Legionella*, *Bartonella*, *Bordetella* and other pathogens. In *H. pylori* CagA toxin and the T4SS machinery responsible for its transfer to the eukaryotic cell are encoded by the so-called cytotoxin associated gene-pathogenicity island (cagPAI), a 40 kbp DNA fragment, flanked by 31 bp direct repeats, which possibly was acquired horizontally from an unknown ancestor and integrated into the chromosomal glutamate racemase gene (Tegtmeyer *et al.,* 2011). CagA is considered as a model for bacterial carcinogenesis (Hatakeyama and

Higashi, 2005). Once translocated into the host cell, CagA is tyrosine-phosphorylated by host cell kinases (Covacci and Rappuoli, 2000), and interacts with more than 20 different human proteins involved in signal transduction (Backert *et al.,* 2010). As a consequence of CagA action, epithelial cells experienced major functional disturbance, such as cell–cell adhesion, signalling, adherence and proliferation (Hatakeyama, 2008). The cagPAI system also delivers bacterial peptidoglycan that triggers the Nod1-response and induction of the nuclear factor-jB pathway (Viala *et al.,* 2004).

T4SSs typically consist of 11 VirB proteins (encoded by *virB1–11*) and coupling protein (VirD4, an NTPase). In *Agrobacterium* all the VirB proteins are categorized in groups: VirB6-10, forming the channel or core components; VirB4 and VirB11, the energetic components and the VirB2, possibly VirB3 and VirB5 forming the pilus-associated components (Backert and Selbach, 2008). The conserved components of the *cag*-PAI-encoded T4SS (Cag-T4SS) have been identified by sequence comparison analysis with those of the VirB⁄D system (Fischer *et al.,* 2011). *cag*-PAI encodes for 27 proteins although the archetypal VirB⁄D system has 12 components. Of the 32 genes carried by the *cag*-PAI, 18 genes are required for translocation of CagA into gastric epithelial cells (Fischer *et al.,* 2001; Bourzac and Guillemin, 2005) and 9 of them have weak homology with the type IV secretion system of *A. tumefaciens*, the best-studied model. The remaining 14 genes are unique to the *H. pylori cag*-PAI, and very little is known about their function and probably they play a role in cag-PAI-dependent induction of the chemokine IL-8 (Tanaka *et al.,* 2003). The proteins forming the core complex, CagT/HP0532, CagX/HP0528, and CagY/HP0527 in *H. pylori* are homologue of VirB7, VirB9, and 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 most likely useful to avoid the immune-response (Rohde *et al.,* 2003), and by the adhesin CagL, which has a role not only as anchors of the type-IV secretion machinery to the host surface through binding to integrin, but also promotes signal transduction (Kwok *et al.,* 2007). Cagγ/HP0523 is a transglycosylase that lyses the murein cell wall to facilitate assembly of the T4SS across the bacterial cell wall (Zhong *et al.,* 2007). Finally, CagF/HP0543 is crucial for the translocation of CagA, is a chaperone-like binds close to the C-terminal secretion signal of the CagA effector protein (Pattis *et al.,* 2007) (Figure 7).

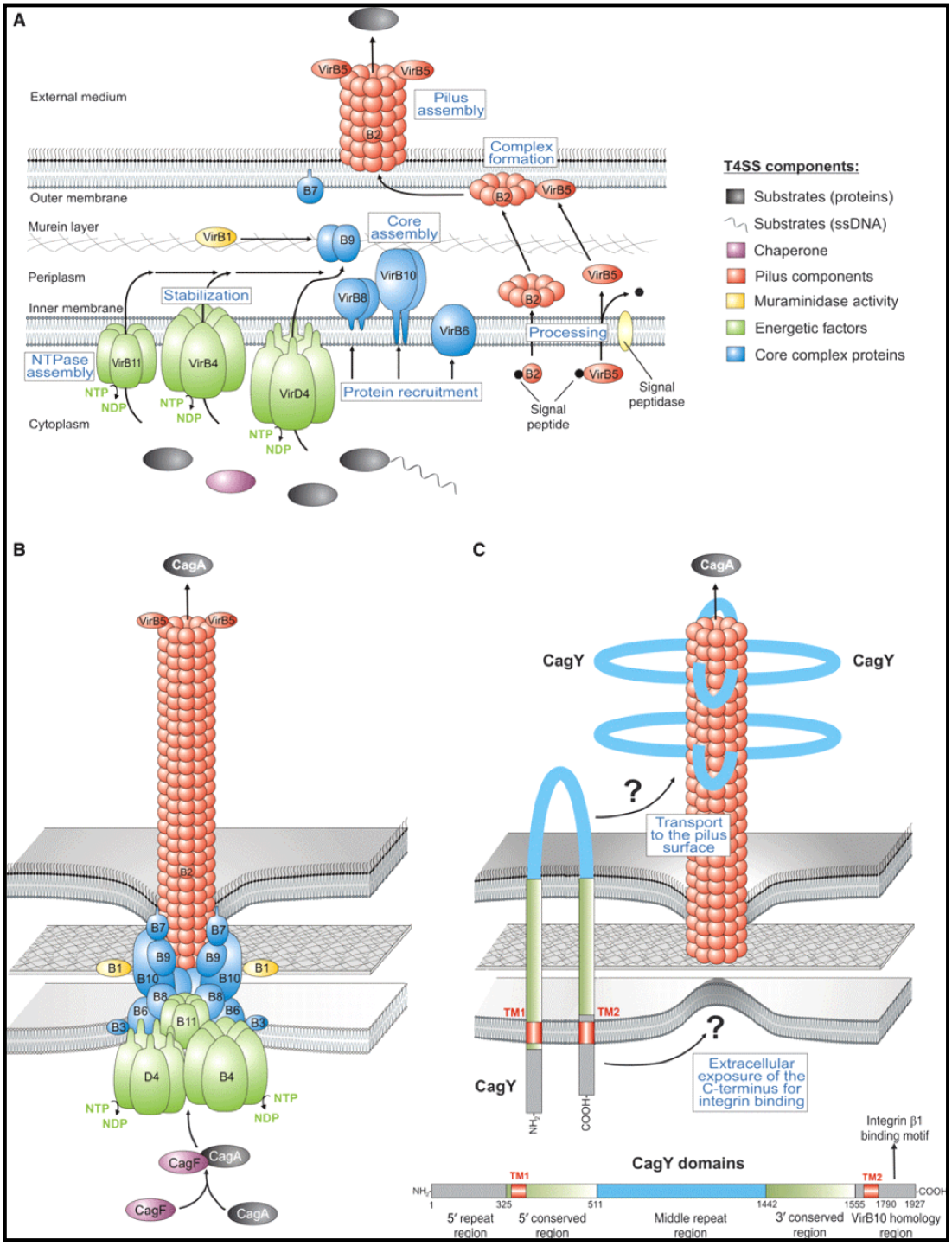


Figure 7. Models for the assembly and assembled structure of T4SS in *A. tumefaciens* and *H. pylori*. (A) The proposed assembly of the prototypical Agrobacterium VirB / VirD4 T4SS machinery is shown. (B) Model-1 for the assembled T4SS machinery in *H. pylori* assuming that all VirB1–11 proteins are encoded by the cagPAI and assemble in a similar fashion as proposed for *A. tumefaciens.* The reported substrates for this T4SS are CagA and peptidoglycan. (C) Model-2 proposes that the T4SS requires essentially the same VirB proteins as in (B), with one major difference. The pilus surface is proposed to be covered with CagY molecules. By contrast to VirB10 in many T4SS, *H. pylori* VirB10 (CagY) is a very large protein carrying two transmembrane domains (TM1 and TM2) to form a hairpin-loop structure in the membranes as depicted (Tegtmeye *et al.,* 2011).

1.4.4. *Lipopolysaccharide*

Many bacterial pathogens use LPS as a powerful signal for the development of an inflammatory response. LPS from *H. pylori* is less proinflammatory than from many other gram-negative species (Khamri *et al.,* 2005). It is comparatively 500-fold less endotoxic than LPSs from *E. coli* or *Salmonella enterica* serovar Typhimurium (Moutiala *et al.,* 1992). LPS from *H. pylori* presents weak ability to stimulate macrophages for the production of proinflammatory cytokines, nitric oxide, and prostaglandins (Peterson *et al.,* 2003). The rough-form LPS of three different *H. pylori* strains have been chemically characterized, showing that Lipid A components are modified with respect to other gram-negative bacteria. In addition, the presence of uncommonly long 3-hydroxy fatty acids (i.e., 3-hydroxyhexadecanoic and 3-hydroxyoctadecanoic) can justify the low biological activity (Molinari *et al.,* 1998). The backbone of Lipid A of *H. pylori* contains a D-glucosamine disaccharide which carries a phosphate group (Moutiala *et al.,* 1992). LPS O-antigens that are commonly expressed by *H. pylori* strains are structurally related to Lewis blood group antigens present on human cells (Monteiro *et al.,* 1998). This may represent a form of molecular mimicry or immune tolerance that allows *H. pylori* LPS antigens to be protected from immune recognition, since there is a similarity to "self" antigens (Algood and Cover, 2006).

1.5. *Structural study of H. pylori*

The genome of *H. pylori* codes for about 1500 ORFs and several complete genomes of different strains are now available. On the other hand, structural data available are relatively limited: 273 files of atomic coordinates are present in the Protein Data Bank (http://www.pdb.org) at August, 2011, and 36 are waiting for release. Among them, 255 structures were solved by X-ray crystallography technique and 18 structures were by solution NMR. Structures present in the protein data bank mainly belongs to five major categories: members of *cag*-PAI, enzymes, proteins interacting with nucleic acids, toxins, unknown function proteins.

The crystal structure of a subdomain of CagA (residues 885–1,005) in complex with microtubule affinity–regulating kinase MARK2 indicates that CagA mimics host substrates. Differences are present between the unbound and the CagA-bound forms of MARK2, and both are structural hallmarks of kinases in their fully activated state. CagA peptide makes MARK2 kinase to its activated conformation, even in the absence of several elements

normally required (no nucleotide or magnesium present, and no phosphorylation of Thr208) (Nešic et. al., 2009). Crystal structures of some TFSS components have been published, including homologues of hexameric NTPases VirB11 (Cagα from *H. pylori*; Yeo *et al.,* 2000; Savvides *et al.,* 2003) and VirD4 (TrwB from *E. coli* R388; Gomis-Ruth and Coll, 2001), the VirB5 protein (TraC from pKM101 plasmid; Yeo *et al.,* 2003), VirB8 (form *B. suis* and *A. tumefaciens*; Terradot *et al.,* 2005, Bailey *et al.,* 2006), and VirB10 (ComB10 from *H. pylori*; Terradot *et al.,* 2005). Till now, with the exception of the ATPase Cagα VirB11 homologue, high-resolution structure data of Cag proteins are available only for CagD, CagS and CagZ. The crystal structure of CagD, determined in two different crystal forms, shows that the protein is a covalent dimer in which each monomer folds as a single domain that is composed of five β-strands and three α-helices. No significant functional clues comes from the structure (Cendron *et al.,* 2009). The same happens for CagS, which do not show any clear evidence of architectural similarity with other known structures (Cendron *et al.,* 2007). CagZ consists of a single compact L-shaped domain, composed of seven α-helices including about 70% of the total residues. Three-dimensional homology searches did not reveal structural homologues, and CagZ can be considered representative of a new protein fold. (Cendron *et al.,* 2004).

The crystal structures of many enzymes of *H. pylori* are known. Most of them are considered possible pharmacological targets and, for some, inhibitors have already been developed. Enzymes, which structure are deposited in PDB are belonging to pathways of lipid metabolism, shikimate pathway, antioxidant enzymes, sugar metabolism, cell wall biosynthesis, vitamin precursor biosynthesis, NAD biosynthesis, peptide synthesis and degradation pathway, urea metabolism, other enzymes (spermidine synthase, ATP dependent protease, ferritin, flavodoxin etc.).

The structure of several proteins interacting with DNA and regulators of gene expression are available for *H. pylori*. Among them are single-stranded DNA-binding protein (PDB CODE: 2VW9), IS608 transposase (TnpA) (PDB CODE: 2VIC), DNA glycosylases (PDB CODE: 1PU7), ArsR DNA binding domain (PDB CODE: 2K4J). Other enzymes regulate gene expression indirectly, without interacting with nucleic acids (HobA, PDB CODE: 2UVP; HP-RR, PDB CODE: 2PLN; Luxs, PDB CODE: 1J6X; apo-NiKR, PDB CODE: 2CA9). Structure of toxins and other *H. pylori* proteins directly involved in pathogenicity (Vacuolating toxin p55 domain, PDB CODE: 2QV3; TIP-alpha N34, PDB CODE: 2WCQ; NAP-A, PDB CODE: 1JI4; MotB, PDB CODE: 3CYP; HpaA, PDB CODE: 3BGH; FIC,

PDB CODE: 2F6S etc.) were deposited along with more than 20 proteins of which no predicted function.

1.6. *Therapeutical consideration against Helicobacter pylori*

The most effective treatment for *H. pylori* infection so far is antibiotic therapy. The incidence of *H. pylori* infection has been decreasing in industrialized countries due to improved sanitation, smaller family sizes and decreased overcrowding (Parsonnet, 1995). However, the prevalence of resistance against the most frequently used antibiotics is rising, especially in developing countries, where most antibiotics can be obtained without special prescription (Frenck and Clemens, 2003). It is estimated that in Europe approximately 10% of *H. pylori* strains are resistant against clarithromycin and about 30% against metronidazole (these numbers increase in developing countries, where the resistance varies between 25 to 50% for clarithromycin and to near complete resistance for metronidazole (Glupczynski *et al.,* 2001; Meyer *et al.,* 2002; Torres *et al.,* 2001). Therefore, new strategies to treat *H. pylori* infections are necessary and desirable. As *H. pylori* uses a metal-cofactored enzyme as a first line defense to colonize the human stomach, the metal metabolism is a potential drug target.

Vaccination. The first-line therapy for *H. pylori* 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 antibiotics and proton pump inhibitors in most patients can successfully eradicate *H. pylori*, increasing antibiotic resistance in the bacterium remains a serious problem (D'Elios and Andersen 2007). Therefore, there still exists a strong rationale for development of effective vaccines against *H. pylori*. It has been suggested that a 10-year vaccination program in the USA starting in 2010 could reduce *H. pylori* prevalence to 0.07% by the end of the 21st century (Rupnow *et al.,* 2001). In contrast, in developing countries due to poor hygiene and overcrowded living conditions a continuous vaccination program would be required to get similar effects. Thus, vaccination could be a cost-effective way to control *H. pylori* infection. The safety and immunogenicity of a number of *H. pylori* vaccines such as recombinant urease (rUrease), killed whole cells, and live vectors expressing *H. pylori* antigens have been examined in a few clinical trials (Table 1.1) (Kabir, 2007).

Table 1.1. Vaccine trials against *H. pylori* in humans (adapted from Kabir, 2007).

Antigen	Adjuvant/ delivery system	No. of subjects	*H. pylori* status	Route	Results
Urease	None	12	Asymptomatic Infection	Oral	No immunogenicity
Urease	LT	26	Asymptomatic	Oral	Immunogenicity to urease, reduction in *H. pylori*
Urease	LT	42	Uninfected	Oral	Immunogenicity to urease
Killed whole cells	Mutant LT	41	Uninfected and infected	Oral	Immunogenic, did not eradicate *H. pylori*
Killed whole	Mutant LT	5	Uninfected	Oral	Immunogenic gastric B Cell response detected
Urease	LT	18	Uninfected	Rectal	Poor immunogenicity to urease
Urease	Attenuated *S. typhi*	8	Uninfected	Oral	No immune response to urease
Urease	Attenuated *S. enterica* serovar *typhimurium*	6	Uninfected	Oral	Weak immune response to urease
Urease	Typhoid vaccine Ty21a T	12	Uninfected	Oral	No antibody to urease, Weak T-cell responses
CagA,VacA, NAP	Aluminum hydroxide	-	Uninfected	Parenteral	Antibody response detected

So far, it is unknown why vaccines are quite effective in a rodent model but not in humans. An explanation might be that *H. pylori* is able to alter the human immune response against itself in humans, but these mechanisms might not be as effective in its non-natural host (Del Giudice and Michetti, 2004).

Inhibition of Enzymes. The availability of whole genome sequences, transcriptomics and proteomics made it possible to identify genes unique for *H. pylori*. Subsequent colonizing or growth experiment mutants indicate essential genes. With further analysis of the regulation and mechanism of action, it is feasible to identify possible targets for inhibitors, which can in turn alter expression of essential genes or block enzyme activity. Possible targets for inhibitors are the hydrogenase and the urease system of *H. pylori*. Both nickel-cofactored systems are lacking in the human metabolism, but are present in the periplasm (hydrogenase and UreI) or cytoplasm (urease) of *H. pylori*. Mutations in both systems were shown to be at least growth deficient (Mobley, 2001; Olson *et al.,* 2001). Both enzymes have unique active centers and a complicated mechanism of activation by assembly. Therefore, the design and testing of therapeutical inhibitors of the hydrogenase and urease active enzymes or its assembly function may be of interest, especially in regard of the increasing resistance of *H. pylori* against antibiotics (Maier, 2003; Maier *et al.,* 1996; Mobley *et al.,* 1995). *H. pylori* has a strong need in nickel ions, since the latter is used as a cofactor for urease. On the other hand, nickel ions are not essential for human metabolism (Denkhaus and Salnikow, 2002). However, several human food sources like coffee, tea, nuts and chocolate are rich in nickel

ions (Sunderman, 1993). As nickel depletion or probably even a nickel shortage is detrimental for *H. pylori* growth (van Vliet *et al.,* 2001a, 2002b), a nickel-deficient diet might help to treat *H. pylori* infection.

We have studied three-dimensional structure of *H. pylori* proteins and characterize their function which are important for stomach colonization and pathogenesis. In particular, we concentrate our efforts on cell wall modification enzymes (peptidoglycan deacetylase), LPS biosynthesis (ADP-L-*glycero*-D-*manno*-heptose-6-epimerase, rfaD), periplasmic substrate binding protein of ABC transporter (ceuE), key enzymes in nitrogen assimilation pathway (Glutamine synthetase) and secreted immunogenic disulfide isomerase (DsbG). In the chapter 2 the three dimensional structure and the enzymatic activity of a putative peptidoglycan deacetylase is described. The crystal structure of the last enzyme of the biosynthesis pathway of core oligosaccharides (L, D-heptose) of the LPS is discussed in details in the chapter 3. The crystal structure of the periplasmic ABC transporter substrate-binding protein is discussed in chapter 4. The crystal structure of the key enzyme in the unique nitrogen assimilation pathway glutamine synthetase is discussed in chapter 5, and in the chapter 6 the cloning, expression and characterization of secreted immunogenic protein DsbG is discussed.

Chapter Two

The Structure of *Helicobacter pylori* HP0310 Reveals an Atypical Peptidoglycan Deacetylase

This chapter is adapted from

Shaik MM, Cendron L, Percudani R, Zanotti G, 2011 The Structure of *Helicobacter pylori* HP0310 Reveals an Atypical Peptidoglycan Deacetylase. PLoS ONE 6(4): e19207. doi:10.1371/journal.pone.001920.

The Structure of *Helicobacter pylori* HP0310 Reveals an Atypical Peptidoglycan Deacetylase

Md Munan Shaik, Laura Cendron, Riccardo Percudani, Giuseppe Zanotti

PLoSONE 6(4): e19207. doi:10.1371/journal.pone.0019207

Abstract

Peptidoglycan deacetlyase (HP0310, HpPgdA) from the gram-negative pathogen *Helicobacter pylori*, has been indicated as the enzyme responsible for a peptidoglycan modification that counteracts the host immune response. HpPgdA has been cloned, purified and expressed in good yield in *E. coli*. It has been crystallized, its structure determined and activity tests *in vitro* performed. The enzyme, which belongs to the polysaccharide deacetylases protein family, is a homo-tetramer. The four polypeptide chains, each folded into a single domain characterized by a non-canonical TIM-barrel fold, are arranged around a four-fold symmetry axis. The active site, one per monomer, contains a heavy ion coordinated in a way similar to other deacetylases. However, the enzyme showed no in vitro activity on the typical polysaccharide substrates of peptidoglycan deacetylases. In striking contrast with the known peptidoglycan deacetylases, HpPgdA does not exhibit a solvent-accessible polysaccharide binding groove, suggesting that the enzyme binds a small molecule at the active site.

Introduction

Peptidoglycan, which is one of the constituents of the protective barrier of gram-negative bacteria, at the same time represents one of the main targets of the innate immune system of the host (Boneca *et al.,* 2007). Some peptidoglycan components are recognized by specific receptors of human cells, Nod1 and Nod2, which initiate the immune response by activating the NF-kB pathway (Strober *et al.,* 2006; Meylan *et al.,* 2006). In order to escape the recognition by the host receptors, some bacteria have developed a mechanism of chemical modification of peptidoglycan, in particular through N-deacetylation (Boneca *et al.,* 2007). In *Helicobacter pylori*, the gram-negative pathogen that affects about half of the human population, this task is accomplished by a peptidoglycan deactelyase coded by gene *hp0310* (Wang *et al.,* 2010; Wang *et al.,* 2009). It was in fact demonstrated that the HP0310 protein was over-expressed under strong oxidative stress conditions and that peptidoglycan of a HP0310 knock-out mutant presented larger acetylation compared to wild-type, along with increased susceptibility to lysozyme degradation. Moreover, expression of HP0310 was induced when *H. pylori* was held in contact with macrophages and a significant increase of IL-10 and TNF-a was observed in mutant-infected mice compared to wild-type.

According to all existing evidence, HP0310 was recognized as a *H. pylori* peptidoglycan deacetylase (HpPgdA). HpPgdA is a soluble protein of 293 amino acids, which has limited sequence similarity with enzymes of the polysaccharide deacetylase family (pfam01522). Among the proteins of the family that have been well characterized is the peptidoglycan deacetylase from *Streptococcus pneumoniae* (SpPgdA), a Zn-dependent enzyme (Blair *et al.,* 2005). It is significantly longer than HpPgdA (463 amino acids) and it shares less than 19% identity with our protein. Moreover, the *H. pylori* enzyme apparently lacks the conserved metal-binding motif (Wang *et al.,* 2009). Nevertheless, reliable HpPgdA homologues can be found in several pathogenic bacteria. For the above reasons, HpPgdA was proposed to be the prototype of a new family of polysaccharide deacetylases (Wang *et al.,* 2009). In addition, in a proteomic analysis of a *H. pylori* strain adapted *in vivo* to induce gastric adenocarcinoma in rodent models (7.13 strain), the loss of a functional HpPgdA resulted in higher levels of CagA translocation compared to the non-carcinogenic strain. This observation suggests that peptidoglycan alteration by HpPdgA could affect the composition of substrates translocated by the *cag* secretion system (Franco *et al.,* 2009).

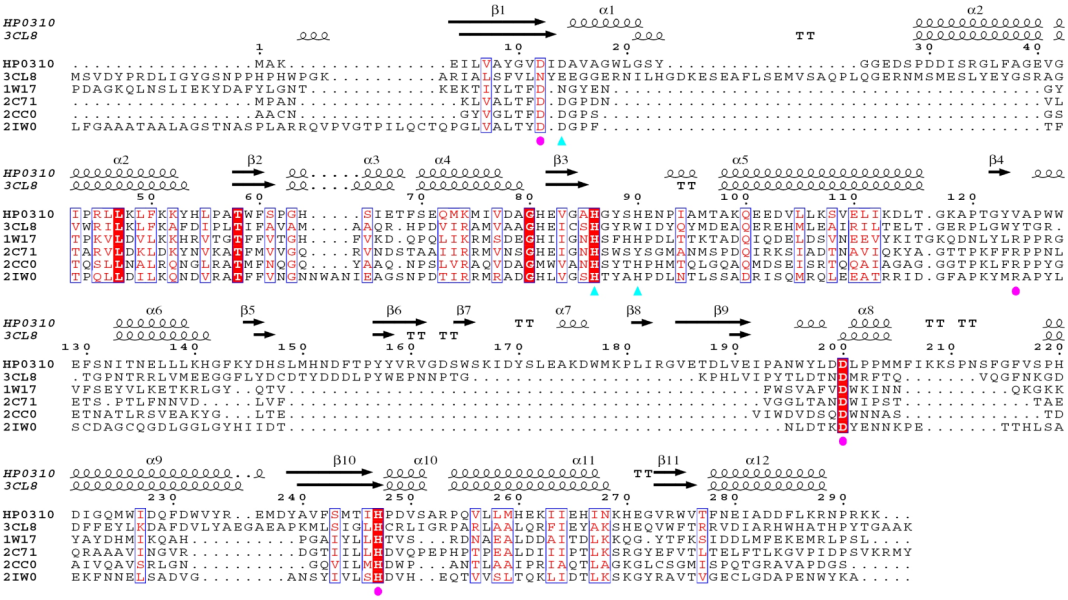
If a deacetylation activity has been already established for the protein, the specific substrate of HpPgdA has yet to be identified. It is also possible that HpPgdA is active on substrates other than peptidoglycan, for example lipopolysaccharides or membrane anchored proteins. In order to elucidate the structural features of the prototype of a new family of deacetylase and to establish a possible enzymatic mechanism, peptidoglycan deacetlyase (HP0310) from *H. pylori* was cloned, expressed and purified in good yield in *E. coli*. The enzyme was crystallized, its structure determined and *in vitro* activity tests were performed.

Results

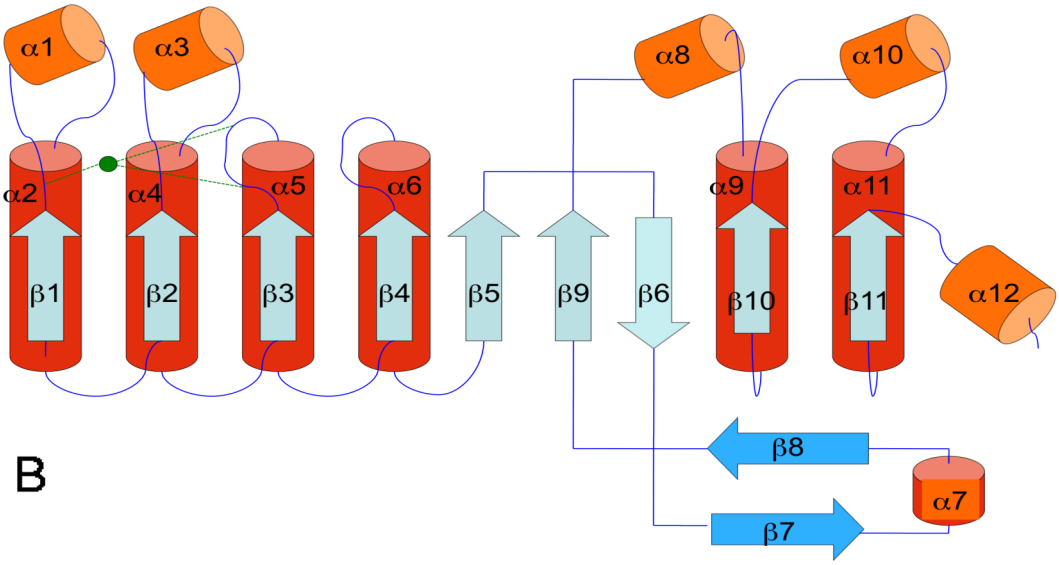
The overall model

The crystal structure, despite the relatively limited resolution, 2.57 Å, appears quite reliable: the electron density for the protein main chain atoms is clearly visible from residue 2 to 291, without gaps and with a good stereochemistry. Only the N-terminal methionine and two residues at the C-terminus are missing in the crystallographic model.

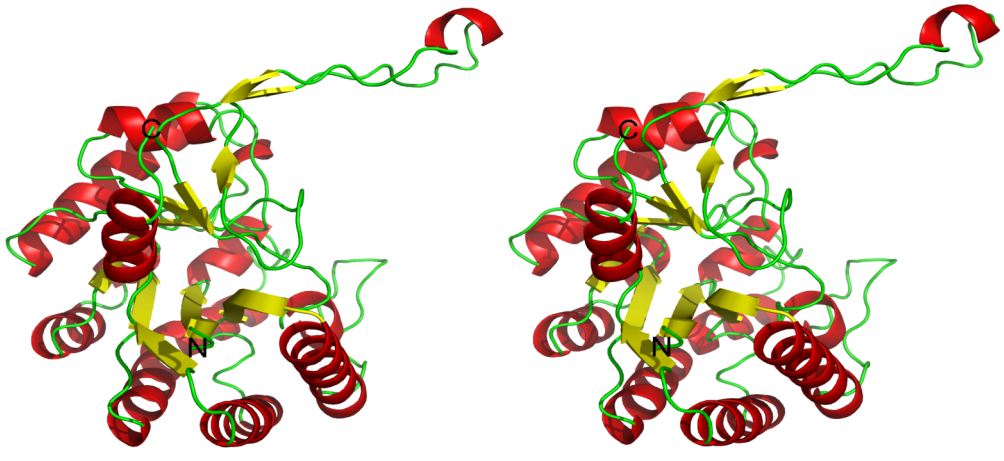
The secondary structure elements of the monomer are illustrated in Fig. 1A, and their topology and tertiary structure organization in Fig. 1B and 1C. The polypeptide folds into a single domain, characterized by a non-canonical TIM-barrel fold (Banner *et al.,* 1976): nine ß-strands, labeled from 1 to 6 and from 9 to 11, are arranged in a central barrel surrounded by six α-helices (labeled α-2, 4, 5, 6, 9, 11). Helices α -1, α-3, α-7, α-8, α-10 α-12 surround the barrel on the two sides. The topology (Fig. 1B) can be compared with those of peptidoglycan deacetylase from *S. pneumoniae* (SpPgdA, see Fig.2A of paper Blair *et al.,* 2005) or from *B. subtilis* (PdaA, Blair and van Aalten, 2004). The more significant difference from the other members of the family is represented by a long arm of about 25 residues, from 161 to 185, which extends away from the molecular core and embraces a nearby enzyme monomer, strengthening the quaternary organization of the protein.



A


α1
α3
α8
α10
α2
α4
α5
α6
α9
α11
β1
β2
β3
β4
β5
β9
β6
β10
β11
α12
β8
α7
β7

B


C
N
C
N

C

Figure 1. Topology and structure of HpPgdA monomer.

A) Amino acid sequence of HP3010 protein aligned with PuuE allantoinase from *P. fluorescens* (Protein Data Bank code 3CL8), peptidoglycan deacetylase from B. subtilis (Protein Data Bank code 1W17), acetyl xylan esterases from *C. thermocellum* (Portein Data Bank code 2C71) and from S. lividans (Protein Data Bank code 2CC0), and chitin deacetylase from *C. lindemuthianum* (Protein Data Bank code 2IW0). Secondary structure elements deriving from Protein Data Bank coordinates are drawn over the alignment. Residues previously reported to be involved in metal-binding or in catalysis are denoted by cyan arrowheads (metal-binding) and magenta circles (catalysis).

B) Topology diagram of HpPgdA. Helices and strands of the modified TIM barrel are in red and cyan, respectively; helices and strands surrounding thebarrel in orange and light blue, respectively. The location of heavy ion is indicated by a green filled circle.

C) Stereo view of a Cartoon of HpPgdA monomer. a-helices are in red, b-strands in yellow, others in green. N and C-terminus are labeled. (This and the following drawing were done using the Pymol program (Delano, 2008).

Quaternary structure

Four monomers are present in the asymmetric unit, arranged around a four-fold rotation axis (Fig. 2). The enzyme in solution is also a tetramer, as confirmed by gel-filtration experiments (Supplementary Figure S2). The four monomers do not present significant conformational differences, and in fact non-crystallographic restraints were applied through all the refinement cycles. The root mean square deviation (r.m.s.d.) between equivalent main chain atoms of different monomers is in the range 0.27-0.28 Å. Interactions among monomers are both hydrophobic and hydrophilic: there are 29 hydrogen-bond interactions between two adjacent monomers, along with several hydrophobic contacts. The surface buried after tetramer formation is 11,420 $Å^2$, which represents nearly one third of the sum of the area of the four isolated monomers, 36,858 $Å^2$. The most relevant stabilization of the quaternary organization of the protein is the long arm that protrudes from one monomer to embrace the other.

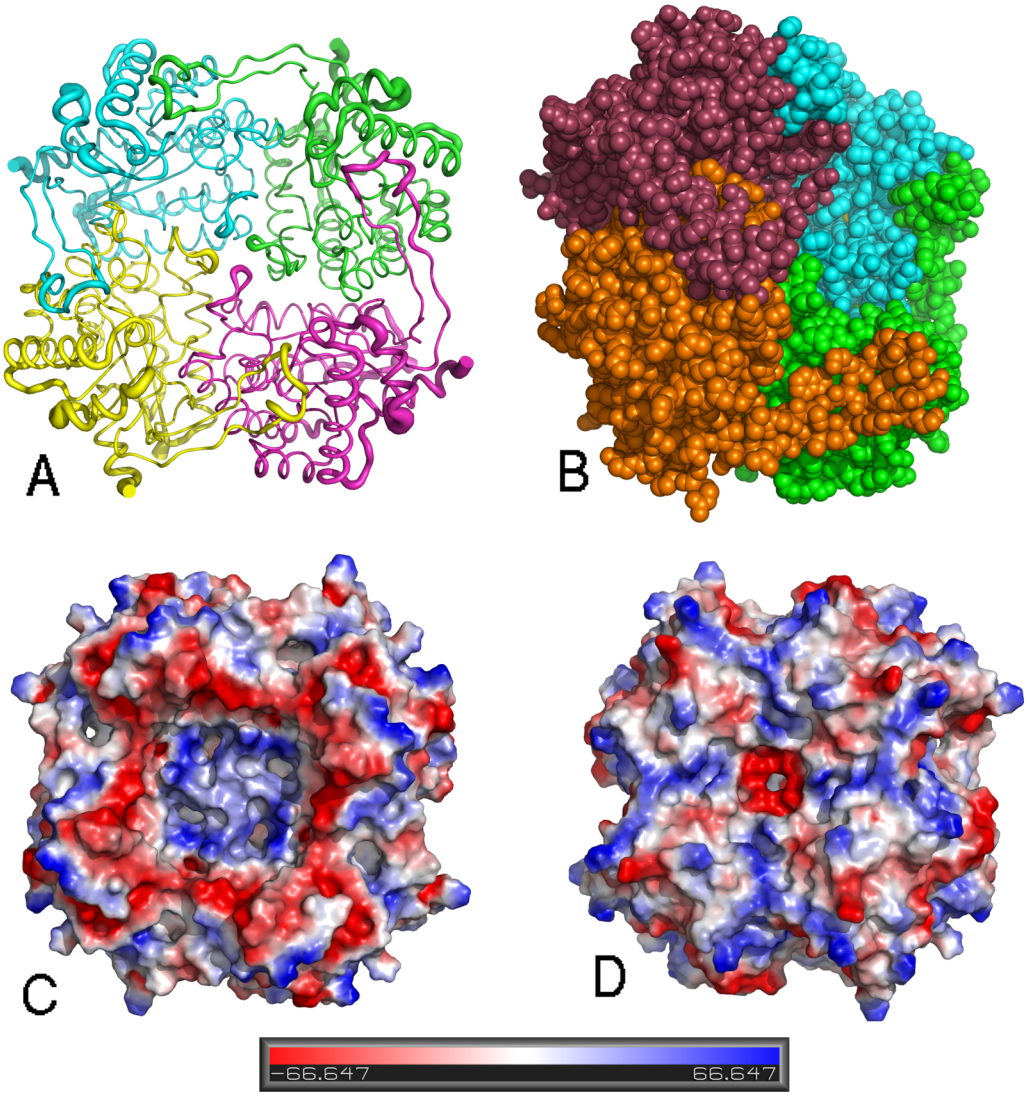


Figure 2. The HpPgdA tetramer.

A) Ribbon drawing of the tetrameric assembly of HpPgdA. The view is approximately perpendicular to the molecular four-fold axis. The diameter of the ribbon tube is proportional to the thermal parameters of the atoms.

B) van der Waals model of HpPgdA tetramer. Each monomer is colored differently, so that it is possible to se the long arm protruding from the core of one monomer to embrace the nearby one in the tetramer. The orientation is slightly rotated with respect to Fig. 2A, so that one heavy ion (yellow) is partially visible in one subunit.

C) And D) Qualitative electrostatic potential surface of HpPgdA tetramer. In the left picture the orientation is similar to that of Fig. 2A and represents the side from which the active site is partially accessible. The ion is totally concealed under the surface. The picture on the right shows the opposite face of the tetramer.

Structure comparison

An exhaustive search with DALI server (Holm, 1993) shows that the two three-dimensional structures most similar to HpPgdA are a *bona fide* allantoinase (Ramazzina *et al.,* 2008) from *Pseudomonas aeruginosa* (PDB ID 1z7a, Z-score 30.5, r.m.s.d. from our structure for 300 amino acids 1.9 Å) and PuuE allantoinase from *Pseudomonas fluorescens* (PDB ID 3cl8 (Ramazzina *et al.,* 2008), Z-score 30.3, r.m.s.d. for 300 superimposed amino acids 2.0 Å). The latter was in fact used as a template to build the model for the molecular replacement. The most relevant differences from our structure are represented by the long arm that in HpPgdA protrudes from the monomer core, by a loop region (205-217) connecting helices α-7 to α-8, which is definitely longer in HpPgdA, and by the loop connecting helices α-1 to α-2, which in HpPgdA includes only five residues, whilst in PuuE is about 27 residues long. It must be considered that the latter area in PuuE is close to the active site and some of the residues may be involved in substrate recognition.

Some structural similarities are also observed with the catalytic domain of 4- α-glucanotransferase (PDB ID 1k1x, Luthy *et al.,* 2002) and α-amilase (PDB ID 2b5d, Dickmanns *et al.,* 2006). Members of the same enzymatic family, such as SpPgdA from S. *pneumoniae* (PDB ID 2c1g, Blair *et al.,* 2005) and PdaA from *B. subtilis* (PDB ID 1w17, Blair and van Aalten, 2004) are structurally less similar, consistent with a lower sequence identity.

The active site

The putative active site of *H. pylori* peptidoglycan deacetylase is represented by a heavy ion coordinated to the N□ nitrogen atoms of His 86 and 90 and to the two oxygen atoms of Asp 14. A fourth coordination site is occupied by a water molecule. The latter is held in place by H-bonds with His 247, Asp 12 and 14 (Fig. 3). Since a similar situation has been observed in two other mono-nuclear Zn-dependent deacetylases, polysaccharide deacetylase from *B. subtilis* (PdaA) (Blair and van Aalten, 2004) and peptidoglycan GlcNac deacetylase from *S. pneumoniae* (SpPgdA) (Blair *et al.,* 2005), the ion was assumed to be a Zn. In both cases, the main difference from our structure is represented by the accessibility to the active site: in HpPgdA the presence of helix α1, the longest helix α2 and the fact that the loop connecting helices α7 to α8 protrudes towards the active site opening, makes it poorly accessible. Moreover, some bulky side chains, in particular Trp 127 and 196, Tyr 23 and Leu 201 obstruct the access to the ion (Fig. 2B). An anomalous difference electron density map was

generated, revealing a peak that was modeled as a zinc ion coordinated octahedrally by Asp14, His90, His86, and a water molecule.

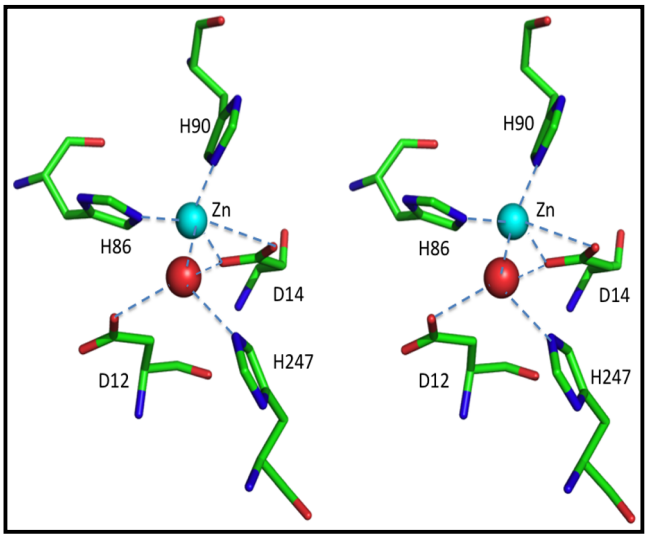


Figure 3. HpPgdA active site. The Zn ion (cyan) is coordinated by Nε of His 90 (distance 2.06 Å), Nε of His 86 (2.12 Å), Oδ1 and Oδ2 of Asp 14 (2.21 Å and 2.98 Å), a water molecule (red sphere, 2.08 Å). Distances between the latter and Oδ2 of Asp 12, Nε of His 247 and Oδ1 of Asp 14 are 2.68 Å, 2.60 Å and 2.99 Å, respectively.

