UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Università degli Studi di Padova

Dipartimento di Biologia

Scuola di Dottorato di Ricerca in Bioscienze e Biotecnologie

Indirizzo: Biotecnologie

Ciclo XXVIII

STRUCTURAL CHARACTERIZATION OF *HELICOBACTER PYLORI* PROTEINS CONTRIBUTING TO STOMACH COLONIZATION

Direttore della Scuola: Ch.mo Prof. Paolo Bernardi

Coordinatore di Indirizzo: Ch.ma Prof.ssa Fiorella Lo Schiavo

Supervisore: Ch.mo Prof. Giuseppe Zanotti

Dottorando: Maria Elena Compostella

31 Gennaio 2016

UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Università degli Studi di Padova

Department of Biology

School of Biosciences and Biotechnology

Curriculum: Biotechnology

XXVIII Cycle

STRUCTURAL CHARACTERIZATION OF *HELICOBACTER PYLORI* PROTEINS CONTRIBUTING TO STOMACH COLONIZATION

Director of the Ph.D. School: Ch.mo Prof. Paolo Bernardi

Coordinator of the Curriculum: Ch.ma Prof.ssa Fiorella Lo Schiavo

Supervisor: Ch.mo Prof. Giuseppe Zanotti

Ph.D. Candidate: Maria Elena Compostella

31 January 2016

Contents

ABBREVIATIONS AND SYMBOLS

26695	*Helicobacter pylori* strain 26695
Å	Angstrom
aa	aminoacid
Abs	Absorption
ADP	Adenosine diphosphate
AEBSF	4-(2-Aminoethyl)-benzenesulfonylfluoride hydrochloride
AGS	Human cultured gastric adenocarcinoma cells
AMP	4-amino-5-aminomethyl-2-methylpyrimidine
ATP	Adenosine triphosphate
ATPase	Adenosine triphosphate hydrolase
B. subtilis	*Bacillus subtilis*
C	Concentration
CA	Carbonic Anhydrase
cag	cytotoxin associated gene (gene)
Cag	Cytotoxin associated gene (associated protein)
CCD	Charge-Coupled Device
CHAPS	3-[(3-Cholamidopropyl)-dimethylammonio]-1-propane sulfonate / N,NDimethyl-3-sulfo-N-[3-[[3α,5β,7α,12α)-3,7,12 trihydroxy-24-oxocholan-24-yl]amino]propyl]-1-propanaminium hydroxide, inner salt
CMC	Critical micelle concentration
CV	Column volume
Da	Dalton
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
DXP	1-deoxy-D-xylulose-5-phosphate
E. coli	*Escherichia coli*
EDTA	Ethylene Diamino Tetracetic Acid
ESRF	European Synchrotron Radiation Facility
FAD	Flavin Adenine Dinucleotide

FAMP	N-formyl-4-amino-5-aminomethyl-2-methylpyrimidine
F(hkl)	Structure factor amplitude
F_{calc}	Calculated structure factor amplitudes
F_{obs}	Observed structure factor amplitudes
FPLC	Fast Protein Liquid Chromatography
Fur	Ferric Uptake Regulator protein
FW	Forward
G27	*Helicobacter pylori* strain G27
Hepes	N-[2-Hydroxyethyl] piperazine-N'-[2-ethanesolfonic] acid
H. pylori	*Helicobacter pylori*
I	Measured intensity of the diffraction spots
IL	Interleukin
IMAC	Immobilized Metal ion Affinity Chromatography
IPTG	Isopropyl-β-D-1-thiogalactopyranoside
IS	Insertion sequence
J99	*Helicobacter pylori* strain J99
LB	Luria Bertani liquid medium
LDAO	LaurylDimethylAmine Oxide (detergent)
MAD	Multiple Anomalous Dispersion
mAU	milli Absorption Unit
MES	2-(N-Morpholin) ethansulfonate
MIR	Multiple Isomorphous Replacement
mRNA	Messenger ribonucleic acid
MS	Mass Spectrometry
MW	Molecular Weight
NAP	Neutrophil activating protein
NFkB	Nuclear factor-kB
O/N	Overnight
OD	Optical Dispersion
OMP	Outer membrane protein
ORF	Open Reading Frame
PAI	Pathogenicity island
PBS	Phosphate Buffer Saline
PCR	Polymerase Chain Reaction

PDB	Protein Data Bank
PEG	Polyethylene glycol
pI	Isoelectric point
PMSF	Phenylmethanesulfonyl fluoride
r.m.s.d.	Root-mean-square deviation
RNA	Ribonucleic acid
RP-HPLC	Reversed Phase-High Performance Liquid Chromatography
RV	Reverse
SAD	Single Anomalous Dispersion
SDS	Sodium dodecyl sulfate
SDS-PAGE	SDS-Polyacrylamide gel electrophoresis
sec	*Escherichia coli* secretory pathway
SEM	Scanning Electron Microscopy
SH	Src homology domain
spp	species
Src	Rous sarcoma virus non-receptor tyrosine kinase
TBS	Tris-buffered saline
TLC	Thin Layer Chromatography
TLR4	Toll-Like Receptor 4
Tris	2-amino-2-(hydroxymethyl)-1,3-propanediol
Triton	Octylphenoxypolyethoxyethanol polyethylene glycol-*p*-isooctylphenyl ether
TTBS	Tween 20 Tris-buffered saline
T3SS	Type III secretion system
T4SS	Type IV secretion system
Ure	Components of the urease complex
VacA	Vacuolating cytotoxin A
*vir*B/D/F/E	Virulence factor B/D/F/E (gene)
VirB/D/F/E	Virulence factor B/D/F/E (associated protein)
σ(I)	Standard deviation of the measured intensities (I)

AMINOACIDS

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

SUMMARY

Helicobacter pylori is a well-characterizDe human pathogen that colonizes the stomach of more than half of the world's population. It is a Gram-negative, microaerophilic, flagellated, spiral shaped bacterium able to establish a life-long chronic infection in the gastric mucosa. Infection with *H. pylori* is generally acquired early in childhood, with a higher prevalence in developing countries, and typically persists for life. As in many chronic infections, most individuals remain asymptomatic with only a small proportion developing clinical disease. *H. pylori* is considered a pathogen as it universally causes progressive inflammation and gastric mucosal damage; in 1994 it was declared a class I human carcinogen by the World Health Organization (WHO). The clinical outcomes associated to *H. pylori* infection include severe gastroduodenal diseases, such as peptic and duodenal ulcers, noncardia gastric adenocarcinoma, and gastric mucosa-associated lymphoid tissue (MALT) lymphoma. For more than 100 years it has been recognized that atrophic gastritis was tightly associated with gastric cancer. The discovery of *H. pylori* in 1983 identified the cause of chronic gastric mucosal inflammation and thus the underlying cause of gastric cancer. As consequence, since its culture from a gastric biopsy, *H. pylori* has been the subject of intense investigations and provoked the interest of many scientists, such as bacteriologists, molecular biologist, gastroenterologists, infectious disease specialists, cancer biologists, epidemiologists, pathologists, and pharmaceutical scientists.

H. pylori has developed a surprising molecular machinery to survive in the unfriendly environment and achieve a successful colonization of the stomach. Since *H. pylori* is not an acidophilus bacterium, it has evolved several specialized mechanisms to survive gastric acid. The pathogen has to resist in the gastric lumen for a short period, enough to enter into the highly viscous mucosa, reach the gastric epithelium, find nutrients and multiply. Some acid-adaptive mechanisms include an acid-activated inner membrane urea channel, UreI, a neutral pH-optimum intrabacterial urease, and periplasmic and cytoplasmic carbonic anhydrases. This acid acclimation system allows to regulate the pH of the periplasm and of the surrounding liquid in acidic medium at levels compatible with survival and growth. A key factor essential for survival and successful colonization is the bacterial motility, mediated by its sheathed unipolar flagella, allowing *H. pylori* to swim in response to a gradient of pH and to stay within the mucus layer, where the pH is generally higher with respect to the lumen. Approximately only 20% of *H. pylori* bacteria in the

stomach adhere to the surface of the gastric epithelial cells; bacterial adhesion involves specialized molecular interactions mediated by adhesins and surface components, which are able to evade the host immune recognition by displaying a high antigenic variation. *H. pylori* is characterized by high genetic variability, not only in gene sequence but also in gene content, evidenced by the availability of complete genome sequences. One of the most striking differences in *H. pylori* strains is the presence or absence of a 40-kb DNA region named *cag* Pathogenicity Island, that encodes a Type IV Secretion System, causing the translocation of CagA toxin, one of the most relevant virulence factor of *H. pylori*. Upon injection into epithelial gastric cells, CagA induces cellular modifications, including alteration of cell structure, motility, cell scattering and proliferation, and tight junctions. A further relevant virulence factor is the vacuolating cytotoxin VacA, which is a secreted, pore-forming toxin able to induce vacuolization in gastric epithelial cells. Almost all *H. pylori* strains contain a *vacA* gene, but the gene sequence is highly variable, causing changes in VacA virulence activity. Therefore, *H. pylori* strains can be classified in subtypes associated with different levels of pathogenic offense during colonization, on the basis of the variability of the virulence factors. However, the various and divergent clinical outcomes deriving from the *H. pylori* infection are dictated by a complex balance between host genetic factors, bacterial virulence determinants, and environmental components. Therefore, understand in detail the host-pathogen relationship is a complex challenge, still incomplete. Despite that the bacterial genome has been completely sequenced, several pathogenic mechanisms have not yet been defined. Moreover, currently *H. pylori* can be eradicated by a triple therapy combining a protonic pump inhibitor and antibiotics; but the increasing antibiotic resistance is the main reason for this treatment failure. Therefore, it becomes necessary to identify new pharmacological targets against the bacterium, in order to overcome the serious problem of the drug-resistance and to develop new antibiotic treatments.

The main purpose of this research project is focused on identification and structural characterization of new potential pharmacological targets of *H. pylori*. In this respect, proteins responsible for colonization and virulence, as well as secreted proteins mediating important pathogen-host interactions, are interesting candidates for structural characterization, in order to deepen their putative function. In particular, the investigations were focused on the periplasmic α-carbonic anhydrase (HPG27_1129), the cytoplasmic β-carbonic anhydrase (HPG27_4), the flagellar protein FliK (HPG27_857), the thiol:

disulfide oxidoreductase HPG27_1020, and two secreted “hypothetical proteins”, namely HPG27_1030 and HPG27_1117.

The research described in this thesis was mostly carried out at the Department of Biomedical Sciences, University of Padova, and at Venetian Institute of Molecular Medicine (VIMM), Padova. The strategy adopted included preliminary bioinformatic analyses, PCR-amplification of the selected genes starting from purified *H. pylori* chromosomal DNA (strain G27), cloning in a His-tag-containing vector and expression of the protein in *E. coli* competent cells. The recombinant proteins were then purified using two chromatography steps, from soluble or insoluble fractions, and concentrated for crystallization trials. The α-carbonic anhydrase was successfully crystallized and the structure was determined by x-ray diffraction. Crystals of β-carbonic anhydrase and HPG27_1117 were also obtained, nevertheless not suitable to x-ray diffraction measurement. To ensure the sample quality, Western blotting, analytical gel-filtration, UV-Vis absorption spectrum, circular dichroism analyzes were performed.