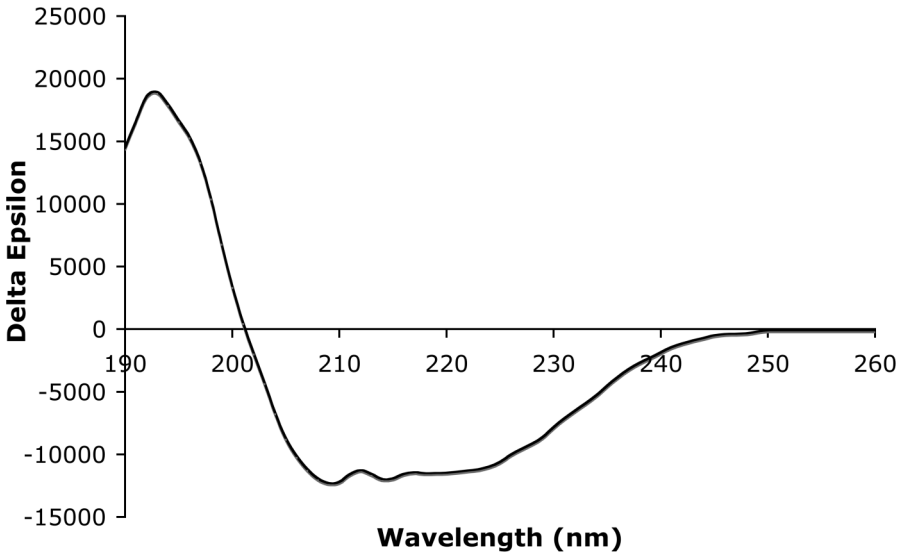


Figure 4. Circular dichroism spectrum of HP0310. Spectrum was measured with J-715 spectropolarimeter (JASCO, Corporation) at 298 K. 10 runs were accumulated with 1 mg/ml protein (Tris 3 mM, NaCl 10 mM, pH 8.0) in a 0.2 mm path cuvette in the wavelength interval 190–260 nm. The CD spectrum was rescaled with respect to a standard solution containing the buffer and the molar ellipticity calculated. The CD spectrum was deconvoluted with software ‘‘CD Spectra Deconvolution’’ (CDNN 2.1) and the secondary structure predicted (36.6% a-helix, 15.6% b-sheet and 47.6% other).

Activity tests

In order to address the substrate specificity of HpPgdA, single components of peptidoglycan were tested in solution, in particular N-acetyl glucosamine (GlcNAc), and chitotriose GlcNAc3. All the activity tests were performed using the standard conditions used for other polysaccharide deacetylases. In order to check that the protein we have purified is properly folded in solution, a circular dichroism spectrum, showing the expected content of □-helix and □-strands, was performed (Fig. 4). The gel-filtration experiment shows that the protein elutes as a single tetrameric species (see Figure S2 of Supplementary Material). Since no activity was detected, other acetylated amines possibly present in *H. pylori* were tested: N-acetyl Putrescine, N-acetyl Spermidine, N-acetyl cadaverine, and some N-acetyl dipeptides (Ac-D-Ala-D-Ala-OH, Ac-D-Ala-D-Ala-OCH3, Ac-D-Ala-L-Ala-OH, Ac-D-Ala-L-Ala-OCH3). These substrates were suggested by comparative gene cluster analyses showing that genes homologous to HpPgdA are physically associated with genes encoding polyamine importers and genes encoding D-Ala-D-Ala peptidases (Supplementary Figure S1). Peptide chains containing D-amino acids are ubiquitous peptidoglycan components, while polyamines have been reported to be present in the peptidoglycan of some bacteria (Kamio and Nakamura,1987). Finally, allantoin was also tested, given the close structural similarity of HpPgdA to PuuE allantoinase. The enzyme, however, was found inactive for all of these compounds at all pH tested (6, 7, and 7.5)

Phylogenetic relationship with other polysaccharide deacetylases

Polysaccharide deacetylase is a very large and functionally variegated protein family to which HpPgdA belongs. The structure of several members of the family is known from structural genomics and dedicated studies. We selected representative members of the polysaccharide deacetylase family with known structure (see Fig. 1A) for a phylogenetic comparison with HpPgdA (Fig 4). Most members of the family act on polysaccharide substrates and are involved in cell wall modifications such as peptidoglycan deacetylases, chitin deacetylases, xylan esterases. A common feature of polysaccharide deacetylases that recognize oligosaccharide substrates is the presence of a large groove at the top of the barrel, with the residues involved in metal binding or catalysis typically exposed at the surface of the active site (Fig. 5). Consistent with the previously reported structural similarities, the phylogenetic analysis shows that HpPgdA has a distant relationship with proteins that bind polysaccharides and a closer relationship with PuuE allantoinase. This member of the polysaccharide deacetylase family has a considerably reduced binding cavity and recognizes

a small substrate molecule (Ramazzina *et al.,* 2008). Similarly to PuuE, additional structural elements located at the top of the barrel in HpPgdA obstruct the access to the active site (Fig. 5). At variance with PuuE, however, HpPgdA has the ability to bind a metal ion at the active site, as observed in most members of the family (see Fig 1A).

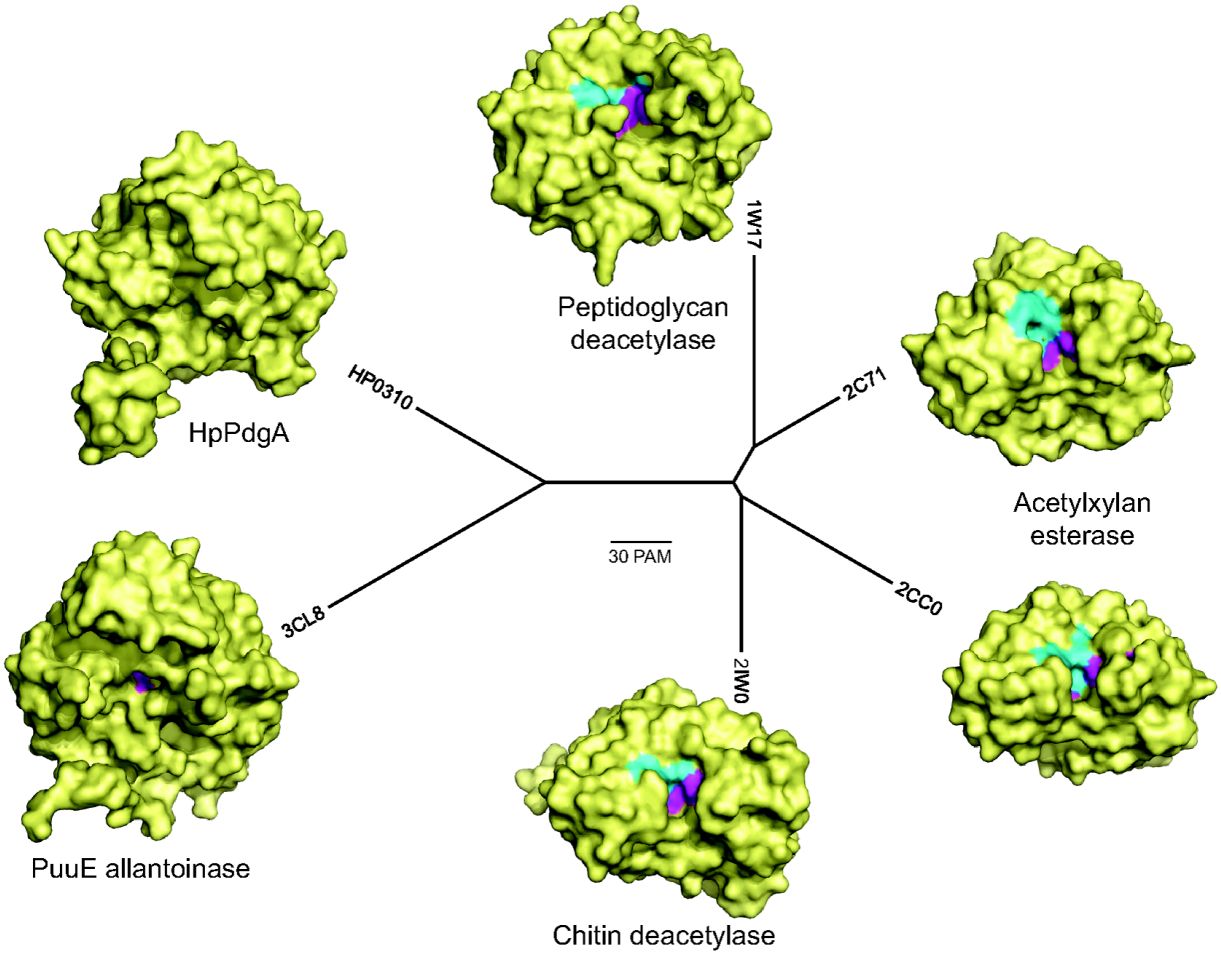


Figure 5. Phylogenetic and structural comparison of the polysaccharide deacetylase family. Unrooted tree showing the phylogenetic relationship of HpPgdA with other polysaccharide deacetylase proteins. Branches are proportional to genetic distance as indicated by the scale bar representing point-accepted mutations (PAM). Proteins are indicated by PDB codes. Monomeric structures are shown in the same orientation (from the top of the barrel) as surfaces, with solvent-accessible residues involved in metal-binding or catalysis (see Fig. 1A) highlighted in cyan and magenta, respectively.

Discussion

HP0310 has been demonstrated to be a peptidoglycan deacetylase in a test performed on a crude extract of peptidoglycan from *H. pylori* (Wang *et al.,* 2009), despite the activity on a specific peptidoglycan component was not identified. Subsequent experiments on a PgdA mutant (Wang *et al.,* 2010; Wang *et al.,* 2009) have supported this finding, though indirectly. The three-dimensional structure of the enzyme is compatible with this function: the putative active site of HpPgdA contains an ion, possibly Zn^{2+}, coordinated in the same way as chitin deacetylase from *C. lindemuthianum* and similar to that of acetyl xylan esterases from *C. thermocellum* and *S. lividans*. Residues involved in catalysis are also conserved (Fig. 1A). The main difference of our enzyme from other members of the family is the accessibility to the active site, that fully justifies the absence of enzymatic activity on large substrates. The presence in HP0310 of two additional helices, α-1 and α-2, and its tetrameric assembly, in particular the long arm protruding from a nearby subunit, obstruct the access to the active site. Other enzymes of the family are monomeric or even dimeric, but in the latter case the active site is far from the dimerization side, as in the *S. lividans* xylan esterase. In addition, some bulky residues, in particular Trp 127, interpose between the ion bound and the solvent. This suggests that the access of the substrate to the active site must be accompanied by at least a movement of the side chain of Trp127 or possibly by a conformational rearrangement of the loop containing residues Trp 127 and 128. It is unlikely, however, that these structural rearrangements can generate an active site cavity large enough to accommodate a large polysaccharide. On the other hand, typical polysaccharide deacetylases do not show activity with monomeric substrates. For example, Chitin deacetylases need at least a chito-oligosaccharide trimer for activity but prefer longer multimers (Tokuyasu *et al.,* 2000; Hekmat *et al.,* 2oo3).

The lack of enzymatic activity on isolated N-acetyl polysaccharides we observed for HpPgdA and the reported deacetylase activity on crude peptidoglycan extracts from *H. pylori* (Wang *et al.,* 2009) could be reconciled if HpPgdA recognizes a peptidoglycan component that is different from the polysaccharide moiety. From the structure analysis, it appears more likely that HpPgdA binds a small molecule at the active site, though we cannot exclude that a significant conformational modification in the ternary and quaternary structure of the enzyme can take place in order to accommodate a large substrate. An enzyme with a similar tetrameric arrangement and a close phylogenetic relation with HpPgdA is PuuE allantoinase, whose natural substrate is in fact a small cyclic imide.

Another feature that distinguishes HpPgdA and its close homologues from classical peptidoglycan deacetylases is the lack of a signal peptide for secretion in the periplasmic space, a characteristic that is expected for an enzyme exclusively acting on cell-wall peptidoglycan. In proteomic analysis, HpPgdA has been found moderately enriched (1.8 ratio) in the extracellular fraction compared to the soluble cell-associated samples (Smith *et al.,* 2007) suggesting an additional cytoplasmic role for the enzyme.

Conclusions

Bacterial peptidoglycan is a complex molecule that undergoes significant processing (Patti *et al.,* 2008). Many fine details of peptidoglycan modifications and of the enzymes active in editing the cell-wall structure are still poorly understood. We have presented evidence that the HpPgdA substrates are not the N-acetylated polysaccharides recognized by the typical peptidoglycan deacetylases. The determination of the HpPgdA structure provides valuable information for the precise identification of its role in peptidoglycan metabolism and for a deeper understanding of those cell-wall modifications that affect the ability of bacteria to evade the hostimmune system response.

Materials and Methods

Cloning, expression, purification

The HP0310gene was amplified by PCR from genomic *H. pylori* G27, using the primers 5'-CACCATGGCAAAAGAAATTTTAGTGGC-3'(forward, Topoisomerase recognition site underlined) and 5'-TTACTATTTTTTTCTAGGGTTTCGTTTTAAG-3' (reverse). It was 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-HP0310 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 30 mM Tris pH 8.0, 150 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 1 ml of Ni^{2+} charged Chelating Sepharose™ (GE Healthcare). After extensive washing using the lysis buffer, supplemented with 20mM imidazole, the protein was eluted from column by linear gradient of 80 to 300mM imidazole. The protein was further purified by Superdex 200™ 10/300 GL (GE Healthcare), equilibrated

with 30 mM Tris pH 8. 0, 150mM NaCl. The His tagged-Hp0310 was eluted as a single peak, roughly corresponding to a tetramer and migrated as a single 37.5 kDa species on SDS-PAGE (theoretical mass: 33.7 Kda plus the 2.8 Kda His-tag). To check the quality of the purified recombinant protein DLS and CD were performed (Figure 4).

Crystallization and structure determination

The purified protein was concentrated to 20 mg/ml and used for crystallization trials, partially automated using an Oryx 8 crystallization robot (Douglas Instruments). The best crystals were obtained at 20°C by vapor diffusion technique using a 18 mg/ml protein stock solution and, as precipitant, a solution containing 0.2 M Ammonium Sulfate, 0.1M Tri Sodium citrate pH 5.6, 15% (w/v) PEG 4000 (PEG II crystal screening solution no. 56 from Molecular Dimension, UK).

Crystals could be processed as orthorhombic, space group $C222_1$, with **a**=154.56 Å, **b**=154.73.44 Å, **c**=158.68 Å. Since **a** and **b** cell parameters are very similar, data were also processed in a tetragonal crystal system, point groups P4 or P422, but in both cases the merging R factor was significantly higher (around 0.30), and the orthorhombic system was selected. The identity between two cell parameters can also induce merohedral twinning in the crystal, according to the law (*k,h,-l*). In the crystal used for structure determination and refinement the twinning fraction was estimated to be 0.19. One tetramer is present in the asymmetric unit, with VM= 4.13 $Å^3$/Da corresponding to a solvent content of 70%. The high solvent content justifies the maximum resolution of 2.57 Å. The data set used in the final refinement was measured at the beamline BM30 of European Synchrotron Radiation Facility, Grenoble, France. It was indexed and integrated with Mosflm software (Leslie, 2006) and merged and scaled with Scala (Evans, 2006), contained in the CCP4 crystallographic package (CCP4, 1994). The structure was solved by molecular replacement using Phaser software (McCoy *et al.,* 2007) starting from a tetrameric model built by the SWISS-PDB server using as a template the crystal structure of PuuE allantoinase (PDB ID code 3CL6, Ramazzina *et al.,* 2008). Refinement was carried on using the simulated annealing procedure contained in CNS (Brunger *et al.,* 1998) in the first stages of refinement and Refmac (Murshudov *et al.,* 1997) and Phenix (Adams *et al.,* 2010) in the subsequent steps. Several steps of manual rebuilding, performed with Coot graphic software (Emsley and Cowtan, 2004), were necessary in order to reach the final structure, which significantly diverges from the starting model. Solvent molecules were added with the automated procedure of Phenix. Since the first

cycles of refinement a maximum of about 14σ was clearly visible in the electron density map. Given the presence of two histidines and one aspartate residues coordinating it, it was interpreted as a Zn ion.

The final model contains 9429 protein atoms, four Zn atoms and 860 water molecules. The crystallographic R factor is 0.2056 (R_{free} 0.2350). Statistics on data collection and refinement are reported in Table 1. Geometrical parameters of the model, checked with Procheck (Laskowski *et al.,* 1993), are as expected or better for this resolution. Buried surface calculations were performed using the Areaimol program (CCP4, 1994).

Bioinformatics

Search and comparison of genetic clusters containing HP0310 and homologous gene were conducted using the Microbesonline web server (http://microbesonline.org). Sequence alignments were visualized with Espript (Gouet *et al.,* 1998) and carried out with Clustalw (Thompson *et al.,* 1994), with manual adjastements based on structure comparisons. Phylogenetic analysis was performed using the neighbor-joining method (Saitou and Nei, 1987) implemented in Clustalw, and the resulting unrooted tree was visualized as radial layout with FigTree (http://tree.bio.ed.ac.uk). The absence of a signal peptide in the sequence of HP0310 and its close homologs was predicted using the SignalP server (http://www.cbs.dtu.dk/services/SignalP).

Table 2.1. Data collection and refinement statistics. A wavelength of 0.9765 Å was used. Rotations of 1° were performed.

Data collection *	
Space group	$C222_1$
Cell dimensions	
a, ***b***, ***c*** (Å)	154.56, 154.73, 158.68
Resolution (Å)	69.54-2.57 (2.71-2.57)
R_{sym} or R_{merge}	0.125 (0.505)
$<I/\sigma(I)>$	6.4 (2.0)
Completeness (%)	98.1 (97.6)
Redundancy	3.8 (3.7)
Refinement	
No. reflections	59318
R_{work} / R_{free}	0.2056 / 0.2350
No. atoms	
Protein	9429
Ion ligand	4
Water	860
R.m.s. deviations	
Bond lengths (Å)	0.007
Bond angles (°)	1.05
Ramachandran plot (%)	
favored	92.6
outliers	1.5
Overall G factor	0.1

- 1 crystal was used to collect all diffraction data. *Highest-resolution shell is shown in parentheses.

Activity tests

Deacetylase activity of HP0310 in solution was assayed in three ways: for substrates without a free amino group before the reaction, a fluorogenic test was used; for all the others, by determining the amount of acetate liberated from the substrate during the reaction using an acetic acid kit (R-biopharm). All experiments were performed using the enzyme as purified from *E. coli*, and repeated after addition of 10mM $ZnCl_2$, and after incubation of the enzyme with EDTA for 1 hour at room temperature to extract the ion bound and subsequent addition of 10mM $ZnCl_2$ or $MgCl_2$.

Flurogenic Test

Enzymatic activity with N-acetyl glucosamine (GlcNAc), chitotriose GlcNAc3 and N-acetyl dipeptide (Ac-D-Ala-D-Ala-OH, Ac-D-Ala-D-Ala-OCH_3, Ac-D-Ala-L-Ala-OH, Ac-D-Ala-D-Ala-OCH_3) was performed according to (Blair *et al.,* 2005), with slight modification. In brief, a reaction mixture consisting of 1 μM peptidoclycan deacetylase, 5 μM cofactors, 50 mM Bis·Tris (pH 7.5), and 2 mM substrate (Sigma) in a total volume of 1 ml, was incubated for 90 min at 37°C. 2 microliters of 0.4 M borate buffer (pH 9.0) was then added and free amines labeled with 3 μl of 2 mg/ml fluorescamine in DMSO for 15 min at room temperature. The labeling reaction was terminated by the addition of 3 μl of DMF/H2O (1:1). Fluorescence was quantified by using an Ascent Fluorescence Reader (Thermo Electron Corporation, Finland), with excitation and emission wavelengths of 355 and 455 nm, respectively. The production of free amine was quantified with a glucosamine standard. All measurements were performed in triplicate.

Acetate test

The reaction mixture (0.5 ml) contained 50 mM HEPES buffer (pH 7.0), 5 mM N-acetyl putrescine, 10μM ZnCl2 and different amounts of the enzyme. The mixture was incubated at 37 °C for 4h and then boiled for 3 min to arrest the reaction. The release was determined with the acetic acid kit. The final calculations were performed by using the formula provided in the kit.

Allontoinase activity

Enzymatic activity with Allantoin was performed according to [32]. Circular dichroism measurements were carried out at 25 °C in a 10-mm path length cuvette with a Jasco J715

spectropolarimeter. HP0310 activity was determined using (R,S)-allantoin (0.5 mM) as substrates in the presence of HP0310 (0.70 µM) in 1ml of PBS, pH 7.5. CD spectra were recorded in the 200-350 nm range for 10 min, one spectrum every min. The last spectrum was collected after 20 min.

Supporting Information

Figure S1 Comparison of the HpPgdA genomic context in *Helicobacter pylori* and other bacteria. The analysis has been performed with the Microbesonline server (http://microbesonline.org) using HP0310 (A) or HP0312 (B) as anchor genes for the comparison. Homologous genes occurring at the same locus in different genomes are colored according to their homology group as indicated in the legend.

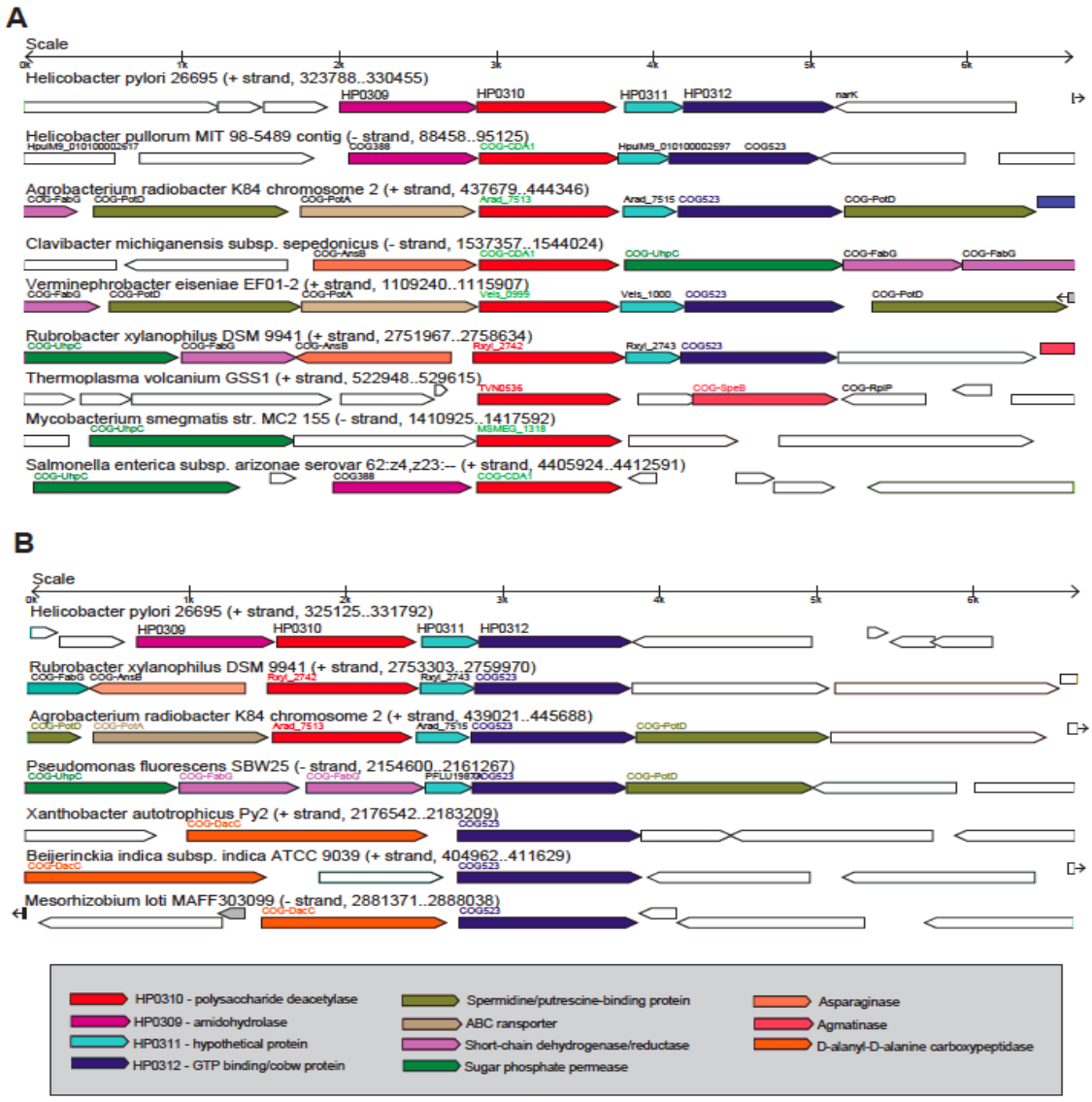
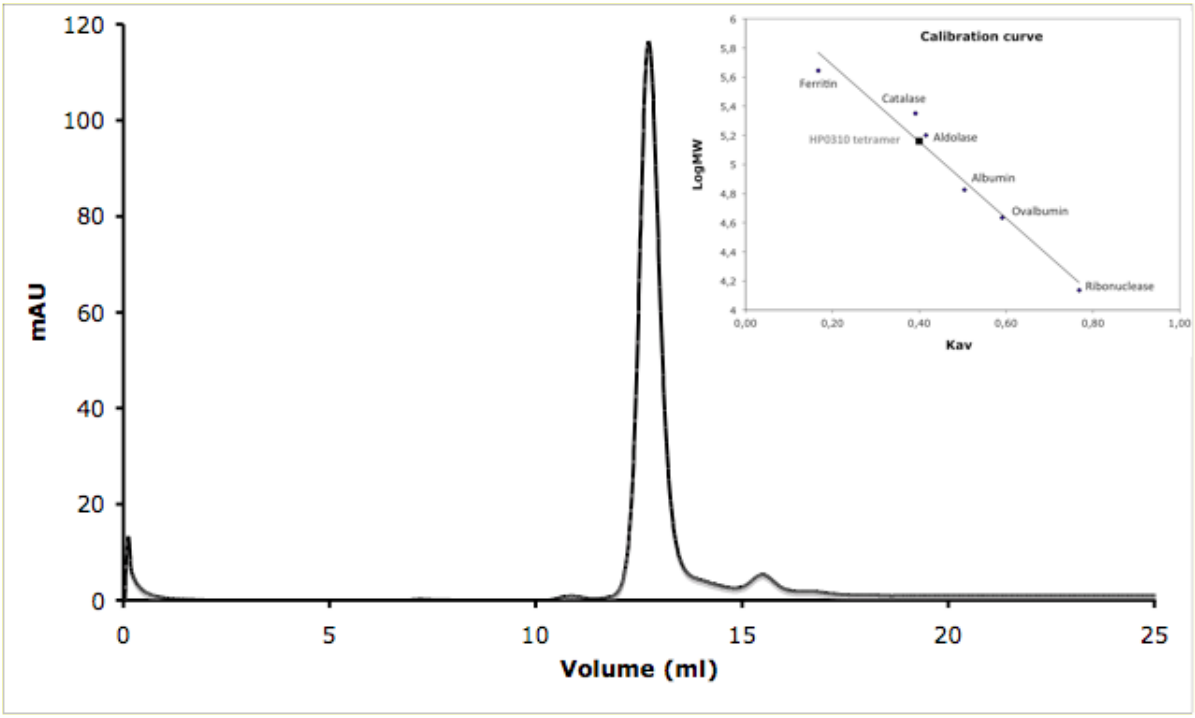


Figure S2 Analytical gel filtration of HpPgdA. Protein was purified by Superdex 200TM 10/300 GL (GE Healthcare), equilibrated with 30 mM Tris pH 8.0, 150 mM NaCl. His tagged-HP0310 eluted as a single peak, roughly corresponding to a tetramer: molecular mass estimated from analytical gel filtration is 144.54 kDa, in good agreement with the calculated one for the tetramer, 149.5 kDa.



Protein Data Bank accession codes. The coordinates and structure factors have been deposited in the Protein Data Bank (http://www.pdb.org) for immediate release with PDB ID code 3QBU.

Acknowledgments

We thank the staffs of beamline BM30 of ESRF, Grenoble, for technical assistance during data collection. We are grateful to Alessandro Moretto for preparing the N-acetyl-dipeptides.

Chapter Three

The crystal structure of ADP-L-*glycero*-D-*manno*-heptose-6-epimerase (HP0859) from *Helicobacter pylori*

This chapter is adapted from

M.M. Shaik, et al., The crystal structure of ADP-L-glycero-D-manno-heptose-6-epimerase (HP0859) from *Helicobacter pylori* , Biochim. Biophys. Acta 1814 (2011) 1641–1647, doi:10.1016/j.bbapap.2011.09.006.

The crystal structure of ADP-L-*glycero*-D-*manno*-heptose-6-epimerase (HP0859) from *Helicobacter pylori*

Md Munan Shaik, Giuseppe Zanotti and Laura Cendron

Biochimica et Biophysica Acta 1814 (2011) 1641–1647. doi:10.1016/j.bbapap.2011.09.006

Abstract

Helicobacter pylori , the human pathogen that affects about half of the world population and that is responsible for gastritis, gastric ulcer and adenocarcinoma and MALT lymphoma, owes much of the integrity of its outer membrane on lipopolysaccharides (LPSs). Together with their essential structural role, LPSs contribute to the bacterial adherence properties, as well as they are well characterized for the capability to modulate the immuno-response. In *H. pylori* the core oligosaccharide, one of the three main domains of LPSs, shows a peculiar structure in the branching organization of the repeating units, which displayed further variability when different strains have been compared. We present here the crystal structure of ADP-L-*glycero*-D-*manno*-heptose-6-epimerase (HP0859, rfaD), the last enzyme in the pathway that produces L-*glycero*-D-*manno*-heptose starting from sedoheptulose-7-phosphate, a crucial compound in the synthesis of the core oligosaccharide. In a recent study, a HP0859 knockout mutant has been characterized, demonstrating a severe loss of lipopolysaccharide structure and a significant reduction of adhesion levels in an infection model to AGS cells, if compared with the wild type strain, in good agreement with its enzymatic role. The crystal structure reveals that the enzyme is a homo-pentamer, and NAD is bound as a cofactor in a highly conserved pocket. The substrate-binding site of the enzyme is very similar to that of its orthologue in *E. coli*, suggesting also a similar catalytic mechanism. The other enzymes of the pathway are also discussed in terms of their three-dimensional structure.

KEYWORDS: *Helicobacter pylori* ; LPS; lipopolysaccharides; SDR family.

1. Introduction

Lipopolysaccharides (LPS), a complex glycolipid of the outer membrane of Gram-negative bacteria, play an important role in maintaining the structural integrity of the bacterium, confer resistance to detergents and cationic antimicrobial peptides, providing a physical barrier against the external world, and prevents complement-mediated cell death (Munford and Varley, 2009). LPS, which consists of three components, the endotoxin lipid A, a core domain and the immunogenic O-antigen repeat polymer (Raetz and Whitfield, 2002), is one of the powerful activator of innate immune responses (Bryant *et al.,* 2010). The LPS core domain, which represents a barrier to antibiotics, is an oligosaccharide containing different sugars. In *Helicobacter pylori ,* the human pathogen that affects half of the world population, LPSs shows a low immunostimulating activity, more similar to that of enterobacteria that establish a long-term chronic infection. In it, the three sugars that compose the core of LPS are 3-deoxy-D-*manno*-oct-2-ulsonic acid, L-*glycero*-D-*manno*-heptose (L,D-Hep) and D-*glycero*-D-*manno*-heptose (Chang *et al.,* 2011). Since the enzymes involved in their biosynthesis are not present in humans, they are potential targets for the development of antibiotics.

Twenty-seven *H. pylori* genes, spread throughout the genome, have been proposed to code for enzymes with a role in the biosynthetic pathways responsible for the synthesis of LPSs (Moran, 2001). In particular, the biosynthesis of L,D-Hep starting from sedoheptulose-7-phosphate is performed in Gram-negative bacteria by a five-steps pathway, illustrated in Scheme 1 (Chang *et al.,* 2011; Kneidinger et al., 2002). After the ketose sugar is isomerized by GmhA enzyme to *manno*-aldose, the latter is converted to D-β-D-heptose-1,7-bisphosphate by the bifunctional enzyme HldE. The previous reaction requires that one ATP molecule is spent in the process. Phosphatase GmhB converts D-β-D-heptose-1,7-bisphosphate into D-β-D-heptose-1-phosphate. The latter is converted to ADP-D-β-D-heptose-1-phosphate by the second activity of the bifunctional enzyme HldE. In the final step, ADP-L-β-D-heptose 1-phosphate is generated thanks to the catalysis of ADP-L-glycero-D-manno-heptose-6-epimerase (HldD or AGME).

In *H. pylori* four enzymes, from HP0857 to HP0860 (Tomb *et al.,* 1997), are responsible for this pathway. In particular, HP0859, which corresponds to AGME, has been recently recognized as essential for LPS synthesis (Chang *et al.,* 2011). A HP0859 knockout mutant of *H. pylori* showed a severe truncation of LPS, which induces a decreased growth rate and a

reduced adhesive capacity of the bacterium to AGS cells. The only crystal structure available of a ADP-L-glycero-D-manno-heptose-6-epimerase is that from *E. coli*, with different ligands bound (EcAGME, PDB ID:1EQ2, 2X6T, 2X86). Its quaternary structure corresponds to a homopentamer (Deacon *et al.*, 2000), each monomer being composed by two domains. EcAGME binds preferentially NADP as a cofactor, but it can also use NAD, resulting in a slower reaction rate (Ni *et al.,* 2001). On the contrary, HP0859 apparently behaves in solution as a homoexamer and evidence of a preferential binding to NADP instead of NAD has been proposed (Chang *et al.,* 2011). Given the relevance of LPS for *H. pylori* infection and persistence, we have decided to clone express, purify and crystallize the HP0859 enzyme. Its structure is discussed below.

Scheme 1. The enzymes involved in the linear pathway of conversion fromD-sedoheptulose 7-phopshate to ADP-L-β-D-heptose. The nomenclature of each *H. pylori* enzyme corresponding to strain 26695 is reported.

2. Materials and Methods

2.1. Molecular cloning of HP0859 gene. HP0859 gene was amplified by PCR from *H. pylori* G27 genomic DNA, using proofreading Phusion DNA polymerase Mastermix (Finnzymes, Finland), with the following primers: CACCATGCGTTATATTGATGATGAATTAG (Forward) and TGCGCGCTGTCCTTTAAAAATCG (Reverse). The amplified fragment was cloned into the pET101vector (Invitrogen) in frame with a C-terminal His-tag, using a TOPO® Cloning kit by Invitrogen to obtain the pET101-HP0859 plasmid. Right insertion in the cloned vector was confirmed by colony PCR and sequencing with T7 forward and reverse primer.

2.2. Overexpression and affinity purification of HP0859 protein

E. coli BL21(DE3) cells (Novagen) were transformed with the pET101-HP0859 plasmid and grown in LB media at 37° C, supplemented with Ampicillin (100 µg/ml). Protein expression was triggered by 0.5 mM isopropyl-β-D-thiogalactoside (IPTG, Inalco) when the culture reached an optical density (OD_{600}) of 0.7. After 4h incubation at 28° C, bacteria were collected and resuspended in a lysis buffer (30 mM Tris, pH 8.0, 150 mM NaCl) and then disrupted by a One Shot Cell breakage system (Constant System Ltd., UK). The lysate was centrifuged to remove cell debris (18,000 g for 25 min) and loaded into a column containing 5 ml of Ni^{2+} charged Chelating Sepharose™ (GE Healthcare, UK). After extensive washing using the lysis buffer, supplemented with 20mM imidazole, the protein was eluted by a linear gradient from 80 to 300 mM imidazole. The protein was further purified by gel filtration, using a Superdex 200™ 16/60 GL (GE Healthcare) equilibrated with 30 mM Tris pH 7.5, 150mM NaCl. HP0859-His tagged eluted as a single peak, roughly corresponding to a pentamer. All the protein samples collected throughout the purification were separated and analyzed on 15% sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE). The resolved gels were stained with 0.25% Coomassie Brilliant Blue R250 reagents.

2.3 Crystallization and structure determination

The purified protein was concentrated to 35 mg/ml and used for crystallization tests, partially automated using an Oryx 8 crystallization robot (Douglas Instruments). Crystals grew in several conditions. The best crystals were obtained at 20°C by vapor diffusion technique using a 20 mg/ml protein stock solution and, as precipitant, a solution containing 0.2 M Ammonium Sulfate, 0.1M Tri Sodium citrate pH 5.6, 15% (w/v), PEG 4000 (PEG II crystal screen, solution n. 56, Molecular Dimension, UK). Crystallization conditions were optimize on a 24 well plate by adding 5 % glycerol to both reservoir and drop precipitant. Crystals could be processed as monoclinic, space group I2, with **a**=130.94 Å, **b**=106.77 Å, **c**=152.22 Å β=108.35^0. One pentamer is present in the asymmetric unit, with V_M= 2.7Å^3/Da corresponding to an approximate solvent content of 54%. The data set used in the final refinement was measured at the beamline BM14 of the European Synchrotron Radiation Facility, Grenoble, France. It was indexed and integrated with software Mosflm (Leslie, 2006) and merged and scaled with Scala (Evans, 2006), contained in the CCP4 crystallographic package (CCP4, 1994). The structure was solved by molecular replacement using software Phaser (McCoy *et al.,* 2007) starting from a monomer model built by the SWISS-PDB server using as a template the crystal structure of ADP-L-glycero-D-

mannoheptose 6-epimerase from *E. coli* (PDB ID:1EQ2 (Deacon *et al.,* 2000)). Refinement was carried on using Phenix (Adams *et al.,* 2010), CNS and Refmac (Murshudov *et al.,* 1997). Several steps of manual rebuilding, performed with graphic software Coot (Emsley and Cowtan, 2004), were necessary in order to reach the final model. Solvent molecules were added with the automated procedure of Phenix. Since the first cycles of refinement, a clearly visible electron density was found in the enzyme active site and it was interpreted as NAD. The restraints dictionary for NAD was prepared from PRODRG2 server (Schuttelkopf and van Aalten, 2004) **(Table 1).**

The final model contains five protein monomers, five NAD and 245 water molecules. The final crystallographic R factor is 0.239 (R_{free} 0.279). Geometrical parameters of the model, checked with software Procheck (Laskowski *et al.,* 1993), are as expected or better for this resolution.

2.4 Molecular modeling and bioinformatics

Theoretical models were built using the SWISS-MODEL homology-modeling server (Kiefer *et al.,* 2009). Models obtained were optimized by a short molecular dynamics with software CNS (Murshudov *et al.,* 1997). Sequence alignments were carried out with Clustalw (Thompson *et al.,* 1994), and visualized with Esprit (Gouet *et al.,* 1994) with manual adjustments based on the 3D structure. Search for genetic clusters containing *hp*0859 gene were conducted using the Microbesonline web server (http://microbesonline.org). Ligplot software was used to plot the interaction of cofactor with the enzymes (Wallace *et al.,* 1995).

Table 1. Data collection and refinement statistics. A wavelength of 1.0 Å was used. Rotations of 1° were performed.

Data collection *	
Space group	I2
Cell dimensions	
a, *b*, *c* (Å), β(Å)	130.94, 106.77, 152.22, 108.35
Resolution (Å)	53.38 – 2.55 (2.69 – 2.55)
R_{sym} or R_{merge}	0.096 (0.408)
$\langle I/\sigma(I)\rangle$	4.8 (1.8)
Completeness (%)	86 (70.2)
Redundancy	2.3 (2.0)
Refinement	
No. reflections	55485
R_{work} / R_{free}	0.239 / 0.279
No. atoms	
Protein	12448
Ligand	220
Water	245
R.m.s. deviations	
Bond lengths (Å)	0.009
Bond angles (°)	1.4
Ramachandran plot (%)	
Allowed	97.5
Generously allowed	2.2
Disallowed	0.4
Overall G factor	0.2

* 1 crystal was used to collect all diffraction data. Highest-resolution shell is shown in parentheses.

3. Results and Discussion

All the genes coding for the enzymes involved in the biosynthesis of L-*glycero*-D-*manno*-heptose are clustered in the *H. pylori* genome, from *hp*0857 to *hp*0860 (Chang *et al.,* 2011). Theoretical models, built by homology modeling, of the aforementioned enzymes are available through the site http://tiresia.bio.unipd.it/zanotti/Helicobacter.

3.1. The crystal structure of HP0859 (AGME)

The crystal structure of *Hp*AGME has been determined at 2.55Å resolution. The polypeptide chain could be fitted for residues from 1 to 324, with the exception of two long loops, residues from 49 to 55 and from 282 to 288, which are possibly flexible and cannot be seen in the map. The electron density for the last 11 amino acids at the C-terminus is also lacking.

The overall folding of the monomer is quite similar to that of AGME from *E. coli* (EcAGME, PDB ID:1EQ2, Deacon *et al.,* 2000), which shares 30% identity with the *H. pylori* enzyme (Fig. 1A). Each monomer, which can be classified as a member of the SDR family (Kavanagh *et al.,* 2008), is constituted by two domains (Fig. 1B): an N-terminal seven-stranded modified Rossmann fold where the NAD cofactor is bound and a smaller C-terminal α/β domain, responsible for the binding of the substrate. Domain I includes residues 1-186, 232-254 and 295-307, domain II residues 187-231, 255-294, 308-324. The numbering scheme of secondary structure elements is reported in Fig. 1B

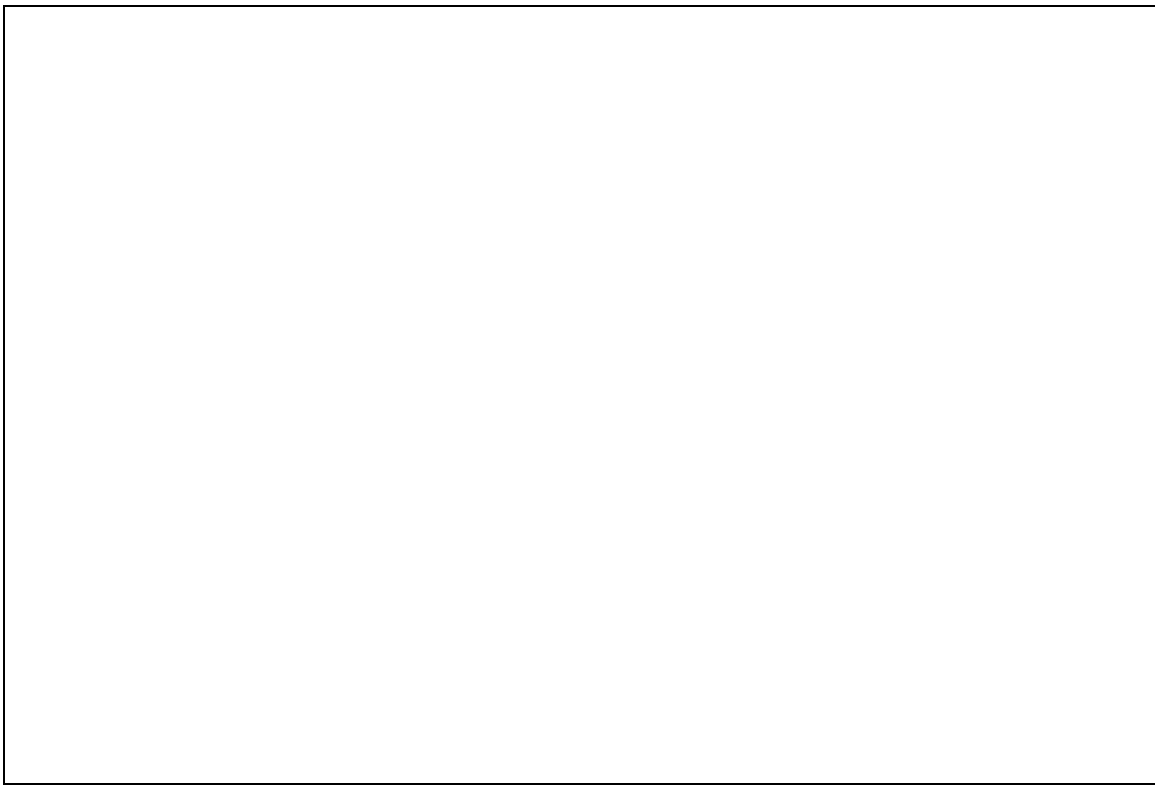
The r.m.s.d. between the core Cα atoms of the monomers of *Hp*AGME and *Ec*AGME is 1.6 Å. Major differences (Fig. 2) are represented by residues 1-10 (we refer to the numbering system of the *H. pylori* enzyme hereafter), not visible in the *E. coli* enzyme structure, and by region 46-60: in fact, strand β2 and helix η1 of the Rossmann fold in *Ec*AGME are connected through a short loop of four residues, whilst in the *H. pylori* enzyme there is an insertion of 10 amino acids, which forms a long loop partially flexible. In addition, residues 46-48 and 56-60, which are visible, assume a different conformation in the two enzymes. This is particularly relevant, since these residues interfere with the binding of the nucleotide (see below). Other structurally similar enzymes of the SDR family are UDP-galactose 4-epimerase, UGE (PDB ID:1NAH and PDB ID:1NAI, Thoden *et al.,* 1996a), dTDP-D-glucose 4,6 dehydrogenase, DGD (PDB ID:1BXK, Thoden *et al.,* 1996b) and GDP-4-keto-6-deoxy-D-mannose epimerase/reductase (GMER, PDB ID:1BWS; Rizzi *et al.,* 1998). The r.m.s.d. of the core Cα atoms of these three structures with the HpAGME monomer are 2.6 Å for the

alignment of 282 residues, 1.8 Å for 280 residues and 2.2 Å for 262 residues, respectively. Consider that the two former enzymes have a preference for binding NAD as a cofactor (Nayar and Bhattacharyya, 1997), whilst the third binds NADP.

3.1.1. Quaternary structure.

The biological unit of the enzyme is that of a homo-pentamer: five identical monomers are arranged around a five-fold symmetry axis (Fig. 3A, B). The five monomers interact through the larger N-terminal domain that constitutes the base of the pentamer, whilst a smaller C-terminal domain protrudes from the core of the assembly. As a consequence, the five substrates are expected to bind on the same side of the pentamer. At variance with *Ec*AGME, whose electrostatic surface is strongly negative (see Fig. 4 of reference, Deacon *et al.,* 2000), the surface of *Hp*AGME shows a slight prevalence of positive charges on the base of the pentamer and a substantial balance of positive and negative charges on the substrate-binding face. This is in line with the theoretical pI of the protein, which is 6.8 for *Hp*AGME and 4.8 for *Ec*AGME.

A



B

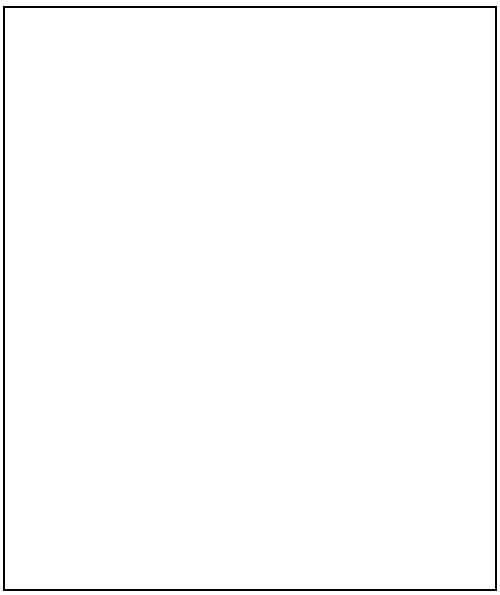


Figure 1. A. Alignment and secondary structure. Amino acid sequence of HpAGME aligned with ADP-L-glycero-D-manno-heptose 6-epimerase from *E. coli* (PDB: 1EQ2), UDP-Nacetylglucosamine 4-epimerase from *Pseudomonas aeruginosa* (PDB: 1SB8), UDP-GalNAc 4-epimerase from *Plesiomonas shigelloides* (PDB: 3LU1), hypothetical protein PH0414 from *Pyrococcus horikoshii* OT3 (PDB: 2HUN), dTDP-D-glucose 4,6-dehydratase (RmlB) from *Streptococcus suis* (PDB: 1KEP), UDP-glucose 4 epimerase (galE-1) from *Archaeoglobus fulgidus* (PDB: 3EHE), DTDP-D-glucose 4,6-dehydratase (RMLB) from *Salmonella enterica* serovar typhimurium (PDB:1G1A). Secondary structure elements for HpAGME are drawn over the alignment. **B.** Cartoon view of the monomer of HpAGME with NAD bound. The N-terminal domain that binds the nucleotide, is colored in light blue; the C-terminal domain, mainly responsible of the binding of the substrate, in marine blue.

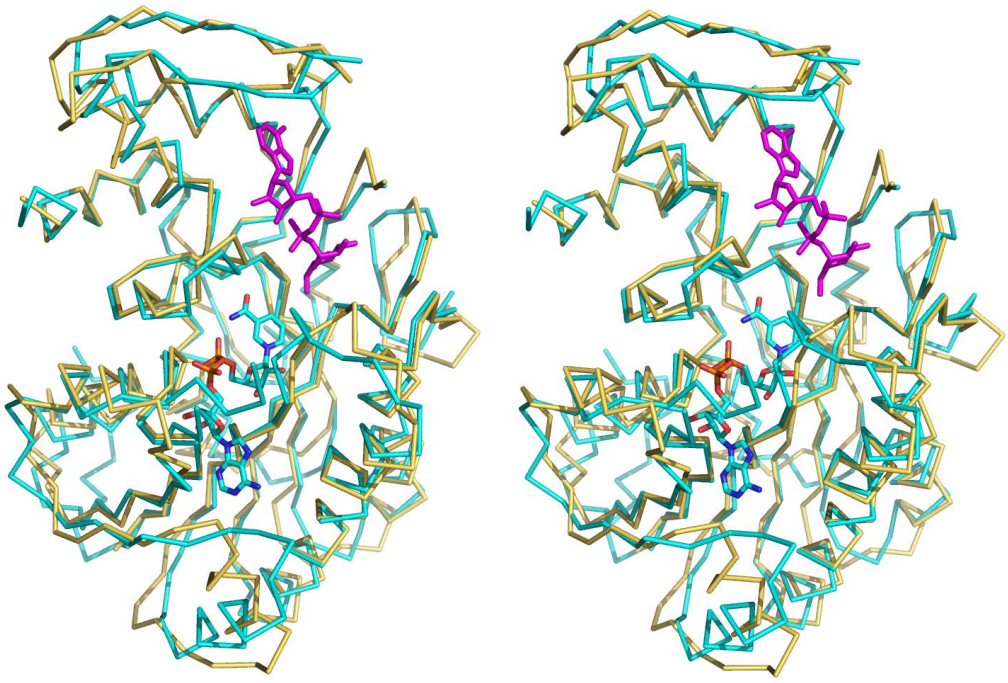


Figure 2. Stereo view of the Cα chain trace of *Hp*AGME (cyan) superposed to that of *Ec*AGME (yellow). NAD bound to the former is also, shown, along with the ADP-glucose bound to the E. coli enzyme (magenta).

3.1.2 The nucleotide binding site.

In the crystal structure of *Hp*AGME expressed in *E. coli,* a molecule of NAD is clearly visible bound to the Rossmann fold of the N-terminal domain (Fig. 4). The nucleotide is bound in an extended conformation and the distances between atom C6 of adenine and atom C2 of nicotinamide range from 15.4 Å to 15.8 Å in different monomers. Theses value are higher than those observed for NADP bound to *Ec*AGME (12.71-13.12 Å), and makes the conformation of our ligand more similar to the conformation of NADP bound to GMER (15.5 Å) or of NAD bound to UGE (14.4 Å).

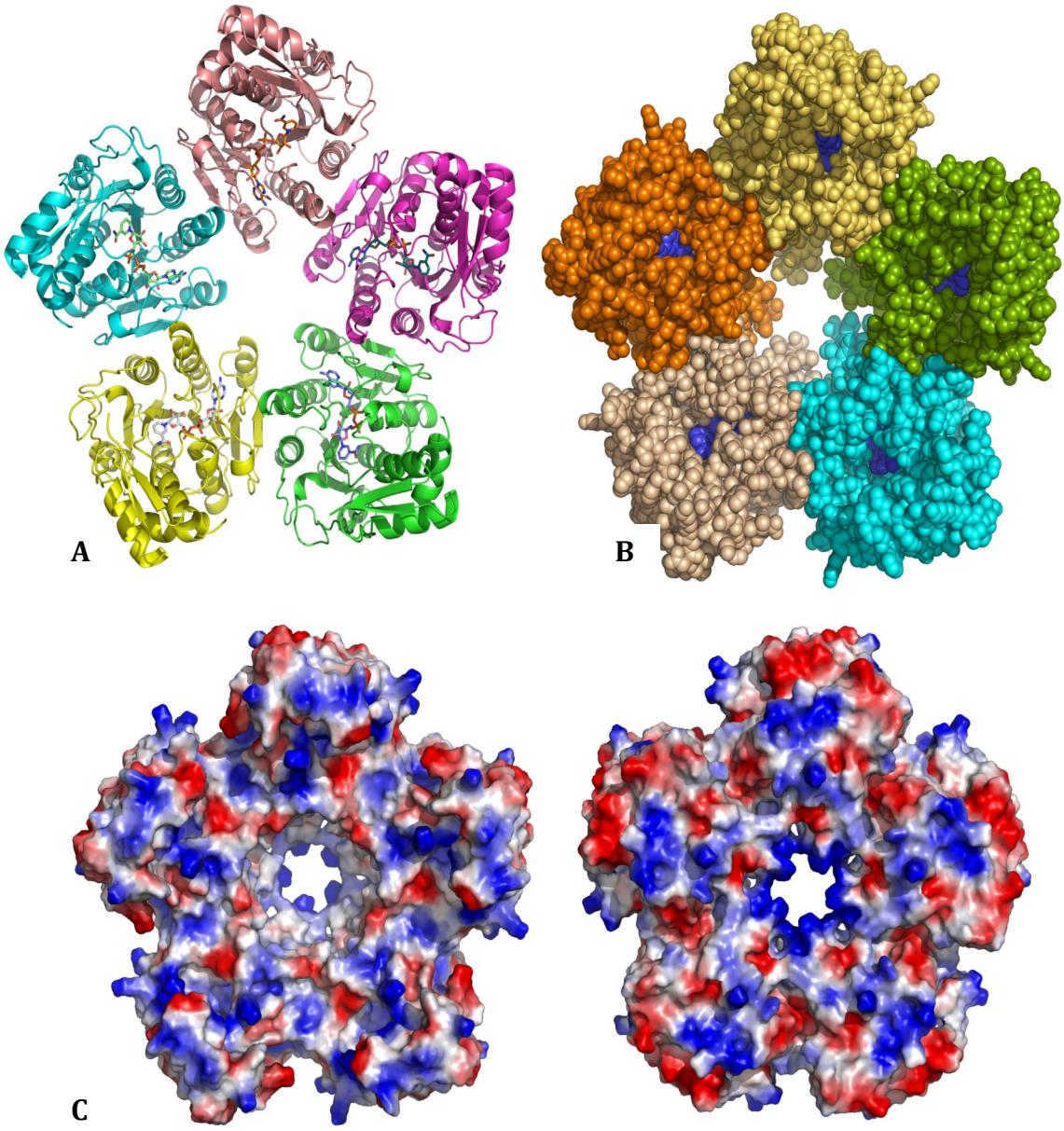


Figure 3. A) Cartoon view of the pentamer of HpAGME. The five monomers are shown in different colors. B) Same as A), but represented with the CPK model. The NAD cofactor is colored blue. C) Qualitative electrostatic potential surface of the pentamer, calculated with software Pymol (Delano, 2007). The base of the pentamer is shown on the left, the substrate-binding face on the right.

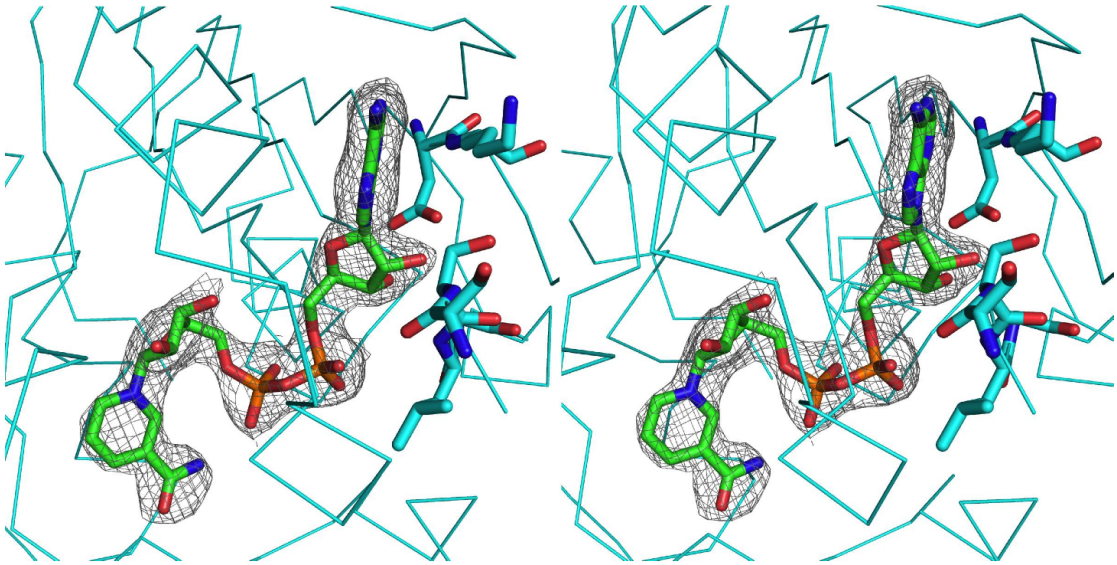


Figure 4. Stereo view of a portion of the nucleotide-binding site of HpAGME. The electron density map calculated with coefficients |2Fobs–Fcalc| is shown only around the nucleotide, contoured at 1.2σ. Side chains of residues 43–44 and 58–60 are shown explicitly.

The nucleotide-binding site in this protein family is characterized by the fingerprint sequence Gly-Gly-X-Gly-X-X-Gly. This sequence in HpAGME corresponds to residues 17 - 23, a loop connecting strand β1 to helix α2. They interact with the two phosphates and the ribose of the nicotinamide portion of NAD (Fig. 4). This interaction is strengthened by a solvent molecule that bridges an oxygen of the N phosphate to the NH group of Gly 23 and to the carbonyl oxygen of Gly 17.

Surprisingly, *Hp*AGME expressed in *E. coli* has a NAD molecule bound per monomer. The preference for NAD^+/NADH in *Hp*AGME seems to be due to region 56-60, which assumes a different conformation in the *H. pylori* crystal structure with respect to *Ec*AGME: in fact, in the former the side chain of Ser58 would clashes with the position occupied by the extra phosphate if NADP was hypothetically bound. Since the loop from 49 to 55 is flexible and cannot be observed in the electron density map, we cannot exclude that the conformation in the disordered region of the *H. pylori* enzyme is due to the presence of NAD bound instead of NADP, but it appears more likely that this conformation determines a preference of *Hp*AGME for NAD compared to NADP. In a recent paper concerning the biochemical characterization of recombinant *Hp*AGME, produced in *E. coli* (Chang *et al.,* 2011), an experiment was conducted supporting the conclusion that NADP is the preferred cofactor of

*Hp*AGME, supporting the hypothesis according to which the flexible loop could assume a different conformation upon NADP binding. In any case, as clearly proven by the structure presented here, *Hp*AGME can naturally acquire NAD from the *E. coli* background and properly accommodate such cofactor in its binding pocket.

A scheme of the interactions of the nucleotide bound to *Hp*AGME, compared to *Ec*AGME, is reported in Fig. 5. It is evident that, whilst the binding site is very well conserved around the nicotinamide moiety and the two phosphates of the ligand, interactions of protein residues with the adenine ring and ribose are not preserved, as expected due to the difference in the cofactor nature. In particular, Asp 76, Ile 77 and Gly 60 in *H. pylori* are the main residues forming hydrogen bonds with the adenine and ribose portions, while in *E. coli* Ile 77 and Asp 76 are involved in hydrogen bonds with the nucleotide, Lys 38, Lys 53, Asn32 and Tyr 38 define the network of interactions that coordinates the phosphorylated ribose moiety of NADP.

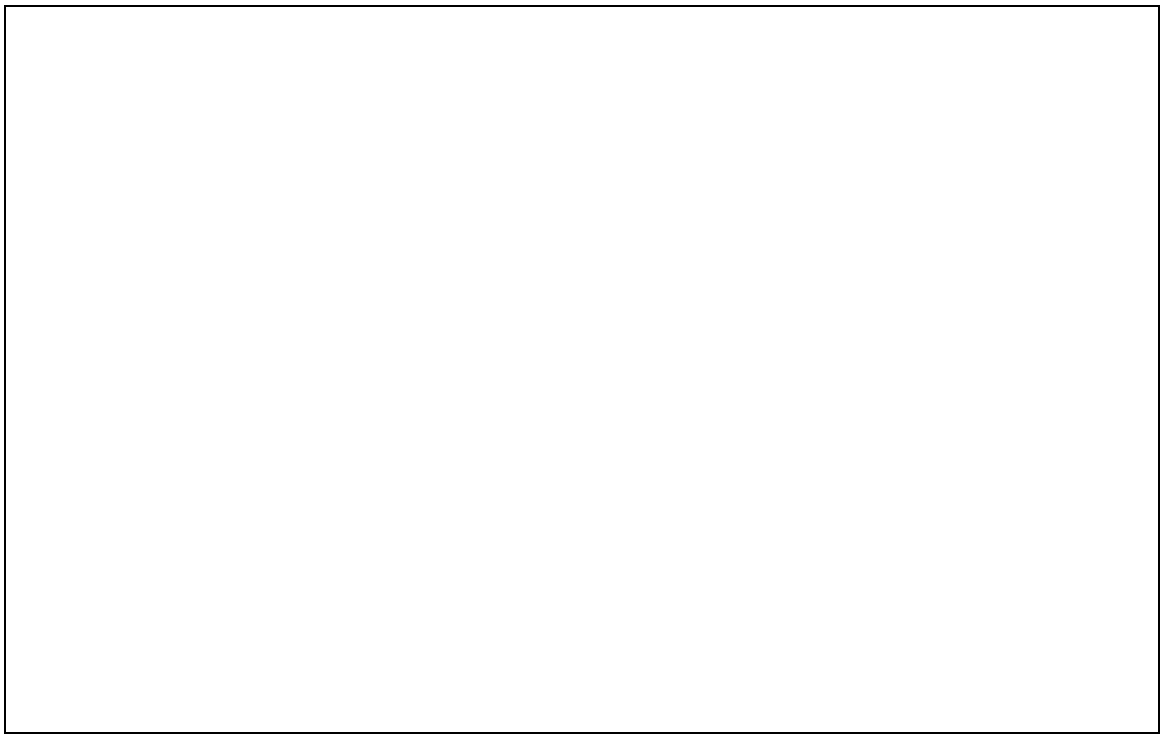


Figure 5. Comparative scheme of the interactions of the nucleotide with AGME. The diagram, drawn with LIGPLOT (Wallace et al., 1995), shows the contacts of cofactor NADP with the EcAGME (PDB: 1EQ2, left) and of NAD with HpAGME (this paper, right). Hydrogen bonds are presented as dashed lines; interatomic distances are in angstroms. "Radiating" curves indicate hydrophobic contacts between the corresponding atoms or residues and the surrounding residues.

3.2. The substrate-binding site

No molecule is bound in the substrate-binding site of *Hp*AGME crystal structure. A superposition of the monomer of the structures of *Ec*AGME, which have ADP-glucose (Deacon *et al.,* 2000) or ADP-β-mannose (Kowatz *et al.,* 2010) bound, allows a discussion of the active site residues. The region of binding of the glucose and of the two phosphates is very well conserved, with all the side chains of residues involved in some interaction present in both enzymes in the same positions. Only the area of binding of the adenine ring presents some conformational differences, possibly due to the absence of the ligand in our structure. This area is essentially hydrophobic, with the exception of Gln225 that replaces Phe207. Another major difference in the amino acid sequence in the area is represented by Met 203, which substitutes Val184 and, being more cumbersome, its side chain becomes quite close to the putative position of the adenine ring. In addition, residues considered relevant for the catalytic mechanism proposed for UGE (Thoden and Holden, 1998) are conserved, in particular Ser138 (116 in *Ec*AGME), Tyr161 (140) and Lys165 (144) appears to be properly placed for catalysis, suggesting a substantially identical catalytic mechanism for these enzymes.

4. Conclusions

The crystal structure of *Hp*AGME demonstrates that the enzyme is a homo-pentamer, as the orthologue enzyme from *E. coli*, and suggests that the catalytic mechanism of the *H. pylori* enzyme is probably identical or very similar to that of the other members of the family. The only significant difference with the *E. coli* enzyme is the use of NAD as a cofactor. Nevertheless, we cannot state from the crystal structure if the enzyme can use only NAD or else NADP.

The other *H. pylori* enzymes involved in the biosynthesis of L,D-Hep are also structurally related to their counterparts, suggesting very similar catalytic mechanism. The only enzyme of the pathway whose folding remains elusive is the kinase domain of HP0858, which is waiting for structural analysis of one of the members of the family to allow a complete description of the pathway.

Acknowledgments

We thank the staff of beamline BM14 of ESRF, Grenoble, for technical assistance during data collection. This work was supported by the University of Padua. Md Shaik's position was supported by CARIPARO grants for foreign Students at the University of Padua.

Protein Data Bank accession codes. The coordinates and structure factors have been deposited in the Protein Data Bank (http://www.pdb.org) for immediate release with PDB ID code 3SXP.

Chapter Four

The crystal structure of *Helicobacter pylori* ceuE (HP1561) reveals an ABC transporter substrate binding protein

This chapter is adapted from

Md Munan Shaik, Laura Cendron and Giuseppe Zanotti. The crystal structure of *Helicobacter pylori* ceuE (HP1561) reveals ABC transporter substrate binding protein (Manuscript in Preparation).

The crystal structure of *Helicobacter pylori* ceuE (HP1561) reveals an ABC transporter substrate binding protein

Md Munan Shaik, Giuseppe Zanotti and Laura Cendron

Abstract

Periplasmic substrate binding protein of ATP binding cassette (ABC) transporter play a essential role in facilitating uptake of essential nutrients of Gram negative pathogens. *Helicobacter pylori* are able to use host heme derived iron. We reported the crystal structure of *H. pylori* ceuE, the protein that binds heme and delivers it to the periplasmic surface of the ABC transporter. The structure reveals a bi-lobed fold with a narrow cleft between the N- and C-terminal domains share a common architecture typical of Class III periplasmic binding proteins. The detailed architecture of the heme pocket is quite different than other reported in ShuT, periplasmic heme binding protein from *Shigella dysenteriae* HmuT, periplasmic heme binding protein from *Yersinia pestis* and PhuT, from *Pseudomonas aeruginosa*. The heme binding affinity is explored by spectroscopic binding studies performed in solution. Fluorescense and Isothermal titration calorimetry suggests that the hemes are bound with lower affinity, which is similar to the affinities observed in other bacterial substrate binding proteins.

KEYWORDS: *Helicobacter pylori*; Iron; Heme; ABC Transporter Substrate Binding Protein SBP, Periplasmic Binding Protein.

1. **Introduction**

H. pylori, a human pathogen that colonizes about half of the world population and is responsible for gastritis and gastroduodenal ulcers, also presents a high risk factor for the development of mucosa-associated lymphoid-like tissue (MALT) lymphoma as well as gastric adenocarcinoma (Peek and Blaser, 2002). The gastric adenocarcinoma, which is the second leading cause of cancer-related death in the world, brings much attention about *H. pylori* (Correa, 1996). The human stomach is a unique ecological niche characterized by a very acid pH, and, as such, it is a hostile environment for most microorganisms, including bacteria. Nevertheless, *H. pylori* is able to survive in it, colonizing the gastric epithelial layer (Algood and Cover, 2006). Trace minerals are essential for life, which act as essential cofactors of enzymes and as organizers of the molecular structures of the cell. Successful pathogenic bacteria must acquire nutrients from the host and one such essential nutrient is iron. *H. pylori,* like many other mucosal pathogens, requires iron to survive and colonize. Iron is particularly depleted at the mucosal surface and the iron acquisition system differs according to the ecological niche of the organism. Expression of bacterial virulence factors in *H. pylori* are also controlled by iron availability (Merrell *et al.,* 2003), and proteins involved in iron metabolism are suggested to represent major virulence determinants (McGee and Mobley, 1999). It has been shown that addition of iron to the apical medium to some extent rescues the defect of cell surface microcolony formation by isogenic ΔcagA mutants, suggesting that along with many other effects by CagA on host cells, one is to facilitate iron acquisition from the host (Tan *et al.,* 2011). *H. pylori* iron acquisition is also of interest, because infection with *H. pylori* is associated with iron-deficiency anemia (Muhsen and Cohen, 2008). Positive association between heme iron intake and Gastric cancer risk in the Europe was reported (Jakszyn *et al.,* 2011). Pathogenic bacteria have developed several mechanisms for acquiring iron from the host: (i) Siderophore-mediated iron uptake involves the synthesis of low molecular weight iron chelators, called siderophores, which compete with the host iron-binding glycoproteins lactoferrin (LF) and transferrin (TF) for iron, (ii) outer membrane proteins mediate iron uptakes; the latter are receptors that recognize the complex of TF or LF with iron, resulting in the internalization of this metal, and (iii) periplasmic heme-binding proteins use heme-compounds released into the circulation after lysis of erythrocytes (Otto *et al.,* 1992).

Host-derived iron sources available in the gastric mucosa are lactoferin-bound iron, heme compounds released from damaged tissues, and iron derived from pepsin-degraded food. *H.*

pylori can use heme as the sole iron source, but it is able also to use human lactoferin and transferrin (Husson *et al.,* 1993; Worst *et al.,* 1995; Velayudhan *et al.,* 2000). *H. pylori* is known to possess several iron uptake systems, but the sources of iron that *H. pylori* utilizes during colonization of the gastric mucosa remain unclear (van Vliet *et al.,* 2001). In *H. pylori*, iron uptake is differently regulated compared with other bacteria: it presents around 13 genes encoding putative iron-uptake proteins in the genome. Some iron-uptake systems are constitutively expressed, whilst other iron uptake systems display the iron- and bacterial iron-responsive regulation Ferric Uptake Regulator (Fur)-mediated repression, a system common to many bacteria (van Vliet *et al.,* 2002). Presence of different iron uptake regulation systems may be a specific adaptation to the conditions in the human stomach, where iron starvation and iron overload can be encountered at relatively short time intervals. *H. pylori* has not been shown to synthesize siderophores, like other mucosal-colonizers that possess siderophore-mediated mechanisms for the uptake of iron (van Vliet *et al.,* 2001; Dhaenens *et al.,* 1999). While the acidity of the gastric lumen releases iron from ingested food (Miret *et al.,* 2003), *H. pylori* is not found in the gastric lumen but rather it colonizes the neutral environment of the epithelial cell surface and the overlying mucus layer (Schreiber *et al.,* 2004).

The uptake of heme as an iron source is a common mechanism by which pathogenic bacteria obtain the iron necessary for their survival and ability to establish an infection (Henderson *et al.,* 1993; Hornung *et al.,* 1996; Mills and Payne, 1995; Stojiljkovic and Hantke 1992). Numerous Gram-negative bacterial pathogens have developed a sophisticated mechanism for recruiting heme iron (Eakanunkul *et al.,* 2005; Tong and Guo, 2007), which has iron-dependent outer membrane receptors specific for heme *via* a TonB-mediated gated pore mechanism (Ridley *et al.,* 2006; Wyckoff *et al.,* 2004; Perry *et al.,* 2003).

Heme transport across the periplasmic space and into the cytoplasm is affected by an active transport system, comprising a soluble periplasmic binding protein, a cytoplasmic permease, and an ATPase (ABC transporter). The ABC transporters are an ubiquitous family of proteins comprising two membrane-spanning transporter domains and the corresponding cytoplasmic domains that, through ATP hydrolysis, drive uptake of the substrate (Davidson, 2002). The fate of heme after entering the cytoplasm is not well understood. Heme oxygenase-like activity is present in *Bacillus cereus* and *Streptococcus mitis* for heme degradation (Engel *et al.,* 1972) and *hugZ* (heme oxigenase) was identified in *H. pylori* responsible for heme

degradation and utilization of iron (Guo *et al.,* 2008). Periplasmic heme binding protein are identified and characterized from several different pathogenic bacteria, PhuT from *Pseudomonas aeruginosa* (Ho *et al.,* 2007), ShuT from *Shigella dysenteriae* (Ho *et al.,* 2007), HemT from *Yersinia enterocolitica* (Stojiljkovic and Hantke, 1994), HmuT from *Yersinia pestis*, ChuX from *Escherichia coli* O157:H7 (Suits *et al.,* 2009), HutB from *Vibrio cholerae*, HugB from *P. shigelloides* and BhuT from *Bordetella avium, B. pertusis, B. bronchiseptia*. A lactoferrin-binding protein has been described for *H. pylori* (Dhaenes *et al.,* 1997), but until now no protein was recognized for heme binding. HP15161 (ceuE) was annotated as an ABC transporter periplasmic iron-binding protein (ceuE). Later, cytoplasmic permease (fecD) and ATPase (fecE) were identified and reported to be involved in iron uptake (Contreras, *et al.,* 2003). Transcription of the genes encoding the components of a cytoplasmic ABC-transporter (ceuE1, ceuE2, fecD and fecE genes) was not repressed by iron, and also not affected by fur mutation (Delany et. al., 2001; van Vliet *et al.,* 2002). ceuE and fecDE genes are likely to encode a novel nickel and cobalt acquisition system in *H. mustelae,* readdressing the apparent imbalance between iron and nickel acquisition genes in the genome (Stoof *et al.,* 2010). To elucidate the structural features of this putative periplasmic ABC transporter binding protein HP1561 from *H. pylori,* the protein was cloned, expressed and purified in good yield in *E. coli*, crystallized, its structure determined and *in vitro* binding activity with different possible metal ions and metal containing compounds were performed.

2. Materials and Methods

2.1. Molecular cloning of HP1561 gene

HP1561 gene was amplified by PCR from *H. pylori* G27 genomic DNA, using proofreading *pfu* DNA polymerase (Finnzymes, Finland), with the following primers: 5'-CACCATGGAAGTCAAAGTTAAGGATTATTTCG (Forward) and 5' CCATAAGAATGGCTCAACTTCTGCGTC (Reverse), which amplify from 33 to 332 amino acids from total length 333 amino acids with N-terminal signal peptide/transmembrane helix deletion (32 amino acids) and one His deleted from C- terminal. The amplified fragment was cloned into the pET101vector (Invitrogen) in frame with a C-terminal His-tag, using a TOPO® Cloning kit by Invitrogen to obtain the pET101-HP1561 plasmid. Right insertion in the cloned vector was confirmed by colony PCR and by sequencing with T7 forward and reverse primer.

2.2. Overexpression and affinity purification of HP1561 protein.

E. coli BL21(DE3) cells (Novagen) were transformed with the pET101-HP1561 plasmid and grown in LB media at 37° C, supplemented with Ampicillin (100 µg/ml). Protein expression was triggered by 1 mM isopropyl-β-D-thiogalactoside (IPTG, Inalco) till the culture reached an optical density (OD_{600}) of 0.7. After 4h incubation at 28° C, bacteria were collected and resuspended in a lysis buffer (30 mM Tris, pH 8.0, 150 mM NaCl) and then disrupted by a One Shot Cell breakage system (Constant System Ltd., UK). The lysate was centrifuged to remove cell debris (18,000 g for 25 min) and loaded into a column containing 5 ml of Ni^{2+} charged Chelating Sepharose™ (GE Healthcare, UK). After extensive washing using the lysis buffer, supplemented with 20mM imidazole, the protein was eluted by a linear gradient from 80 to 300 mM imidazole. The protein was further purified by gel filtration, using a Superdex 200™ 16/60 GL (GE Healthcare) equilibrated with 30 mM Tris pH 8.0, 150mM NaCl. HP1561-His tagged eluted as a single peak. All the protein samples collected throughout the purification were separated and analyzed on 15% sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE). The resolved gels were stained with 0.25% Coomassie Brilliant Blue R250 reagents.

To prepare selenomethionyl protein, plasmid pET101 containing the HP1516 gene was transformed into the methionine auxotrophic *E. coli* strain B834. The transformed bacteria were grown in M9 minimal medium supplemented with 0.4% (w/v) glucose, salts and all the amino acids except Met, substituted by Se-Met (50mg/L). About 30 minutes before induction with 0.3 mM IPTG, a further solution of Se-Met plus Leu, Ile, Val, Phe, Lys and Thr was added to the medium to inhibit the *E. coli* methionine pathway and to force the incorporation of Se-Met. The Se-Met derivative of HP1561 was purified as the native protein, with enhanced content of reducing agent in buffer (5 mM DTT), to prevent oxidation.

2.3 Circular Dichroism (CD) and UV Analysis

Circular Dichroism spectrum was measured with J-715 Spectropolarimeter (JASCO, Corporation) at 298K. 10 runs were accumulated with 1mg/ml protein (Tris 3mM, NaCl 10mM, pH 8.0) in a 0.2mm path cuvette in the wavelength interval 190-260nm. The CD spectrum was rescaled with respect to a standard solution containing the buffer and the molar ellipticity calculated (Figure 1). The CD spectrum was deconvoluted with software "CD Spectra Deconvolution" (CDNN 2.1) and the secondary structure predicted.

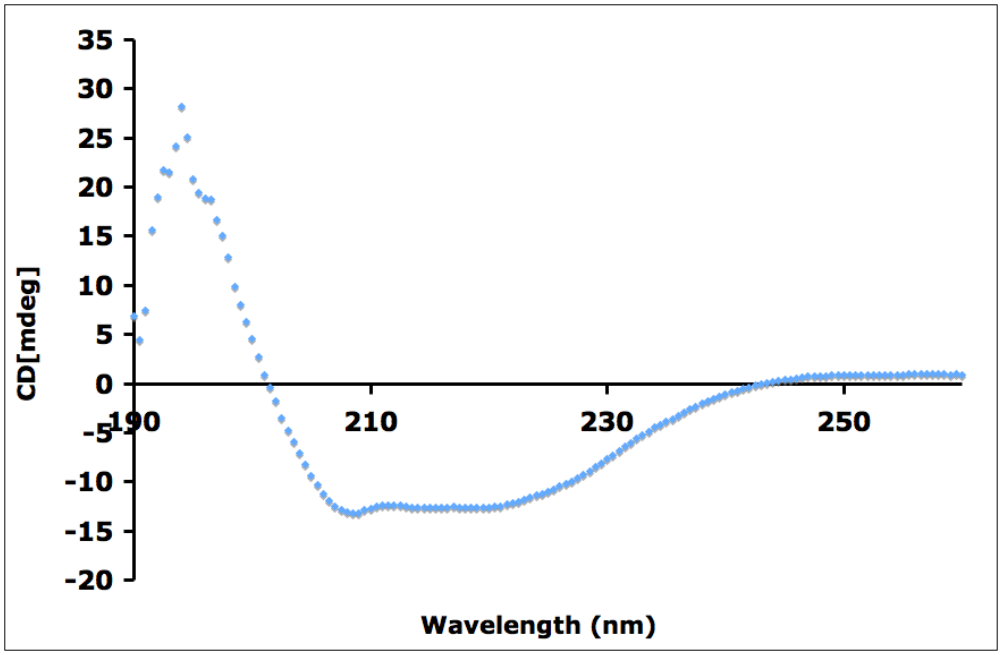


Figure 1. Circular Dichroism spectrum of HP1561 was measured with J-715 spectropolarimeter (JASCO, Corporation. The CD spectrum was deconvoluted with software "CD Spectra Deconvolution" (CDNN 2.1) and the secondary structure predicted.

2.4 Molecular mass determination

The native molecular mass of HP1561 protein was determined by analytical gel filtration using an analytical column Superdex 200™ 10/300 GL (GE Healthcare), equilibrated with 30 mM Tris pH 8.0, 150mM NaCl. His tagged-HP1561 eluted as a single peak, roughly corresponding to a monomer: molecular mass estimated from analytical gel filtration is 34.2 kDa, in good agreement with the calculated one for the monomer, 36.39 kDa (33.69 kDa plus the 2.7 kDa His-tag) (Figure 2). The system was operated on an AKTA FPLC instrument (GE Healthcare).

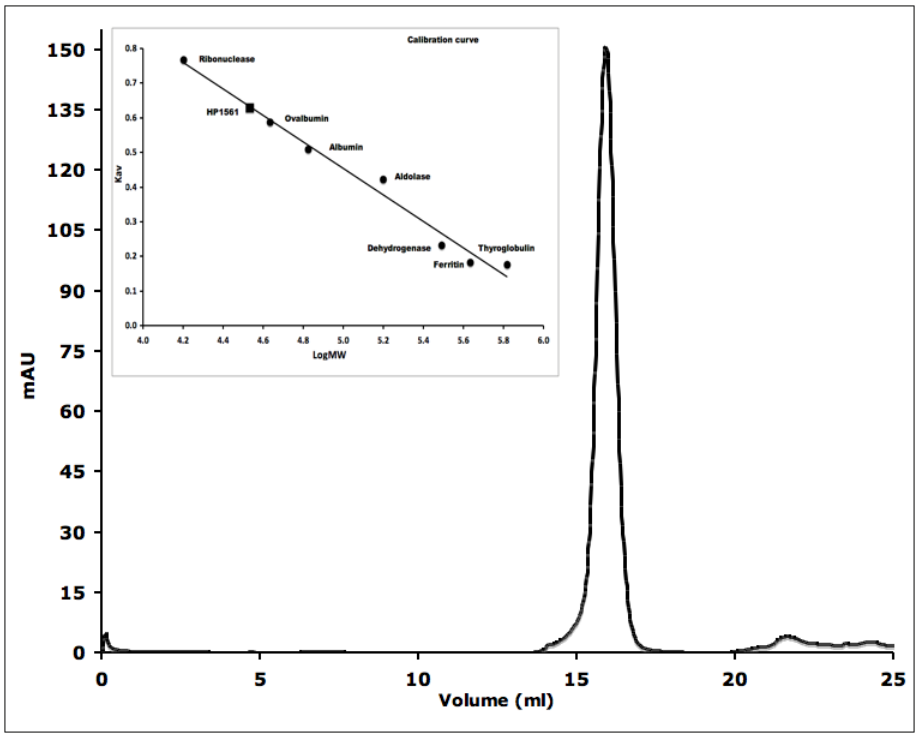


Figure 2. Analytical gel filtration of HP1561 by analytical column Superdex 200™ 10/300 GL (GE Healthcare). His tagged-HP1561 eluted as a single peak.