Structural peculiarities and possible functional implications of α-carbonic anhydrase are described in Chapter III. This periplasmic protein plays a key role in the complex balance of urea and bicarbonate aimed to the survival in the stomach, catalyzing the reversible conversion of carbon dioxide to bicarbonate; thus, it is fundamental in buffering the pH of the periplasm. *H. pylori* α-carbonic anhydrase was cloned as recombinant protein lacking the N-terminal secretion signal, expressed in *E. coli* cells and purified; crystals were obtained by vapor-diffusion technique and the structure was determined at 1.52 Å by molecular replacement, based on a model built from α-carbonic anhydrase from *Sulfurihydrogenibium yellowstonense* (Di Fiore et al., 2013; PDB accession code: 4G7A). The protein structure shares many features with other members of the α-carbonic anhydrase family, showing a central ten-stranded β-sheet surrounded by three α-helices and by the remainder of the protein chain. Structural peculiarities are presented by the active site, since the glutamic acid residue (position 127) interacting with three catalytic histidine residues is substituted by a serine residue and the absent negative charge is replaced by a chloride ion captured from the external medium. The definition of the structural details of the protein allows to investigate new specific inhibitors as potential antibiotics against *H. pylori*. Moreover, cocrystallization trials were performed to investigate the molecular binding of inhibitor compounds to the active site; but cocrystals suitable to x-ray diffraction measurement have not been obtained yet.

The pathogen encodes a further carbonic anhydrase, namely the cytoplasmic β-carbonic anhydrase, whose investigations are described in Chapter IV. The enzyme is hypothesized to catalyze the same conversion for the carbon dioxide molecules that do not freely diffuse out of the inner membrane, contributing to buffer the pH of the cytoplasm and survival in the gastric acid environment. The β-carbonic anhydrase was cloned as 6-His-tag recombinant protein and expressed in *E. coli* competent cells, exhibiting a limited yield of soluble protein, the most relevant limit encountered, likely owing to an improper folding by *E. coli* cells. The purification was performed both from the soluble and from the insoluble fractions, adopting various chromatographic techniques. Higher quality protein sample was obtained *via* immobilized-metal ion affinity chromatography, although the final yield of purified protein was impaired by the low affinity for the Ni-NTA resin. The purified protein was concentrated for crystallization trials, but crystals obtained were not suitable to x-ray diffraction measurement.

In Chapter V the investigations on the flagellar protein FliK are reported. As mentioned before, bacterial motility mediated by unipolar flagella is an essential factor to minimize the exposure to the acid environment and to achieve a successful colonization of the gastric mucosa. In *H. pylori* more than 50 putative proteins are predicted to be involved in expression, secretion and assembly of the flagellar apparatus. It is composed of three structural elements: a basal body, an external helically shaped filament, and a hook that serves as a joint. FliK is responsible for the hook length control and in *fliK* mutants it has been observed that long hooks of unregulated length, named polyhooks, are formed, impairing the bacterial motility. Preliminary bioinformatics analyzes have evidenced that the flagellar protein exhibits an overall unstructured nature, with a limited folded region located at the C-terminal domain. Flik was cloned as 6-His-tag recombinant protein and several expression attempts were performed, adopting various *E. coli* strains and varying the conditions. Nevertheless, FliK exhibited an improper production by *E. coli* cells and degradation processes, likely ascribed to the high disorder level of the sequence. Strategies to overcome the limits of successful expression could be the cloning as single domains, or selecting more sophisticated system of expression, able to properly fold the protein.

Since the formation of disulfide bonds plays a key role also in bacterial virulence, many bacteria possess an oxidative protein-folding machinery to properly assemble their proteins, including *H. pylori*. The thiol:disulfide oxidoreductase HPG27_1020, whose experimental procedures are reported in Chapter VI, is a thioredoxin-fold protein which

plays a role in the cytochrome *c* maturation, as well as in oxidized protein proper folding. Therefore, it provides essential function in *H. pylori* and represents a possible pharmacological target. Since its N-terminal region encode an export signal, the protein was cloned as 6-His-tag recombinant protein lacking of 24 N-terminal aminoacids. The recombinant HPG27_1020 protein was successfully expressed in *E. coli* cells, exhibiting a significant amount of soluble protein (approximately 60%). The researches were forcedly interrupted since meantime the x-ray structure of the thiol:disulfide oxidoreductase from *H. pylori* 26695, namely HP0377, has been determined and published. Their aminoacid sequences show a high degree of identity (96%), therefore the investigation has not longer been considered innovative.

In Chapter VII cloning, expression, purification and crystallization trials concerning two secreted “hypothetical proteins”, namely HPG27_1030 and HPG27_1117, are described. Recently, several secreted proteins were identified by proteomic analysis of *H. pylori* secretome; they represent attractive subjects of structural and functional investigations, since they could mediate important pathogen-host interactions and, thus, represent potential target for antibiotics and vaccine development. HPG27_1030 was successfully cloned as 6-His-tag recombinant protein, expressed in *E. coli* cells and purified by two chromatography steps. A significant amount of soluble purified protein was achieved, but the protein exhibited instability in solution and a clear tendency to aggregation, resulting in a limited final concentration of purified sample for crystallization trials. HPG27_1117 was cloned, expressed and purified as before. The most relevant limits encountered were the low yield of expression and the tendency to degradation. Nevertheless, purified protein was concentrated for crystallization trials and crystals were obtained by vapor-diffusion technique; but the crystals diffracted at a limited resolution and crystals suitable to x-ray diffraction measurement have not been obtained yet. To overcome the common problem of instability and degradation of these secreted proteins, changings in the buffer composition could improve the stability in solution and enhance the final yield of purified product for crystallization trials.

Concluding, identification of some new bacterial features have made possible to increase the overall knowledge about *H. pylori* and its peculiar mechanisms aimed to survival and virulence. On the basis of these findings, new investigations can be approached, in order to widely understand the pathophysiological mechanisms of this peculiar pathogen and to develop new eradication treatments.

SOMMARIO

Helicobacter pylori è un microorganismo patogeno ben caratterizzato, che colonizza lo stomaco di più di metà della popolazione mondiale. È un batterio Gram-negativo, microaerofilo, flagellato, spiraliforme, in grado di instaurare un'infezione cronica della mucosa gastrica, che può durare tutta la vita se non trattata. L'infezione da *H. pylori* è generalmente acquisita in età infantile, con un tasso di prevalenza maggiore nei paesi in via di sviluppo, e tipicamente persiste per tutto il corso della vita. Come nel caso di molte infezioni croniche, la maggior parte degli individui risulta asintomatica, mentre solo una limitata porzione sviluppa patologie correlate. *H. pylori* è considerato un microorganismo patogeno poiché causa universalmente un'infiammazione progressiva e danni tissutali alla mucosa gastrica; nello specifico, nel 1994 *H. pylori* è stato dichiarato un agente carcinogeno di classe I per l'uomo da parte della World Health Organization (WHO). Gli esiti clinici conseguenti all'infezione da *H. pylori* comprendono patologie gastrointestinali particolarmente severe, quali ulcere peptica e duodenale, adenocarcinoma gastrico non cardia e MALT linfoma (mucosa-associated lymphoid tissue lymphoma). Da più di 100 anni è riconosciuto che la gastrite atrofica è strettamente associata al cancro del tessuto gastrico. La scoperta dell'esistenza di *H. pylori* nel 1983 ha identificato la causa dell'infiammazione cronica della mucosa gastrica e quindi la causa fondamentale del cancro allo stomaco. Di conseguenza, sin dalla sua scoperta a partire da biopsie di tessuto gastrico, *H. pylori* è al centro di intense investigazioni e suscita l'interesse di molti studiosi, quali batteriologi, biologi molecolari, gastroenterologi, infettivologhi, biologi specializzati in patologie cancerose, epidemiologi, patologi e farmacologi.

Per sopravvivere nell'ambiente estremamente inospitale dello stomaco e potervi realizzare una colonizzazione efficace, *H. pylori* ha sviluppato una sorprendente macchina molecolare. Poiché non è un batterio acidofilo, *H. pylori* ha evoluto molti espedienti specializzati per sopravvivere all'acidità gastrica. Innanzitutto, il patogeno deve resistere alle condizioni estreme del lume gastrico solo per un breve periodo, sufficiente per penetrare nella mucosa altamente viscosa, raggiungere l'epitelio gastrico, recuperare nutrienti e moltiplicarsi. Alcuni dei meccanismi coinvolti nell'adattamento alle condizioni acide prevedono il canale per l'urea, UreI, localizzato nella membrana interna e attivato da un pH acido, l'ureasi citoplasmatica, caratterizzata da un optimum di attività a pH neutro, e due anidrasi carboniche, localizzate nel citoplasma e nel periplasma. Questo sistema di adattamento all'acidità gastrica permette di regolare il pH del

periplasma e anche del liquido circostante nonostante l'ambiente acido, a livelli compatibili con la sopravvivenza e la crescita. Inoltre, un fattore cruciale per la sopravvivenza e una colonizzazione efficace del tessuto gastrico è rappresentato dalla motilità del batterio, resa possibile da flagelli unipolari e rivestiti da una guaina di difesa; grazie a quali *H. pylori* è in grado di nuotare in risposta a un gradiente di pH e di rimanere all'interno dello strato di muco gastrico, dove il pH è generalmente maggiore rispetto al lume dello stomaco. Circa solo il 20% dei microorganismi nello stomaco aderisce alla superfice delle cellule epiteliali gastriche; in particolare, l'adesione batterica vede coinvolte interazioni molecolari specializzate, mediate da adesine e altre componenti della superficie batterica, che sono in grado di eludere il riconoscimento da parte del sistema immunitario dell'ospite grazie a una elevata variabilità antigenica. Infatti, *H. pylori* è caratterizzato da una sorprendente variabilità genetica, non solo per quanto riguarda la sequenza dei geni, ma anche nel contenuto genico; la disponibilità delle sequenze genomiche complete ha reso possibile rilevare questa elevata variabilità in *H. pylori*. Soprattutto, una delle differenze più evidenti tra i ceppi di *H. pylori* è la presenza o meno di un frammento di DNA cromosomico di 40 kb chiamato isola di patogenicità *cag*, che codifica per un sistema di secrezione di tipo IV, responsabile della traslocazione della tossina CagA, uno dei più importanti fattori di virulenza di *H. pylori*. In seguito all'iniezione all'interno delle cellule epiteliali gastriche, CagA induce una serie di modificazioni cellulari, tra le quali alterazioni della struttura cellulare, della motilità, della proliferazione e della migrazione cellulari, della struttura delle giunzioni cellulari occludenti. Un ulteriore importante fattore di virulenza è la citotossina vacuolizzante VacA, che consiste in una tossina secreta, in grado di formare pori nelle membrane e indurre vacuolizzazione nelle cellule epiteliali gastriche. Quasi tutti i ceppi di *H. pylori* contengono il gene che codifica VacA, ma la sequenza genica è altamente variabile, causando perciò cambiamenti nell'intensità dell'attività di VacA. Perciò, in base alla variabilità dei fattori di virulenza, i ceppi di *H. pylori* possono essere classificati in sottotipi, ciascuno dei quali è associato a differenti livelli di patogenicità in seguito a colonizzazione. Oltre a quanto riportato, gli esiti clinici vari e divergenti derivanti dall'infezione da *H. pylori* dipendono da un intricato bilancio tra variabilità genetica dell'ospite, fattori di virulenza batterica e componenti ambientali. Perciò, la comprensione dettagliata della relazione tra ospite e patogeno è una sfida complessa, ancora da chiarire nella sua interezza. Nonostante che il genoma da più ceppi di *H. pylori* sia stato completamente sequenziato, molti dei meccanismi di patogenicità non sono ancora stati definiti. Inoltre, l'attuale trattamento di eradicazione di