2.5 Crystallization and structure determination. The purified protein was concentrated to 50 mg/ml and used for crystallization tests, partially automated using an Oryx 8 crystallization robot (Douglas Instruments). Crystals grew in several conditions. The best native crystals were obtained at 20°C by vapor diffusion technique using a 20 mg/ml protein stock solution and, as precipitant, a solution containing 0.1 SPG Buffer pH 8.0, 25% (w/v), PEG 1500 (The PACT Suit, solution n. 5, Qiagen, USA). Crystals could be processed as orthorhombic, space group $P2_12_12_1$, with unit cell dimension **a**=67.37 Å, **b**=87.02 Å, **c**=105.90 Å. Two monomers are present in the asymmetric unit, with V_M= 2.3Å^3/Da, corresponding to an approximate solvent content of 47%. A diffraction data set was measured at the beamline ID29 of the European Synchrotron Radiation Facility (ESRF), Grenoble, France. Crystals of Se-Met derivative were grown in the same condition and a SAD dataset was measured at the beamline ID14-4 of ESRF. The Se-Met derivative crystal belongs to the same space group and can be considered isomorphous with the native protein. To explore the metal coordination in the metal binding site, HP1561 crystals was grown in the presence of 10mM $NiSO_4$ (Aldrich) in precipitant solution (0.5 M Sodium Formate, 0.1 M Bis Tris Propane, pH 7.5, 20 % (w/v) PEG 3350) (Qiagen, USA), by sitting-drop vapor diffusion at 20^{0}C. These

crystals belong to the monoclinic $P2_1$ space group with unit cell dimensions a= 60.18 Å, b=76.95 Å, c=72.78 Å, β=94.52 Å. Nickel anomalous data was measured at the beamline ID 23-1 of the ESRF at the wavelength of 1.48520 Å. All the datasets were indexed and integrated with software Mosflm (Leslie, 2006) and merged and scaled with Scala (Evans, 2005), contained in the CCP4 crystallographic package (Collaborative Computational Project, Number 4 1994). The structure was solved by experimental phasing using the selenium dataset, in which position of selenium was explored by Autosol (Terwilliger *et al.,* 2009) and further model building with extension to high resolution native dataset with Autobuild (Terwilliger *et al.,* 2008) present in the Phenix package (Adams *et al.,* 2010). Model building was further extended with Buccaneer (Cowtan, 2006) and manual rebuilding with graphic software Coot (Emsley, 2004). Refinement was carried on using Phenix (Adams *et al.,* 2010), and Refmac (Murshudov, 1997). The final crystallographic R factor is 0.198. (R_{free} 0.239). The complex with Nickel was solved by molecular replacement using software Phaser (McCoy, 2007) starting from a monomer model of HP1561. Solvent molecules were added with the automated procedure of Phenix. Geometrical parameters of the models, checked with software Procheck (Laskowski, 1993), are as expected or better for this resolution. Data collection and refinement statistics are presented in the table 1.

Table 1. Data collection and refinement statistics.

Data set	Native	Selenium SAD	Nickel SAD
Heavy Atom	-	Selenium	Nickel
Wavelength	0.97627	0.97920	1.48520
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_1$
Cell dimensions			
A, ***b***, ***c*** (Å)	**a**=67.37, **b**=87.02, **c**=105.90	**a**=67.650, **b**=87.440, **c**=106.150	**a**= 60.18, **b**=76.95, **c**=72.78, β=94.52
Resolution (Å)	47.59-1.65 (1.74-1.65)	67.65-3.00 (3.16-3.00)	52.79-2.30 (2.42-2.30)
R_{sym} or R_{merge}	0.054 (0.510)	0.163 (0.463)	0.078 (0.226)
$<I/\sigma(I)>$	10.2(2.3)	14.2 (7.2)	8.8 (5.2)
Completeness (%)	97.0 (96.8)	100 (100)	92.0 (96.20)
Anomalous Completeness (%)	-	100 (100)	70.5 (73.1)
Redundancy	4.1 (4.2)	14.1 (14.4)	2.5 (2.5)
Refinement			
No. reflections	73090		27148
R_{work} / R_{free}	0.198/ 0.239		0.193/ 0.268
No. atoms	5205		5003
Protein	4798		4765
Ion ligand	-		6
Water	409		200
R.m.s. deviations			
Bond lengths (Å)	0.027		0.019
Bond angles (°)	2.206		1.740
Ramachandran plot (%)			
Preferred	533 (98.7)		531 (98.9)
Allowed	7(1.3)		4 (0.7)
Outliers	0(0.0)		2 (0.4)

2.5. Fluorescence measurements

Fluorescence spectra were masured with a Perkin Elmer LS50B (USA) Fluorescence Spectrometer. 1 ml of protein sample in quartz-cuvette was incubated for 10 minutes at 25^0C under constant stirring before stepwise addition of the substrate (or buffer as control). Samples were incubated for 3 min before the fluorescence signal was collected for 120 seconds to obtain an averaged value at each substrate concentration.

For fluorescence titration experiments, around 0.5 μM of purified HP1561 (Tris 30mM, NaCl 150mM, pH 7.5) was used, and a solution of substrate 0.1-15μM (Hemin or Vitamin B12) was added stepwise. For the intrinsic protein fluorescence measurements, the excitation wavelength was 295 nm with slit widths of 5 nm, for emission a scan was made in the range of 300 to 400 nm. Corrections for background fluorescence changes were made by titrations with buffer.

2.6. Isothermal titration calorimetry (ITC)

Hemin binding to HP1561 was measured by VP-ITC Microcalorimeter (GE Healthcare) at 25^0C. Hemin was solubilized with 1M NaOH and diluted to 20μM with buffer (Tris 30mM, NaCl 150mM, pH 8.2). 2 ml of 20 μM hemin was added to the cell. To determine the binding constant, HP1561 concentrated to 350 μM (in the same buffer as the Hemin) was added stepwise. Typically, 25 injections of 10μl volume were made with intervals of 500 seconds between each addition. The first titration in each experiment was 1 μl of heme instead of 10μl, which was subsequently deleted in the data analysis; data were analyzed using the Origin 7 software provided with the instruments.

2.7 Bioinformatics.

Signal sequence analysis in HP1561 protein was carried out with SignalIP 3.0 server (Bendtsen *et al.,* 2004). Sequence alignments were carried out with ClustalW (Thompson *et al.,* 1994), and visualized with Esprit (Gouet *et al.,* 1999) with manual adjustments based on the 3D structure.

3. Results and Discussion

3.1 The overall Model

The crystal structure of *H. pylori* ceuE (HP1561) was determined at 1.65Å resolution using phases obtained from a SAD experiment. The overall fold of HP1561 shares structural features in common with PhuT and ShuT, two heme transport proteins from *Pseudomonas aeruginosa* (PDB ID: 2R79) and *Shigella dysenteriae* (PDB ID: 2R7A), respectively (Ho *et al.,* 2007), BtuF, a vitamin B12 transport protein from *E. coli* (PDB ID: 1N2Z, Borths *et al.,* 2002), FhuD, periplasmic siderophores binding protein from *E. coli* (PDB ID: 1K7S, Clarke *et al.,* 2002) and HmuT, periplasmic heme binding protein from *Yersinia pestis* (PDB ID: 3NU1, Mattle *et al.,* 2010). They all classify as class III periplasmic binding protein. HP1561 has two structurally similar globular domains, a N-terminal domain (34-165) and a C-terminal domain (187-335), each consisting of a central five-stranded β-sheet surrounded by α-helices (known also as a Rossmann-like fold). The two domains are topologically similar and a long, rigid α-helix (α10, 166-186) acts as an interdomain linker (Figure 3 and 4). The qualitative surface charge distributions are shown in Figure 5.

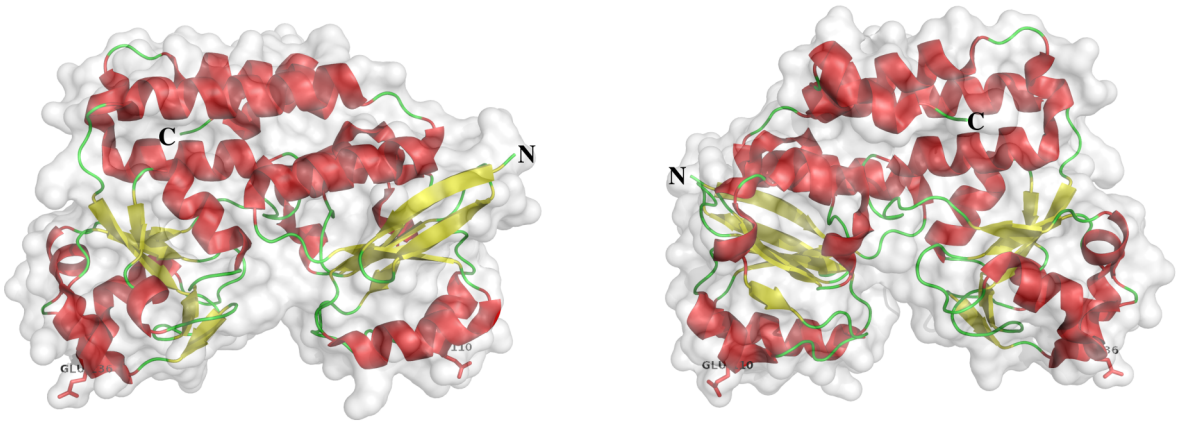


Figure 3. Overall Cartoon model of HP1561 monomer. α-helices are in red, β-strands in yellow, others in green. N and C-terminus are labeled. Figure on the right is seen from the opposite side, after 180^0 rotation (Drawing produced using Pymol program).

HP1561 structure does not only demonstrate the similarities of the protein fold with the other members of the family, but also reveals that HP1561 features two conserved, surface-exposed glutamates (E110 and E236), one in each lobe (Figure 9), as in the others, interacting with two positively charged ‘‘pockets’’ on the periplasmic surface of its cognate binding protein

of an ABC transporter. Two molecules, oriented face-to-face facing the binding cavity, are present in the asymmetric unit. Despite the fact that the protein is a monomer in solution, as clearly demonstrated by analytical gel filtration (Figure 1).

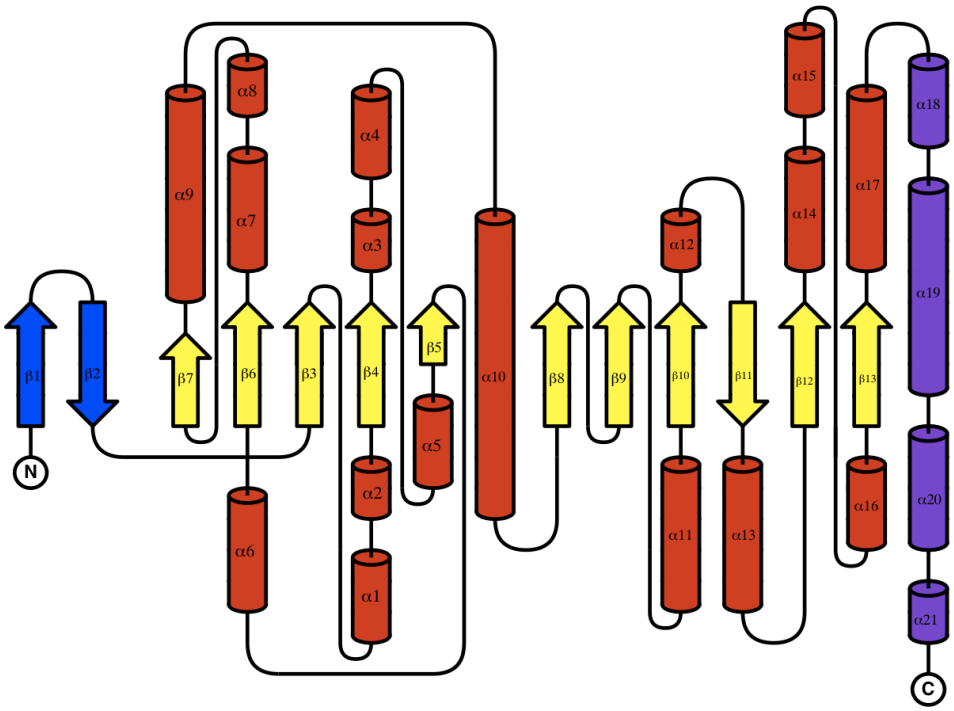


Figure 4. Topology diagram of HP1561. Helices and strands belonging to the Rossman fold are in red and yellow, respectively; Two antiparallel β-sheet in the N-terminal and four helices in the C-terminal domains, which are not present in other members of this family, are shown in blue and dark violet, respectively. (This drawing was done using the Topdraw program present in CCP4).

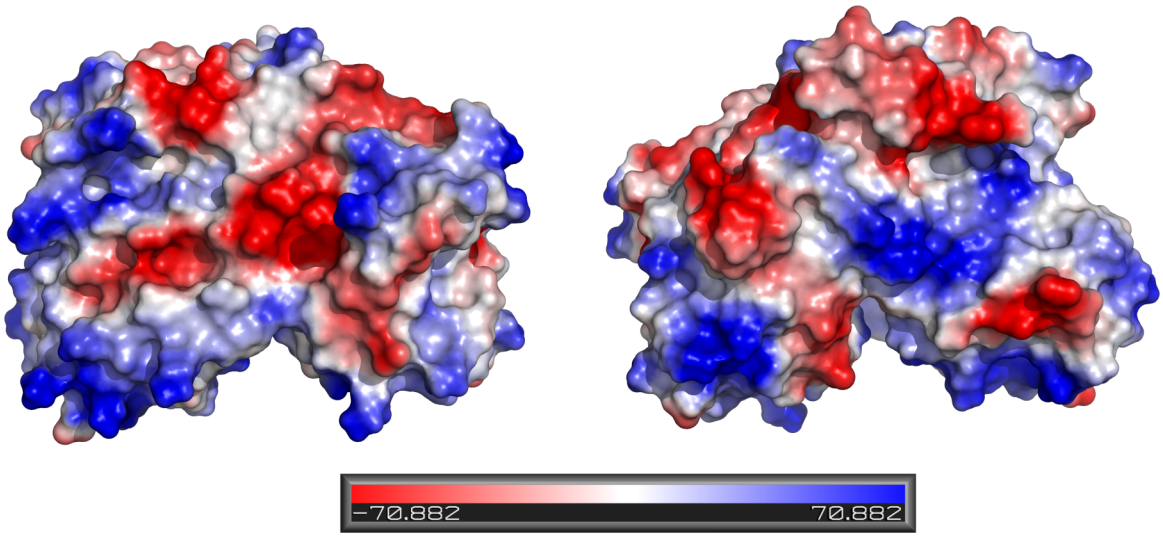


Figure 5. Qualitative electrostatic potential surface of HP1561. The picture on the right shows the opposite face of the tetramer. Figures are on the same orientation as of Figure 2.

Our structure does not have hemin included in crystallization solutions. HP1561 was also crystalized in the presence of $NiSO_4$ in the crystallization solution, since the homologous gene from *H. mustelae* was characterized to encode for a novel nickel and cobalt acquisition system (Stoof *et al.,* 2010). Three nickels were identified on the nickel anomalous map, one of which is present in the potential substrate-binding site, and other two are visible close to the protein surface Figure 6. None of the three Ni^{2+} ions are characterized by a specific coordination. They are in general close to one protein residue (for example, Nickel in the substrate-binding site interacts with Glu69) but the other potential coordination sites are occupied by acetate or water molecules.

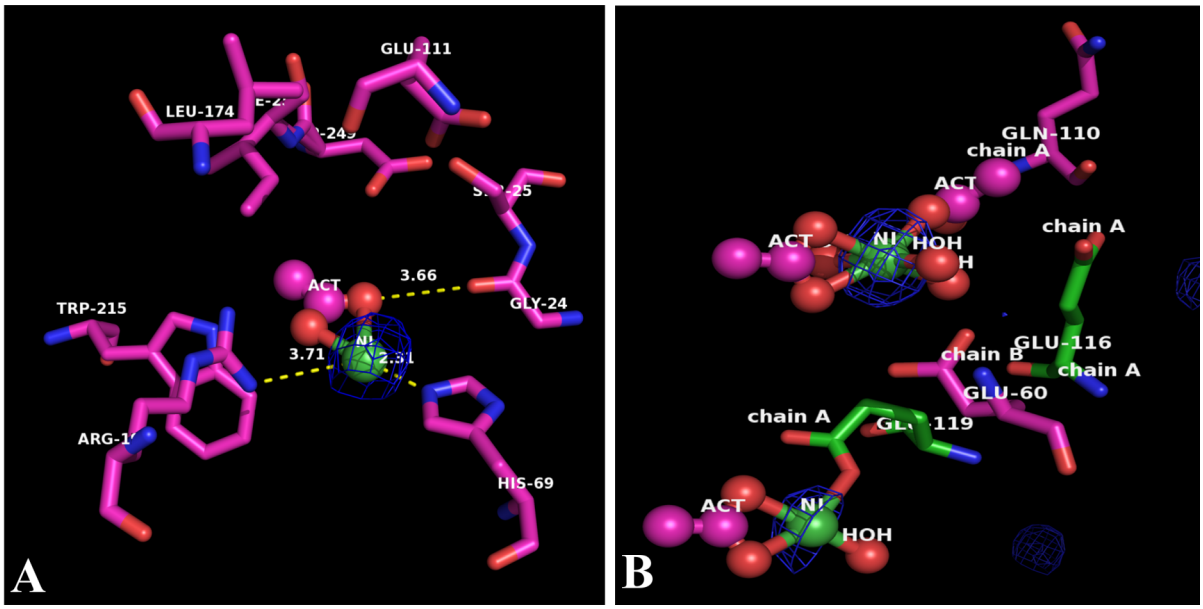


Figure 6. Nickel binding by HP1561. A. Nickel present in the substrate-binding site, which was coordinated by acetate and His 69. B. Unspecific Nickel binding in the surface of HP1561. One Nickel in the surface is coordinated by Gln 110, two acetate and two water molecules and Glu 60 from another chain, making a bridge between two chains. Glu 119 coordinates the second one, along with acetate and one molecule of water.

3.2 The heme binding cleft

Based on the structures deposited in the PDB, heme-binding proteins were classified by heme coordination motif. Heme binding can be quite variable, including often interactions of the Fe ion with His/His or His/Met, but even single Cys, His or Tyr can binds (Reedy *et al.,* 2007). In the crystal structure of PhuT and ShuT, heme is embedded in a cleft between the N- and C-terminal domains and is pentacoordinate, the fifth coordination site being represented by Tyr71 or Tyr67, respectively. In the structure of HP1561, assuming a similar way of binding, Tyr is replaced by Ser101, which is located in the loops between shhet β5 and helix α6. A tyrosine, Tyr79 (present in helix α3) is present in the cleft, but it is quite far. An arginine (Arg73) in the proximal pocket of PhuT and Lys69 in the case of the ShuT forms a H-bonds with the Fe ion, and this position in HP1561 is occupied by His103. In PhuT, Arg228 is stacked against the distal surface of heme, which is directly hydrogen bonded to one of the propionate group. In Hp1561, Ser252 took this position and there is another His197 present in the loop between β8 and β9, where the side chain of this histidine is shifted and extended to the position of Arg228 of PhuT (Figure 7A,B,C).

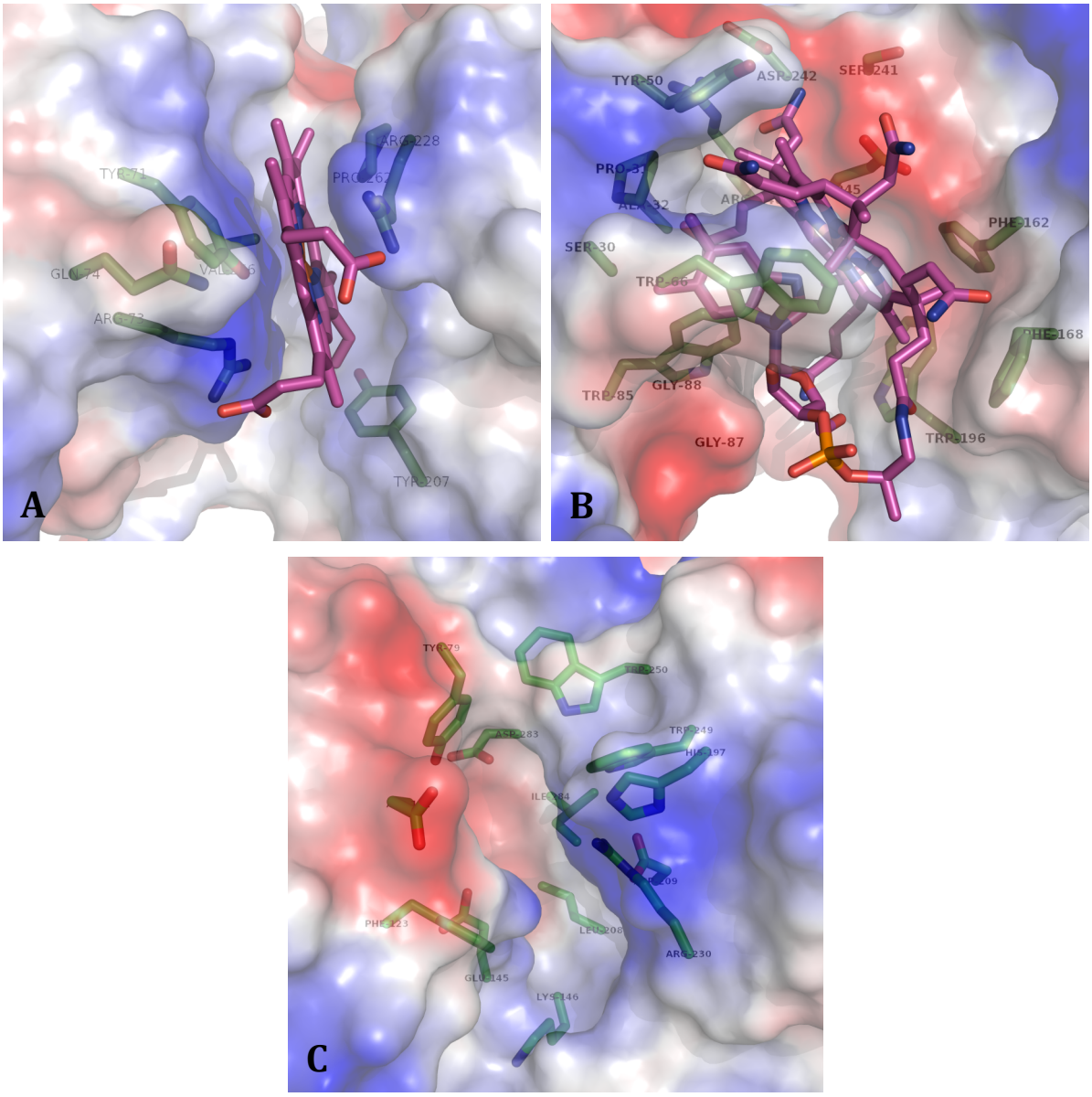


Figure 7. Comparison of substrate binding cleft. A. Heme binding site in PhuT, Periplasmic Heme Binding Protein from *Pseudomonas aeruginosa* (PDB: 2R79, Ho *et al.,* 2007). B. BtuF, periplasmic Vitamin B12 binding protein from *Escherichia coli* (PDB: 1N2Z, Borths et al, 2002). C. Putative residues for heme binding. Residues involved in substrate binding of PhuT and BtuF were identified plotting the ligand with software LigPlot. Electrostatic surface charge calculation performed with Pymol program).

A major differences in the heme-binding cleft between PhuT and HP1561 is that in the former there is a loop 10 residues long (260-270) which extends inside the heme binding cleft from the bottom, making the heme-binding site more compact. In HP1561 and ShuT, the residues corresponding to that loop are 278-286 and 250-260, respectively. Exceptionally, two hemes was reported to be bound in the heme-binding cleft of the similar homologue members HmuT from *Yersinia pestis* (PDB ID: 3NU1), which also presents a more extended groove to accommodate the heme (Mattle *et al.,* 2010).

3.3 Structure comparison

Although almost all of the secondary structure elements described for the BtuF, PhuT and ShuT are conserved in HP1561, in the latter there are some extra features present. Two antiparallel β-sheet (34-49) present in the N-terminal domain are absent in PhuT (PDB: 2R79), ShuT (PDB: 2R7A) and BtuF (PDB: 1N2Z), but present in the putative Fe(III) ABC transporter from *Thermotoga maritima* (PDB: 2ETV), in siderophore binding protein FeuA from *Bacillus subtilis* (PDB: 2XUZ), in siderophore Receptor HtsA from *Staphylococcus aureus* (PDB: 3Li2, 3EIW) and in ferric enterobactin binding protein from *Campylobacter jejuni* (PDB: 2CHU). Three additional helices, α2 (69-71), α15 (262-266) and α16 (168-273) are observed in HP1561, similarly to PhuT and ShuT. They are absent in BtuF. Amino acids residues 79-96, that together constitute three small consecutive helices (α3, α4 and α5), extended outside compared to BtuF, PhuT and ShuT. In the C-terminal domain there are additional four helices α18(302-306), α19(309-323), α20(325-330) and α21(333-334), which are present only in 2ETV along with two additional β-sheets. In fact, HP1561 is quite longer in sequence (335 amino acids) if compared to ButF (266), PhuT (308) and ShuT (277), but smaller than 2ETV (356). Even if the structure of HP1561 enjoys many similarities with siderophore-binding protein, *H. pylori* is unable to synthesize siderophores for iron uptake (van Vliet *et al.,* 2001; Dhaenens *et al.,* 1999).

Two loops at the surface of the N-terminal domain of the structure of HP1561 display quite high B-factors. These two loops are composed of residues 101-116 and 126-138 (Figure 8). The elevated B-factor was reported to be present in other members of this class III periplasmic binding proteins. Three loops with similar features are present at the surface of the C-terminal domain of the crystal structure of siderophore receptor HtsA from *Staphylococcus aureus* (PDB: 3LI2, 3EIW), which is described as open form of the protein, whilst residues in this three loops have lower B-factors in a close form of the protein,

crystallized in a different space group (See Figure 1 in Grig *et al.,* 2010). Referring to the crystal structure of HtsA, the *H. pylori* structure can be defined as an open form. Similar high B-factors in the residues at the surface of these two loops are present also in the periplasmic heme-binding protein ShuT and PhuT. In addition, the C-terminal domain of ShuT has a small-extended loop (residues 168-173) that presents very high B-factor and which extends over the substrate binding site. It can be speculated that this loop is flexible and acts as a gate for the substrate binding site. This extended extra loop is absent in the structure of HP1561 and makes the gate open for the entrance of the substrate. As to the release of substrate to the ABC transporter, SBP undoubtedly undergoes significant conformational changes (Borths *et al.,* 2002), but the long rigid interconnecting helix between the two domains in our protein is less likely to permit a strong domain movement during substrate binding and release. The flexibility of the exposed surface loops in the N-terminal domain might play a role in substrate translocation.

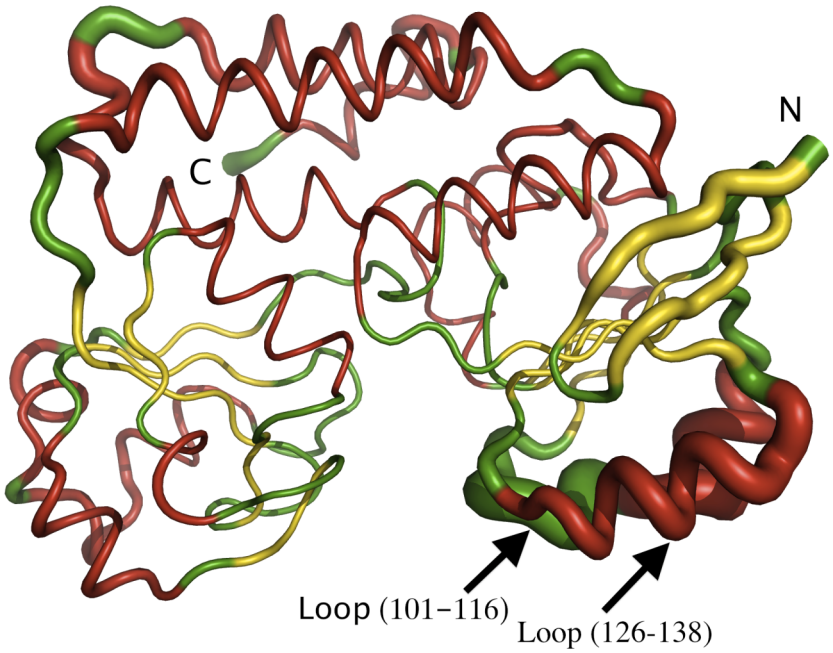


Figure 8. Ribbon drawing of the monomer of HP1561. The diameter of the ribbon tube is proportional to the thermal parameters of the atoms.

An extensive search with DALI server (Holm and Sander, 1993) shows that the most similar three-dimensional structures to HP1561 with other members of this family are reported in Table 2. Sequence alignment shows that the sequence similarity is very low between the members of the family, but, along with the two conserved, surface-exposed glutamates (E110 and E236), one in each lobe, three other residues (Pro 117, Glue 118 and Pro 243) also well conserved (Figure 9). Pro 117 and Glu118 are present in a loop between helix α6 and sheet β4. Pro 243 is present in a loop between helix α13 and sheet β11. In some way they probably denote the structural features of this family.

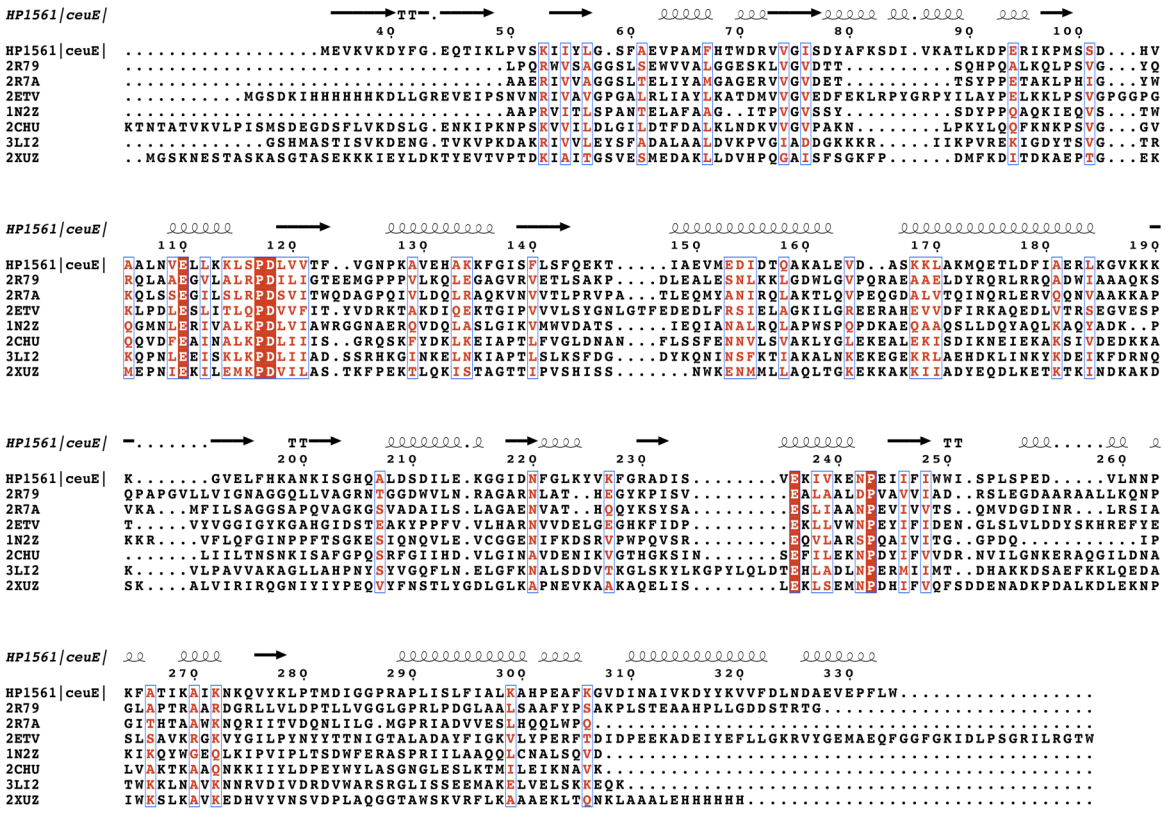


Fig. 9. Amino acid sequence of HP1561 protein aligned with PhuT, Periplasmic Heme Binding Protein from *Pseudomonas aeruginosa* (PDB: 2R79, Ho *et al.,* 2007); ShuT, Periplasmic Heme Binding Protein from *Shigella dysenteriae* (PDB: 2RG7, Ho *et al.,* 2007); putative Fe (iii) ABC transporter (tm0189) from *Thermotoga maritima* (PDB: 2ETV), Siderophore Receptor HtsA from *Staphylococcus aureus* (PDB: 3Li2; Grig *et al.,* 2010); BtuF, periplasmic Vitamin B12 binding protein from *Escherichia coli* (PDB: 1N2Z, Borths et al, 2002); FeuA from *Bacillus subtilis* (PDB: 2XUZ, Peuckert *et al.,* 2011); CeuE Ferric Enterobactin Binding Protein, from *Campylobacter jejuni* (PDB: 2CHU, Müller *et al.,* 2006). Secondary structure elements of HP1561 are drawn over the alignment.

Table 2. Comparison of structural similarities to HP1561 with other members of periplasmic substrate binding family.

Organism	Function	Z-score	Root mean square deviation (r.m.s.d.) Å corresponding residues are in parenthesis	PDB ID	References
Thermotoga maritima	Putative Fe(III) ABC transporter	22.2	4.2 (292)	2ETV	-
Pseudomonas aeruginosa	PhuT, periplasmic heme binding protein	22.1	2.5 (246)	2R79	Ho *et al.,* 2007
Bacillus subtilis	Siderophore binding protein	21.3	3.0 (261)	3GFV	Zawadzk *et al.,* 2009
Shigella dysenteriae	ShuT, periplasmic heme binding protein	21.1	2.7 (238)	2RG7	Ho *et al.,* 2007
Staphylococcus aureus	HtsA, Siderophore receptor	20.8	3.1 (254)	3LI2	Grig *et al.,* 2010
E. coli	BtuF, Vitamin B12 binding protein	20.5	2.3 (231)	1N2Z	Borths et al, 2002
B. subtilis	FeuA	20.4	2.6 (252)	2XUZ	Peuckert *et al.,* 2011
Campylobacter jejuni	CeuE, ferric enterobactin binding protein	20.2,	3.0 (253)	2CHU	Müller *et al.,* 2006
Yersinia pestis	HmuT, periplasmic heme binding protein	19.1	3.3 (231)	3NU1	Mattle *et al.,* 2010

A gene duplication is present for this gene in two strains of *H. pylori,* 26695 and J99. In 26695, HP1562 share 86% sequence identity with HP1561, whilst in J99 jhp1470 shares 87% identity with jhp1469. In strain G27 this gene duplication is not present. Homology model of HP1562 was built from SwissModel server and absolutely no change was observed in the heme-binding groove.

3.4 Binding Assay

HP1561 belongs to a family of periplasmic substrate-binding proteins involved in transport to the cytoplasm of a wide range of substrates through the ABC transporters. The 3D structure presents a similarity to the sub-family of periplasmic heme-binding proteins, periplasmic siderophore receptor and Vitamin B12 binding proteins. Since *H. pylori* is unable to produce siderophores, the binding assays were performed with Hemin and Vitamin B12. Binding of Hemin (solubilized with 1M NaOH and pH adjusted to 8.0 with Tris buffer) and Vitamin B12 to purified HP1561 was assessed. Hemin binding was measured with two different methods: (i) quenching of intrinsic protein fluorescence and (ii) isothermal titration calorimetry (ITC). The intrinsic protein fluorescence of HP1561 decreases after the addition of Hemin until saturation is achieved (Figure 10 A). This decrease reflects the number of tryptophan residues

in HP1561 (4 in total, of which 2 are present in the binding cavity). The dissociation constants for hemin binding to HP1561 can be evaluated as Kd 1.47 ±0.08 μM (Figure 10 B,C). Saturation was not achieved with vitamin B12 even when the concentration was increased to 15 μM, but the decrease of the intrinsic protein fluorescence can be interpreted as a weak binding. Dissociation constant for vitamin B12 was calculated as 4.47±0.25μM. ITC was consequently used to determine the thermodynamic parameters contributing to the ligand binding of hemin only. A representative measurement, corrected for blanks, was fitted assuming a one site binding model, yielding a dissociation constant of 3.30±1.4μM (Figure 10 D). From the heme protein database, the weakest heme dissociation constant was reported to be 2.5 mM and the tightest one 300 pM, indicating that our value is still inside the interval. Heme binding affinity constants are quite various and they probably depend on the family of proteins that bind heme and their functional role (Reedy *et al.,* 2007). It must be considered that a weak bionding constant is in line with the role of this proteins, since they have to bind and also release heme. On the other hand, the dissociation constants for ferric siderophore-receptor complexes present higher values if compared to those of the heme transporter protein (Grig *et al.,* 2010).

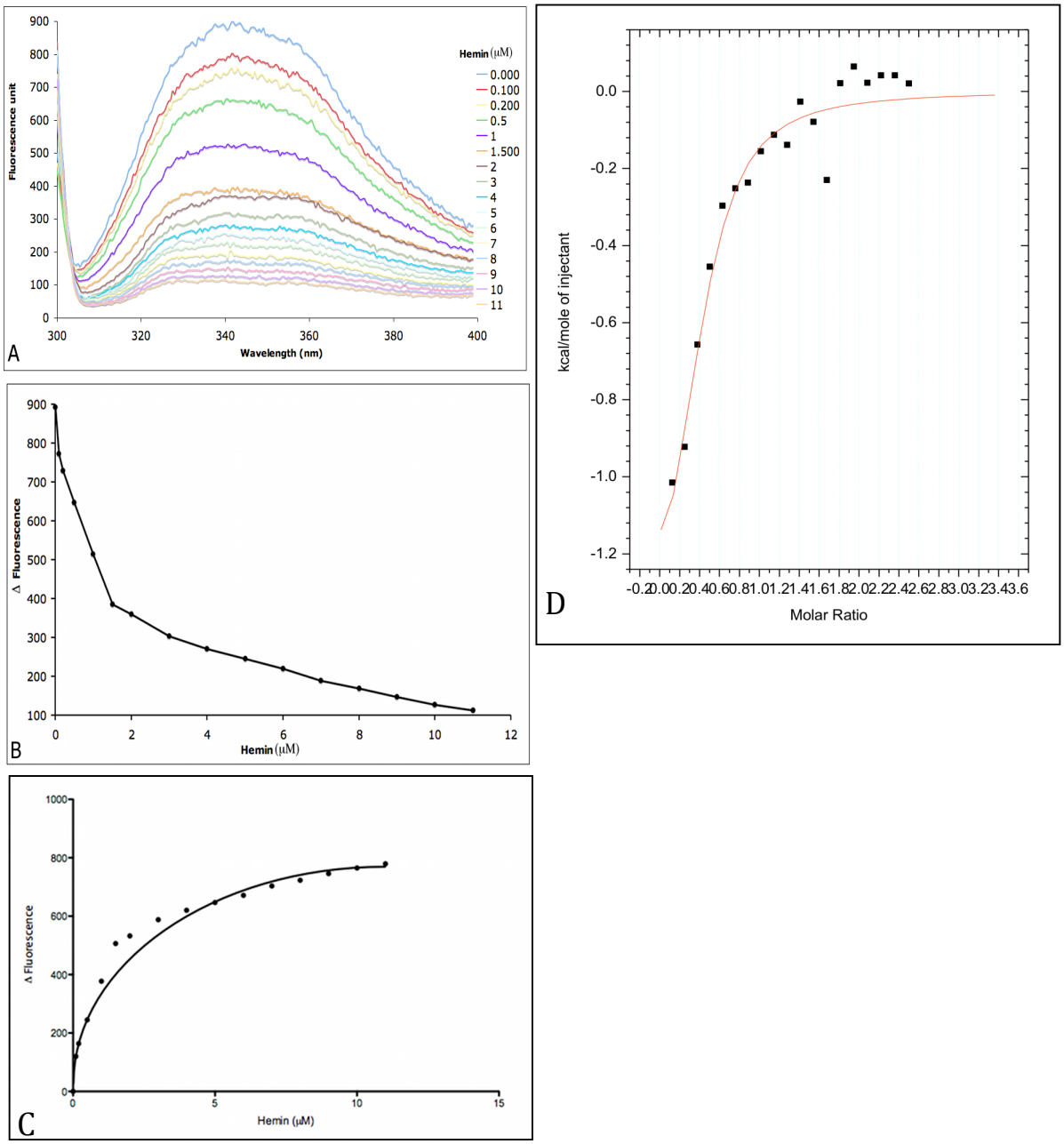


Figure 10. Fluorescence titration and Isothermal Titration Calorimetry. Binding was measured using intrinsic protein Fluorescence (A, B), A. Fluorescence scan of 5μM of HP1561 with different concentration of Hemin (0.1 to 11μM), scan range from 300 to 400nm wavelength. B. The fluorescent measurements were plotted as the absolute change from the initial fluorescence signal in the absence of substrate (panel A & B); Fluorescence titration was made by adding different concentration of Hemin (0.1 to 11μM) to HP1561 and the data were collected at 340nm fixed wavelength for 120 seconds. C. Fluorescence data fitted for a single site substrate binding with GraphPad Prism 5 to calculate the dissociation constant. In

all the cases the excitation was at 295nm. D. Hemin binding to HP1561 was measured by VP-ITC Microcalorimeter (GE Healthcare) at 250C. Titration of HP1561 was made with Hemin by the addition of HP1561 to the ITC-cell containing 20μM of Hemin at 25 injections of 10μl HP1561 (350μM) with interval of 500seconds. The integrated heat peaks were fitted to a one site-binding model with Origin 5 software.

4. Conclusion

Pathogenic bacteria are adapted to use different host-derived iron sources and *H. pylori*, which colonize the gastric mucosa, has developed specific and peculiar iron transport pathways. Periplasmic heme-binding proteins have been identified and characterized from several different pathogenic bacteria. *H. pylori* can use heme as the sole iron source (Husson *et al.,* 1993; Worst *et al.,* 1995; Velayudhan *et al.,* 2000), but until now no protein was identified for heme binding. We have demonstrated that HP1561 (ceuE) in *H. pylori* functions as an ABC transporter-associated substrate binding protein for heme, and that this protein is not involved in Nickel or Cobalt acquisition. The coordinating residues for heme are not well conserved among the members of this family, suggesting that heme coordination is different. The determination of the HP1561 structure provides valuable information of its role in heme transport and for a deeper understanding of ABC transported mediated transport of substrates.

Acknowledgments

We thank the staff of beamline BM14 and ID29 of ESRF, Grenoble, for technical assistance during data collection. This work was supported by the University of Padua. Md Shaik's position was supported by CARIPARO grants for foreign Students at the University of Padua. Thanks to Elena Ravazzolo for her cordial help during ITC experiment.

Protein Data Bank accession codes. The coordinates and structure factors have been deposited in the Protein Data Bank (http://www.pdb.org) for immediate release with PDB ID code 3UI8, 3UI9.

Chapter Five

Crystal Structure of Glutamine Synthetase (HpGS) from *Helicobacter pylori*

Introduction

Glutamine synthetase (GS, EC 6.3.1.2) catalyzes the synthesis of glutamine, a central intermediate in nitrogen metabolism, from ATP, glutamate, and ammonia in a divalent metal ion dependent reaction (Colombo and Villafrancas, 1996). Three types of glutamine synthetase were identified: GS I, expressed mainly in prokaryotes (*Streicher et al.,* 1975), GS II, found in eukaryotes and in some bacteria belonging to Rhizobiaceae, Frankiaceae, and Streptomycetaceae (Behrmann *et al.,* 1990; Hillemann *et al.,* 1993) and GS III, found in *Bacteroides fragilis* and in *Butyrivibrio fibrisolvens* (Crespo *et al.,* 1998). While GS II also is common to both eukaryotes and some bacteria, GS I has been found only in eukaryotes (Kumada *et al.,* 1993). Glutamine is used to synthesize many nitrogenous compounds (alanine, glycine, serine, histidine, tryptophan, CTP, AMP, carbamoyl-phosphate, and glucosamine 6-phosphate) in which its amide group serves as the nitrogen donor (Reitzer, 1996). GS plays a central role in nitrogen assimilation and the expression of the corresponding gene and its activity are highly regulated by the use of multiple promoters (s70 and s54), positive activators (NRI and NRII), post-translational adenylation, and allosteric inhibition by nitrogenous compounds. In *E. coli*, the enzymatic reaction catalyzed by GS I is inhibited by nine end products of glutamine metabolism (serine, alanine, glycine, AMP, CTP, tryptophan, histidine, carbamoyl phosphate and glutamine-6-phosphate) (Woolfolk and Stadtman, 1967). The posttranslational modification of GS by adenylation to a tyrosine residue (consensus sequence NLYDLP), is generally a key mechanism for turning off the activity of the enzyme or makes the enzyme more sensitive to other feedback inhibitors (Shapiro and Stadtman, 1968).

Helicobacter pylori , the etiologic agent of gastritis and peptic ulcer in humans, colonizes the stomach of more than half of the world's population (Rothenbacher and Brenner, 2003). It produces urease as one of its most abundant protein components (10–15% of total proteins by weight) (Phadnis *et al.,* 1996), which hydrolyzes the host urea and liberate ammonia. Urease has buffering properties and is essential to the protection of *H. pylori* against gastric acidity, neutralizing the bacterial microenvironment. Ammonia, which is also a preferred nitrogen source for *H. pylori*, is assimilated into protein and other nitrogenous compounds (Williams *et al.,* 1996). The only single nitrogen incorporation pathway which is reported to be present in *H. pylori* (Reitzer, 1996) is mediated by GS encoded by gene *Hp0512* (*glnA*). The *H. pylori* glutamine synthetase has been shown to complement an *E. coli* GS deficient mutant. It was impossible to obtain a *glnA*-knockout mutant, indicating that Hp0512 is an essential gene

(Garner *et al.,* 1998). GS from *H. pylori* lacks the highly conserved target tyrosine residue within the adenylation motif and no consensus sequence for σ54-dependent transcription factor has been found upstream of the *H. pylori glnA* gene. This apparent lack of control seems to be necessary to handle the large amount of ammonia deriving from the urease activity, not regulated or regulated only through a feedback inhibition (Garner *et al.,* 1998). GS from *H. pylori* presents 49% and 47% identity with the corresponding *Salmonella typhimorium* and *E. coli* homologs respectively. The crystal structure of the former is reported (Yamashita *et al.,* 1989) and its feedback inhibitory mechanism well studied (Liaw *et al.,* 1993). The crystal structure of several other GS from different organism is known, including *Mycobacterium tuberculosis* (Krajewski *et al.,* 2005), *Saccharomyces cerevisiae* (He *et al.,* 2009), Zea mays (Unno *et al.,* 2006) and *Canis familiaris* (Krajewski *et al.,* 2008).

The catalytic mechanism of this enzyme has been quite well established, along with the differential regulation with different end product of the metabolic pathway and allosteric regulation by adenylation. Absence of allosteric regulation site (adenylation site) and key enzyme in the single nitrogen assimilation pathways makes *Hp*GS an interesting subject for structural study. In order to elucidate the structural features of the synthetase and possible regulatory mechanisms, glutamine synthetase (*Hp*GS) from *H. pylori* was cloned, expressed and purified in good yield in *E. coli*, crystallized, and its structure determined.

Materials and Methods

Molecular cloning of HP512 gene

HP0512 gene was amplified by PCR from *H. pylori* G27 genomic DNA, using *pfu* proofreading DNA polymerase (Finnzymes, Finland), with the following primers: CACCATGATAGTAAGAACTCAAAATAGTGAAAG (Forward, Topo isomerase recognition site underlined) and TTATTAGCATGAATAAGTGGTGATAAATTC (Reverse). The amplified fragment was cloned into the pET151vector (Invitrogen) in frame with a N-terminal His-tag, using a TOPO® Cloning kit by Invitrogen to obtain the pET151-HP0512 plasmid. The correct insertion in the cloned vector was confirmed by colony PCR and sequencing with T7 forward and reverse primer.

Overexpression and affinity purification of HP0512 protein

E. coli BL21(DE3) cells (Novagen) were transformed with the pET151-HP0512 plasmid and grown in LB media at 37° C, supplemented with Ampicillin (100 µg/ml). Protein expression

was triggered by 1 mM isopropyl-β-D-thiogalactoside (IPTG, Inalco) when the culture reached an optical density (OD_{600}) of 0.7. After 4h incubation at 28° C, bacteria were collected and resuspended in a lysis buffer (30 mM Tris, pH 8.0, 150 mM NaCl) and then disrupted by a One Shot Cell breakage system (Constant System Ltd., UK). The lysate was centrifuged to remove cell debris (18,000 g for 25 min) and loaded into a column containing 5 ml of Ni^{2+} charged Chelating Sepharose™ (GE Healthcare, UK). After extensive washing using the lysis buffer, supplemented with 20mM imidazole, the protein was eluted by a linear gradient from 80 to 300 mM imidazole. The protein was further purified by gel filtration, using a Superdex 200™ 16/60 GL (GE Healthcare) equilibrated with 30 mM Tris pH 7.5, 150mM NaCl. HP0512-His tagged eluted as a single peak. All the protein samples collected throughout the purification were separated and analyzed on 15% sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE). The resolved gels were stained with 0.25% Coomassie Brilliant Blue R250 reagents.

Molecular mass determination

The native molecular mass of HP0512 protein was determined by analytical gel filtration by analytical column Superose 6 10/300 GL (GE Healthcare), equilibrated with 30 mM Tris pH 8.0, 150mM NaCl. His tagged-HP0512 eluted as a single peak, (Figure 1). The molecular mass estimated from analytical gel filtration is 416.87 kDa, which differs from the calculated one for the dodecamer, 654 kDa. This discrepancy is probably due to the two plates-like hexameric ring structure of the 12 monomers. The presence of single species in solution was also proved by gel filtration. The system was operated on an AKTA FPLC instrument (GE Healthcare).

Circular Dichroism (CD) and UV Analysis

The CD spectrum was measured with a J-715 spectropolarimeter (JASCO, Corporation) at 298K. 10 runs were accumulated with 1mg/ml protein (Tris 3mM, NaCl 10mM, pH 8.0) in a 0.2mm path cuvette in the wavelength interval 190-260nm. The CD spectrum was rescaled with respect to a standard solution containing the buffer and the molar ellipticity calculated (Figure 2). The CD spectrum was deconvoluted with software “CD Spectra Deconvolution” (CDNN 2.1).

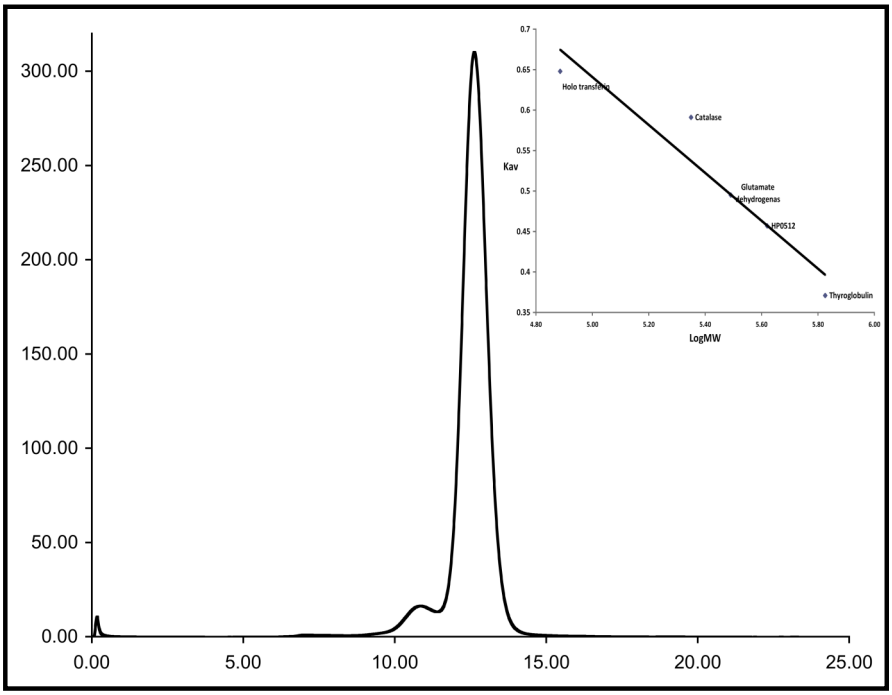


Figure 1. Analytical gel filtration of HpGS. Protein was purified by Superose 6 10/300 GL (GE Healthcare), equilibrated with 30 mM Tris pH 8.0, 150 mM NaCl. His tagged-HpGS eluted as a single peak.

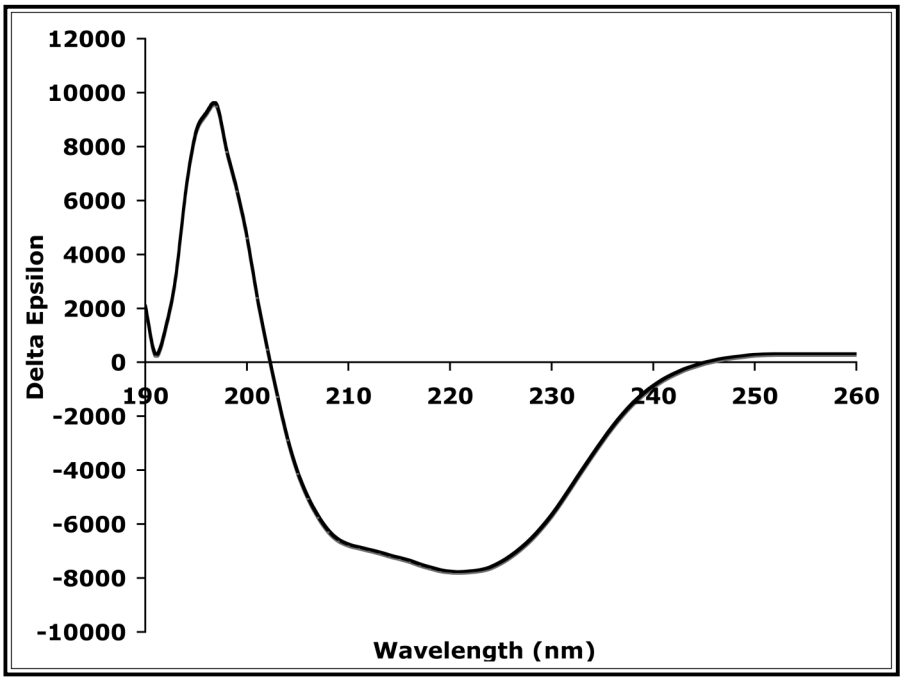


Figure 2. Circular dichroism spectrum of HPGS. Spectrum was measured with J-715 spectropolarimeter (JASCO, Corporation) at 298 K.

Crystallization and structure determination

The purified protein was concentrated to 20 mg/ml and used for crystallization tests, partially automated using an Oryx 8 crystallization robot (Douglas Instruments). Crystals grew in several conditions. The best crystals were obtained at 20°C by vapor diffusion technique using a 10 mg/ml protein stock solution and, as precipitant, a solution containing 0.1M Tri pH 8.5, 12% (w/v), PEG 4000 (PEG II crystal screen, solution n. 30, Qiagen, USA). Crystals could be processed as triclinic, space group P3, with **a**=**b**=132.81 Å, **c**=105.57 Å. One tetramer is present in the asymmetric unit, with V_M= 2.98Å^3/Da corresponding to an approximate solvent content of 50%. The data set used in the final refinement was measured at the beamline ID23-1 of the European Synchrotron Radiation Facility, Grenoble, France. Frames were indexed and integrated with software Mosflm (Leslie, 2006) and data merged and scaled with Scala (Evans, 2005), contained in the CCP4 crystallographic package (Collaborative Computational Project, Number 4, 1994). The structure was solved by molecular replacement using software Phaser (McCoy *et al.,* 2007) starting from a monomer model built by the SWISS-PDB server using as a template the crystal structure Glutamine synthetase from *S. typhimurium* (PDB ID:2GLS, Yamashita *et al.,* 1989). The refinement was carried on using Phenix (Adams *et al.,* 2010), and Refmac (Murshudov *et al.,* 1997). Manual rebuilding was performed with graphic software Coot (Emsley and Cowtan, 2004). Data collection and preliminary refinement statistics are summarized in Table 1.

The final model contains four protein monomers. The final crystallographic R factor is 0.2490 (R_{free} 0.3100). Geometrical parameters of the model were checked with software Procheck (Laskowski *et al.,* 1993).

2.6 Bioinformatics

Search and comparison of genetic clusters containing *HP*GS and homologous genes were conducted using the Microbesonline web server (http://microbesonline.org). Structure similarity search was performed with the Dali server (Holm and Rosenstrom, 2010). Sequence alignments were visualized with Espript (Gouet *et al.,* 1998) and carried out with Clustalw (Thompson *et al.,* 1994), with manual adjastements based on structure comparisons. Phylogenetic analysis was performed using the neighbor-joining method (Saitou and Nei, 1987) implemented in Clustalw, and the resulting unrooted tree was visualized as radial layout with FigTree (http://tree.bio.ed.ac.uk). The absence of a signal peptide in the sequence

of *HP*GS and its close homologs was predicted using the SignalP server (http://www.cbs.dtu.dk/services/SignalP).

Table 1. Data collection and preliminary refinement statistics. A wavelength of 0.9765 Å was used. Rotations of 1° were performed.

Data collection *	
Space group	P3
Cell dimensions	
a, b, c (Å)	132.81, 132.81, 105.57
Resolution (Å)	52.78-3.5 (3.69-3.50)
Rsym or Rmerge	0.432 (0.873)
<I /σ(I)>	2.1 (1.1)
Completeness (%)	96.7 (97.9)
Redundancy	3.4 (3.4)
Refinement	
No. reflections	25394
Rwork / Rfree	0.2490 / 0.3100
No. atoms	
Protein	14992
Ion ligand	-
Water	-
R.m.s. deviations	
Bond lengths (Å)	0.013
Bond angles (°)	1.887
Ramachandran plot (%)	
Favored	60.6
Allowed	22.1
Outliers	17.3

Results and Discussion

Overall structure

The crystal structure of the monomer, despite the relatively limited resolution, 3.5 Å, and the high R_{merge} in higher resolution shells, appears quite reliable and comparable to other Glutamine synthase structures. The electron density for the protein main chain atoms is clearly visible from residues 12 to 480. The eleven N-terminal and one residue at the C-terminus are missing in the crystallographic model. The secondary structure elements of the monomer are illustrated in Figure 3, and their topology in Figure 4. The structure of the monomer consists of a smaller N-terminal domain (residues 1–112) and a larger C-terminal domain (residues 115–480). The N-terminal domain consists of three α helices (α1-3) and the C-terminal domain contains twelve α helices (α4-15) and ten β-strands (β1-10), in a total of fifteen α helices and ten β-strands per monomer (Figure 4).

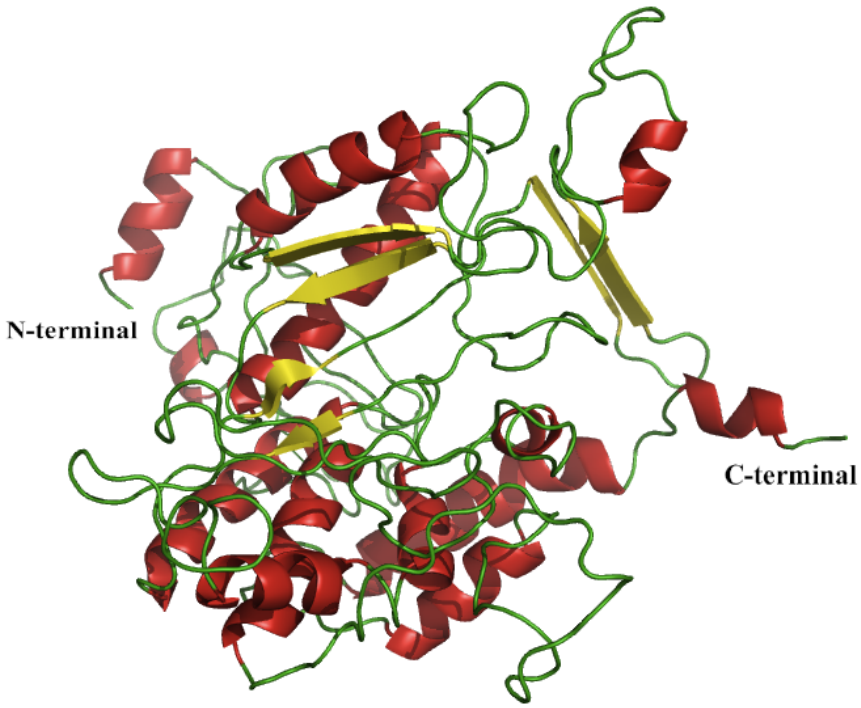


Figure 3. Cartoon view of *HP*GS monomer. □-helices are in red, □-strands in yellow, others in green. N and C-terminus are labeled. (This figure was produced with Pymol program (Delano, 2008).

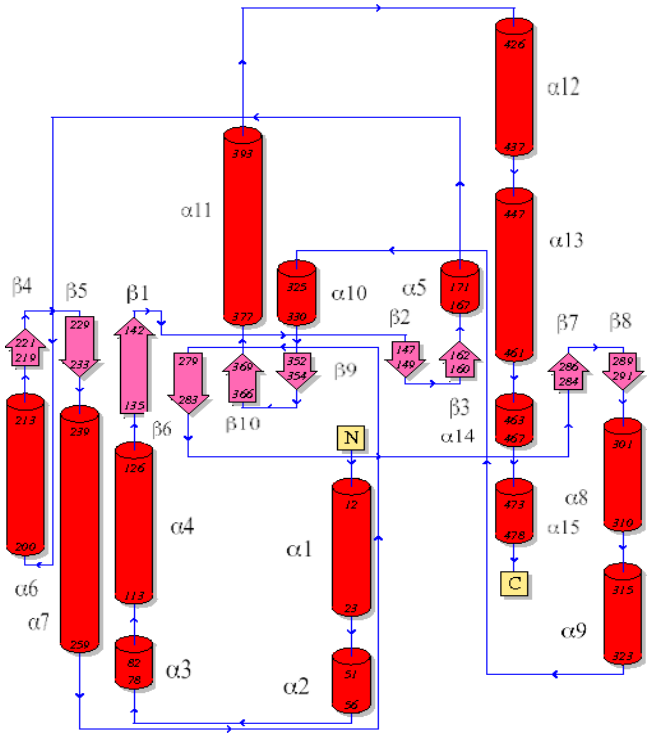
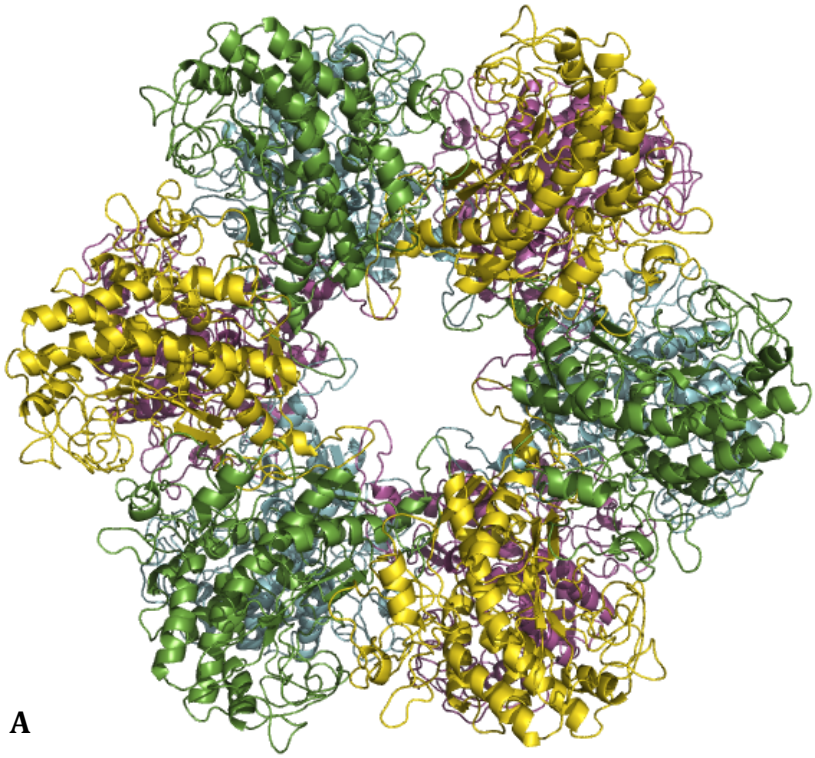


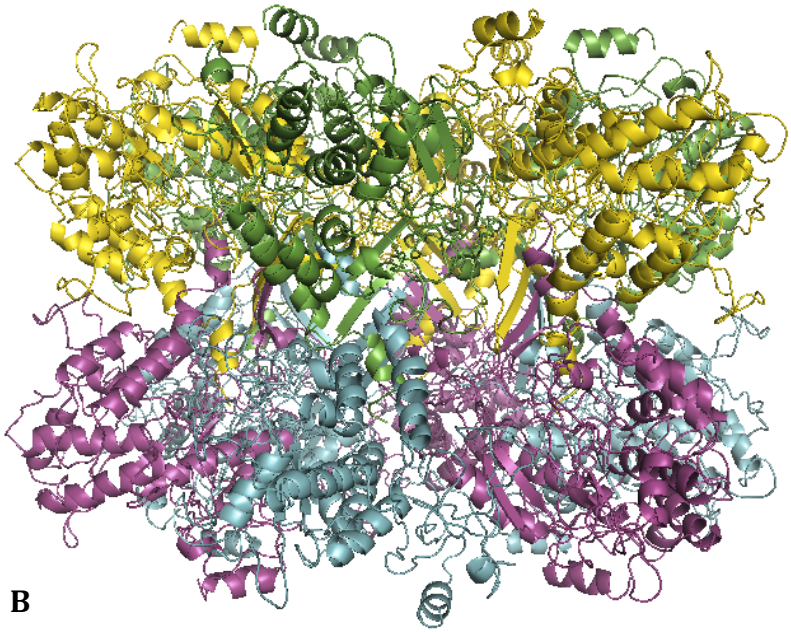
Figure 4. Topology diagram of *HP*GS. Helices and strands are in red and magenta, respectively; loops are in blue. (This figure was produced with PDBSum server).

The tertiary structure

The *HP*GS dodecamer is held together mainly by hydrophobic and hydrogen bonding interactions between the two hexameric rings, as in the case of GS from *S. typhimurium* (Almassy *et al.,* 1996). The N-terminal helix sits above the hexameric ring and is exposed to solvent. The C-terminal helix, called the `helical thong,' is inserted into a hydrophobic hole in the eclipsed subunit on the opposite hexameric ring. The helical thong is visible as the large rod-like structure extension from the bottom of the isolated subunit. In addition, the central channel of the dodecamer is lined by six four-stranded L-sheets, each built from an antiparallel loop (residues 147-162) contributed by subunits in opposite rings. The 12 thongs and six sheets give the dodecamer additional adhesion (Figure 5). At variance of bacterial GS, Eukaryotic GS are decamers. The maize GS crystal structure is a decamer, composed of two face-to-face pentameric rings of identical subunits, with a total of 10 active sites, each formed between every two neighboring subunits within each ring (Unno *et al.,* 2006).



A



B

Figure 5. The dodecameric assembly of *HP*GS. A. Top view of the overall dodecameric structure, B. Side view of the overall dodecameric structure.

The structure of the dodecamer exposes several loops, which are believed to have functional relevance. One loop, which consists of hydrophilic residues 167-185, protrudes into the central channel of the dodecamer and is a potential site for proteolysis (Lei *et al.,* 1979). Another loop is the adeadenylylation loop, so called since it contains the tyrosyl residue 397 in *St*GS and *Mt*GS. The latter is covalently modified by addition of AMP (Eisenberg *et al.,* 2000). This loop sits just outside the bottom entrance of the bifunnel.

Active site

Each monomer's active site can be described as a 'bifunnel', in which ATP and glutamate bind at the opposite ends. The ATP binding site is at the top of the bifunnel, where it opens to the external 6-fold surface of GS. At the junction of the bifunnel two divalent cation binding sites are present (Eisenberg *et al.,* 2000). As the metals are missing in our structure, the cation binding site can be analyzed only by comparing the structure of *HP*GS with that of S. *typhimurium* (PDB: 2GLS). The outer metal binding site is coordinated by Glu 139, His 279, His 281 and Glu 367, whereas the inner metal binding site by Glu 141, Tyr 190, His 221, Glu 223 and Glu 230. Both cations are reported to be important to stabilize the structure, as well as for catalysis. The inner one is possibly involved in phosphoryl transfer (Hunt *et al.,* 1975), while the outer ion stabilizes an active GS (Shapiro *et al.,* 1968) and plays a role in binding the glutamate (Hunt and Ginsburg, 1980). The types of residues coordinating the ions were found to be conserved in all GS characterized to far. In GS from S. *typhimurium* (PDB: 2GLS), the affinity for the metal ion at the inner site is 50 times greater than at the outer site (Hunt *et al.,* 1980). This can be attributed to the greater negative charge toward the bottom half of the bifunnel, in the vicinity of the inner cation site. In the structure of the *HP*GS there are two serine residues present toward the bottom of the bifunnel, in close proximity of the inner ion binding site. This could suggest that the binding affinity could also be higher than in other GS (Figure 6). The ATP-binding motif GDNGSG (residues 272 to 277) exactly matches the sequence the sequence found in other GS.

Significant differences are present in our structure compared with GS from Maize. In the latter, in fact, the active sites (20 Å deep) are formed between two neighboring monomers in a ring where its opening is roughly parallel to the 5-fold axis. Three Mn^{2+} ions lie at the middle of the cleft, whereas only two Mn^{2+} ions are present in bacterial GS (Unno *et al.,* 2006).

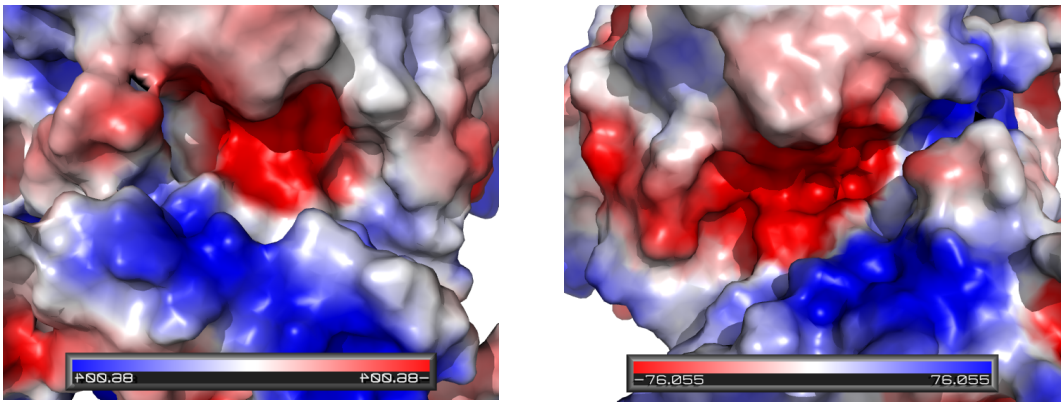


Figure 6. Charge distribution in the Active site. Left, *Hp*GS, right, *St*GS, A and B are the inner and outer metal binding site respectively.

Structure comparison

Bacterial GS molecules are dodecamers formed by two face-to-face hexameric rings of subunits, with 12 active sites formed between monomers (Almassy *et al.,* 1996; Gill *et al.,* 1999). This is also true in the case of *Hp*GS. The overall fold of the *Hp*GS structure is similar to those of GS from *M. tuberculosis and S. typhimurium*. The topology of *Hp*GS can be compared to those of GS from *S. typhimurium*, *M. tuberculosis*, maize and human. Sequence identities and r.m.s.d. between HpGS and GS from other species are reported in Table 2, which indicate that the GS from both bacterial and mammalian originates from common ancestor and diverged later (Figure 7).

Table 2. Sequence identities and r.m.s.d. between HpGS and GS from other species.

Organism	**Sequence identity (%)**	**Root mean square deviation (r.m.s.d.) Å corresponding residues are in parenthesis**	**PDB ID**	**References**
S. typhimurium	49	1.2 (463)	2GLS	Yamashita *et al.,* 1989
M. tuberculosis	40	1.7 (455)	1HTO	Gill *et al.,* 2002
Synechocystis sp.	43	1.5 (454)	3NG0	-
Canis Familiaris	18	3.3 (329)	2UU7	Krajewski *et al.,* 2008
Zea mays	19	3.3 (320)	2D3A	Unno *et al.,* 1006
Homo sapiens	18	3.4 (320)	2OJW	Krajewski *et al.,* 2008

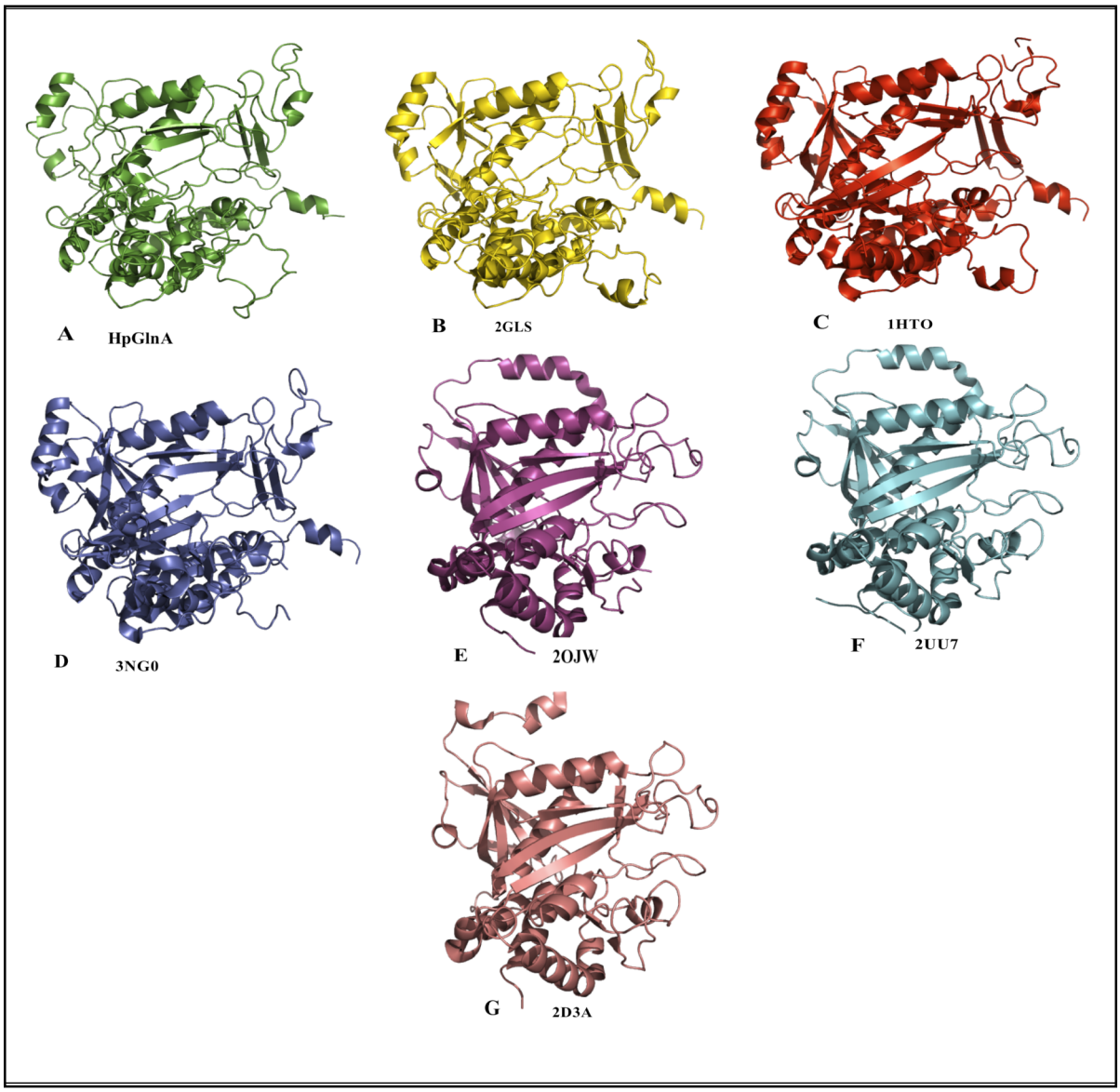


Figure 7. Cartoon view of GS from different origin. A. GS from *H. pylori* (*Hp*GS), B. GS from *S. typhimurium*, C. GS from *Mycobacterium tuberculosis*, D. GS from *Synechocystis sp.* E. GS from Human, F. GS from *Canis Familiaris*, G. GS from Zea mays.

The C-terminal residues (residues 393–480) of *Hp*GS are missing in the in maize and human GS, they play a role as the adenylylation site in the other bacterial GS, where they also contribute to intimate interactions between two hexamer rings. The C-terminal residues (residues 393–480) of *Hp*GS contribute to intimate interactions between two hexamer rings (helical thong). Most of residues for binding substrate and divalent ions are well conserved between bacterial GS and maize/human hence the enzyme reaction mechanism of *Hp*GS is likely to be essentially the same as that of the proposed for bacterial GS or eukaryotic GS (Figure 8). When aligned, the most similar homolog, *S. typhimurium* GS, was found to have

48.3% amino acid sequence identity and 65.3% amino acid sequence similarity (identical residues and conservative replacements) (Figure 8).

Adenylation site and regulation of HpGS

The adenylation site found in all of the most similar homologs of *Hp*GS contains the consensus sequence NLYDLP. The latter is replaced in *H. pylori* by NLFKLT (residues 405 to 410). Since the Tyr407 residue is the well-conserved target of adenylation (Shapiro and Stadtman, 1968) (Figure 9) and the *H. pylori* glutamine synthetase lacks that residue, no adenylation can occurs for this enzyme within this motif (Garner *et al.,* 1998). Another tyrosine containing sequence, NPYLAF (residues 376 to 381), well conserved among homolog enzymes, is present upstream of the consensus adenylation site. Consistent with the observation that *H. pylori* lack adenylation site, homolog proteins of GlnB (regulatory protein PII), GlnD (uridylyltransferase/uridyl-removing enzyme) and GlnE (adenylyltransferase) which are all required for posttranslational modification of glutamine synthetase to carry out adenylation in other gram-negative enteric bacterial species do not appear to be present in the *H. pylori* genome (Garner *et al.,* 1998).

The lack of such control is an exception rather than the rule among bacterial glutamine synthetases, since the enzymes from most species have rigorously conserved the Tyr containing adenylation site (conserved in all 10 of the most closely related GlnA homologs). But in the protein database search revealed some bacterial species where the GS doesn't goes for adenylation as they lack this converved Tyr residue, Phe: *Clostridium acetobutylicum* (Usdin *et al.,* 1986), *Lactobacillus delbrueckii* (Ishino *et al.,* 1992), and *Methanococcus voltae* (accession no. P21154).

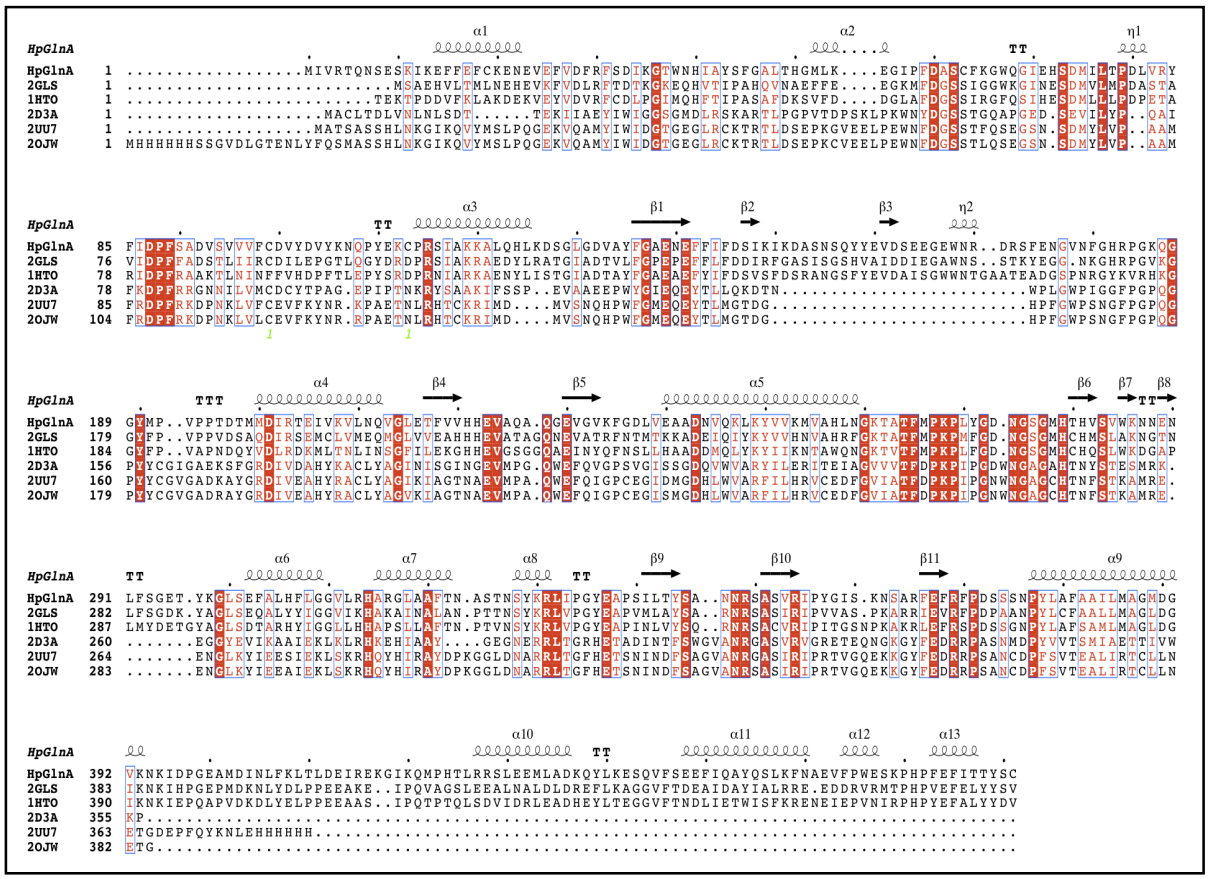


Figure 8. Amino acid sequence of *HP*GS protein aligned with GS from *S. typhimurium* (PDB: 2GLS, Yamashita *et al.,* 1989), GS from *Mycobacterium tuberculosis* (PDB: 1HTO, Gill *et al.,* 2002), GS from *Canis Familiaris* (PDB: 2UU7, Krajewski *et al.,* 2008), GS from Zea mays (PDB: 2D3A, Unno *et al.,* 1006), GS from human (PDB: 2OJW, Krajewski *et al.,* 2008).