H. pylori prevede una tripla terapia che combina un inibitore di pompa protonica e due antibiotici; ma la crescente diffusione di antibiotico resistenza è il principale motivo del fallimento di questa terapia. Perciò si rende necessario identificare nuovi target farmacologici contro questo patogeno, al fine di superare il preoccupante problema della farmaco resistenza e di sviluppare nuovi trattamenti antibiotici.

Lo scopo principale di questo progetto di ricerca verte sull'identificazione e la caratterizzazione strutturale di nuovi potenziali target farmacologici di *H. pylori*. A questo proposito, proteine responsabili di colonizzazione e virulenza, così come proteine secrete che mediano le rilevanti interazioni tra ospite e patogeno, sono ritenute interessanti candidati per la caratterizzazione strutturale, allo scopo di approfondire la loro funzione presunta. In dettaglio, le indagini di questo progetto di ricerca si sono concentrate sull'α-anidrasi carbonica (HPG27_1129), con localizzazione periplasmatica, la β-anidrasi carbonica (HPG27_4), con localizzazione citoplasmatica, la proteina flagellare FliK (HPG27_857), l'ossidoreduttasi HPG27_1020 e infine due "proteine ipotetiche" secrete, di funzione sconosciuta, cioè HPG27_1030 e HPG27_1117.

Il lavoro di ricerca descritto in questa tesi è stato eseguito presso il Dipartimento di Scienze Biomediche dell'Università di Padova e presso l'Istituto Veneto di Medicina Molecolare (VIMM) di Padova. La strategia adottata prevedeva analisi bioinformatiche preliminari, amplificazione del gene di interesse tramite PCR a partire da DNA cromosomico purificato di *H. pylori* (ceppo G27), clonaggio in vettori in fusione con un 6-His-tag ed espressione in cellule competenti di *E. coli*. Di seguito, Le proteine ricombinanti sono state purificate tramite procedimenti che prevedono due passaggi cromatografici, sia dalla frazione solubile che da quella insolubile, e quindi concentrate per le prove di cristallizzazione. α-anidrasi carbonica è stata cristallizzata con successo e la struttura è stata determinata tramite diffrazione a raggi X. Inoltre, sono stati ottenuti cristalli anche di β-anidrasi carbonica e di HPG27_1117, però non adatti per la misura di dati di diffrazione a raggi X di buona risoluzione. Per assicurare la qualità del campione di proteina, sono state eseguite analisi quali Western blotting, gel-filtrazione analitica, spettro di assorbimento UV-Vis, spettro di dicroismo circolare.

Le peculiarità strutturali e le possibili implicazioni funzionali di α-anidrasi carbonica sono descritte nel Capitolo III. Questa proteina periplasmatica svolge un ruolo chiave nell'intricato bilancio di urea e bicarbonato volto alla sopravvivenza del batterio nello stomaco, poiché catalizza la conversione reversibile dell'anidride carbonica in bicarbonato; perciò, essa è fondamentale nel regolare il pH del periplasma, dove è

localizzata. α-anidrasi carbonica da *H. pylori* è stata clonata come proteina ricombinante mancante del segnale N-terminale di secrezione, è stata espressa in cellule di *E. coli* e infine purificata; cristalli sono stati ottenuti mediante il metodo a diffusione di vapore e la struttura è stata determinata a 1.52 Å tramite molecular replacement, basandosi su un modello costruito a partire da α-anidrasi carbonica di *Sulfurihydrogenibium yellowstonense* (Di Fiore et al., 2013; codice PDB: 4G7A). La struttura della proteina condivide molte caratteristiche con altri membri della famiglia delle α-anidrasi carboniche, in quanto presenta un β-foglietto centrale costituito da 10 filamenti, circondato da 3 α-eliche e dalla rimanente catena polipeptidica. Alcune peculiarità strutturali sono presentate dal sito attivo, poiché il residuo di acido glutammico (posizione 127) che interagisce con i tre residui catalitici di istidina è sostituito da un residuo si serina nella stessa posizione e la carica negativa mancante è rimpiazzata da uno ione cloro catturato dal mezzo esterno. La determinazione dei dettagli strutturali di questa proteina permette di ricercare nuovi specifici inibitori che possano agire come potenziali antibiotici contro *H. pylori*. Inoltre, sono state eseguite delle prove di cocristallizzazione con inibitori sulfamidici, per investigare i dettagli strutturali delle interazioni dei composti inibitori col sito attivo; ma cocristalli di qualità adatta per la misura dei dati di diffrazione a raggi X non sono stati ancora ottenuti.

Il microorganismo patogeno codifica anche un'ulteriore anidrasi carbonica, cioè β-anidrasi carbonica localizzata nel citoplasma, le cui indagini sono descritte nel Capitolo IV. Si ipotizza che questo enzima catalizzi la stessa conversione per quanto riguarda le molecole di anidride carbonica che non diffondono liberamente al di fuori della membrana interna; perciò contribuisce alla regolazione del pH del citoplasma e alla sopravvivenza nell'ambiente gastrico estremamente acido. β-anidrasi carbonica è stata clonata come proteina ricombinante con un 6-His-tag ed espressa in cellule competenti di *E. coli*; però il principale limite incontrato è stato una limitata resa di proteina solubile, probabilmente dovuta a un'impropria organizzazione tridimensionale da parte delle cellule di *E. coli*. La purificazione è stata eseguita sia a partire dalla frazione solubile sia da quella insolubile, adottando tecniche cromatografiche variegate. Il campione di proteina di migliore qualità è stato ottenuto per mezzo della cromatografia di affinità per ioni metallici immobilizzati, sebbene la resa finale di proteina purificata sia stata compromessa a causa della moderata affinità per la resina Ni-NTA. La proteina purificata è stata concentrata per le prove di cristallizzazione, ma i cristalli ottenuti non sono di qualità adatta per la misura dei dati di diffrazione a raggi X.

Nel Capitolo V è riportato il lavoro di ricerca sulla proteina flagellare FliK. Come menzionato in precedenza, la motilità batterica mediata dai flagelli unipolari è un fattore essenziale per minimizzare il contatto con l'ambiente acido e realizzare una colonizzazione efficiente della mucosa gastrica. In *H. pylori* si prevede che più di 50 proteine siano coinvolte nell'espressione, secrezione e assemblaggio dell'apparato flagellare. Quest'ultimo è composto di tre elementi strutturali; un corpo basale, un filamento esterno a forma elicoidale e un uncino che serve ad unione. FliK è responsabile del controllo della lunghezza dell'uncino e si è osservato che in mutanti mancanti del gene di FliK si formano lunghi uncini di lunghezza incontrollata, chiamati "polyhooks", che compromettono la motilità batterica. Analisi bioinformatiche preliminari hanno evidenziato come questa proteina flagellare presenti una struttura globale altamente disordinata, con una limitata regione strutturata localizzata a livello del dominio C-terminale. FliK è stata clonata come proteina ricombinante con un 6-His-tag e numerosi tentativi di espressione sono stati eseguiti, facendo uso di differenti ceppi di *E. coli* e variando le condizioni. Nonostante ciò, si sono riscontrati un'impropria produzione di FliK da parte delle cellule di *E. coli* e un'evidente degradazione della proteina, probabilmente entrambi gli eventi dovuti all'elevato grado di disordine della sequenza amminoacidica. Alcune strategie per risolvere questo limite dell'espressione potrebbero essere il clonaggio dei singoli domini oppure l'utilizzo di sistemi di espressione più sofisticati, in grado di strutturare correttamente la proteina.

Poiché la formazione dei ponti disolfuro riveste un ruolo chiave anche nella virulenza batterica, molti batteri posseggono sistemi molecolari per l'assemblaggio delle proteine nel corretto stato ossidativo, tra cui anche *H. pylori*. L'ossidoreduttasi HPG27_1020, le cui procedure sperimentali sono riportate in Capitolo VI, è una proteina con un'organizzazione simile alla tioredoxina che riveste un ruolo cruciale nella maturazione del citocromo *c*, così come nell'assemblaggio corretto di proteine ossidate. Perciò, questa proteina fornisce funzioni essenziali per *H. pylori* e rappresenta un possibile target farmacologico. Poiché la regione N-terminale codifica un segnale di secrezione, la proteina è stata clonata come proteina ricombinante con un 6-His-tag e mancante dei 24 amminoacidi N-terminali. HPG27_1020 ricombinante è stata espressa con successo in cellule di *E. coli*, mostrando una quantità significativa di proteina nella frazione solubile (circa il 60%). Però le ricerche sono state obbligatoriamente interrotte, in quanto nel frattempo è stata determinata e pubblicata la struttura dell'ossidoreduttasi da *H. pylori*

26695, cioè HP0377. Poiché la loro sequenza amminoacidica presenta un elevato grado di identità (96%), le indagini sono state considerate non più innovative.

Nel Capitolo VII sono descritti il clonaggio, l'espressione, la purificazione e le prove di cristallizzazione per quanto riguarda due "proteine ipotetiche" secrete, cioè HPG27_1030 e HPG27_1117. Recentemente numerose proteine secrete sono state identificate tramite analisi proteomica del secretoma di *H. pylori*; queste rappresentano interessanti soggetti di indagini strutturali e funzionali, poiché potrebbero mediare importanti interazioni tra ospite e patogeno e, quindi, concorrere come potenziali target per lo sviluppo di antibiotici e vaccini. HPG27_1030 è stata clonata con successo come proteina ricombinante con un 6-His-tag, espressa in cellule di *E. coli* e purificata tramite due passaggi cromatografici. È stato possibile ottenere una quantità molti rilevante di proteina solubile, questa ha esibito un'elevata instabilità in soluzione e una chiara tendenza all'aggregazione, portando perciò a una limitata concentrazione finale di campione purificato per le prove di cristallizzazione. HPG27_1117 è stata clonata, espressa e purificata come riportato sopra. I limiti più rilevanti che sono stati incontrati sono una bassa resa di espressione e la tendenza alla degradazione del campione. Nonostante ciò, la proteina purificata è stata concentrata per le prove di cristallizzazione e sono stati ottenuti cristalli utilizzando il metodo di diffusione di vapore; ma questi hanno diffranto ad una risoluzione troppo limitata e non è stato possibile ottenere cristalli di qualità adatta per le misure di diffrazione a raggi X. Per superare il problema comune dell'instabilità e della degradazione di queste proteine secrete, cambiamenti nella composizione dei tamponi di purificazione potrebbe migliorare la stabilità in soluzioni e così la resa finale di prodotto purificato per le prove di cristallizzazione.

In conclusione, grazie all'individuazione di alcune nuove peculiarità di questo patogeno è stato possibile accrescere la conoscenza in merito a *H. pylori* e i suoi meccanismi peculiari volti alla sopravvivenza e alla virulenza. Questi primi risultati costituiscono la base per nuove investigazioni, al fine di apprendere nel modo più completo possibile i meccanismi patofisiologici di questo peculiare microorganismo e di sviluppare nuovi trattamenti per l'eradicazione.

Chapter I

INTRODUCTION

1.1 *HELICOBACTER PYLORI*

Helicobacter pylori is an important and one of the most common and successful human pathogens. It affects approximately half of the world's population and is responsible for severe gastric diseases (Fig. 1.1; Rothenbacher and Brenner 2003). It is a Gram-negative spiral bacterium (2.4-4 µm long and 0.5-1 µm wide) able to colonize the human stomach, a unique ecological niche. *H. pylori* has adapted itself to surviving in this unfriendly environment; after the first settlement that usually occurs early in life, it is able to establish a life-long chronic infection.

The presence of spiral microorganisms in the human stomach was observed for the first time over one hundred years ago by Walery Jaworski and then was confirmed in animals by Giulio Bizzozero, but was not really taken seriously until the discoveries of two Australian scientists in the late XX century (Konturek 2003). In 1983 Barry Marshall and Robin Warren isolated a curved bacillus from the stomach epithelium of patients with gastritis and ulcer diseases and were able to demonstrate a strong association between the presence of this microorganism and the finding of inflammation in gastric biopsies (Marshall and Warren 1984). The flagellated, spiral-shaped bacterium was initially classified as *Campilobacter pylori*, due to its curved morphology; then the name was changed to *Helicobacter pylori*, considering its structural and genetic features. Since the description of this bacterium and its association with gastritis and peptic ulceration shown by Marshall and Warren, the interest in this microorganism has continued to grow. Its association with the development of gastric cancer and mucosa-associated lymphoid tissue (MALT) lymphoma has served only to increase the interest in this area (Taylor 1999). Moreover, the importance of relationship between *H. pylori* and gastric diseases

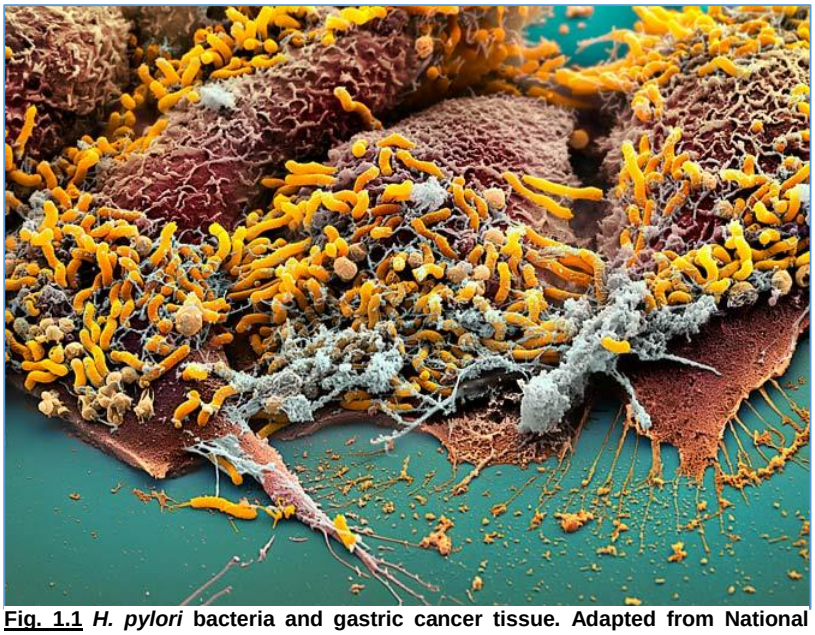

Fig. 1.1 *H. pylori* bacteria and gastric cancer tissue. Adapted from National Geographic web site.

has been supported by Nobel Prize in Physiology or Medicine (Marshall and Warren, 2005 award).

The infection shows appreciable differences between countries and racial or ethnic groups, with a stronger incidence in the developing regions and a tendency to disappear in industrialized ones. Risk factors for *H. pylori* contagion include household crowding, low socio-economic status, country of origin and ethnicity. The family is the core unit of *H. pylori* transmission and it is usually acquired within the first few years of childhood (Covacci et al. 1999). The infection occurs most likely by fecal-oral, gastric-oral or waterborne transmission or by improperly cleaned endoscopic equipment. Indeed *H. pylori* has been found in vomitus, saliva and diarrheal stools. Colonization by *H. pylori* is specific to the stomach of humans and primates; no other natural reservoir has been demonstrated (De Reuse and Bereswill 2007). Once the stomach is colonized, the microorganism persists chronically if untreated. Subsequent infections by other *H. pylori* strains appear to be rare among European and North American population, but they are more common in developing countries (Frenck and Clemens 2003). However, transiently infecting *H pylori* strains, although unable to colonize, provide genetic material to the resident strain. Bacteria isolated from the same patient at intervals of several years have identical DNA fingerprints; however, continuous evolution occurs within the stomach of the infected person, owing to nucleotide mutations, excision of the *cag*-pathogenicity island (*cag*PAI), transposition of insertion elements, recombination with DNA from incoming strains that do not establish a chronic infection and horizontal transfer of new genes.

According to several phylogeographical studies and thanks to genetic comparisons between human and *H. pylori* populations, it has been revealed that humans have been colonized by *H. pylori* since their migration out of Africa, about 100,000 years ago (Fig. 1.2; Covacci et al., 1999). Subsequent spread of the bacterium all over the world can be attributed to human migratory fluxes, such as the prehistoric colonization of Polynesia and the Americas, the Neolithic introduction of

Fig. 1.2 World map indicating the direction of human migrations (arrows) and time range (years since migrations happened). Adapted from Covacci et al., 1999.

farming to Europe, the Bantu expansion within Africa, and the slave trade (Falush et al. 2003). Especially, *H. pylori* can be divided into seven populations and subpopulations with distinct geographical distributions (Tab. 1.1); they derived their gene pools from ancestral populations arising in Africa, Central Asia and East Asia (De Reuse and Bereswill 2007). Close associations between *H. pylori* subtypes and human subpopulations in one continent or even ethnic subgroups within small geographic regions have been detected. Therefore, studies of the population genetics of the microorganism can provide information about the prehistoric and modern migrations of human populations (Yamaoka 2009). Since *H. pylori* is present in populations as an ancestral host, this long-standing relationship suggests that the parasite could provide also beneficial effects. Its infection appears protective in case of esophageal diseases, childhood-onset asthma or rhinitis or atopic dermatitis, and other commensals proliferation; thus there is no advantage to eradicate *H. pylori*, because of development of drug-resistance and alteration of the balance of microbiota (Malnick et al. 2014). Moreover, the long permanence of each strain within the same person and the family-linked mode of transmission suggest that the evolution of *H. pylori* is linked to the social behavior of humans. For most of history, humans have been socially organized and, consequently, their genetic traits segregated in the communities. It is likely that during the social evolution, while mutations accumulated and segregated in the human genes, a co-segregation of the genes of *H. pylori* occurred (Covacci et al. 1999). A deeper understanding of the human-bacterium relationship is required to elucidate the role of *H. pylori* in human life and to formulate efficient preventive and therapeutic strategies.

H. pylori population	*H. pylori* subpopulations	Geographic location or human population
hpAfrica1	hspWAfrica	W. Africa
	hspSAfrica	S. Africa
hpAfrica2		S. Africa
hpNEAfrica		Ethiopia, N. Nigeria, Somalia, Sudan
hpAsia2		Bangladesh, N. India, Malaysia, Thailand
hpEastAsia	hspAmerind	Native Americans
	hspEAsia	East Asians
	hspMaori	Taiwanese Aboriginals, Melanesians, Polynesians
hpEurope		Europe, Middle East, India, Iran
hpSahul		Australian Aboriginals, Papua New Guineans

Tab. 1.1 An overview of *H. pylori* populations. Adapted from Bridge and Scott, 2013.

1.2 GENETIC VARIABILITY

1.2.1 GENOME COMPARISON

The scientific investigation of *H. pylori* and infection-related diseases was greatly promoted by the release of the genome sequences and the genetic analysis (De Reuse and Bereswill 2007). Complete genome sequence provides information for biological mechanisms, evolution, drug discovery, vaccine development. *H. pylori* is the first bacterium to have more completely sequenced genomes from unrelated strains, of which the first two identified were 26695 and J99. *H. pylori* 26695 was isolated in the mid-1980s in the United Kingdom from a patient with gastritis; whereas strain J99 was isolated in 1994 in the United States from a duodenal ulcer patient. The comparison of these two genomes have provided a framework for understanding the level and mechanisms of genetic variability in this gastroduodenal pathogen (Alm and Trust 1999). Both genomes were sequenced using a random shotgun approach from libraries of cloned chromosomal fragments of ~2.5 kb.