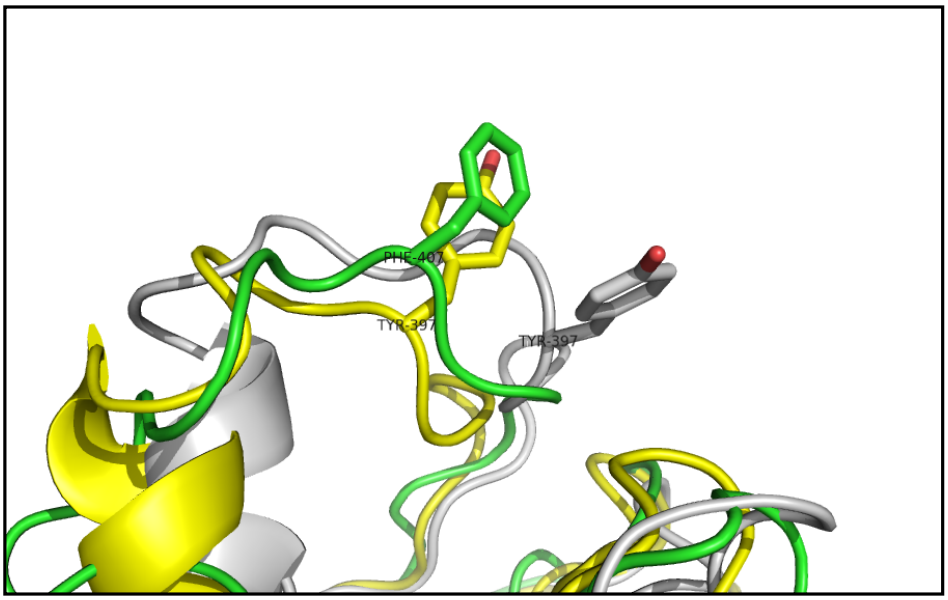


Figure 9. Comparison of Adenylation loop. S. *typhimurium* (PDB: 2GLS) (yellow, adenylation site Tyr 397); *M. tuberculosis* (PDB: 1HTO) (white, adenylation site Tyr 397); *Hp*GlnA in green (corresponding residue Phe 407).

To search physically in the genome of the *H. pylori* for other regulatory genes upstream or downstream of the gene for *Hp*GlnA, no s70 or s54 promoter sequence is readily identifiable, nor is a rho-independent transcriptional terminator located downstream of the end of the gene. Analysis of sequences upstream (573 bp) and downstream (451 bp) of *Hp*GS did not reveal any homologs related to ammonia assimilation or nitrogen regulation (Garner *et al.,* 1998). *H. pylori* appears to lack the proteins required for transcriptional and posttranslational regulation of glutamine synthetase. As far as the feedback regulation of GS is concerned, unadenylated GS from *S. typhimurium is* also fully inhibited by Glycine, alanine, and serine, that compete with the substrate for the active site (Liaw et al, 1993). In this respect *Hp*GS, lacking the adenylation site, could be regulated in a similar way by the end products of the glutamine metabolism.

Conclusion

*Hp*GS is a dodecameric enzyme containing 12 active sites, which seem not to be regulated by a covalent modification as in the case of other characterized bacterial GS. *Hp*GS remain unregulated following translation of the enzyme, and the only possible regulation mechanism is by feedback inhibition by products of the glutamine metabolism. Since *H. pylori* is

confined to the gastric mucosa of humans, an environment that does not present drastic changes in nitrogen availability, sophisticated levels of regulation of GS would not be so important for their survival. On the other hand, urease makes constant supply of available ammonia as long as there is a supply of urea. This pathway of nitrogen assimilation appears to be required for viability, and thus glutamine synthetase appears to represent an unusually critical enzyme in *H. pylori.*

Chapter Six

A phylogenetic-based characterization of DsbG from *Helicobacter pylori*

Introduction

To obtain the correct tertiary structure of proteins with multiple cysteines, most often the disulfide isomerization reaction is required. Enzymes that catalyze the previous reaction belongs to the family named thiol:disulfide oxidoreductases. In eukaryotes, the protein disulfide isomerase (PDI) is present in the endoplasmic reticulum. In prokaryotes, related proteins involved in disulfide isomerization have been identified located in the membranes or in the periplasm and are called Dsb's, that stands for "disulfide bond formation" (Bardwell, 1991). They share the active site sequence ccaracterized by two vicinal cysteines in the Cys-X-X-Cys, and they can exist either in the reduced form, or in the oxidized form with an intramolecular disulfide bond between both half-cysteine residues (Raina and Missiakas, 1997). In *Escherichia coli* six Dsb family members (DsbA, B, C, D, E and G) involved in correct introduction of disulfide bonds into proteins in the periplasm have been identified and they have been well characterized through a combination of genetic and biochemical approaches (Raina and Missiakas, 1997; Collet and Bardwell, 2002; Kadokura *et al.,* 2003). These enzymes catalyze the formation of disulfide bridges, the isomerization (shuffling) of incorrectly present disulfide bonds and reduction (removal) of inappropriate disulfide bonds (Messens and Collet, 2006; Kadokura *et al.,* 2003; Nakamoto *et al.,* 2004). In general, the family of thiol:disulfide oxidoreductases has two main features: an active site containing two cysteines arranged in a CXXC motif and a common tertiary structure known as the thioredoxin-like fold, despite very low sequence homology (Fabianek *et al.,* 2000).

DsbA is a periplasmic enzyme, serves as the direct donor of disulfides to secreted proteins, thereby catalyzing their oxidative folding. DsbB is an integral membrane protein, that recycles the oxidized DsbA by oxidizing reduce DsbA. DsbB generates protein disulfides de novo, via the reduction of quinones (Bader *et al.,* 1999; Kobayashi *et al.,* 1999). DsbC repairs incorrectly formed disulfide bridges and it is similar to eukaryotic PDI both in structure and mechanism (Gleiter and Bardwell, 2008). An inner-membrane protein, DsbD, is involved in transporting electrons across the inner membrane from the cytoplasmic thio-redoxin. DsbG, which is not fully characterized yet, may be responsible for maintaining the proper redox balance in the periplasm (Raczko *et al.,* 2005) or possibly isomerization of incorrect disulfides (Kaakoush *et. al.,* 2007; Missiakas *et al.,* 1994; Andersen *et al.,* 1997; Shevchik *et al.,* 1994). This is different from the situation in eukaryots, where a single disulfide isomerase (PDI) is thought to play roles in both isomerization and oxidation (Gleiter and Bardwell, 2008).

Recently it has been established that Dsb proteins are essential for correct folding or assembly of a number of pathogenic determinants, including toxins, adhesins, components of the TTSS (type III secretion system) and many other proteins. *dsb* gene mutations decrease the rate of disulfide bond formation and very often result in attenuation of pathogens (Peek and Taylor, 1992; Yu, 1998; Yu and Kroll, 1999; Stenson and Weiss, 2002).

H. pylori, the gram-negative pathogen that affects about half of the human population, induces persistence infection in gastroduodena. Pathogenesis of *H. pylori* depends on the persistence of the infection, specific virulence factors, and the inflammatory response of the host (Kusters *et al.,* 2006). In spite of the many investigations about the factors involved in *H. pylori* gastric colonization and virulence, there are very few investigation on the contribution of *H. pylori* thiol:disulfide reductases to bacterial pathogenesis. Secreting proteins is a way to interact with host by many pathogenic bacteria. In *H. pylori,* type IV secretion pathways are well identified for pathogenesis. Among the secreted proteins, many of the proteins either residing in or transiting through the periplasmic space form disulfide bridges after they translocate the inner membrane. In *H. pylori* strain 26695, protein encoded by hp0595 gene has been experimentally identified as a DsbB-like protein (DsbI) (Raczko *et al.,* 2005), and the protein encoded by gene hp0377, based on the sequence alignment with *E. coli*, has been identified as a putative DsbC (Kaakoush *et al.,* 2007). HP0231 has a sequence similarity with *E. coli* DsbG and contain the CXXC motif. Apparently DsbA is absent in *H. pylori,* suggesting the existence of a novel Dsb oxidising system. HP0231 protein is secreted via the general secretory pathway of extracellular degradative enzymes by gram-negative bacteria and enriched more than 10-fold compared with UreB (*Kim et al.,* 2002; Hueck, 1998). HP0231 has already been identified as an immunogenic protein recognized by patient sera (Haas *et al.,* 2002) and potential involvement in virulence and colonization are reported (Godlewska *et al.,* 2006). This protein is also considered as a noble antigen with low homology to other organisms, since it confers protective immunity against *H. pylori* in the mouse infection model (Sabarth *et. al.,* 2002). Live vaccines (*Salmonella enterica* serovar Typhi Ty21a expressing *H pylori* urease or HP0231) against *H pylori* tested in human volunteers revealed evidence for T cell-mediated immunity against *H pylori* infection in humans (Aebischer *et. al.,* 2008). There is a great attention in the function and mechanism of Dsb proteins and in the identification and characterization of new members of the thiol-oxidoreductases family, which could be involved in pathogenesis. In the present

investigation, we report the identification and characterization of a new DsbG disulfide isomerase from *H. pylori* defined according to phylogenetic, structural and functional criteria.

Materials and Methods

Molecular cloning of HP0231 gene

H. pylori genomic DNA was applied as the template for amplification of HP0231 gene by polymerase chain reaction (PCR) using proofreading Phusion DNA polymerase Mastermix (Finnzymes, Finland) with primers CACC ATG ATA TTA AGA GCG AGT GTG TTG AG (Forward, topo cloning sequence underlined) and CTT TCG CCA TTT CTT CTT GGA TTA AAG (Reverse). The amplified fragment was then cloned into two different vectors, pET101vector (Invitrogen) in frame with a C-terminal His-tag flanked by a TEV proteolysis site and pET151 vector (Invitrogen) in frame with an N-terminal His-tag flanked by a TEV proteolysis site using a TOPO® Cloning kit by Invitrogen to obtain the pET101-HP0859 plasmid.

Expression of HP0231

E. coli BL21(DE3) cells, harboring the pET101-*HP0231* or pET151-*HP0231* plasmid, were grown in LB medium supplemented with 100 µg/mL ampicillin at 37°C. Protein expression was induced by 0.5 mM isopropyl-β-D-thiogalactopyranoside (IPTG) with an incubation temperature of 28°C. Cells were collected after 4 h of induction by centrifuging at 6000 rpm for 20 minutes and pellet stored at -80°C.

Affinity Purification of HP0231

The bacterial pellet was resuspended in 30 mM Tris pH 7.5, 150 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 1 ml of Ni^{2+} charged Chelating Sepharose™ (GE Healthcare). After extensive washing using the lysis buffer, supplemented with 20mM imidazole, the protein was eluted from the column by linear gradient of 40mM to 300mM imidazole. The protein was further purified by Superdex 200™ 10/300 GL (GE Healthcare), equilibrated with 30 mM Tris pH 8. 0, 150mM NaCl.

Refolding of HP0231

The bacterial pellet was resuspended in 30 mM Tris pH 7.5, 150 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 at 15000rpm for 25 min and the pellet collected. The pellet was washed by resuspension in buffer (30 mM Tris pH 7.5, 150 mM NaCl). Inclusion bodies were collected as pellet by centrifugation at 2000rpm for 20 minutes. Inclusion bodies were solubilized using denaturation buffer (6M GuanidineHCl, 50 mM Tris, pH 7.5, 150 mM NaCl) over night at room temperature. Refolding was performed by dilution (25 times) to the refolding buffer (50 mM Tris, pH 7.5, 200mM Arginine, 10% Glycerol, 2mM of Reduced Glutathione and 0.2mM Oxidized Glutathione) at 4°C. The refolded solution was then centrifuged at 6500 rpm for 15 minutes to remove the aggregates. The refolded supernatant was loaded into a column containing 5 ml of Ni^{2+} charged Chelating Sepharose™ (GE Healthcare) equilibrated with refolding buffer. After extensive washing using the refolding buffer, the column was further washed with Tris buffer (30 mM Tris pH 7.5, 150 mM NaCl) supplemented with 20mM imidazole. The protein was eluted from the column by linear gradient of 40 to 300mM imidazole. The protein was further purified by Superdex 200™ 10/300 GL (GE Healthcare), equilibrated with 30 mM Tris pH 7.5, 150mM NaCl.

Protein electrophoresis and immunoblotting

The collected protein samples were separated on 15% sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE). The resolved gels were stained with 0.25% Coomassie Brilliant Blue R250 reagents or transferred to 0.2μm nitrocellulose membrane (Whatman Protran BA83 Nitrocellulose, Whatman GmbH) for immunodetection. After probing with a mouse anti-His primary antibody (Calbiochem, Germany), followed by a Alkaloine phosphatase-coupled goat anti-mouse secondary antibody (Calbiochem, Germany), the substrate BCIP (5-bromo-4-chloro-3-indolyl-phosphate) was used in conjunction with NBT (nitro blue tetrazolium); BCIP/NBT (SigmaFAST, Sigma-aldrich, Chemie GmbH) was applied for detection.

Molecular mass determination

The native molecular mass of HP0231 protein was determined by analytical gel filtration by analytical column Superdex 200™ 10/300 GL (GE Healthcare), equilibrated with 30 mM Tris pH 7.5, 150mM NaCl. His tagged-HP0231 eluted as a single peak, roughly corresponding to a dimer. Molecular mass estimated from analytical gel filtration is 58 kDa,

in good agreement with the calculated one for the dimer, 64.31 kDa (29.45 Kda plus the 2.7 Kda His-tag for each monomer). The system was operated on an AKTA FPLC instrument (GE Healthcare).

ELISA

The test was performed on serum from patients infected by *H. pylori* with the aim of detecting the presence of antibodies that interact with HP0231. The antigen (HP0231;10 ng/μl in PBS) was incubated in a 96 wells (one hundred μl/well) at room temperature over night. Subsequently, the excess antigen was removed by two washes with 200 μl per well of 1X PBS and the plate incubated with 100 μl/well of PBS, 1% BSA for 45 minutes at room temperature. After this phase of saturation, performed to avoid nonspecific antibody adsorption, the wells were washed two times with 200 μl/well of 1X PBS. Then 100 μl/well of diluted 1:100 in 1X PBS serum was added, incubated at room temperature for one hour; the unbound antibody was removed with washing buffer (200 μl / well of PBS + 0.05% Tween 20 + 1% BSA) four times. Each well was then incubated for one hour at room temperature with the secondary antibody (human IgG), diluted 1:5000 in PBS 1% BSA. Followed by 10 times washing with 200 μl/well of washing buffer to remove excess secondary antibody. Finally, the presence of the complex was detected by adding to each well 100 μl of TMB; peroxidase converts the substrate into a blue colored product, which was measured in a spectrophotometer at 370 nm. This measurement was made at time 0 (T0) and after 20 minutes.

Bioinformatics

Sequence analysis

The protein sequence of *H. pylori* (HPDsbG) was retrieve from the NCBI database with accession no. NP_207029.1 Protein properties and amino acid sequence analyses were performed using Protparam (Gasteiger *et al.,* 2005), a tool present in the ExPASy Proteomics Server. Searches for sequences similar to HPDsbG were done with the BLAST tool available at National Center for Biotechnology Information (NCBI, www.ncbi.nlm.nih.gov). The conserved domains in the enzymes were analyzed using the Conserved Domain Database (www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml) (Marchler-Bauer *et al.,* 2011) and InterProScan. The presence of a signal peptide sequence was predicted from SignalIP 3.0 (Bendtsen *et al.,* 2004). ClustalW (Li, 2003) from EMBL-EBI server was used to perform multiple alignments. Aligned sequences were edited manually and visualized by Espript

(Gouet *et al.,* 1999). Based on the alignment, a phylogenetic tree was constructed according to the neighbor-joining method (Saitou and Nei, 1987) and the resulting unrooted tree was visualized as radial tree layout with FigTree v1.3.1 (Papadopoulos and Agarwala, 2007) program (http://tree.bio.ed.ac.uk). The coordinates and sequence file of *E. coli* DsbG (1V57.pdb (Heras *et al*., 2004)) were obtained from the RCSB Protein Data Bank (PDB, www.rcsb.org/pdb/home/home.do (Berman *et al.,* 2000).

Homology Modeling

FFAS server, a profile–profile alignment and fold-recognition tool, was used to identify the better template (Rychlewski *et al.,* 2000). The homology model of HPDsbG was built using I-TASSER server (Roy *et al.,* 2010) selecting the template PDB option and putting all other options as default. An initial structural model was generated for HPDsbG and subjected to energy minimization with Chiron, a server that rapidly minimizes steric clashes in proteins using short discrete molecular dynamics (DMD) simulations (Ramachandran *et al.,* 2011) to improve the van der Waals contacts and correct the stereochemistry of the model. The model was further improves with YASARA Energy Minimization Server (Krieger *et al.,* 2009). The qualities of the model were checked using the SAVES server (http://nihserver.mbi.ucla.edu/SAVES_3/saves.php), PSQS server (http://www1.jcsg.org/psqs/psqs.cgi) and ProSA web server (Wiederstein & Sippl (2007). The quality scores of the selected structural models (Table 1) suggested that the model of HPDsbG was reliable. Comparison of the three dimensional structure of HPDsbG was made with DaliLite v. 3 (Holm and Rosenstrom, 2010).

Table 1. Statistics of Model valiadation

PROCHECK (number of amino acids in the parenthesis)	
Most favoured regions	(182) 85.4%
Additional allowed regions	(29) 13.6%
Generously allowed regions	(1) 0.5%
Disallowed regions	(1) 0.5%
Errat	97.74 %
Verify 3D	64.22%
Prove	94.8 %
PSQS	-0.1859
ProSA Server (Z-Score)	-6.60
Close Contacts (within 2.2 Å):	0
Ideal Geometry	
RMS deviation for bond angles	1.8 °
RMS deviation for bond lengths	0.010 Å

Results and Discussion

Expression, purification and characterization

HPDsbG encodes for a protein of 265 amino acids, with calculated molecular mass of the mature protein of approximately 29.5 kDa. With a predicated isoelectric point of 9.09, *HP0231* gene had been previously cloned in frame with a N-terminal His-tag into *pET151* expression vector. *E. coli* BL21(DE3) cells harboring *pET151- HP0231* were used for expression and solubility trials. N-His_6-tag HP0231 solubility was tested in two different buffers with different expression conditions, but in both cases the majority of the protein was found in the pellet. To purify N-His_6-tag HP0231 in high yield, a 2 liters culture was set up and an affinity IMAC-Ni^{2+} chromatography was performed. The SDS-PAGE, loaded with fraction samples from chromatographic steps, showed that very small amount of protein was bound to the nickel matrix. To increase N-His_6-tag HP0231 solubility, trials were performed using different buffers in combination with varying pH, but the protein yield was not improved significantly.

Since the N-His_6-tag HP0231 could not be purified in sufficient amount by affinity IMACNi^{2+} chromatography or other standard chromatography methods, to solve this problem, two new strategies were planned: the cloning of a C-terminal His-tag construct and

the purification of insoluble fraction of HP0231 from inclusion bodies in *E. coli*, followed by protein refolding. The new C-terminal His-tag protein construct was prepared (C-His6- tag HP0231) and expressed in *E. coli*. Affinity chromatography IMAC-Ni^{2+} revealed that the new C-terminal His-tag construct was slightly more solubile, but the amount of protein bound to the affinity resin was comparable to the N-terminal His-tag construct. The final yield was even lower than the previous case (Figure 1A). Probably the protein was degraded from the C-terminus and, losing the His-tag, it was unable to bind the column.

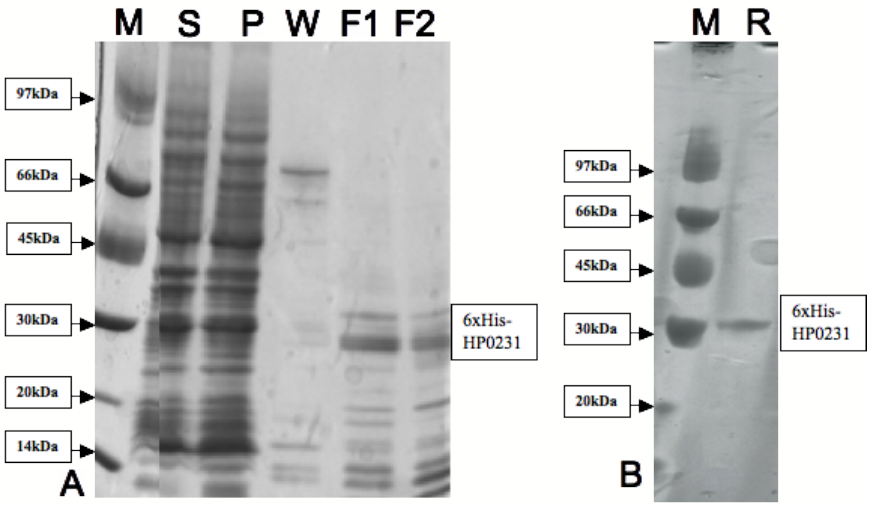


Figure 1. Purification of HP0231. A. Affinity purification of C-terminal 6xHis-HP0231; M-low molecular weight marker, S-supernatant, P-pellet, W-wash fraction, F1 and F2- affinity fraction. B. Refolding and followed by affinity purification, M-low molecular weight marker, R-affinity fraction after refolding. 15% SDS gel under denaturing condition.

A N-His_6-tag HP0231 refolding method was set up (Figure 1B). The refolding protocol was optimized, and the yield was increased. HP0231 protein obtained at the end of the purification process was enough to procede to further characterizations trial. The protein was found as a dimer in solution, as in other members of this protein family (Heras *et al.,* 2004), as demonstrated by analytical gel filtration (Figure 2B)**.** Western-blotts studies showed that most of the N-His6-tag HP0231 protein is a dimer even in the presence of reducing agent (Figure 2A).

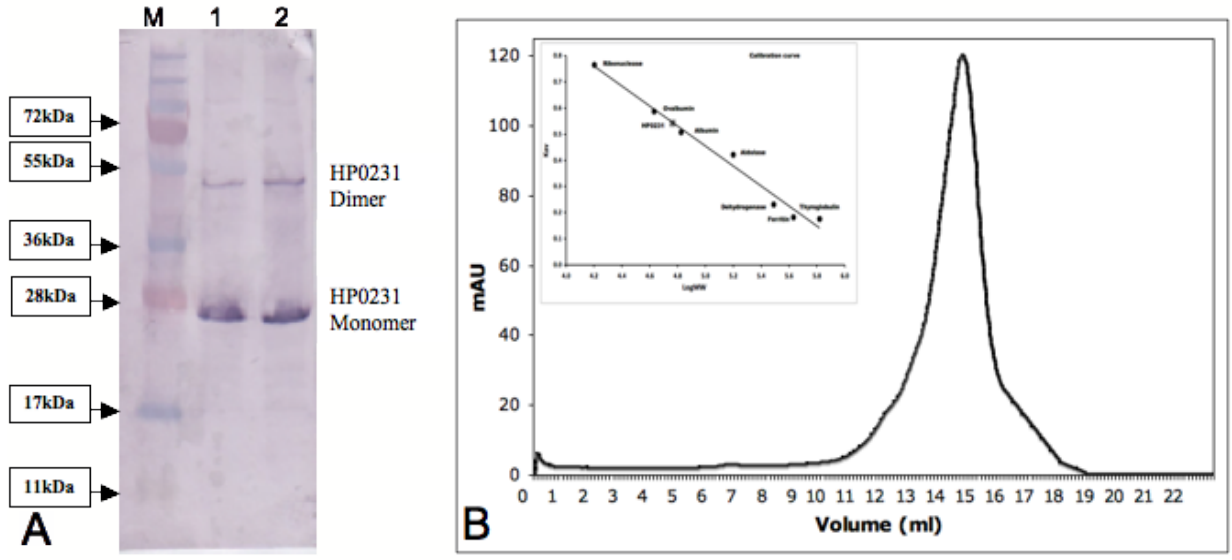


Figure 2. Characterization of HP0231. A. Western blotting of 6xHis-HP0231 against 6xHis-tag; M-prestain molecular weight marker, 1-HP0231 without reducing agent, 2- HP0231 with reducing agent. B. Analytical gel filtration of HP0231; Affinity purification of C-terminal 6xHis-HP0231 by analytical column Superdex 200™ 10/300 GL with standard proteins.

Immunogenicity and role in pathogenesis

Blood sera from 22 patients (*H. pylori* infected, infection was confirmed by other standard methods) and 23 normal controls were explored in terms of the presence of antibodies against purified DsbG. The presence of such antibodies was detected significantly in the patient sera (Figure 3B). In fact, *H. pylori* expresses a number of antigenic proteins and, in previous study, to search for potential antigen candidates from *H. pylori*, at least 59 antigens were recognized by *H. pylori*-infected patients (Haas *et al.,* 2002; Kimmel *et al.,* 2000; McAtee *et al.,* 1998). 48 of these antigens had a staining intensity that is higher than an arbitrary cutoff, equivalent to 0.1% of the total staining intensity (Jungblut *et al.,* 2000). HP0231 was recognized by *H. pylori* positive patient sera in a systematic, proteome-based approach used to detect candidate antigens of *H. pylori* for diagnosis, therapy and vaccine development, and to investigate potential associations between specific immune responses and manifestations of disease (Haas *et al*., 2002). HP0231 conferred protective immunity in the mouse *Helicobacter* infection model with levels of protection generally considered the gold standard for *Helicobacter* immunization (Sabarth *et al.,* 2002). HP0231 protein is involved in the colonization by *H. pylori* (Kaakoush et al., 2007). Rabbit anti-HP0231 antibody was produced to detect HP0231 protein of *H. pylori* and was tested for the development of a stool antigen detection kit for *H. pylori* (Lee *et al.,* 2006).

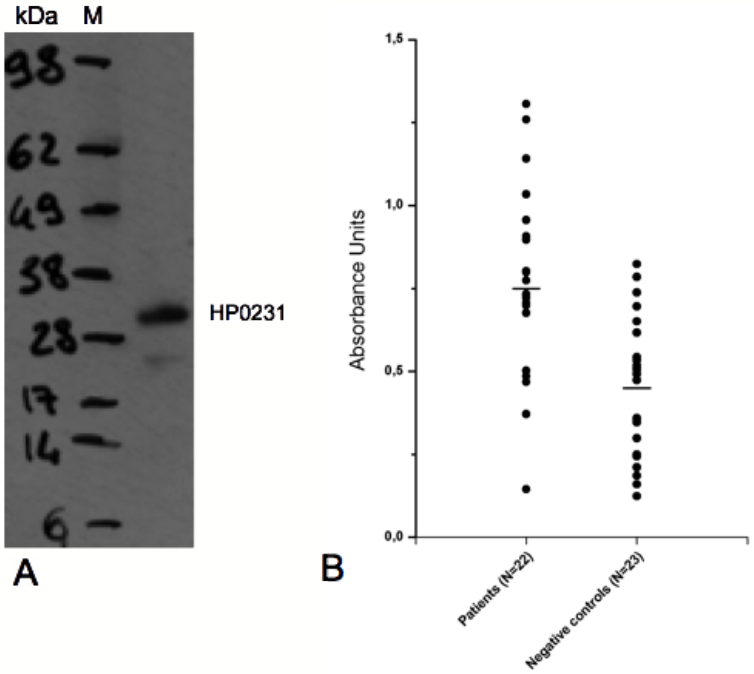


Figure 3. Immunogenicity of HP0231. A. Western blott of HP0231 with patient sera; B. Detection of antibody against HP0231 in the patient and control by ELISA.

Isomerase activity and Chaperone-like activity

The precise function and the substrate of DsbG have not been characterized yet, but the protein was found inactive in the insulin-reduction assay, used as a test for the disulfide reductase family (Hiniker and Bardwell, 2004). Overexpression of DsbG can partially rescue the formation of DsbC mutant in *E. coli* (Bessette et al., 1999). DsbG acts as a thiol-disulfide isomerase with narrower substrate specificity than DsbC (Heras *et al.,* 2004). DsbG presents significant chaperone-like activity: expression and solubility of single-chain Fv antibodies (scFv) was significantly increased when the gene fuse with dsbG-scFv forms a polycistronic co-expression system (Zhang *et al.,* 2002). Overexpression of the isomerases DsbG assists in the folding of a number of eukaryotic proteins, indicating that the periplasm of *E. coli* can be altered to improve the folding of eukaryotic disulfide-containing proteins (Bessette *et al.,* 2001). Many peptides and proteins with therapeutic applications contain disulfides. The proper formation of these disulfides is one of the rate limiting steps in the commercial production of these proteins. Further optimization of the periplasm for thiol-disulfide exchange is clearly of great interest for the commercial production of disulfide containing proteins. (Bessette *et al.,* 2001)

Structure Modeling of HP0231

DsbG protein belongs to Thioredoxin-like superfamily (cl00388), as predicted by Conserved Domain Database search and InterProScan, since it contains the thioredoxin (TRX) fold with a redox active CXXC motif. There is a signal sequence of 26 amino acids present at the N-terminus of DsbG. Non-Ordinary Secondary Structure (NORS) was predicted for HPDsbG and alpha helix 46.8%, beta Strand 18.9%, loop 34.3%. Protein Localization prediction in prokaryotes was done using LOCtree (Predict Protein server) , that suggests that HPDsbG is a periplasmic protein.

The result of a protein-protein BLAST search for a suitable template structure related to the target sequence (AAB40805.1, PDB ID: 1V57, Heras *et al.,* 2004) showed that *E. coli* periplasmic disulfide isomerase/thiol-disulphide oxidase presents the best sequence similarity (18%). I-TASSER server built the model, which was refined with Chiron energy minimization server that improves the model geometry, Ramachandan plot and torsion angles. The output model was further improved with several cycle of energy minimization from YASARA Energy Minimization Server (Krieger *et al.,* 2009). Ramachandran plot analysis shows that in the final model 85.4% of the residues are in the core region, 13.6% in the allowed and 0.5% in the generously allowed regions. One residue was found in the disallowed region of the Ramachandran plot.

The DsbG Fold

The structure of DsbG is folded in to two domains: a N-terminal dimerization domain and a C-terminal catalytic domain, which presents a thioredoxin fold (Martin, 1995). A linker connects the two domains. The N-terminal dimerization domain in each DsbG subunit (residues 33–102) has a cystatin-like fold found in cysteine protease inhibitors (Bode *et al.,* 1988). The fold consists of an α-helix (α1), followed by a four-stranded antiparallel β-sheet (β1–β4) (Figure 4). The dimerization interface is formed by β-sheet interactions between the β4 strands of the two subunits. The linker (residues 103–126) that connects the dimerization domain with the C-terminal catalytic domain forms a curved helix, α2 (Figure 4). This long linker places the two catalytic domains far apart and separates the N-terminal domain from the catalytic domain (Figure 4). The helical linker between the two domains substantially increases the dimensions of the binding-site groove, where negatively charged residues are predominating. The catalytic domain (residues 127–261) incorporates a thioredoxin (TRX) fold with a helical insert (α4). As in all redox TRX-like proteins, a CXXC redox active center

is located at the N terminus of the first helix (α4) (Figure 4).

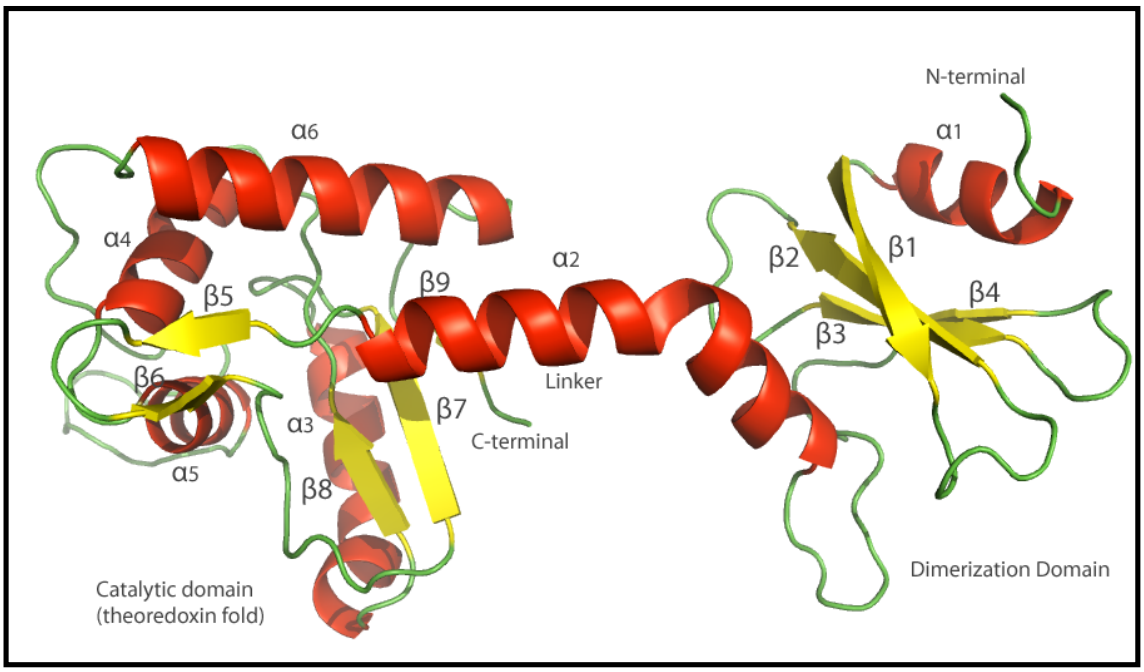


Figure 4. Homology Model of HP0231. Helices and strands are colored in red and yellow, respectively, whereas loops are in light green. All helices and strands are numbered consequently.

The homology model of DsbG presents 18% sequence similarity to *E. coli* DsbG (PDB ID: 1V57, Heras *et al.,* 2004), and our model was built based on it. The V-shaped cleft has been postulated as the binding site for unfolded proteins in DsbC (McCarthy *et al.,* 2000) and DsbG (Heras *et al.,* 2004). It is possible to hypothesize that a similar situation holds also for HP0231. The surface charge distribution in the region of the catalytic cleft is comparatively similar to that of DsbG from *E. coli*. The redox potentials for DsbG from *E. coli* was measured and indicates that the reduced form is more favored than the respective oxidized form of the protein (Heras *et al.,* 2004).

Phylogenetic analysis

The TRX fold is the core scaffold in proteins with very different functions (Martin, 1995). Among them, the TRX-like oxidoreductases (including TRX and Dsb proteins) control the cellular redox environment, which is critical for the folding and stability of other proteins (Heras *et al.,* 2004). A GenBank Blast search with the amino acid sequence of HP0231 (DsbG) (Accession no. NP_207029.1) within the genus *Helicobacter* revealed that the Protein disulfide isomerase/thiol-disulfide oxidase are present in most of the *Helicobacter*

species. Among them, *H. mustelae* (Accession no. YP_003516087.1) presents 44% sequence similarity, 43% *H. felis* (Accession no. YP_004073329.1), 31% *H. bilis* (Accession no. ZP_04581782.1), 34% *H. hepaticus* (Accession no. NP_860672.1), 33% *H. pullorum* (Accession no. ZP_04809358.1), 29% *H. winghamensis* (Accession no. ZP_04582733.1), 32% *H. canadensis* (Accession no. ZP_04871051.1), 43% *H. bizzozeronii* (Accession no. YP_004608326.1) and 34% *H. cinaedi* (Accession no. ZP_07805739.1). Sequence allingment based on PAM protein weight matrix showed that CPXC motif sequence are well conserved among all the members of the family (Figure. 5). The unrooted tree obtained was visualized as a radial tree layout, showing that *H. winghamensis, H. pullorum* and *H. canadensis* can be considered as a group that is relatively far in the tree. In all the other analyzed *Helicobacter* species, DsbG proteins cluster in a single group (Figure 6A).

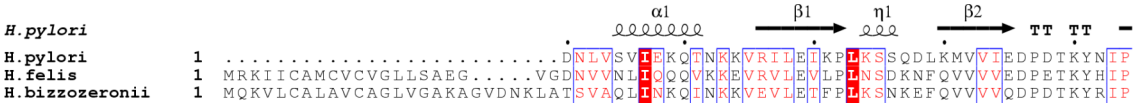


Figure 5. Amino acid sequence of HP0231 protein aligned with protein disulfide isomerase from, A. Different *Helicobacter* species, B. Different genus. Secondary structure elements deriving from Homology Model of HP0231 are drawn over the alignment.

A GenBank Blast search with the amino acid sequence of HP0231 revealed that the Protein disulfide isomerase/thiol-disulfide oxidase are present in many other organisms, with sequence identity of 31% with *Wolinella succinogenes* (Accession no. NP_907869.1), 33% with *Campylobacter concisus* (Accession no. YP_001466211.1), 22% with *Hydrogenobaculum sp.* (Accession no. YP_002121307.1), 31% with *Neisseria polysaccharea* (Accession no. ZP_06864860.2), *Ralstonia solanacearum* (Accession no.

YP_003751285.1), *Hippea maritima* (Accession no. YP_004339974.1), *Thiomonas intermedia* (Accession no. YP_003642442.1), *Bordetella avium* (Accession no. YP_787226.1), *Vibrio cholera* (Accession no. ZP_05417704.1), *Nautilia profundicola* (Accession no. YP_002606495.1). The TRX-like redox proteins share no overall sequence homology, yet they all contain a CXXC motif and a cis-Pro loop at the active site, which is conserved in DsbG from *H. pylori* (Figure 5B and 6B).

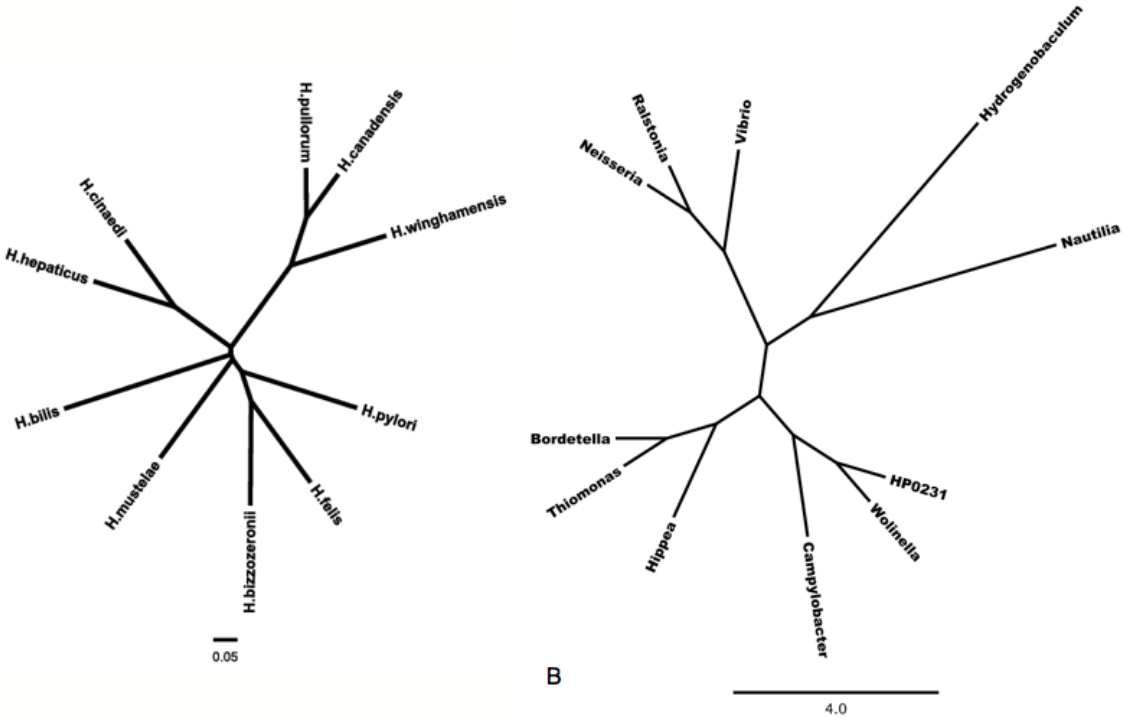


Figure 6. Phylogenetic analysis of HP0231. Unrooted tree showing the phylogenetic relationship of HP0231 with other DsbG proteins. Branches are proportional to genetic distance as indicated by the scale bar representing point-accepted mutations (PAM). Proteins are indicated by their Genus and species name. A. The HP0231 (DsbG) protein with the putative DsbG sequence from different species of *Helicobacter*. B. With different genus.