1.2.1.1 HELICOBACTER PYLORI 26695

H. pylori 26695 genome was the first to be sequenced, in August 1997. It consists of a circular chromosome of 1,667,867 bp and includes 1590 predicted coding sequences (Fig. 1.3); these predicted genes have an average size of 945 bp, similar to that observed in other prokaryote, and 1091 among them were assigned biological roles. More than 70% of the predicted proteins in H. pylori have a calculated isoelectric point (pI) greater than 7.0, compared to ~40% in *Haemophilus influenzae* and *Escherichia coli*. The basic aminoacids, arginine and lysine, occur twice as frequently in *H. pylori* proteins, perhaps reflecting an adaptation to the very acid environment of the stomach. The average content of G+C is 39% and five regions within the genome have a significantly different G+C composition. Two of them contain one or more copies of the insertion sequence IS605 and are flanked by a 5S ribosomal RNA sequence at one end and a 521 bp repeat near the other. These two regions are also notable because they contain genes involved in DNA processing. Two distinct insertion sequence (IS) elements have been discovered. An insertion sequence is the simplest autonomous transposable elements and transposition is responsible of several consequences, among which the transfer of

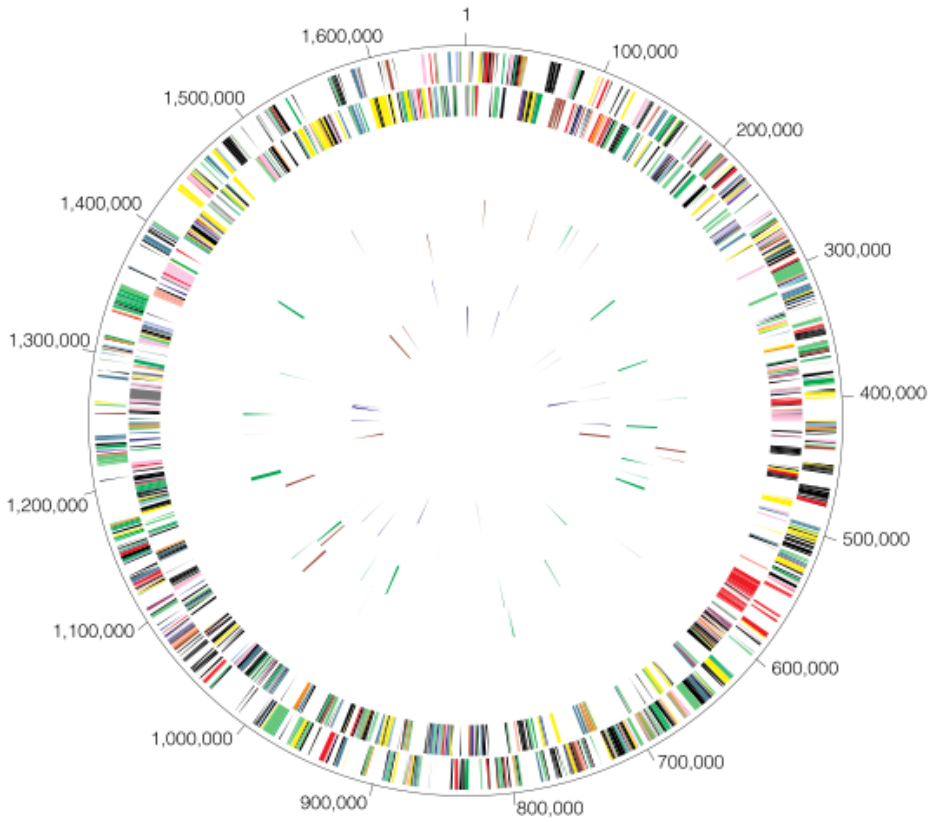

Fig.1.3 Circular representation of the *H. pylori* 26695 chromosome. Outer concentric circle: predicted coding regions on the plus strand; second concentric circle: predicted coding regions on the minus strand; third and fourth concentric circles: IS elements (red) and other repeats (green) on the plus and minus strand, respectively; fifth and sixth concentric circles: tRNAs (blue), rRNAs (red), and sRNAs (green) on the plus and minus strand, respectively. Adapted from Tomb et al., 1997.

antibiotic resistance genes. There are five full-length copies of the IS605 and two of a newly identified element designated IS606. In addition, there are eight partial copies of IS605 and two partial copies of IS606 (Tomb et al. 1997). These elements have some peculiar hallmarks, compared to the insertion sequences belonging to other bacteria. Both elements have two open reading frames (*orfA* and *orfB*) and encode two divergently transcribed transposases, named TnpA and TnpB. IS606 has less than 50% nucleotide identity with IS605 and the IS606 transposases have 29% aminoacid identity with their IS605 counterpart (Ronning et al. 2005). The typical eubacterial block of replication genes, *dnaA-dnaN-recF-gyrB*, hasn't been recognized in 26695 genome sequence. The *dnaA* gene is located ~600 kb away from the *dnaN-gyrB* genes, while the *recF* gene is missing. The *dnaC* gene encoding DnaC protein, which delivers the DnaB helicase to the prepriming complex, is absent. Moreover, an origin of DNA replication, *oriC*, has not been precisely identified from the genome sequence. In many eubacterial chromosomes *oriC* is located close to the *dnaA* gene. In 26695 genome it was detected by protein-DNA interaction techniques and supposed to be located 1.4 Kb upstream of the *dnaA* gene. The presence of typical repeated sequences confirms the location of *oriC* and *dnaA* gene. Bacterial *oriC* consists of repeated DnaA binding motifs, named DnaA boxes, and AT-rich regions. Five DnaA boxes have been found next to the start of the *dnaA* gene.

Comparison of these five DnaA boxes from the putative *H.pylori oriC* region allowed to propose the consensus sequence: $T^{T}/_{C}$ATTCACA (Zawilak et al. 2001). Surprisingly, in silico analysis allowed to identify an additional replication origin region (*oriC2*), separated from the original one (*oriC1*) by the *dnaA* gene, therefore downstream of the gene. Both regions are required for the initiation of *H. pylori* chromosome replication, which indicates a bipartite structure of *H. pylori oriC*, being the first such origin discovered in a Gram-negative bacterium. *oriC2* is bound exclusively as a supercoiled DNA, indicating the importance of the DNA topology in the replication initiation (Donczew et al. 2012).

1.2.1.2 HELICOBACTER PYLORI J99

General comparative features of the *H. pylori* genomes

Genome features	*H. pylori* 26695	*H. pylori* J99
Size (base pairs)	1,667,867	1,643,831
(G + C) content (%)	39	39
Regions of different (G + C) content	8*	9†
AGTGATT repeats at bp = 1	26	2‡
vacA genotype	*s1b/m1*	*s1a/m1*
Open reading frames		
Per cent of genome (coding)	91.0	90.8
Predicted number	1,590	1,495
Functionally classified	875§	895
Conserved with no fucntion	275§	290
H. pylori specific	345§	367
Number with signal sequencell	517	502
Average length (base pairs)	954	998
Per cent AUG initiation codons	81.8	82.7
Per cent GUG initiation codons	9.7	6.7
Per cent UUG initiation codons	8.1	10.4
Per cent other initiation codons	0.4¶	0.2#
Insertion elements☆		
Complete IS*605* copies	5	0
Partial IS*605* copies	8	5
Complete IS*606* copies	2	1
Partial IS*606* copies	2	2(4**)
RNA elements		
Per cent of genome (stable RNA)	0.75	0.75
23S–5S rRNA	2††(3‡‡)	2††
16S rRNA	2§§	2§§
tRNAs	36	36

Tab. 1. 2 General comparative features of the *H. pylori* genomes. Adapted from Alm et al., 1998

H. pylori J99 genome is 24,036 bp smaller and contains 57 fewer predicted open reading frames (91% of the genome) than strain 26695 (Tab. 1.2; Alm and Trust 1999). Common features are the average content of G+C (both 39%), the hard identifiable origin of replication, the average length of coding sequence, the relative frequency of the different initiation codons, the location of the strain-specific genes. Both genomes present genes encoding for two 16S and two 23S-5S ribosomal RNA copies in the same relative locations, but strain 26695 has a further, orphan gene for 5S rRNA and one structural RNA gene (Alm et al. 1998). Each strain contains a set of specific genes, about 6-7% of the entire coding capacity, which are absent from the other (89 in J99 and 117 in 26695) and represent a variable gene pool. In both strains almost half of these specific genes are clustered into one locus called hypervariable plasticity zone, which has a G+C content of 35%, indicating that this region contains fragments that have been acquired by horizontal gene transfer. This is supported by the finding that in *H. pylori* 26695 genome the plasticity region contains DNA that had previously been characterized as part of a

plasmid, pHPM186 (Suerbaum 2000). This region is continuous in strain J99, but is split into two regions in 26695. The majority (approximately 60%) of the strain-specific genes are also *H. pylori* specific. To relatively few of these strain-specific genes can be attributed a function (25 and 26 for J99 and 26695, respectively) with the majority (approximately 60%) encoding DNA restriction/modification enzymes in both strains. It has been demonstrated that restriction/modification systems reduce the efficiency of DNA exchange between bacterial strains belonging to different clonal lineages. Therefore, they help to stabilize clonal groupings by reducing the efficiency of intraspecific transformation (Suerbaum 2000). Several strain-specific genes encode products that are likely to be able to alter the complexion of the bacterial cell envelope and subsequently may alter the interaction with the host immune system. 1406 genes in J99 genome have orthologues in 26695 genome and the extent of gene order conservation was determined by comparing each of the J99 genes to its orthologous 26695 partner (if present) with respect to its immediate neighboring gene flanking each side. Two genomes have 84.7% of their genes (1267 genes) preserved in the same genetic order, on both sides. There are 161 genes (10.8%) where the gene order is disrupted on one side by the insertion/deletion of a strain-specific gene while maintaining the gene order on the other side. 40 genes (2.7%) are flanked by strain-specific genes on both sides, although only one (JHP1295) even possesses an orthologue in strain 26695 at all. Surprisingly, only 27 genes (1.8%) have the same neighboring gene on one side and are flanked on the other side by a gene common to both strains that is out of order due to an organizational rearrangement such as a translocation and/or inversion. In addition, one of the 23S-5S rRNA loci is associated with the plasticity zone, and is flanked by strain-specific genes in both strains. The average nucleotide identity for all orthologues is 92.6%. However, the average nucleotide identity for the orthologues with a predicted function is higher, 94.0%. Within the genes that have been assigned a predicted function, the genes (and corresponding proteins) predicted to be involved in DNA restriction and modification display the highest level of divergence with only 90.6% identity. Therefore, the *H. pylori* specific genes have the lowest level of identity. There were nine strings of conserved genes over 50 genes in length, representing 46% of the total number of common genes, with the longest being 133 genes. The limited gene shuffling observed is consistent with a low level of evolutionary divergence within *H. pylori*. J99 genome has 14 *Not*I sites, compared to 7 in 26695. All 7 of the 26695 *Not*I sites are also found in the corresponding orthologues in strain J99. Nucleotide changes in orthologous genes, which fail to affect the sequence of

the encoded protein, are responsible for 6 of the additional 7 *Not*I sites in J99 genome, whereas the last difference is due to a single aminoacid change (Alm and Trust 1999). Both genomes present more than 25 homopolymeric tracts and dinucleotide repeats, that can be subjected to frequent length changes due to slipped-strand mispairing. These repeats are identified in a subset of genes, named contingency genes, that are hypermutable because of slippage within DNA repeats. This mechanism results in frequent shifting into and out of frame (relative to the translational start), leading to an on–off switching of the associated gene products (Saunders et al. 1998). Repeat lengths in some J99 genes differ from those in 26695 genes, indicating that such genes may be differently expressed in the two strains. This variation in gene expression facilitates adaptation to changing host environments (Alm et al. 1998).

1.2.2 CORE GENOME

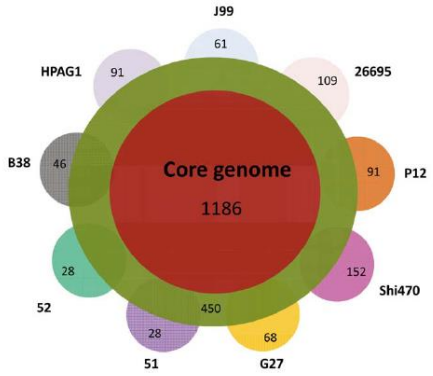

Fig. 1.4 Sketch map of core genome for protein-coding genes. Adapted from Lara-Ramírez et al., 2011.

Inter-strain diversity is common to many other bacterial species. It has been defined that a particular bacterial species presents a core set of genes, named core genome, and some auxiliary genes. The core genome contains genes that are common to all or nearly all of the strains. It determines the properties that are characteristic of that species. The auxiliary genes are present in some strains; they are determinants of the biological properties unique to some of the strains. The core genome of *H. pylori* was firstly investigated by Nina Salama in 2000 and was found to be composed by 1281 genes (Salama et al. 2000). Since a limited number of *H. pylori* strains was considered and they were only isolated from Western individuals, the core set of genes was again examined by Helga Gressmann in 2005 and it consist of 1111 genes, based on 56 representative *H. pylori* strains. In 2011 data were revised by Edgard Eduardo Lara-Ramírez, using new bioinformatics tools, and the core genome was found to be composed by 1186 genes (Fig. 1.4; Lara-Ramírez et al. 2011). By sequence analysis it has been detected that core genes are located in potential genome rearrangement sites, in the hypervariable plasticity zone and in the *cag*PAI. The auxiliary set of genes in *H. pylori* amounts to 22%-27% of the genome and they encode for functionally unknown proteins, *cag* proteins, outer

membrane proteins and proteins of DNA metabolism (Dong et al. 2009). Furthermore, in *H. pylori* genome species-specific genes can be detected. *H. pylori* is classified into the Epsilon subdivision of *Proteobacteria*, as well as *Helicobacter hepaticus*, *Campylobacter jejuni*, *Wolinella succinogenes*. The availability of the sequenced genomes has led to define species-specific genes as well as genes exchanged and shared by members of this bacterial group. Whole genome clustering of *H. pylori* and *C. jejuni* demonstrated that 648 *H. pylori* genes are species-specific. The fact that most of the strain-specific genes are species-specific supports the assumption that genes of the flexible gene pool are exchanged among *Helicobacter* species, but are not transferred to bacteria of other genus (De Reuse and Bereswill 2007).

1.2.3 MECHANISMS GENERATING GENETIC VARIABILITY

Concluding, *H. pylori* appears unique among bacteria regarding the diversity between strains. Nowadays several *H. pylori* genomes have been sequenced and have revealed significant differences each other in size, nucleotide sequence, protein and gene arrangement. The most unusual feature of this diversity is the very high number of unique nucleotide sequences for every gene. Most of these heterogeneities are clustered in the hypervariable plasticity zone. This heterogeneity can be analyzed at two different levels: genotypic variation among strains and variations in *H. pylori* populations within an individual host. Corresponding alleles in different *H. pylori* strains typically present an average identity in nucleotide sequence of 92 to 99%, but several genes exhibit a much higher level of genetic diversity. In addition, there is considerable variation among strains in gene content (McClain et al. 2009). Strain-specific genetic diversity is involved in adaptation to the changing microenvironment and to the individual hosts, microorganism's ability to establish a lifelong chronical infection and to cause different diseases. Despite this diversity, there is a great deal of conservation of proteins; the presence of such a great degree of homology among open reading frames present in different strains suggests that the key cellular processes, including DNA replication, are conserved among these strains (Nitharwal et al. 2011). The high level of genetic diversity in *H. pylori* strains is probably the consequence of multiple factors, including a high rate of mutation, a high rate of intra-species genetic recombination, a long evolutionary history. Vertical genetic transfer mechanisms are the most frequent. Point mutations, DNA rearrangement, as insertions, deletions or inversions of genes and intergenic regions,

provide mechanisms to adapt to unfavorable conditions. Genetic diversity can arise also by acquisition of heterologous DNA from other species. There are several areas of *H. pylori* genome, including the plasticity zones, the *cag*PAI and several DNA restriction/modification genes, which significantly differ in G+C content from the remainder of the genome (Alm and Trust 1999) and attest events of horizontal transfer of genetic materials. The inter-strain diversity of *H. pylori* is extended by plasmids, that are involved in the same phenomenon of horizontal transfer, by sequence changes affecting phase-variable genes, in which mutations can modulate gene expression, and by insertion sequences, that act as small transposable elements.

1.2.3.1 MUTAGENESIS

Owing to their haploid genotype and mode of replication, bacteria are by default clonal; genetic diversity firstly occurs by sequential acquisition of mutations. The spontaneous mutation rate of the majority of *H. pylori* strains lies between 10^{-5} and 10^{-7} (Björkholm et al. 2001); thus, it is several orders of magnitude higher than the average mutation rate of *E. coli*. The bacterial chromosome is exposed to mutagenic effects at all times, such as alkylation or oxidation of single nucleotides resulting in mispairing. These damaged nucleotide bases can be removed by a number of repair mechanisms, involving different glycosylases and endonucleases. Most of the information about DNA repair mechanisms in *H. pylori* is deduced from the two completely sequenced genomes of strains 26695 and J99. Damaged bases can be repaired by glycosylases that belong to the base excision repair pathway. All glycosylases can excise a damaged base resulting in an apurinic/apyrimidinic site, while some of them additionally nick the DNA deoxyribosephosphate backbone through an apurinic/apyrimidinic lyase activity. Finally, the gap is refilled by the action of RecJ, polymerase I and ligase (Krokan, Standal, and Slupphaug 1997). The *H. pylori* chromosome contains orthologues of the glycosylase genes *ung*, *nth*, *mutY* and *magIII*, whereas several other genes that protect the chromosome of *E. coli* from mutational influences appear to be absent (some examples are *tag*, *alkA* and *mutM*). *H. pylori nth* gene encodes the endonuclease III enzyme, which removes oxidized pyrimidine bases and also displays apurinic/apyrimidinic lyase activity. Initially *magIII* gene was also annotated as an endonuclease III, but later shown to encode an unusual 3-methyladenineDNA glycosylase that protects *H. pylori* from the effect of alkylating agents, but lacks apurinic/apyrimidinic lyase activity. Repair pathways

recognizing mismatches in the double-helix structure (such as *mutH*, *mutS*, *mutL*, belonging to the methyl-directed mismatch repair system) were absent from the *H. pylori* genome (Alm et al. 1998). This evidence suggests that *H. pylori* has a less stringent control of replicative errors and explains the higher mutation rate. However, the *H. pylori* genome contains a gene with weak homology to the *E. coli mutS* gene, but it was demonstrated that it belongs to the *mutS*2 subfamily, that usually have no function in methyl-directed mismatch repair system; its major role is to repair oxidative DNA lesions, particularly 8-oxoguanine (Wang et al. 2005). Furthermore, *H. pylori*, like other bacterial species that include *Campylobacter jejuni* and *Helicobacter hepaticus*, uses hypermutable homopolymeric or dinucleotide repeat sequences to control gene expression. Such mutations which occur at a very high frequency can switch genes on and off, if the repeats are located within coding regions, or affect gene regulation by changing promoter activity. Many repeats are located in genes important for survival, such as the flagellar gene *fliP*, or for adaptation to the individual human host, such as lipopolysaccharide biosynthesis genes, restriction/modification systems, composition of the antigenic structure of outer cell surface (Kraft and Suerbaum 2005). The best studied examples of the phase-variation phenomenon are the fucosyltransferase genes, involved in the final stages of lipopolysaccharide biosynthesis. The two copies of the α-1,3 fucosyltransferase genes and the single copy of the α-1,2 fucosyltransferase gene in *H. pylori* are involved in the synthesis of Lewis X and Y antigens of lipopolysaccharide, and may contribute to antigenic mimicry and autoimmune disease. There is a long homopolymeric C tract in the 5' coding region of each of the four *H. pylori* α-1,3 fucosyltransferase genes and in the central domain of the α-1,2 fucosyltransferase gene. The serotype of lipopolysaccharide-phase variants is correlated with the varying length of the homopolymeric tract (Alm and Trust 1999). Phase-variation provides a fast and effective possibility to react to environmental changes and to alter the appearance of the micro-organism to the host immune system. Finally, the *H. pylori* genome contains numerous repetitive sequences of different lengths that permit intra-genomic deletions or rearrangements. It was demonstrated that the deletion of fragments between repeats of up to 100 bp was RecA-independent and that deletion frequencies increased with the increasing length of the repeats; exemplificative genes include *cagY* and *cagA*, located on the *cag*PAI, as well as *aimA*, encoding an amidase involved in peptidoglycan biosynthesis, and genes involved in the fucosylation of lipopolysaccharide (Suerbaum and Josenhans 2007). Another mechanism to generate genetic variability in *H. pylori* is

provided by the insertion sequences (IS), as result of their translocation. They are segments of DNA that can move from one position on a chromosome to a different position on the same chromosome or on a different chromosome; they encode proteins implicated in the transposition activity, such as transposase and regulatory proteins. In bacteria transposition is a relevant phenomenon, because is also responsible for the emergence of antibiotic resistance; as can be evidenced, transposable elements can carry antibiotic resistance genes. In *H. pylori* more insertion sequences have been discovered; the first two identified are IS605 and IS606, which are present in one-third of tested strains (Kersulyte et al. 1998). Insertion sequences can be classified into families based on the general features of their DNA sequences and associated transposases; IS605 and IS606 belong to the widespread IS200/IS605 bacterial family. This group differs from classical insertion sequences, because its members transpose using obligatory single-strand DNA intermediates, instead of double-strand ones, and because they carry subterminal palindromic structures, instead of inverted repeats, and insert 3'-end to specific AT-rich tetra- or penta-nucleotide sequences, without duplicating (or deleting) the target site (He et al. 2015). Other members of this family identified in *H. pylori* are also IS607, IS608 and IS609. Another peculiarity of these IS elements is their chimeric feature, because they contain two transposition-related genes (*orfA* and *orfB*, with *orfA* upstream of *orfB*), which encode for two transposase (TnpA and TnpB, respectively). Exception is IS609, which encode for two additional open reading frames (*orf1* and *orf2*, besides *orfA* and *orfB*,), whose function is unknown. The IS200/IS605 family elements can be divided into two subfamilies, based on *orfA* homologies. In one subfamily, including IS607, *orfA* encodes a putative serine recombinase; in the other subfamily, including IS605, IS606 and IS608, *orfA* encodes a transposase homologous to that encoded by IS200 in *E. coli*, whose product is distinct from serine recombinase proteins (Kersulyte et al. 2004). On the other side, *orfB* shows strong homology with the putative gene of IS1341, that is evidenced in Gram-positive species. Transposition plays an important role in genomic evolution and facilitates the horizontal transfer of genetic material. It is also responsible for gain of advantages by bacteria; emergence of antibiotic resistance is the most common.