Conclusion

Dsb proteins control the formation and rearrangement of disulfide bonds during the folding of secreted and membrane proteins in bacteria. DsbG, a member of this family, has disulfide bond isomerase and chaperone activity. Dsb proteins play important roles in bacterial pathogenicity. To better understand the role of Dsb proteins in *Helicobacter pylori*, we have structurally and functionally characterized *H. pylori* DsbG (HP0231). HP0231 was recognized by *H. pylori* positive patient sera and is folded in to two domains: a N-terminal

dimerization domain and a C-terminal catalytic domain, which presents a thioredoxin fold. As in all redox TRX-like proteins, a CXXC redox active center is located at the catalytic domain. Our work contributes to our understanding of the role of Dsb proteins in the pathogenicity of *H. pylori*.

Conclusions

Conclusions

The main focus of this study was to determine the three dimensional structure and to characterize the function of *H. pylori* proteins important for stomach colonization and pathogenesis. In conclusion, we have fully determined the following crystal structures:

- peptidoglycan deacetylase, a cell wall modification enzyme,
- ADP-L-*glycero*-D-*manno*-heptose-6-epimerase, rfaD, an enzyme involved in LPS biosynthesis,
- CeuE, a periplasmic substrate binding protein of ABC transporter,
- Glutamine synthetase, a key enzyme in nitrogen assimilation pathway.,

In addition, DsbG, a secreted immunogenic disulfide isomerase, has been characterized in solution.

Peptidoglycan deacetlyase, which belongs to the polysaccharide deacetylases protein family, is a homo-tetramer, each domain beinbg characterized by a non-canonical TIM-barrel fold. *Hp*PgdA does not exhibit a solvent-accessible polysaccharide binding groove, suggesting that the enzyme binds a small substrate at the active site.

The crystal structure of rfaD reveals that the enzyme is a homo-pentamer, and NAD is bound as a cofactor in a highly conserved pocket. The substrate-binding site of the enzyme is very similar to that of its orthologue in *E. coli*, suggesting also a similar catalytic mechanism.

The periplasmic ABC transporter substrate-binding protein shares the common features of other bacterial periplasmic ABC transporter heme- or vitamin B12-binding protein. The crystal structure along with different techniques (Fluorescence and ITC) confirms that it is a periplasmic heme binding protein responsible for heme uptake.

Glutamine synthetase is a dodecamer, organized in two hexamers. The dodecameric enzyme contains twelve active sites, in which ATP and glutamate bind at opposite ends. The absence of the classical adenylation site, present in all the homologs of this family, indicates a different mechanism of control for this enzyme.

DsbG (HP0231) was purified to high homogeneity and characterized employing immunological techniques and blot with patient blood sera.

Appendix

Appendix A: Abbreviations and Symbols

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
cag	cytotoxin associated gene (gene)
Cag	Cytotoxin associated gene (associated protein)
CCD	Charge-Coupled Device
HpPgdA	*Helicobacter pylori* Peptidoglycan deacetlyase
GlcNAc3	Chitotriose
GlcNAc	N-acetyl glucosamine
DMSO	Dimethyl sulfoxide
Da	Dalton
DTT	DiThioThreitol
DMF	Dimethylformamide
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
FPLC	Fast Protein Liquid Chromatography
Fur	Ferric Uptake Regulator protein
HEPES	N-[2-Hydroxyethyl] piperazine-N'-[2-ethanesolfonic] acid
Nod1 and Nod2	Nucleotide-binding oligomerization domain-containing protein 1 and 2
IL	Interleukine
TNF	Tissue necrosis factor
SpPgdA	Streptococcus pneumoniae peptidoglycan deacetylase
PdaA	peptidoglycan deacetylase
H. pylori	*Helicobacter pylori*

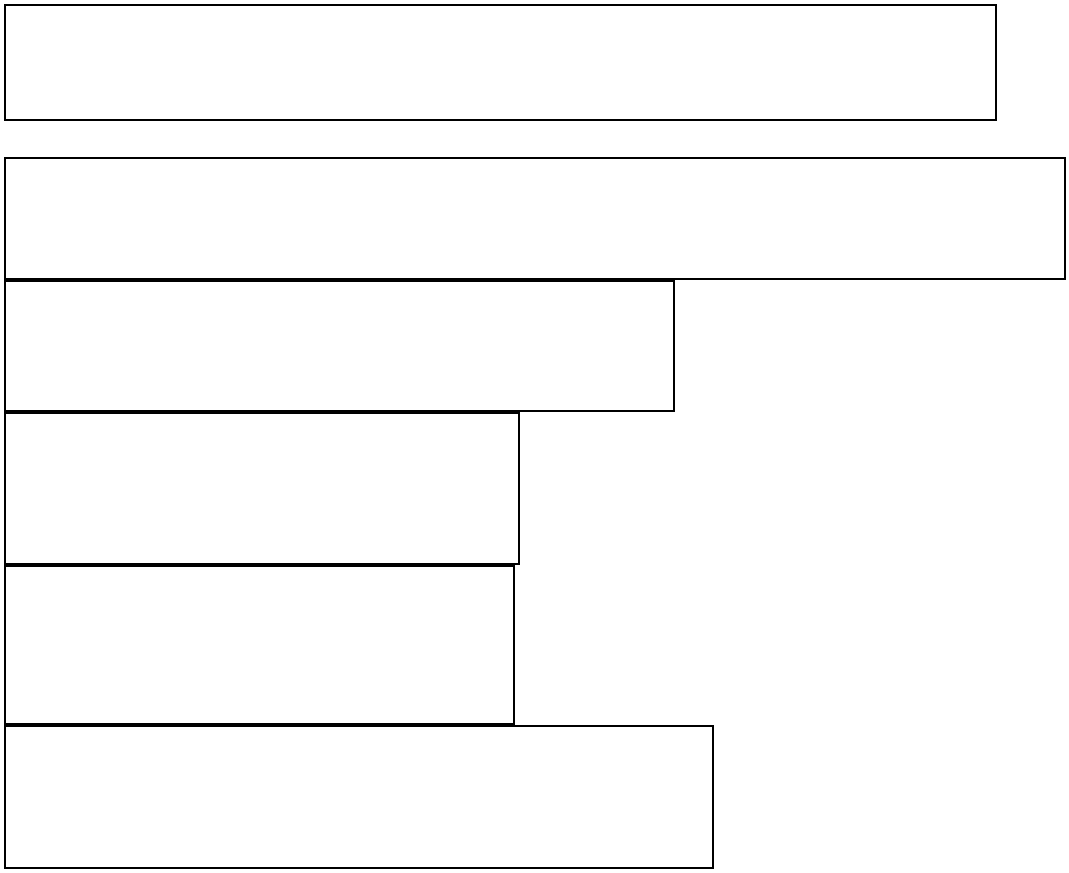
I	Measured Intensity of the diffraction spots
IMR	Imidazole ribotide
IPTG	IsoPropyl-β-D-ThioGalactopyranoside
LB	Luria Bertani liquid medium
MAD	Multiple Anomalous Dispersion
mAu	milli Absorption unit
LPS	Lipopolysaccharides
MES	2-(N-Morpholin) ethansulfonate
MIR	Multiple Isomorphous Replacement
MS	Mass Spectrometry
MW	Molecular Weight
PAM	Point-accepted mutations
NAD	Nicotinamide adenine dinucleotide
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
TLC	Thin Layer Chromatography
TLR4	Toll-Like Receptor 4
Tris	2-Amino-2-(hydroxymethyl)-1,3-propanediol
NADP	Nicotinamide adenine dinucleotide phosphate
σ(I)	Standard deviation of the measured Intensities (I)
*Hp*AGME	*H. pylori* ADP-L-glycero-Dmanno- heptose-6-epimerase;

*Ec*AGME	*E. coli* ADP-L-glycero-D-manno-heptose-6-epimerase
L,D-Hep	L-glycero-D manno-heptose
D,D-Hep	D-glycero-D-manno-heptos
PBP	Periplasmic Binding Protein
SBP	Substrate binding Protein
ABC transporter	ATP-binding cassette transporter
LF	Lactoferrin
TF	Transferrin
Se-Met	Selenomethionine
CD	Circular dichroism
ITC	Isothermal titration calorimetry
GS	Glutamine synthetase
CTP	Cytosine Tri-Phosphate
glnA	Glutamine synthetase gene
*Hp*GS	*H. pylori* Glutamine synthetase
*St*GS	*Salmonella typhimorium* Glutamine synthetase
*Mt*GS	*Mycobacterium tuberculosis* Glutamine synthetase
PDI	Protein disulfide isomerase
Dsb	Disulfide bond formation
BCIP	5-bromo-4-chloro-3'-indolyphosphate
NBT	Nitro-blue tetrazolium
BSA	Bovine serum albumin
PBS	Phosphate buffer saline

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



References

Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, et al. (2010) PHENIX: A comprehensive python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr 66(Pt 2):213–221.

Aebischer T, Bumann D, Epple HJ, Metzger W, Schneider T, Cherepnev G, Walduck AK, Kunkel D, Moos V, Loddenkemper C, Jiadze I, Panasyuk M, Stolte M, Graham DY, Zeitz M, Meyer TF. (2008) Correlation of T cell response and bacterial clearance in human volunteers challenged with *Helicobacter pylori* revealed by randomised controlled vaccination with Ty21a-based Salmonella vaccines. Gut. 57(8):1065-72.

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

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

Almassy RJ, Janson CA, Hamlin R, Xuong NH, Eisenberg D. (1986) Novel subunit-subunit interactions in the structure of glutamine synthetase. Nature. 323(6086):304-9.

Amedei A, Cappon A, Codolo G, Cabrelle A, Polenghi A, Benagiano M, Tasca E, Azzurri A, D'Elios MM, Del Prete G, de Bernard M. (2006) The neutrophil-activating protein of *Helicobacter pylori* promotes Th1 immune responses. J Clin Invest. 116 (4):1092-101.

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

Andersen CL, Matthey-Dupraz A, Missiakas D, Raina S. (1997) A new *Escherichia coli* gene, dsbG, encodes a periplasmic protein involved in disulphide bond formation, required for recycling DsbA/DsbB and DsbC redox proteins. Mol Microbiol. 26:121–32.

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

Andrzejewska J, Lee SK, Olbermann P, Lotzing N, Katzowitsch E, Linz B, Achtman M, Kado CI, 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. J Bacteriol. 188 (16) :5865-77.

Aras RA, Fischer W, Perez-Perez GI, Crosatti M, Ando T, Haas R, Blaser MJ. (2003a) Plasticity of repetitive DNA sequences within a bacterial (Type IV) secretion system component. J Exp Med. 198 (9):1349-60.

Aras RA, Kang J, Tschumi AI, Harasaki Y, Blaser MJ. (2003b) Extensive repetitive DNA facilitates prokaryotic genome plasticity. Proc Natl Acad Sci U S A. 100 (23):13579-84.

Atherton JC, Cao P, Peek RM Jr, Tummuru MK, Blaser MJ, Cover TL. (1995) Mosaicism in vacuolating cytotoxin alleles of *Helicobacter pylori* Association of specific vacA types with cytotoxin production and peptic ulceration. J Biol Chem. 270 (30):17771-7.

Backert S, Meyer TF. (2006) Type IV secretion systems and their effectors in bacterial pathogenesis. Curr Opin Microbiol. 9(2):207-17.

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

Backert S, Tegtmeyer N, Selbach M. (2010) The versatility of *Helicobacter pylori* CagA effector protein functions: The master key hypothesis. Helicobacter. 15(3):163-76.

Bäckhed F, Rokbi B, Torstensson E, Zhao Y, Nilsson C, Seguin D, Normark S, Buchan AM, Richter-Dahlfors A. (2003) Gastric mucosal recognition of *Helicobacter pylori* is independent of Toll-like receptor 4. J Infect Dis. 187 (5):829-36.

Bader M, Muse W, Ballou DP, Gassner C, Bardwell JC. (1999) Oxidative protein folding is driven by the electron transport system. Cell 98:217–27.

Bailey S, Ward D, Middleton R, Grossmann JG, Zambryski PC. (2006) *Agrobacterium tumefaciens* VirB8 structure reveals potential protein-protein interaction sites. Proc Natl Acad Sci U S A. 103(8):2582-7.

Baltrus DA, Amieva MR, Covacci A, Lowe TM, Merrell DS, Ottemann KM, Stein M, Salama NR, Guillemin K. (2009) The complete genome sequence of *Helicobacter pylori* strain G27. J Bacteriol. 191(1):447-8.

Banner DW, Bloomer A, Petsko GA, Phillips DC, Wilson IA. (1976) Atomic coordinates for triose phosphate isomerase from chicken muscle. Biochem Biophys Res. Commun. 72(1):146–155.

Bardwell JCA, McGovern K, Beckwith J. (1991) Identification of a protein required for disulfide bond formation in vivo. Cell 67:581–89.

Behrmann I, Hillemann D, Pühler A, Strauch E, Wohlleben W. (1990) Overexpression of a *Streptomyces viridochromogenes* gene (glnII) encoding a glutamine synthetase similar to those of eucaryotes confers resistance against the antibiotic phosphinothricyl-alanyl-alanine. J Bacteriol. 172(9):5326-34.

Bendtsen JD, Nielsen H, Heijne G and Brunak S. (2004) Improved prediction of signal peptides: SignalP 3.0. J. Mol. Biol. 340:783-795.

Bereswill S, Schönenberger R, Thies C, Stähler F, Strobel S, Pfefferle P, Wille L, Kist M. (2000) New approaches for genotyping of *Helicobacter pylori* based on amplification of polymorphisms in intergenic DNA regions and at the insertion site of the cag pathogenicity island. Med Microbiol Immunol. 189(2):105-13.

Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, and Bourne PE. (2000) The Protein Data Bank. Nucleic Acids Res. 28:235–242.

Bessette PH, Cotto JJ, Gilbert HF. and Georgiou G. (1999) In vivo and in vitro function of the *Escherichia coli* periplasmic cysteine oxidoreductase DsbG. J. Biol. Chem. 274:7784–7792.

Bessette PH, Qiu J, Bardwell JC, Swartz JR, Georgiou G. (2001) Effect of sequences of the active-site dipeptides of DsbA and DsbC on in vivo folding of multidisulfide proteins in *Escherichia coli*. J Bacteriol. 183:980–8.

Beswick EJ, Bland DA, Suarez G, Barrera CA, Fan X, Reyes VE. (2005a) *Helicobacter pylori* binds to CD74 on gastric epithelial cells and stimulates interleukin-8 production. Infect Immun. 73(5):2736-43.

Beswick EJ, Das S, Pinchuk IV, Adegboyega P, Suarez G, Yamaoka Y, Reyes VE. (2005b) *Helicobacter pylori*-induced IL-8 production by gastric epithelial cells up-regulates CD74 expression. J Immunol. 175(1):171-6.

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

Blair DE, Schuttelkopf AW, MacRae JI, van Aalten DM. (2005) Structure and metal-dependent mechanism of peptidoglycan deacetylase, a streptococcal virulence factor. Proc Natl Acad Sci U S A 102(43):15429–

15434.

Blair DE, van Aalten DM. (2004) Structures of bacillus subtilis PdaA, a family 4 carbohydrate esterase, and a complex with N-acetyl-glucosamine. FEBS Lett. 570(1–3):13–19.

Blaser MJ, Atherton JC. (2004) *Helicobacter pylori* persistence: biology and disease. J Clin Invest. 113(3):321-33.

Blaser MJ. (1994) *Helicobacter pylori* phenotypes associated with peptic ulceration. Scand J Gastroenterol Suppl. 205:1-5.

Blaser MJ. (1998) *Helicobacter pylori* and gastric diseases. BMJ. 316 (7143) :1507-10.

Blaser MJ. (2006) Who are we? Indigenous microbes and the ecology of human diseases. EMBO Rep. 7(10):956-60.

Bode W, Engh R, Musil D, Thiele U, Huber R, Karshikov A, Brzin J, Kos J, Turk V. (1988) The 2.0 A X-ray crystal structure of chicken egg white cystatin and its possible mode of interaction with cysteine proteinases. EMBO J. 7(8):2593–2599.

Boncristiano M, Paccani SR, Barone S, Ulivieri C, Patrussi L, Ilver D, Amedei A, D'Elios MM, Telford JL, Baldari CT. (2003) The *Helicobacter pylori* vacuolating toxin inhibits T cell activation by two independent mechanisms. J Exp Med. 198(12):1887-97.

Boneca IG, Dussurget O, Cabanes D, Nahori MA, Sousa S, et al. (2007) A critical role for peptidoglycan N-deacetylation in listeria evasion from the host innate immune system. Proc Natl Acad Sci U S A 104(3): 997–1002.

Borén T, Falk P, Roth KA, Larson G, Normark S. (1993) Attachment of *Helicobacter pylori* to human gastric epithelium mediated by blood group antigens. Science. 262 (5141):1892-5.

Borths EL, Locher KP, Lee AT, Rees DC. (2002) The structure of *Escherichia coli* BtuF and binding to its cognate ATP binding cassette transporter. Proc Natl Acad Sci U S A. 99(26):16642-7.

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

Brünger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta Crystallogr D Biol Crystallogr. 54(Pt 5):905-21.

Bryant CE, Spring DR, Gangloff M, Gay NJ. (2010) The molecular basis of the host response to lipopolysaccharide. Nat Rev Microbiol. 8(1):8-14.

Bury-Moné S, Thiberge JM, Contreras M, Maitournam A, Labigne A, De Reuse H. (2004) Responsiveness to acidity via metal ion regulators mediates virulence in the gastric pathogen *Helicobacter pylori*. Mol Microbiol. 53 (2):623-38.

Carlsohn E, Nyström J, Bölin I, Nilsson CL, Svennerholm AM. (2006) HpaA is essential for *Helicobacter pylori* colonization in mice. Infect Immun. 74 (2):920-6.

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

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. J Mol Biol. 386 (1):204-17.

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. J Mol Biol. 340 (4):881-9.

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. 69 (2):440-3.

Chang PC, Wang CJ, You CK, Kao MC. (2011) Effects of a HP0859 (rfaD) knockout mutation on lipopolysaccharide structure of *Helicobacter pylori* 26695 and the bacterial adhesion on AGS cells. Biochem Biophys Res Commun. 405(3):497-502.

Clarke TE, Ku SY, Dougan DR, Vogel HJ, Tari LW. (2000) The structure of the ferric siderophore binding protein FhuD complexed with gallichrome. Nat Struct Biol. 7(4):287-91.

Clyne M, Labigne A, Drumm B. (1995) *Helicobacter pylori* requires an acidic environment to survive in the presence of urea. Infect Immun. 63 (5) :1669-73.

Collaborative Computational Project, Number 4. (1994) The CCP4 suite: Programs for protein crystallography. Acta Crystallogr D Biol Crystallogr 50(Pt 5): 760–763.

Collet JF and Bardwell JC. (2002) Oxidative protein folding in bacteria. Mol Microbiol 44:1–8.

Colombo G and Villafranca JJ. (1986) Amino acid sequence of *Escherichia coli* glutamine synthetase deduced from the DNA nucleotide sequence. J Biol Chem. 261(23):10587-91.

Contreras M, Thiberge JM, Mandrand-Berthelot MA, Labigne A. (2003) Characterization of the roles of NikR, a nickel-responsive pleiotropic auto regulator of *Helicobacter pylori*. Molecular Microbiology. 49 (4):947–963.

Cooke CL, Huff JL, Solnick JV. (2005) The role of genome diversity and immune evasion in persistent infection with *Helicobacter pylori*. FEMS Immunol Med Microbiol. 45 (1):11-23.

Correa P. (1996) *Helicobacter pylori* and gastric cancer: state of the art. Cancer Epidemiol. Biomarkers Prev. 5:477–481.

Covacci A and Rappuoli R. (2000) Tyrosine-phosphorylated bacterial proteins: Trojan horses for the host cell. J Exp Med. 191(4):587-92.

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

Cover TL and Blaser MJ. (1992) Purification and characterization of the vacuolating toxin from *Helicobacter pylori*. J Biol Chem. 267 (15):10570-5.

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

Cowtan K. (2006) The Buccaneer software for automated model building. 1. Tracing protein chains. Acta Crystallogr D Biol Crystallogr. 62(Pt 9):1002-11.

Crespo JL, García-Domínguez M, Florencio FJ. (1998) Nitrogen control of the glnN gene that codes for GS type III, the only glutamine synthetase in the cyanobacterium *Pseudanabaena sp*. PCC 6903. Mol Microbiol. 30(5):1101-12.

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

D'Elios MM, Montecucco C, de Bernard M. (2007) VacA and HP-NAP of *Helicobacter pylori*-associated

gastric inflammation. Clin Chim Acta. 381 (1) :32-8.

Davidson AL. (2002) Mechanism of coupling of transport to hydrolysis in bacterial ATP-binding cassette transporters. J. Bacteriol. 184:1225-1233.

de Reuse H, Bereswill S. (2007) Ten years after the first *Helicobacter pylori* genome: comparative and functional genomics provide new insights in the variability and adaptability of a persistent pathogen. FEMS Immunol Med Microbiol. 50 (2):165-76.

Deacon AM, Ni YS, Coleman WG Jr, Ealick SE. (2000) The crystal structure of ADP-L-glycero-D-mannoheptose 6-epimerase: catalysis with a twist. Structure. 8(5):453-62.

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

Del Giudice, G., and P. Michetti. (2004) Inflammation, immunity and vaccines for *Helicobacter pylori*. Helicobacter 9 (Suppl 1):23-8.

Del Prete G, De Carli M, Lammel RM, D'Elios MM, Daniel KC, 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. 86 (1):250-7.

Delano WL. (2008) The PyMOL Molecular Graphics System. DeLano Scientific LLC, Palo Alto, CA, USA. http://www.pymol.org.

Delany I, Pacheco AB, Spohn G, Rappuoli R, Scarlato V. (2001) Iron-dependent transcription of the frpB gene of *Helicobacter pylori* is controlled by the Fur repressor protein. J Bacteriol. 183 (16):4932-7.

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

Denkhaus, E, and Salnikow K. (2002) Nickel essentiality, toxicity, and carcinogenicity. Crit. Rev. Oncol. Hematol. 42:35-56.

Devi SM, Ahmed I, Francalacci P, Hussain MA, Akhter Y, Alvi A, Sechi LA, Mégraud F, Ahmed N. (2007) Ancestral European roots of *Helicobacter pylori* in India. BMC Genomics. 8:184.

Devi SM, Ahmed I, Khan AA, Rahman SA, Alvi A, Sechi LA, Ahmed N. (2006) Genomes of *Helicobacter pylori* from native Peruvians suggest admixture of ancestral and modern lineages and reveal a western type cag-pathogenicity island. BMC Genomics. 7:191.

Dhaenens L, Szczebara F, and Husson MO. (1997). Identification, characterization, and immunogenicity of the lactoferrin-binding protein from *Helicobacter pylori.* Infect. Immun. 65:514-518.

Dhaenens L, Szczebara F, Nieuwenhuyse SV, Husson MO. (1999) Comparison of iron uptake in different Helicobacter species Res. Microbiol. 150 475−481.

Dickmanns A, Ballschmiter M, Liebl W, Ficner R. (2006) Structure of the novel alpha-amylase AmyC from *Thermotoga maritima*. Acta Crystallogr D Biol Crystallogr 62(Pt 3):262–270.

Dossumbekova A, Prinz C, Mages J, Lang R, Kusters JG, Van Vliet AH, Reindl W, Backert S, Saur D, Schmid RM, Rad R. (2006) *Helicobacter pylori* HopH (OipA) and bacterial pathogenicity: genetic and functional genomic analysis of hopH gene polymorphisms. J Infect Dis. 194 (10):1346-55.

Dreiseikelmann B. (1994) Translocation of DNA across bacterial membranes. Microbiol Rev. 58 (3):293-316.

Dunn BE, Phadnis SH. (1998) Structure, function and localization of *Helicobacter pylori* urease. Yale J Biol Med. 71 (2):63-73.

Dunn BE, Vakil NB, Schneider BG, Miller MM, Zitzer JB, Peutz T, Phadnis SH. (1997) Localization of

Helicobacter pylori urease and heat shock protein in human gastric biopsies. Infect Immun. 65 (4):1181-8.

Eakanunkul S, Lukat-Rodgers GS, Sumithran S, Ghosh A, Rodgers KR, Dawson JH, and Wilks A. (2005) Characterization of the periplasmic heme-binding protein shut from the heme uptake system of *Shigella dysenteriae*. Biochemistry 44:13179-13191.

Eaton KA, Brooks CL, Morgan DR, Krakowka S. (1991) Essential role of urease in pathogenesis of gastritis induced by *Helicobacter pylori* in gnotobiotic piglets. Infect Immun. 1991 Jul; 59 (7):2470-5.

Eisenberg D, Gill HS, Pfluegl GM, Rotstein SH. (2000) Structure-function relationships of glutamine synthetases. Biochim Biophys Acta. 1477(1-2):122-45.

Emsley P, Cowtan K. (2004) Coot: Model-building tools for molecular graphics. Acta Crystallogr Sect D-Biol Crystallogr 60:2126–2132.

Engel RR, Matsen JM, Chapman SS, Swartz S. (1972) Carbon monoxide production from heme compounds by bacteria. J Bacteriol. 112:1310-1315.

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

Evans P. (2006) Scaling and assessment of data quality. Acta Crystallogr D Biol Crystallogr. 62:72–82.

Fabianek RA, Hennecke H and Thony-Meyer L. (2000) Periplasmic protein thiol:disulfide oxidoreductases of *Escherichia coli*. FEMS Microbiol Rev. 24:303-316.

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

Fan X, Crowe SE, Behar S, Gunasena H, Ye G, Haeberle H, Van Houten N, Gourley WK, Ernst PB, Reyes VE. (1998) The effect of class II major histocompatibility complex expression on adherence of *Helicobacter pylori* and induction of apoptosis in gastric epithelial cells: a mechanism for T helper cell type 1-mediated damage. J Exp Med. 187 (10):1659-69.

Fischer W, Hofreuter D, Haas R. (2001) Natural Transformation, Recombination, and Repair. In: Mobley HLT, Mendz GL, Hazell SL, editors. *Helicobacter pylori*: Physiology and Genetics. Washington (DC): ASM Press; Chapter 22.

Fischer W, Windhager L, Rohrer S, Zeiller M, Karnholz A, Hoffmann R, Zimmer R, Haas R. (2010) Strain-specific genes of *Helicobacter pylori*: genome evolution driven by a novel type IV secretion system and genomic island transfer. Nucleic Acids Res. 38 (18):6089-101.

Fischer W. (2011) Assembly and molecular mode of action of the *Helicobacter pylori* Cag type IV secretion apparatus. FEBS J. 278 (8):1203-12.

Forman D and Burley VJ. (2006) Gastric cancer: global pattern of the disease and an overview of environmental risk factors. Best Pract Res Clin Gastroenterol. 20 (4):633-49.

Foryst-Ludwig A and Naumann M. (2000) p21-activated kinase 1 activates the nuclear factor kappa B (NF-kappa B)-inducing kinase-Ikappa B kinases NF-kappa B pathway and proinflammatory cytokines in *Helicobacter pylori* infection. J Biol Chem. 275 (50):39779-85.

Franco AT, Friedman DB, Nagy TA, Romero-Gallo J, Krishna U, et al. (2009) Delineation of a carcinogenic *Helicobacter pylori* proteome. Mol Cell Proteomics 8(8):1947–1958.

Frenck RW Jr. and Clemens J. (2003) *Helicobacter* in the developing world. Microbes Infect. 5:705-13.

Galmiche A, Rassow J, Doye A, Cagnol S, Chambard JC, Contamin S, de Thillot V, Just I, Ricci V, Solcia E, Van Obberghen E, Boquet P. (2000) The N-terminal 34 kDa fragment of *Helicobacter pylori* vacuolating cytotoxin targets mitochondria and induces cytochrome c release. EMBO J. 19 (23):6361-70.

Gangwer KA, Mushrush DJ, Stauff DL, Spiller B, McClain MS, Cover TL, Lacy DB. (2007) Crystal structure of the *Helicobacter pylori* vacuolating toxin p55 domain. Proc Natl Acad Sci U S A. 104 (41):16293-8.

Garner RM, Fulkerson J Jr, Mobley HL. (1998) *Helicobacter pylori* glutamine synthetase lacks features associated with transcriptional and posttranslational regulation. Infect Immun. 66(5):1839-47.

Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A. (2005) Protein Identification and Analysis Tools on the ExPASy Server; (In) John M. Walker (ed): The Proteomics Protocols Handbook, Humana Press. pp. 571-607.

Gebert B, Fischer W, Weiss E, Hoffmann R, Haas R. (2003) *Helicobacter pylori* vacuolating cytotoxin inhibits T lymphocyte activation. Science. 301 (5636):1099-102.

Gewirtz AT, Yu Y, Krishna US, Israel DA, Lyons SL, Peek RM Jr. (2004) *Helicobacter pylori* flagellin evades toll-like receptor 5-mediated innate immunity. J Infect Dis. 189 (10):1914-20.

Ghose C, Perez-Perez GI, Dominguez-Bello MG, Pride DT, Bravi CM, Blaser MJ. (2002) East Asian genotypes of *Helicobacter pylori* strains in Amerindians provide evidence for its ancient human carriage. Proc Natl Acad Sci U S A. 99 (23):15107-11.

Gill HS, Pfluegl GM, Eisenberg D. (1999) Preliminary crystallographic studies on glutamine synthetase from *Mycobacterium tuberculosis*. Acta Crystallogr D Biol Crystallogr. 55(Pt 4):865-8.

Gill HS, Pfluegl GM, Eisenberg D. (2002) Multicopy crystallographic refinement of a relaxed glutamine synthetase from *Mycobacterium tuberculosis* highlights flexible loops in the enzymatic mechanism and its regulation. Biochemistry. 41(31):9863-72.

Girardin SE, Boneca IG, Carneiro LA, Antignac A, Jéhanno M, Viala J, Tedin K, Taha MK, Labigne A, Zähringer U, Coyle AJ, DiStefano PS, Bertin J, Sansonetti PJ, Philpott DJ. (2003) Nod1 detects a unique muropeptide from gram-negative bacterial peptidoglycan. Science. 300 (5625):1584-7.

Gleiter S and Bardwell JCA. (2008) Disulfide bond isomerization in prokaryotes Biochim Biophys Acta. 1783(4):530–534. doi:10.1016/j.bbamcr.2008.02.009.

Glupczynski Y, Megraud F, Lopez-Brea M, and Andersen LP. (2001) European multicentre survey of in vitro antimicrobial resistance in *Helicobacter pylori*. Eur. J. Clin. Microbiol. Infect. Dis. 20:820-3.

Gobert AP, Cheng Y, Wang JY, Boucher JL, Iyer RK, Cederbaum SD, Casero RA Jr, Newton JC, Wilson KT. (2002) *Helicobacter pylori* induces macrophage apoptosis by activation of arginase II. J Immunol. 168 (9):4692-700.

Gobert V, Gottar M, Matskevich AA, Rutschmann S, Royet J, Belvin M, Hoffmann JA, Ferrandon D. (2003) Dual activation of the Drosophila toll pathway by two pattern recognition receptors. Science. 302 (5653) :2126-30.

Godlewska R, Dzwonek A, Mikula M, Ostrowski J, Pawlowski M, Bujnicki JM & Jagusztyn-Krynicka EK (2006) *Helicobacter pylori* protein oxidation influences the colonization process. Int J Med Microbiol 296:321–324.

Gomis-Rüth FX and Coll M. (2001) Structure of TrwB, a gatekeeper in bacterial conjugation. Int J Biochem Cell Biol. 33 (9):839-43.

Gorden J and Small PL. (1993) Acid resistance in enteric bacteria. Infect Immun. 61 (1):364-7.

Gouet P, Courcelle E, Stuart DI, Metoz F. (1999) ESPript: Analysis of multiple sequence alignments in PostScript. Bioinformatics 15(4):305–308.

Grigg JC, Cooper JD, Cheung J, Heinrichs DE, Murphy ME. (2010) The Staphylococcus aureus siderophore receptor HtsA undergoes localized conformational changes to enclose staphyloferrin A in an arginine-rich binding pocket. J Biol Chem. 285(15):11162-71.

Guo Y, Guo G, Mao X, Zhang W, Xiao J, Tong W, Liu T, Xiao B, Liu X, Feng Y, Zou Q. (2008) Functional identification of HugZ, a heme oxygenase from *Helicobacter pylori*. BMC Microbiol. 8:226.

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

Haas G, Karaali G, Ebermayer K, Metzger WG, Lamer S, ZimnyArndt U, Diescher S, Goebel UB, Vogt K, Roznowski AB, Wiedenmann BJ, Meyer TF, Aebischer T and Jungblut P. (2002) Immunoproteomics of *Helicobacter pylori* infection and relation to gastric disease. Proteomics 2:313-324.

Hatakeyama M, Higashi H. (2005) *Helicobacter pylori* CagA: a new paradigm for bacterial carcinogenesis. Cancer Sci. 96 (12):835-43.

Hatakeyama M. (2008) SagA of CagA in *Helicobacter pylori* pathogenesis. Curr Opin Microbiol. 11 (1):30-7.

He YX, Gui L, Liu YZ, Du Y, Zhou Y, Li P, Zhou CZ. (2009) Crystal structure of *Saccharomyces cerevisiae* glutamine synthetase Gln1 suggests a nanotube-like supramolecular assembly. Proteins. 76(1):249-54.

Hekmat O, Tokuyasu K, Withers SG. (2003) Subsite structure of the endo-type chitin deacetylase from a Deuteromycete, Colletotrichum lindemuthianum: an investigation using steady-state kinetic analysis and MS. Biochem J. 374:369–380.

Henderson DP, and Payne SM. (1993) Cloning and characterization of the *Vibrio cholerae* genes encoding the utilization of iron from haemin and haemoglobin. Mol. Microbiol. 7:461-469.

Heras B, Edeling MA, Schirra HJ, Raina S, Martin JL. (2004) Crystal structures of the DsbG disulfide isomerase reveal an unstable disulfide. Proc Natl Acad Sci U S A. 101(24):8876-81.

Hillemann D, Dammann T, Hillemann A, Wohlleben W. (1993) Genetic and biochemical characterization of the two glutamine synthetases GSI and GSII of the phosphinothricyl-alanyl-alanine producer, *Streptomyces viridochromogenes* Tü494. J Gen Microbiol. 139(8):1773-83.

Hiniker A and Bardwell JC. (2004) In vivo substrate specificity of periplasmic disulfide oxidoreductases J. Biol. Chem. 279:12967-12973.

Ho WW, Li H, Eakanunkul S, Tong Y, Wilks A, Guo M, Poulos TL (2007) Holo- and apo-bound structures of bacterial periplasmic heme-binding proteins. J Biol Chem. 282(49):35796-802..

Hofreuter D and Haas R. (2002) Characterization of two cryptic *Helicobacter pylori* plasmids: a putative source for horizontal gene transfer and gene shuffling. J Bacteriol. 184 (10):2755-66.

Holm L and Rosenström P. (2010) Dali server: conservation mapping in 3D. Nucl. Acids Res. 38, W545-549.

Holm L and Sander C. (1993) Protein structure comparison by alignment of distance matrices. J Mol Biol. 233(1):123-38.

Hornung JM, Jones HA, and Perry RD. (1996) The hmu locus of *Yersinia pestis* is essential for utilization of free haemin and haemsprotein complexes as iron sources. Mol. Microbiol. 20:725-739.

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

Hunt JB, Ginsburg A. (1980) Mn2+ and substrate interactions with glutamine synthetase from *Escherichia coli*.

J Biol Chem. 255(2):590-4.

Hunt JB, Smyrniotis PZ, Ginsburg A, Stadtman ER. (1975) Metal ion requirement by glutamine synthetase of *Escherichia coli* in catalysis of gamma-glutamyl transfer. Arch Biochem Biophys. 166(1):102-24.

Husson MA, Legrand D, Spik G, and Leclerc H. (1993). Iron acquisition by *Helicobacter pylori*: importance of human lactoferrin. Infect. Immun. 61:2694-2697.

Ishino Y, Morgenthaler P, Hottinger H, and Soll D. (1992) Organization and nucleotide sequence of the glutamine synthetase (glnA) gene from *Lactobacillus delbrueckii* subsp. bulgaricus. Appl. Environ. Microbiol. 58:3165-3169.

Jakszyn P, Agudo A, Lujan-Barroso L, Bueno-de-Mesquita HB, Jenab M, Navarro C, Palli D, Boeing H, Manjer J, Numans ME, Igali L, Boutron-Ruault MC, Clavel-Chapelon F, Morois S, Grioni S, Panico S, Tumino R, Sacerdote C, Quirós JR, Molina-Montes E, Castaño JM, Barricarte A, Amiano P, Khaw KT, Wareham N, Allen NE, Key TJ, Jeurnink SM, Peeters PH, Bamia C, Valanou E, Trichopoulou A, Kaaks R, Lukanova A, Bergmann MM, Lindkvist B, Stenling R, Johansson I, Dahm CC, Overvad K, Olsen A, Tjonneland A, Skeie G, Broderstad AR, Lund E, Michaud DS, Mouw T, Riboli E, González CA. (2011) Dietary intake of heme iron and risk of gastric cancer in the European prospective investigation into cancer and nutrition (EURGAST- EPIC) study. Int J Cancer. doi: 10.1002/ijc.26263.

Jiang Q, Hiratsuka K, Taylor DE. (1996) Variability of gene order in different *Helicobacter pylori* strains contributes to genome diversity. Mol Microbiol. 20 (4):833-42.

Jungblut PR, Bumann D, Haas G, Zimny-Arndt U, Holland P, Lamer S, Siejak F, Aebischer A, and Meyer TF. (2000) Comparative proteome analysis of *Helicobacter pylori*. Mol. Microbiol. 36:710-725.

Kaakoush NO, Kovach Z, and Mendz GL. (2007) Potential role of thiol:disulfide oxidoreductases in the pathogenesis of *Helicobacter pylori* , FEMS Immunol Med Microbiol 50:177-183.

Kabir S. (2007) The current status of *Helicobacter pylori* vaccines: a review. Helicobacter. 12 (2):89-102.

Kadokura H, Katzen F. and Beckwith J. (2003) Protein disulfide bond formation in prokaryotes. Annu Rev Biochem 72:111-135.

Kamio Y, Nakamura K. (1987) Putrescine and cadaverine are constituents of peptidoglycan in Veillonella alcalescens and Veillonella parvula. J Bacteriol. 169(6):2881-2884.

Kang J, Blaser MJ. (2006) Bacterial populations as perfect gases: genomic integrity and diversification tensions in *Helicobacter pylori*. Nat Rev Microbiol. 4 (11):826-36.

Karnholz A, Hoefler C, Odenbreit S, Fischer W, Hofreuter D, Haas R. (2006) Functional and topological characterization of novel components of the comB DNA transformation competence system in *Helicobacter pylori*. J Bacteriol. 188 (3):882-93.

Kavanagh KL, Jörnvall H, Persson B, Oppermann U. (2008) Medium- and short-chain dehydrogenase/reductase gene and protein families: the SDR superfamily: functional and structural diversity within a family of metabolic and regulatory enzymes. Cell Mol Life Sci. 65(24):3895-906.

Kavermann H, Burns BP, Angermuller K, Odenbreit S, Fischer W, Melchers K, Haas R. (2003) Identification and characterization of *Helicobacter pylori* genes essential for gastric colonization. J Exp Med. 197 (7):813-22.

Khamri W, Moran AP, Worku ML, Karim QN, Walker MM, Annuk H, Ferris JA, Appelmelk BJ, Eggleton P, Reid KB, Thursz MR. (2005) Variations in *Helicobacter pylori* lipopolysaccharide to evade the innate immune component surfactant protein D. Infect Immun. 73 (11):7677-86.

Kim N, Weeks DL, Shin JM, Scott DR, Young MK, Sachs G. (2002) Proteins released by *Helicobacter pylori* in vitro. J Bacteriol. 184(22):6155-62.