1.2.3.2 RECOMBINATION

Additionally, recombination after natural transformation plays a key role in generating allelic diversity in *H. pylori* populations. This microorganism presents a natural transformation competence for uptake of exogenous chromosomal or plasmid DNA. *H. pylori* is able to differentiate between homologous and heterologous DNA and will not integrate DNA from other *Helicobacter* species, or other genera into its chromosome. The only known exception is the successful transformation with DNA from the related species *Helicobacter acinonychis*, the phylogenetically closest relative of *H. pylori*. Uptake of genomic DNA from the same strain was one to two logs more efficient than uptake of DNA from an unrelated *H. pylori* strain (Kraft and Suerbaum 2005). Chromosomal DNA uptake is mediated by a transport system related to the type 4 secretion system (T4SS), named the ComB system. This conjugative apparatus consists of a nearly complete set of T4SS components with a similar gene cluster organization, which were named according to their orthologous proteins of the *Agrobacterium tumefaciens* VirB/VirD4 system, considered the prototype for type 4 secretion. The ComB system presents all T4SS core components, except for the homologues to VirB1, VirD4, VirB5 factors, as well as VirB11 ATPase (Karnholz et al. 2006). Also the uptake of plasmid DNA, by a conjugative mechanism, contributes to the genetic variability. Many *H. pylori* strains carry cryptic plasmids, that differ in size (from 2 to 100 Kb) and in gene content. The smaller *H. pylori* plasmids encode for an origin of replication, replication genes (*repA* or *repB*) and, occasionally, a small open reading frame, with unknown function. The larger ones carry a number of additional ORFs with unknown function (Fernandez-Gonzalez and Backert 2014). The uptake of plasmid DNA is restricted by an inter-strain transformation barrier. In fact, most strains of *H. pylori* contain a large number of restriction/modification systems (RM systems), described as a defense strategy against invasion by foreign DNA. Many of these are pseudogenes and predicted to be inactive. Most restriction/modification genes predicted to be functional are only found in selected *H. pylori* strains. According to the variable number of these genes, *H. pylori* chromosomal DNA shows highly variable restriction patterns when digested with methylation-sensitive restriction endonucleases. The large number of RM genes homologous to those in other bacterial species and their strain-specificity suggest that *H. pylori* may have horizontally acquired these genes. Furthermore, the large amount of these genes (approximately 3% of the genome), their maintenance in the genome, and their association with genomic rearrangements are consistent with the proposal that these systems can act as selfish genetic elements (Alm

and Trust 1999). The barrier to transformation with heterologous plasmid DNA is related to the number and activity of the RM systems present in the *H. pylori* strain. HpyII, a type IIs RM system, has been demonstrated to present a barrier to uptake of chromosomal DNA fragments > 1 Kb from unrelated donor strains (Aras 2002). Additional genes that have been shown to be involved in transformation competence of *H. pylori* include *dprA*, *comE3*, *comH*, the nuclease *nucT*, a VirB4 homolog, HP0017, and genes HP0015, HP1089, HP1424 and HP1473, which were identified in a large-scale transposon shuttle mutagenesis screen, but their function is still unknown (Kraft and Suerbaum 2005).

Exogenous DNA must be integrated into *H. pylori* chromosome by homologous (or site-specific) recombination or replicated as plasmid. Thus, homologous recombination is essential for bacterial evolution and genome plasticity and, at the same time, it helps to maintain genetic barriers between species by selecting for DNA with sufficient homology (Fischer, Hofreuter, and Haas 2001). The unusually short length of imported fragments of foreign DNA together with the high frequency of recombination leads to mosaic structures and finally to unique alleles of each gene and therefore to unique strains isolated from every patient (Kraft and Suerbaum 2005). The mechanisms of recombination repair have been intensively studied in *E. coli*. Two homologous recombination initiation pathways coexist in *E. coli*: the RecBCD pathway is essential for the repair of double-stranded DNA breaks and for resolving regressed forks; the RecFOR pathway is needed for post-replication gap repair and for replication restart after UV damage (Kuzminov 1999). These initiation pathways metabolize the DNA break to generate single-stranded DNA on which the recombinase RecA is loaded and cooperatively forms a nucleoprotein filament. The filament is then aligned with a homologous duplex and promotes strand exchange. The product of this reaction is a branched DNA molecule, named the Holliday junction, that is processed by the RuvABC complex (Orillard et al. 2011). From the sequenced *H. pylori* genomes it has been possible to detect only a limited amount of genes related to recombinational repair, including a RecA homologue. This *H. pylori* recombination enzyme is required for DNA pairing and homologous strand exchange; it plays a key role because a *recA*-negative mutant is almost deficient in recombination and recombinational repair and exhibits reduced acid tolerance, resistance to UV radiation, and increased susceptibility to metronidazole. Moreover, RecA is subjected to posttranslational modifications; mutants that produce unmodified RecA maintain transformation competence and resistance to UV light and acid, but show an increased susceptibility to metronidazole(Kraft and Suerbaum

2005). Despite the limited number of genes related to recombinational repair, both pathways can be identified in *H. pylori*. An AddAB class of helicase-nuclease enzymes, related to the *E. coli* RecBCD, was demonstrated to be functional in *H. pylori*. By bioinformatics analysis a novel RecO orthologue was identified (Marsin et al. 2008), suggesting the presence of the RecRO pathway in *H. pylori*. Whereas the RecO protein can displace ssDNA-binding protein and bind to ssDNA, RecR is the key component for loading RecA onto ssDNA (Wang, Leja, and Maier 2011). Even genes coding for homologs of *ruvABC*, involved in the Holliday junctions formation, are present in *H. pylori*; it was demonstrated that a *ruvC* mutant lacking the Holliday junction resolvase was unable to establish chronical infection in a mouse model (Robinson et al. 2005). This finding remarks the role of genetic variation on bacterial fitness and adaptation.

1.2.4 *HELICOBACTER PYLORI* AS A "QUASI SPECIES"

A large study comparing mutation and recombination rates between different pathogenic bacteria confirmed that *H. pylori* stands out as the bacterial species with the highest population recombination rate (Suerbaum and Josenhans 2007). Moreover, genetic evolution occurs even during the colonization of the host, a phenomenon named microevolution. *H. pylori* exhibits substantial genetic diversity among bacteria that colonize human tissues. Genetic rearrangement and recombination events may occur among *H. pylori* strains during chronic infections in different parts of the gastric tissue. These genetic variations, mainly recombination between different superinfecting strains, can lead to development of diverse *H. pylori* strains in a single host. Several studies indicated that mixed *H. pylori* infections may occur in single individual (Farzi et al. 2015). Thanks to its high genetic variability, *H. pylori* was recognized as a "quasi species", that is a well-defined distribution of mutants that is generated by a mutation-selection process (Nowak 1992). Its population structure appears to be clonal only over a short period of time after infection, due to frequent recombination; thus, it has a quasi-panmictic feature. A Bayesian mathematical model predicted that up to 50% of the genome of an *H. pylori* strain could be exchanged by recombination over four decades of infection. Furthermore, the large amount of hypermutable genes leads to differ each large *H. pylori* population in many subpopulation, with a specific combination of active and inactive phase-variable genes; the term "quasi species" explains also this phenomenon (Suerbaum and Josenhans 2007). Genetic variation of *H. pylori* during chronic colonization has been

extensively investigated, using either sequential strains, isolated from patients during repeated endoscopies, or multiple isolates taken from one or multiple locations in the same stomach (Suerbaum and Achtman 2004). Dangeruta Kersulyte in 1999 was the first to demonstrate the presence of recombinant exchanges between two *H. pylori* strains during mixed infection in the stomach of a single patient from Lithuania. Some strains from the patient possessed the *cag* pathogenicity island, while others lacked it; this evidence resulted from transfer of DNA from a *cag*- strain and following homologous recombination (Dangeruta Kersulyte, Chalkauskas, and Berg 1999). *H. pylori* genetic diversity is investigated by several methodologies; some of them are genotyping (Chiurillo et al. 2013), plasmid profiling (S. I. Smith et al. 2002), Repetitive Extragenic Palindromic PCR (rep-PCR; Kidd et al. 2001), Amplified Fragment Length Polymorphism analysis (AFLP; Kuipers et al. 2000), Random Amplification of Polymorphic DNA (RAPD; Farzi et al. 2015), microarray analysis (Salama et al. 2000), Multilocus Sequence Typing (MLST; Gunaletchumy et al. 2014). These researches confirm the high recombination rate in vivo and the exchange of genetic material with transient superinfecting *H. pylori* or other bacteria. Furthermore, frequent genetic variations have important implications in pathogenesis; the high variety in clinical outcomes of *H. pylori* infection in different individuals suggests a correlation between strain specific properties and their pathogenic effects. Genotyping based on important *H. pylori* virulence factors, such as *cagA*, *vacA* ($s_{1/2}$, $m_{1/2}$), *iceA* (A_1, A_2), *recA*, provides information on pathogenic diversities and clinical outcomes (Farzi et al. 2015).

1.2.5 CLASSIFICATION OF *HELICOBACTER PYLORI* STRAINS

On the basis of the variability of the virulence factors, *H. pylori* strains have been grouped on two broad families: type I and type II strains; this classification permit to distinguish more virulent population from less virulent ones. In fact, type I strains are characterized by expression of cytotoxin-associated gene A (CagA) and vacuolating cytotoxin A (VacA) aggressive variant and by vacuolating activity; instead type II strains lack *cagA* gene and vacuolating activity, although VacA non toxic variants have been found in the chromosome (Covacci et al. 1997). Other discriminative feature is the presence of the *cag*-pathogenicity island (*cag*PAI), which encodes for virulence factors unique to *H. Pylori* strains, including CagA, and is found only in type I strains. The virulence associated with this chromosomic region can vary because of the splitting of *cag*PAI, in *cag*I and *cag*II,

by insertion of an IS605 sequence (Fig. 1.5; Orodovsky, Appuoli, and Ovacci 1996). Type II strains lack *cag*PAI and IS605 sequences. As evidence, only type I strains have been observed to be involved in severe gastroduodenal diseases. Clearly, pathogenicity of *H. pylori* cannot be due only to VacA and *cag*PAI genes, among which CagA gene; other functions and relative differences in both type strains have been investigated, such as mechanisms of adherence (Su et al. 1998), interactions between pathogen and immune response of the host (Maria et al. 2011), effects on gastric cells (Kawahara et al. 2001). Moreover, allelic polymorphism of the most important virulence factors, that are VacA and CagA, is evaluated to classify *H. pylori* strains and expect the pathogenic consequences of a population; *vacA*+ and *cagA*+ strains are the most virulent and cause severe diseases. A further level of polymorphism of *vacA* and *cagA* lies on multiple alleles of these genes, which show differences in the sequence and can be present or not. This variety results in toxins with varying degrees of virulence (Bridge and Scott Merrell 2013).