Kimmel B, Bosserhoff A, Frank R, Gross R, Goebel W, and Beier D. (2000) Identification of immunodominant antigens from *Helicobacter pylori* and evaluation of their reactivities with sera from patients with different gastroduodenal pathologies. Infect. Immun. 68:915-920.

Kivi M, Tindberg Y, Sörberg M, Casswall TH, Befrits R, Hellström PM, Bengtsson C, Engstrand L, Granström M. (2003) Concordance of *Helicobacter pylori* strains within families. J Clin Microbiol. 41 (12):5604-8.

Kneidinger B, Marolda C, Graninger M, Zamyatina A, McArthur F, Kosma P, Valvano MA, Messner P. (2002) Biosynthesis pathway of ADP-L-glycero-beta-D-manno-heptose in *Escherichia coli*. J Bacteriol. 184(2):363-9.

Kobayashi T, Ito K. (1999) Respiratory chain strongly oxidizes the CXXC motif of DsbB in the *Escherichia coli* disulfide bond formation pathway. Embo J. 18:1192-8.

Kowatz T, Morrison JP, Tanner ME, Naismith JH. (2010) The crystal structure of the Y140F mutant of ADP-L-glycero-D-manno-heptose 6-epimerase bound to ADP-beta-D-mannose suggests a one base mechanism. Protein Sci. 19(7):1337-43.

Krajewski WW, Collins R, Holmberg-Schiavone L, Jones TA, Karlberg T, Mowbray SL. (2008) Crystal structures of mammalian glutamine synthetases illustrate substrate-induced conformational changes and provide opportunities for drug and herbicide design. J Mol Biol. 375(1):217-28.

Krajewski WW, Jones TA, Mowbray SL. (2005) Structure of *Mycobacterium tuberculosis* glutamine synthetase in complex with a transition-state mimic provides functional insights. Proc Natl Acad Sci U S A. 102(30):10499-504.

Krieger E, Joo K, Lee J, Lee J, Raman S, Thompson J, Tyka M, Baker D, Karplus K (2009) Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: Four approaches that performed well in CASP8 Proteins. 77 (Suppl 9):114-22.

Kumada Y, Benson DR, Hillemann D, Hosted TJ, Rochefort DA, Thompson CJ, Wohlleben W, Tateno Y. (1993) Evolution of the glutamine synthetase gene, one of the oldest existing and functioning genes. Proc Natl Acad Sci U S A. 90(7):3009-13.

Kusters JG, van Vliet AHM & Kuipers EJ (2006) Pathogenesis of *Helicobacter pylori* Infection. Clin Microbiol Rev. 19:449-490.

Kwok T, Zabler D, Urman S, Rohde M, Hartig R, Wessler S, Misselwitz R, Berger J, Sewald N, König W, Backert S. (2007) *Helicobacter* exploits integrin for type IV secretion and kinase activation. Nature. 449 (7164):862-6.

Lara-Ramírez EE, Segura-Cabrera A, Guo X, Yu G, García-Pérez CA, Rodríguez-Pérez MA. (2011) New implications on genomic adaptation derived from the *Helicobacter pylori* genome comparison. PLoS One. 6 (2):e17300.

Laskowski RA, Macarthur MW, Moss DS, Thornton JM. (1993) Procheck - a program to check the stereochemical quality of protein structures. J Appl Crystallogr. 26:283–291.

Leduc D, Gallaud J, Stingl K, de Reuse H. (2010) Coupled amino acid deamidase-transport systems essential for *Helicobacter pylori* colonization. Infect Immun. 78(6):2782-92.

Lee J, Lee SY, Lee JH. (2006) Production of Antibody against *Helicobacter pylori* HP0231. Korean J Lab Med. 26(2):98-102.

Lei M, Aebi U, Heidner EG, Eisenberg D. (1979) Limited proteolysis of glutamine synthetase is inhibited by glutamate and by feedback inhibitors. J Biol Chem. 254(8):3129-34.

Leslie AGW. (2006) The integration of macromolecular diffraction data. Acta Crystallogr D Biol Crystallogr. 62:48–57.

Leunk RD, Johnson PT, David BC, Kraft WG, Morgan DR. (1988) Cytotoxic activity in broth-culture filtrates of *Campylobacter pylori*. J Med Microbiol. 26 (2):93-9.

Liaw SH, Pan C, Eisenberg D. (1993) Feedback inhibition of fully unadenylylated glutamine synthetase from *Salmonella typhimurium* by glycine, alanine, and serine. Proc Natl Acad Sci U S A. 90(11):4996-5000.

Linz B, Balloux F, Moodley Y, Manica A, Liu H, Roumagnac P, Falush D, Stamer C, Prugnolle F, van der Merwe SW, Yamaoka Y, Graham DY, Perez-Trallero E, Wadstrom T, Suerbaum S, Achtman M. (2007) An African origin for the intimate association between humans and *Helicobacter pylori*. Nature. 445 (7130):915-8.

Lupetti P, Heuser JE, Manetti R, Massari P, Lanzavecchia S, Bellon PL, Dallai R, Rappuoli R, Telford JL. (1996) Oligomeric and subunit structure of the *Helicobacter pylori* vacuolating cytotoxin. J Cell Biol. 133(4):801-7.

Luthy L, Grutter MG, Mittl PR. (2002) The crystal structure of *Helicobacter pylori* cysteine-rich protein B reveals a novel fold for a penicillin-binding protein. J Biol Chem. 277(12):10187–10193.

Mahdavi J, Sondén B, Hurtig M, Olfat FO, Forsberg L, Roche N, Angstrom J, Larsson T, Teneberg S, Karlsson KA, Altraja S, Wadström T, Kersulyte D, Berg DE, Dubois A, Petersson C, Magnusson KE, Norberg T, Lindh F, Lundskog BB, Arnqvist A, Hammarström L, Borén T. (2002) *Helicobacter pylori* SabA adhesin in persistent infection and chronic inflammation. Science. 297(5581):573-8.

Maier RJ, Fu C, Gilbert J, Moshiri F, Olson J, and Plaut AG. (1996) Hydrogen uptake hydrogenase in *Helicobacter pylori*. FEMS Microbiol. Lett. 141:71-6.

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

Malfertheiner P, Megraud F, O'Morain C, Bazzoli F, El-Omar E, Graham D, Hunt R, Rokkas T, Vakil N, Kuipers EJ. (2007) Current concepts in the management of *Helicobacter pylori* infection: the Maastricht III Consensus Report. Gut. (6):772-81.

Marchler-Bauer A, Lu S, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C, Fong JH, Geer LY, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Jackson JD, Ke Z, Lanczycki CJ, Lu F, Marchler GH, Mullokandov M, Omelchenko MV, Robertson CL, Song JS, Thanki N, Yamashita RA, Zhang D, Zhang N, Zheng C, Bryant SH (2011) "CDD: a Conserved Domain Database for the functional annotation of proteins". Nucleic Acids Res. 39 (Database issue): D225-9. doi:10.1093/nar/gkq1189.

Marcus EA, Scott DR. (2001) Cell lysis is responsible for the appearance of extracellular urease in *Helicobacter pylori*. Helicobacter. 6(2):93-9.

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

Martin JL. (1995) Thioredoxin-a fold for all reasons. Structure. 3(3):245-50.

Mattle D, Zeltina A, Woo JS, Goetz BA, Locher KP. (2010) Two stacked heme molecules in the binding pocket of the periplasmic heme-binding protein HmuT from *Yersinia pestis*. J Mol Biol. 404(2):220-31.

McAtee CP, Lim MY, Fung K, Velligan M, Fry K, Chow T. and Berg DE. (1998) Identification of potential

diagnostic and vaccine candidates of *Helicobacter pylori* by two-dimensional gel electrophoresis, sequence analysis, and serum profiling. Clin. Diagn. Lab. Immunol. 5:537-542.

McCarthy AA, Haebel PW, Torronen A, Rybin,V, Baker EN, and Metcalf P. (2000) Crystal structure of the protein disulfide bond isomerase, DsbC, from *Escherichia coli*. Nat. Struct. Biol. 7:196–199.

McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. (2007) Phaser crystallographic software. J Appl Crystallogr. 40(Pt 4):658-674.

McGee DJ and Mobley HL. (1999). Mechanisms of *Helicobacter pylori* infection: bacterial factors. Curr. Top. Microbiol. Immunol. 241:155–180.

Merrell DS, Thompson LJ, Kim CC, Mitchell H, Tompkins LS, et al. (2003) Growth phase-dependent response of *Helicobacter pylori* to iron starvation. Infect Immun 71:6510-6525.

Messens J, Collet JF. (2006) Pathways of disulfide bond formation in Esche-richia coli, Int. J. Biochem. Cell Biol. 38(2006):1050-1062.

Meyer JM, Silliman NP, Wang W, Siepman NY, Sugg JE, Morris D, Zhang J, Bhattacharyya H, King EC and Hopkins RJ. (2002) Risk factors for *Helicobacter pylori* resistance in the United States: the surveillance of *H. pylori* antimicrobial resistance partnership (SHARP) study, 1993-1999. Ann. Intern. Med. 136:13-24.

Meylan E, Tschopp J, Karin M. (2006) Intracellular pattern recognition receptors in the host response. Nature. 442(7098):39-44.

Mills M and Payne SM. (1995) Genetics and regulation of heme iron transport in *Shigella dysenteriae* and detection of an analogous system in *Escherichia coli* O157: H7. J. Bacteriol. 177:3004-3009.

Miret S, Simpson RJ, McKie AT. (2003) Physiology and molecular biology of dietary iron absorption. Annu Rev Nutr. 23:283-301.

Missiakas D, Georgopoulos C, Raina S. (1994) The *Escherichia coli* dsbC (xprA) gene encodes a periplasmic protein involved in disulfide bond formation. Embo J. 13:2013-20.

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

Mobley HLT. (2001) Urease. *In* H. L. T. Mobley, G. L. Mendz, and S. L. Hazell (ed.), *Helicobacter pylori*: Physiology and Genetics. ASM Press, Washington DC.

Molinari M, Salio M, Galli C, Norais N, Rappuoli R, Lanzavecchia A, Montecucco C. (1998) Selective inhibition of Ii-dependent antigen presentation by *Helicobacter pylori* toxin VacA. J Exp Med. 187(1):135-40.

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

Monteiro MA, Chan KH, Rasko DA, Taylor DE, Zheng PY, Appelmelk BJ, Wirth HP, Yang M, Blaser MJ, Hynes SO, Moran AP, Perry MB. (1998) Simultaneous expression of type 1 and type 2 Lewis blood group antigens by *Helicobacter pylori* lipopolysaccharides Molecular mimicry between *H. pylori* lipopolysaccharides and human gastric epithelial cell surface glycoforms. J Biol Chem. 273(19):11533-43.

Montemurro P, Nishioka H, Dundon WG, 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. Eur J Immunol. 32(3):671-6.

Moran AP. (2001) Molecular structure, biosynthesis, and pathogenic roles of lipopolysaccharides, in: H.L.T. Mobley, G.L. Mendz and S.L. Hazell(eds.), *Helicobacter pylori*: Physiology and Genetics, ASM Press, Washington (DC).

Muhsen K and Cohen D. (2008) *Helicobacter pylori* infection and iron stores: a systematic review and meta-analysis. Helicobacter 13:323-340.

Müller A, Wilkinson AJ, Wilson KS, Duhme-Klair AK. (2006) An [{Fe(mecam)}2]6- bridge in the crystal structure of a ferric enterobactin binding protein. Angew Chem Int Ed Engl. 45(31):5132-6.

Munford R, Lu M, Varley A. (2009) Chapter 2: Kill the bacteria...and also their messengers? Adv Immunol. 103:29-48.

Muotiala A, Helander IM, Pyhälä L, Kosunen TU, Moran AP. (1992) Low biological activity of *Helicobacter pylori* lipopolysaccharide. Infect Immun. 60 (4):1714-6.

Murshudov GN, Vagin AA, Dodson EJ. (1997) Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr Sect D-Biol Crystallogr. 53:240–255.

Nakamoto H. Bardwell JC. (2004) Catalysis of disulfide bond formation and isomerization in the *Escherichia coli* periplasm, Biochim. Biophys. Acta. 1694:111-119.

Nayar S, Bhattacharyya D. (1997) UDP-galactose 4-epimerase from *Escherichia coli*: existence of a catalytic monomer. FEBS Lett. 409(3):449-51.

Nesić D, Miller MC, Quinkert ZT, Stein M, Chait BT, Stebbins CE. (2010) *Helicobacter pylori* CagA inhibits PAR1-MARK family kinases by mimicking host substrates. Nat Struct Mol Biol. 17(1):130-2.

Ni Y, McPhie P, Deacon A, Ealick S, Coleman WG Jr. (2001) Evidence that NADP+ is the physiological cofactor of ADP-L-glycero-D-mannoheptose 6-epimerase. J Biol Chem. 276(29):27329-34.

O'Rourke J, Bode G. (2001) Morphology and Ultrastructure. In: Mobley HLT, Mendz GL, Hazell SL, editors. *Helicobacter pylori*: Physiology and Genetics. Washington (DC): ASM Press; Chapter 6.

O'Toole PW, Kostrzynska M, Trust TJ. (1994) Non-motile mutants of *Helicobacter pylori* and *Helicobacter mustelae* defective in flagellar hook production. Mol Microbiol. 14(4):691-703.

O'Toole PW, Lane MC, Porwollik S. (2000) *Helicobacter pylori* motility. Microbes Infect. 2(10):1207-14.

Odenbreit S, Püls J, Sedlmaier B, Gerland E, Fischer W, Haas R. (2000) Translocation of *Helicobacter pylori* CagA into gastric epithelial cells by type IV secretion. Science. 287(5457):1497-500.

Odenbreit S, Till M, Hofreuter D, Faller G, Haas R. (1999) Genetic and functional characterization of the alpAB gene locus essential for the adhesion of *Helicobacter pylori* to human gastric tissue. Mol Microbiol. 31(5):1537-48.

Oh JD, Kling-Bäckhed H, Giannakis M, Xu J, Fulton RS, Fulton LA, Cordum HS, Wang C, Elliott G, Edwards J, Mardis ER, Engstrand LG, Gordon JI. (2006) The complete genome sequence of a chronic atrophic gastritis *Helicobacter pylori* strain: evolution during disease progression. Proc Natl Acad Sci U S A. 103(26):9999-10004.

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

Otto BR, Verweij-van Vught AM, MacLaren DM. (1992) Transferrins and heme-compounds as iron sources for pathogenic bacteria. Crit Rev Microbiol. 18(3):217-33.

Papadopoulos JS and Agarwala R. (2007) Bioinformatics. 23:1073-79.

Parsonnet J. (1995) The incidence of *Helicobacter pylori* infection. Aliment Pharmacol Ther. 9 (Suppl 2):45-51.

Patti GJ, Chen J, Schaefer J, Gross ML. (2008) Characterization of structural variations in the peptidoglycan of vancomycin-susceptible *Enterococcus faecium*: Understanding glycopeptide-antibiotic binding sites using mass spectrometry. J Am Soc Mass Spectrom 19(10):1467–1475.

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. 153(Pt 9):2896-909.

Peck B, Ortkamp M, Diehl KD, Hundt E, Knapp B. (1999) Conservation, localization and expression of HopZ, a protein involved in adhesion of *Helicobacter pylori*. Nucleic Acids Res. 27(16):3325-33.

Peek JA and Taylor RK. (1992) Characterization of a periplasmic thiol : disulfide interchange protein required for the functional maturation of secreted virulence factors of Vibrio cholerae. Proc Natl Acad Sci U S A 89:6210–6214.

Peek RM Jr. (2002) New insights into microbially initiated gastric malignancies: beyond the usual suspects. Gastroenterology. 123(5):1739-40; discussion 1740-1.

Peek RM and Blaser MJ. (2002) *Helicobacter pylori* and gastrointestinal tract adenocarcinomas. Nat Rev Cancer. 2:28-37

Perry RD, Shah J, Bearden SW, Thompson JM, Fetherston JD. (2003) *Yersinia pestis* TonB: role in iron, heme, and hemoprotein utilization. Infection and immunity. 71:4159-4162. doi: 10.1128/IAI.71.7.4159-4162.2003.

Peterson RA, Hoepf T, Eaton KA. (2003) Adoptive transfer of splenocytes in SCID mice implicates CD4+ T cells in apoptosis and epithelial proliferation associated with *Helicobacter pylori*-induced gastritis. Comp Med. 53(5):498-509.

Peuckert F, Ramos-Vega AL, Miethke M, Schwörer CJ, Albrecht AG, Oberthür M, Marahiel MA. (2011) The siderophore binding protein FeuA shows limited promiscuity toward exogenous triscatecholates. Chem Biol. 18(7):907-19.

Pflock M, Finsterer N, Joseph B, Mollenkopf H, Meyer TF, Beier D. (2006a) Characterization of the ArsRS regulon of *Helicobacter pylori*, involved in acid adaptation. J Bacteriol. 188(10):3449-62.

Pflock M, Kennard S, Finsterer N, Beier D. (2006b) Acid-responsive gene regulation in the human pathogen *Helicobacter pylori*. J Biotechnol. 126(1):52-60.

Phadnis SH, Parlow MH, Levy M, Ilver D, Caulkins CM, Connors JB, Dunn BE. (1996) Surface localization of *Helicobacter pylori* urease and a heat shock protein homolog requires bacterial autolysis. Infect Immun. 64(3):905-12.

Pinto-Santini D, Salama NR. (2005) The biology of *Helicobacter pylori* infection, a major risk factor for gastric adenocarcinoma. Cancer Epidemiol Biomarkers Prev. 14 (8):1853-8.

Polk DB, Peek RM Jr. (2010) *Helicobacter pylori*: gastric cancer and beyond. Nat Rev Cancer. 10 (6):403-14.

Raczko AM, Bujnicki JM, Pawlowski M, Godlewska R, Lewandowska M & Jagusztyn-Krynicka EK (2005) Characterization of new DsbB-like thioloxidoreductases of Campylobacter jejuni and *Helicobacter pylori* and classification of the DsbB family based on phylogenomic, structural and functional criteria. Microbiology 151:219–231.

Raetz CR, Whitfield C. (2002) Lipopolysaccharide endotoxins. Annu Rev Biochem. 71:635-700.

Raina S. and Missiakas D. (1997) Making and breaking the disulfide bondsAnnu. Rev. Microbiol. 51:179-202.

Ramachandran S, Kota P, Ding F. and Dokholyan, NV. (2011) PROTEINS: Structure, Function and

Bioinformatics. 79:261-270.

Ramazzina I, Cendron L, Folli C, Berni R, Monteverdi D, et al. (2008) Logical identification of an allantoinase analog (PuuE) recruited from polysaccharide deacetylases. J Biol Chem. 283(34):23295–23304. 10.1074/jbc.M801195200.

Reedy CJ, Elvekrog MM, Gibney BR. (2008) Development of a heme protein structure-electrochemical function database. Nucleic Acids Res. 36(Database issue):D307-13.

Reitzer LJ. (1996) Ammonia assimilation and the biosynthesis of glutamine, glutamate, aspartate, asparagine, L-alanine, and D-alanine, p. 391–407. In F. C. Neidhardt, R. Curtiss III, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella*: cellular and molecular biology, 2nd ed. ASM Press, Washington, D.C.

Ridley KA, Rock JD, Li Y, Ketley JM. (2006) Heme utilization in *Campylobacter jejuni*. J Bacteriol. 188:7862-7875. doi: 10.1128/JB.00994-06.

Rizzi M, Tonetti M, Vigevani P, Sturla L, Bisso A, Flora AD, Bordo D, Bolognesi M. (1998) GDP-4-keto-6-deoxy-D-mannose epimerase/reductase from *Escherichia coli*, a key enzyme in the biosynthesis of GDP-L-fucose, displays the structural characteristics of the RED protein homology superfamily. Structure. 6(11):1453-65.

Rohde M, Püls J, Buhrdorf R, Fischer W, Haas R. (2003) A novel sheathed surface organelle of the *Helicobacter pylori* cag type IV secretion system. Mol Microbiol. 49(1):219-34

Rost B, Yachdav G and Liu J. (2004) The PredictProtein Server. Nucleic Acids Research 32(Web Server issue):W321-W326.

Rothenbacher D, Brenner H. (2003) Burden of *Helicobacter pylori* and *H. pylori*-related diseases in developed countries: recent developments and future implications. Microbes Infect. 5(8):693-703.

Roy A, Kucukural A, Zhang Y. (2010) I-TASSER: a unified platform for automated protein structure and function prediction. Nature Protocols, vol 5:725-738.

Rupnow MF, Shachter RD, Owens DK, Parsonnet J. (2001) Quantifying the population impact of a prophylactic *Helicobacter pylori* vaccine. Vaccine. 20 (5-6):879-85.

Rychlewski L, Jaroszewski L, Li W. and Godzik A. (2000) Comparison of sequence profiles. Strategies for structural predictions using sequence information. Protein Science. 9:232-241.

Sabarth N, Hurwitz R, Meyer TF. and Bumann D. (2002) Multiparameter selection of *Helicobacter pylori* antigens identifies two novel antigens with high protective efficacy. Infect Immun. 70:6499–6503.

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

Saitou N, Nei M. (1987) The neighbor-joining method: A new method for reconstructing phylogenetic trees. Mol Biol Evol. 4(4):406-425.

Salama NR, Otto G, Tompkins L, Falkow S. (2001) Vacuolating cytotoxin of *Helicobacter pylori* plays a role during colonization in a mouse model of infection. Infect Immun. 69(2):730-6.

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

Savvides SN, Yeo HJ, Beck MR, Blaesing F, Lurz R, Lanka E, Buhrdorf R, Fischer W, Haas R, Waksman G.

(2003) VirB11 ATPases are dynamic hexameric assemblies: new insights into bacterial type IV secretion. EMBO J. 22(9):1969-80.

Schreiber S, Bücker R, Groll C, Azevedo-Vethacke M, Garten D, Scheid P, Friedrich S, Gatermann S, Josenhans C, Suerbaum S. (2005) Rapid loss of motility of *Helicobacter pylori* in the gastric lumen in vivo. Infect Immun. 73 (3) :1584-9.

Schreiber S, Konradt M, Groll C, Scheid P, Hanauer G, et al. (2004) The spatial orientation of *Helicobacter pylori* in the gastric mucus. Proc Natl Acad Sci U S A. 101:5024-5029.

Schüttelkopf AW, van Aalten DM. (2004) PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. Acta Crystallogr D Biol Crystallogr. 60(Pt 8):1355-63.

Sewald X, Fischer W, Haas R. (2008) Sticky socks: *Helicobacter pylori* VacA takes shape. Trends Microbiol. 16 (3):89-92.

Shapiro BM, Stadtman ER. (1968) 5'-adenylyl-O-tyrosine. The novel phosphodiester residue of adenylylated glutamine synthetase from *Escherichia coli*. J Biol Chem. 243(13):3769-71.

Shevchik VE, Condemine G, Robert-Baudouy J. (1994) Characterization of DsbC, a periplasmic protein of Erwinia chrysanthemi and *Escherichia coli* with disulfide isomerase activity. Embo J. 13:2007-12.

Smith TG, Lim JM, Weinberg MV, Wells L, Hoover TR. (2007) Direct analysis of the extracellular proteome from two strains of *Helicobacter pylori*. Proteomics 7(13):2240-2245.

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

Stenson TH and Weiss A. (2002) DsbA and DsbC are required for secretion of pertussis toxin by Bordetella pertussis. Infect Immun 70:2297-2303.

Stojiljkovic I and Hantke K. (1992) Haemin uptake system of *Yersinia enterocolitica:* similarities with other TonB-dependent systems in gram-negative bacteria. EMBO J. 11:4359-4367.

Stojiljkovic I and Hantke K. (1994) Transport of haemin across the cytoplasmic membrane through a haemin-specific periplasmic binding-protein-dependent transport system in *Yersinia enterocolitica*. Mol Microbiol. 13(4):719-32.

Stoof J and Kuipers EJ, (2010a) van Vliet AH.Characterization of NikR-responsive promoters of urease and metal transport genes of *Helicobacter mustelae*. Biometals. 23(1):145-59.

Stoof J, Kuipers EJ, Klaver G, van Vliet AH. (2010b). An ABC transporter and a TonB ortholog contribute to *Helicobacter mustelae* nickel and cobalt acquisition. Infect Immun. (10):4261-7.

Streicher SL, Bender RA, Magasanik B. (1975) Genetic control of glutamine synthetase in *Klebiella aerogenes*. J Bacteriol. 121(1):320-31.

Strober W, Murray PJ, Kitani A, Watanabe T (2006) Signalling pathways and molecular interactions of NOD1 and NOD2. Nat Rev Immunol 6(1): 9-20.

Suarez G, Reyes VE, Beswick EJ. (2006) Immune response to H pylori. World J Gastroenterol. 12 (35):5593-8.

Suerbaum S and Josenhans C. (1999) Virulence factors of *Helicobacter pylori*: implications for vaccine development. Mol Med Today. 5(1):32-9.

Suerbaum S and Josenhans C. (2007) *Helicobacter pylori* evolution and phenotypic diversification in a changing host. Nat Rev Microbiol. 5(6):441-52.

Suerbaum S and Michetti P. (2002) *Helicobacter pylori* infection. N Engl J Med. 347(15):1175-86.

Suerbaum S, Smith JM, Bapumia K, Morelli G, Smith NH, Kunstmann E, Dyrek I, Achtman M. (1998) Free recombination within *Helicobacter pylori*. Proc Natl Acad Sci U S A. 95 (21):12619-24.

Suits MD, Lang J, Pal GP, Couture M, Jia Z. (2009) Structure and heme binding properties of *Escherichia coli* O157:H7 ChuX. Protein Sci. 18(4):825-38.

Sunderman FW Jr. (1993) Biological monitoring of nickel in humans. Scand. J. Work Environ. Health 19 (Suppl 1):34-8.

Sundrud MS, Torres VJ, Unutmaz D, Cover TL. (2004) Inhibition of primary human T cell proliferation by *Helicobacter pylori* vacuolating toxin (VacA) is independent of VacA effects on IL-2 secretion. Proc Natl Acad Sci U S A. 101(20):7727-32.

Takata T, Aras R, Tavakoli D, Ando T, Olivares AZ, Blaser MJ. (2002) Phenotypic and genotypic variation in methylases involved in type II restriction-modification systems in *Helicobacter pylori*. Nucleic Acids Res. 30(11):2444-52.

Takeda K, Kaisho T, Akira S. (2003) Toll-like receptors. Annu Rev Immunol. 21:335-76.

Tan S, Noto JM, Romero-Gallo J, Peek RM Jr, Amieva MR. (2011) *Helicobacter pylori* Perturbs Iron Trafficking in the Epithelium to Grow on the Cell Surface. PLoS Pathog. 7(5)e1002050. doi:10.1371/journal.ppat.1002050

Tanaka J, Suzuki T, Mimuro H, Sasakawa C. (2003) Structural definition on the surface of *Helicobacter pylori* type IV secretion apparatus. Cell Microbiol. 5(6):395-404.

Tegtmeyer N, Wessler S, Backert S. (2011) Role of the cag-pathogenicity island encoded type IV secretion system in *Helicobacter pylori* pathogenesis. FEBS J. 278(8):1190-202.

Telford JL, Ghiara P, Dell'Orco M, Comanducci M, Burroni D, Bugnoli M, Tecce MF, Censini S, Covacci A, Xiang Z. (1994) Gene structure of the *Helicobacter pylori* cytotoxin and evidence of its key role in gastric disease. J Exp Med. 179(5):1653-58.

Terradot L, Bayliss R, Oomen C, Leonard GA, Baron C, Waksman G. (2005) Structures of two core subunits of the bacterial type IV secretion system, VirB8 from *Brucella suis* and ComB10 from *Helicobacter pylori*. Proc Natl Acad Sci U S A. 102(12):4596-601.

Terwilliger TC, Adams PD, Read RJ, McCoy AJ, Moriarty NW, Grosse-Kunstleve RW, Afonine PV, Zwart PH, Hung LW. (2009) Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard.Acta Crystallogr D Biol Crystallogr. 65(Pt 6):582-601.

Terwilliger TC, Grosse-Kunstleve RW, Afonine PV, Moriarty NW, Zwart PH, Hung LW, Read RJ, Adams PD. (2007) Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. Acta Crystallogr D Biol Crystallogr. 64(Pt 1):61-9.

Thiberge JM, Boursaux-Eude C, Lehours P, Dillies MA, Creno S, Coppée JY, Rouy Z, Lajus A, Ma L, Burucoa C, Ruskoné-Foumestraux A, Courillon-Mallet A, De Reuse H, Boneca IG, Lamarque D, Mégraud F, Delchier JC, Médigue C, Bouchier C, Labigne A, Raymond J. (2010) From array-based hybridization of *Helicobacter pylori* isolates to the complete genome sequence of an isolate associated with MALT lymphoma. BMC Genomics. 11:368.

Thoden JB, Frey PA, Holden HM. (1996a) High-resolution X-ray structure of UDP-galactose 4-epimerase complexed with UDP-phenol. Protein Sci. 5(11):2149-61.

Thoden JB, Frey PA, Holden HM. (1996b) Molecular structure of the NADH/UDP-glucose abortive complex of UDP-galactose 4-epimerase from *Escherichia coli*: implications for the catalytic mechanism.

Biochemistry. 35(16):5137-44.

Thoden JB, Holden HM. (1998) Dramatic differences in the binding of UDP-galactose and UDP-glucose to UDP-galactose 4-epimerase from *Escherichia coli*. Biochemistry. 37(33):11469-77.

Thompson JD, Higgins DG, Gibson TJ. (1994) CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22(22):4673-4680.

Tillier ER, Collins RA. (2000) Genome rearrangement by replication-directed translocation. Nat Genet. 26 (2):195-7.

Tokuyasu K, Mitsutomi M, Yamaguchi I, Hayashi K, Mori Y. (2000) Recognition of Chitooligosaccharides and Their N-Acetyl Groups by Putative Subsites of Chitin Deacetylase from a Deuteromycete, Colletotrichum lindemuthianum. Biochemistry. 39: 8837-8843.

Tomb JF, White O, Kerlavage AR, Clayton RA, Sutton GG, Fleischmann RD, Ketchum KA, Klenk HP, Gill S, Dougherty BA, Nelson K, Quackenbush J, Zhou L, Kirkness EF, Peterson S, Loftus B, Richardson D, Dodson R, Khalak HG, Glodek A, McKenney K, Fitzegerald LM, Lee N, Adams MD, Hickey EK, Berg DE, Gocayne JD, Utterback TR, Peterson JD, Kelley JM, Cotton MD, Weidman JM, Fujii C, Bowman C, Watthey L, Wallin E, Hayes WS, Borodovsky M, Karp PD, Smith HO, Fraser CM, Venter JC. (1997) The complete genome sequence of the gastric pathogen *Helicobacter pylori*. Nature. 388 (6642) :539-47.

Tonello F, Dundon WG, Satin B, Molinari M, Tognon G, Grandi G, Del Giudice G, Rappuoli R, Montecucco C. (1999) The *Helicobacter pylori* neutrophil-activating protein is an iron-binding protein with dodecameric structure. Mol Microbiol. 34(2):238-46.

Tong Y and Guo M. (2007) Cloning and characterization of a novel periplasmic heme-transport protein from the human pathogen *Pseudomonas aeruginosa*. J. Biol. Inorg. Chem. 12:735-750

Torres, J., M. Camorlinga-Ponce, G. Perez-Perez, A. Madrazo-De la Garza, M. Dehesa, G. Gonzalez-Valencia, and O. Munoz. (2001) Increasing multidrug resistance in *Helicobacter pylori* strains isolated from children and adults in Mexico. J. Clin. Microbiol. 39:2677-80.

Unno H, Uchida T, Sugawara H, Kurisu G, Sugiyama T, Yamaya T, Sakakibara H, Hase T, Kusunoki M. (2006) Atomic structure of plant glutamine synthetase: a key enzyme for plant productivity. J Biol Chem. 281(39):29287-96.

Usdin KP, Zappe H, Jones DT, and Woods DR. (1986) Cloning, expression, and purification of glutamine synthetase from *Clostridium acetobutylicum*. Appl. Environ. Microbiol. 52:413-419.

Van de Bovenkamp JH, Mahdavi J, Korteland-Van Male AM, Büller HA, Einerhand AW, Borén T, Dekker J. (2003) The MUC5AC glycoprotein is the primary receptor for *Helicobacter pylori* in the human stomach. Helicobacter. 8(5):521-32.

van Doorn LJ, Figueiredo C, Sanna R, Plaisier A, Schneeberger P, de Boer W, Quint W. (1998) Clinical relevance of the cagA, vacA, and iceA status of *Helicobacter pylori*. Gastroenterology. 115(1):58-66.

van Vliet AH, Ernst FD, Kusters JG. (2004) NikR-mediated regulation of *Helicobacter pylori* acid adaptation. Trends Microbiol. 12(11):489-94.

van Vliet AH, Kuipers EJ, Waidner B, Davies BJ, de Vries N, Penn CW, Vandenbroucke-Grauls CM, Kist M, Bereswill S, and Kusters JG. (2001a) Nickel-responsive induction of urease expression in *Helicobacter pylori* is mediated at the transcriptional level. Infect. Immun. 69:4891-7.

van Vliet AH, Poppelaars SW, Davies BJ, Stoof J, Bereswill S, Kist M, Penn CW, Kuipers EJ, and Kusters JG.

(2002a) NikR mediates nickel-responsive transcriptional induction of urease expression in *Helicobacter pylori*. Infect.Immun. 70:2846-52.

van Vliet AH, Stoof J, Vlasblom R, Wainwright SA, Hughes NJ, Kelly DJ, Bereswill S, Bijlsma JJ, Hoogenboezem T, Vandenbroucke-Grauls CM, Kist M, Kuipers EJ, Kusters JG. (2002b) The role of the Ferric Uptake Regulator (Fur) in regulation of *Helicobacter pylori* iron uptake. Helicobacter. 7(4):237-44.

van Vliet AHM, Bereswill S, Kusters JG. (2001b) Ion metabolism and transport. In: Mobley HLT, Mendz GL, Hazell SL, eds. *Helicobacter pylori*: physiology and genetics. Washington, D.C. ASM Press. pp 193-206.

Velayudhan J, Hughes NJ, McColm AA, Bagshaw J, Clayton CL, Andrews SC, and Kelly DJ. (2000) Iron acquisition and virulence in *Helicobacter pylori*: a major role for FeoB, a high-affinity ferrous iron transporter. Mol. Microbiol. 37:274-286.

Viala J, Chaput C, Boneca IG, Cardona A, Girardin SE, Moran AP, Athman R, Mémet S, Huerre MR, Coyle AJ, DiStefano PS, Sansonetti PJ, Labigne A, Bertin J, Philpott DJ, Ferrero RL. (2004) Nod1 responds to peptidoglycan delivered by the *Helicobacter pylori* cag pathogenicity island. Nat Immunol. 5(11):1166-74.

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

Wallace AC, Laskowski RA, Thornton JM. (1995) LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. Protein Eng. 8(2):127-34.

Walz A, Odenbreit S, Stühler K, Wattenberg A, Meyer HE, Mahdavi J, Borén 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. 9 (6):1582-92.

Wang G, Maier SE, Lo LF, Maier G, Dosi S, Maier RJ. (2010) Peptidoglycan deacetylation in *Helicobacter pylori* contributes to bacterial survival by mitigating host immune responses. Infect Immun. 78(11):4660-6.

Wang G, Olczak A, Forsberg LS, Maier RJ. (2009) Oxidative stress-induced peptidoglycan deacetylase in *Helicobacter pylori*. J Biol Chem. 284(11):6790-6800.

Weeks DL, Eskandari S, Scott DR, Sachs GA. (2000) H+-gated urea channel: the link between *Helicobacter pylori* urease and gastric colonization. Science. 287(5452):482-5.

Wen S, Moss SF. (2009) *Helicobacter pylori* virulence factors in gastric carcinogenesis. Cancer Lett. 282(1):1-8.

Wiederstein and Sippl. (2007) ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Research. 35:W407-W410.

Williams CL, Preston T, Hossack M, Slater C, McColl KE. (1996) *Helicobacter pylori utilises* urea for amino acid synthesis. FEMS Immunol Med Microbiol. 13(1):87-94.

Windle HJ, Fox A, Ní Eidhin D, Kelleher D. (2000) The thioredoxin system of *Helicobacter pylori*. J Biol Chem. 275(7):5081-9.

Wirth HP, Yang M, Peek RM Jr, Tham KT, Blaser MJ. (1997) *Helicobacter pylori* Lewis expression is related to the host Lewis phenotype. Gastroenterology. 113(4):1091-8.

Woolfolk CA, Stadtman ER. (1967) Regulation of glutamine synthetase 3. Cumulative feedback inhibition of glutamine synthetase from *Escherichia coli*. Arch Biochem Biophys. 118(3):736-55.

Worst DJ, Otto BR and Graaff J.de. (1995) Iron-repressible outer membrane proteins of *Helicobacter pylori* involved in heme uptake. Infect. Immun. 63:4161-4165.

Wyckoff EE, Schmitt M, Wilks A, Payne SM. (2004) HutZ is required for efficient heme utilization in *Vibrio cholerae*. Journal of bacteriology. 186:4142-4151. doi: 10.1128/JB.186.13.4142-4151.

Yamaoka Y, Kwon DH, Graham DY. (2000) A M(r) 34,000 proinflammatory outer membrane protein (oipA) of *Helicobacter pylori*. Proc Natl Acad Sci U S A. 97(13):7533-8.

Yamaoka Y, Ojo O, Fujimoto S, Odenbreit S, Haas R, Gutierrez O, El-Zimaity HM, Reddy R, Arnqvist A, Graham DY. (2006) *Helicobacter pylori* outer membrane proteins and gastroduodenal disease. Gut. 55(6):775-81.

Yamaoka Y, Orito E, Mizokami M, Gutierrez O, Saitou N, Kodama T, Osato MS, Kim JG, Ramirez FC, Mahachai V, Graham DY. (2002) *Helicobacter pylori* in North and South America before Columbus. FEBS Lett. 517(1-3):180-4.

Yamashita MM, Almassy RJ, Janson CA, Cascio D, Eisenberg D. (1989) Refined atomic model of glutamine synthetase at 3.5 A resolution. J Biol Chem. 264(30):17681-90.

Yeo HJ, Savvides SN, Herr AB, Lanka E, Waksman G. (2000) Crystal structure of the hexameric traffic ATPase of the *Helicobacter pylori* type IV secretion system. Mol Cell. 6(6):1461-72.

Yeo HJ, Yuan Q, Beck MR, Baron C, Waksman G. (2003) Structural and functional characterization of the VirB5 protein from the type IV secretion system encoded by the conjugative plasmid pKM101. Proc Natl Acad Sci U S A. 100(26):15947-52.

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

Yoshiyama H, Nakamura H, Kimoto M, Okita K, Nakazawa T. (1999) Chemotaxis and motility of *Helicobacter pylori* in a viscous environment. J Gastroenterol. 34(Suppl 11):18-23.

Yu J. (1998) Inactivation of DsbA, but not DsbC and DsbD, affects the intracellular survival and virulence of *Shigella flexneri*. Infect Immun. 66:3909–3917.

Yu J. and Kroll JS. (1999) DsbA: a protein-folding catalyst contributing to bacterial virulence. Microbes Infect. 1:1221-1228.

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*. J Mol Biol. 323(1):125-30.

Zawadzka AM, Kim Y, Maltseva N, Nichiporuk R, Fan Y, Joachimiak A, Raymond KN. (2009) Characterization of a *Bacillus subtilis* transporter for petrobactin, an anthrax stealth siderophore. Proc Natl Acad Sci U S A. 106(51):21854-9.

Zhang Z, Li ZH, Wang F, Fang M, Yin CC, Zhou ZY, Lin Q, Huang HL. (2002) Overexpression of DsbC and DsbG markedly improves soluble and functional expression of single-chain Fv antibodies in *Escherichia coli*. Protein Expr Purif. 26(2):218-28.

Zhong Q, Shao SH, Cui LL, Mu RH, Ju XL, Dong SR. (2007) Type IV secretion system in *Helicobacter pylori*: a new insight into pathogenicity. Chin Med J (Engl). 120(23):2138-42.

Acknowledgements

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. Timothy J. Hobley that gave me the opportunity to spend 1 months at DTU in Copenhagen and that contributed to enrich my scientific knowledge and experience. I also want to say thanks to Prof. Elspeth Graham, Oxford University to give me a opportunities to attend the BCA protein crystallography course in Oxford and that helps me to build my knowledge in protein crystallography.

I want to extend my thanks to prof. Roberto Battistutta as well for his helpfulness and Kindness. 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, Nicola, Sandra, Ilenia, Sara, Marco and the undergraduate students that worked with me: Enrico.

I would like to thank also all the people from VIMM: Elisa, Rosa, Allesandro, Graziano and all the PhD students from the other labs that I had the pleasure to know during this PhD.

From all the friends and nice people that I have met in Padua, I want to thank first of all Abhijeet, Maulilio, Azad, Varun, Sajid and many more for their friendship and to make me feel like at home.

I finally want to thank my parents and my brother that have loved me always and my beloved wife Monira, to whom I share everything and the support she gave me everyday.