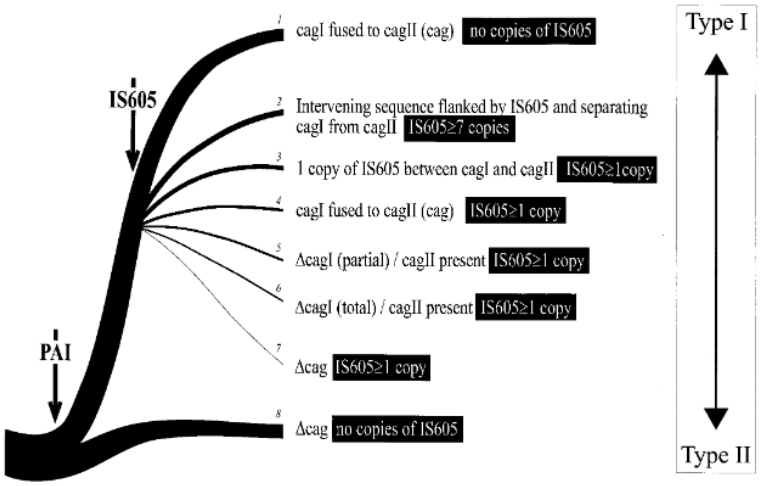

Fig. 1. 5 Evolutionary tree that describes the hypothetical emergence of *H. pylori* type I as a result of an event of gene conversin (PAI acquisition). Presumably, it is only after the integration of an IS605 that subpopulations of intermediate strains with attenuated virulence are differentiated. *cag* retention is indicated by arrows. Type II strains are distinguishable from type I strains with a complete cag deletion by the absence of IS605. Adapted from Orodovsky et al., 1996.

1.3 EPIDEMIOLOGY

Bacteria are diffused everywhere on Earth, in soil, water, air and also in bodies of humans and animals. Every mucosal and cutaneous surface of the human body is colonized by microbes, overall named microbiota. Most of these microorganisms are not harmful for the host; on the contrary, they also perform useful functions and participate to maintain health of the human. A classical exemplification is the intestinal flora, which is the largest bacterial ecosystem of the human body and consists in a wide variety of microorganisms. Instead, the stomach is inhabited by a little amount of microbes (mainly *Streptococcus*, *Staphylococcus*, *Lactobacillus*, *Peptostreptococcus* and types of yeast; Willey et al. 2011), usually transient, due to the very acidic environment; only *H. pylori* is able to establish a chronic colonization and, when present, is numerically predominant on other microorganisms. *H. pylori* is one of the most common human pathogens and approximately 50% of the world's population is infected, reaching up to 90% in developing countries. The prevalence of infection varies widely, depending on geographic area, age, race, ethnicity and socio-economic status; as can be demonstrated, it is higher in developing countries, because of poor socio-economic conditions, crowding and lack of hygienic practices. The infection is acquired early in childhood and transmission probably occurs by person-to-person passage; also unclean water sources seem to be implicated in infection transmission. Humans are the main reservoir; but *H. pylori* has been found to colonize also the stomach of some animals, particularly those living in a human environment, for example cows, sheep, goats (Momtaz et al. 2014). However, the possibility of zoonotic transmission has to be further investigated.

1.3.1 INCIDENCE AND PREVALENCE OF *HELICOBACTER PYLORI* INFECTION

H. pylori infection is ubiquitous and infects both males and females. The role of gender as risk factor is still debated, although some studies demonstrate a male predominance in *H. pylori* infection and occurrence of related diseases in adults; such predominance is not confirmed in children (de Martel and Parsonnet 2006). The infection occurs worldwide, but there are significant differences in the prevalence of infection between countries and

also between different regions of the same country (Fig. 1.6). In general, the prevalence

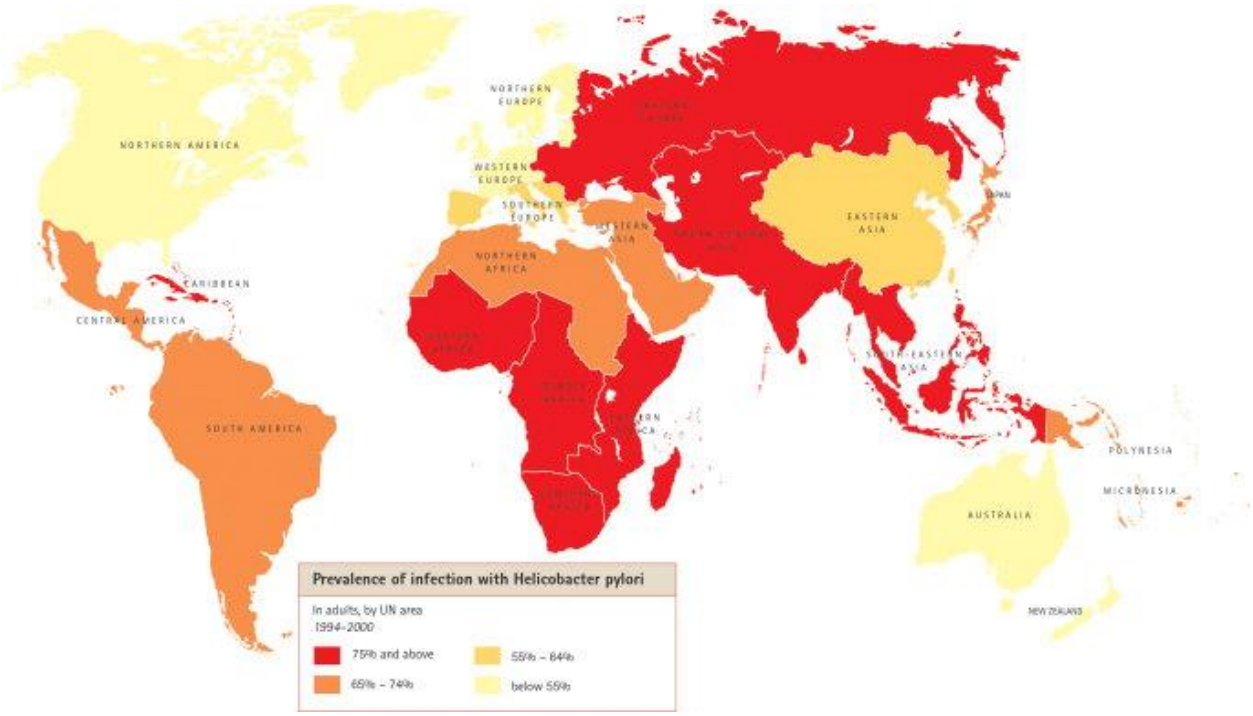

Fig. 1. 6 Mean prevalence of infection with *H. pylori* in adults. Adapted from Parkin, 2006

of the infection is higher in developing countries than in developed ones, because of the poor socio-economic status and overcrowded conditions (Bardhan 1997). The prevalence of *H. pylori* infection in a community is related to three factors: firstly, the rate of acquisition of infection with *H. pylori* (that is the incidence); secondly, the rate of loss of the infection; thirdly, the prolonged persistence of the bacterium in the gastroduodenal mucosa between infection and eradication. Variation in the prevalence of *H. pylori* is mainly attributed to the great differences between communities in the rate of acquisition of *H. pylori* in childhood, under the age of 10 years (Pounder and Ng 1995). Over this age, however, the rate of acquisition of infection in both countries is similar, in the order of 0.5% to 2% per annum (Mitchell 1999). The pattern of infection is an early childhood acquisition of *H. pylori* (30%-50%, by the age of 5 years), that reaches over 90% during adulthood in developing countries. On the contrary, the infection in developed countries is less common in young children (10%-20%) and reaches up to 60% in older ages (Fig. 1.7; Salih 2009). This different trend and the observation of the decreasing of the infection rate in the developed countries during the past few decades suggest that *H. pylori* acquisition and related diseases are a birth cohort-related phenomenon (Sipponen 1995). All individuals are infected in childhood and the decreased levels of *H. pylori* infection associated with younger age groups, particularly in developed countries, are due to gradual improvements in medical care, sanitation, alimentation and living conditions;

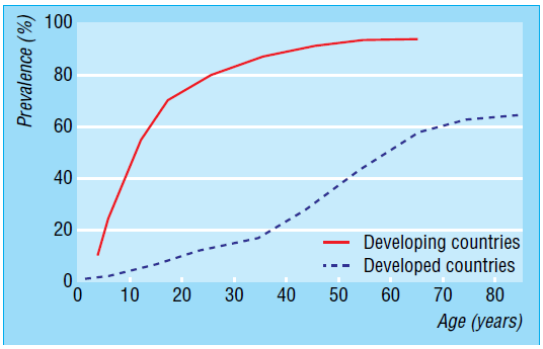

Fig. 1.7 Prevalence of *H. pylori* by age in developing and developed countries. Adapted from Logan and Walker, 2012.

older individuals may have been born in periods when risk of infection was greater. Nevertheless, the acquisition of *H. pylori* is decreasing in developing countries at a slower rate than in developed ones and the prevalence of the infection remains still high; the risk of acquisition can be minimized by implementation of household hygiene, boiling water for drinking purposes and proper cleaning of vegetables (Salih 2009).

The rate of *H. pylori* infection depends not only on socio-economic status or hygienic practice. Marked differences in the prevalence were observed and reported among various ethnic and racial groups. For example, in Malaysia, the increased risk of *H. pylori* infection in Chinese and Indians was suggested as an inherent ethnic genetic predisposition (Goh 1997). In New Zealand, ethnicity was suggested as a risk factor among different groups in the populations. *H. pylori* infection was most prevalent in Pacific Islanders, intermediate in Maori, and least prevalent in Europeans (Fraser et al. 1996). On the contrary, in a study conducted in USA, the prevalence of *H. pylori* infection was almost identical between Hispanic and African Americans, but significantly higher than that among Caucasians. However, ethnicity was ruled out as a major factor and the observed variance was attributed to socioeconomic conditions. Additionally, a study of monozygotic and dizygotic twins (Malaty et al. 1994) suggested also that genetic factors might have some influence on the incidence of *H. pylori* infection (Khalifa, Sharaf, and Aziz 2010). Therefore, also ethnic and genetic predispositions have a relevant role in the occurrence of *H. pylori* infection.

1.3.2 SOURCE AND TRANSMISSION

The sources and the mode of transmission of *H. pylori* is one of the most controversial topics in the investigation about this pathogen. Transmission mainly occurs by person-to-person passage and the failure to consistently isolate from reservoirs other than humans confirms this hypothesis. Besides the well-known reservoir of *H. pylori*, the human stomach, several other potential sources of infection have been highlighted (Mégraud and Broutet 2000). The possibility that *H. pylori* may be a zoonosis first arose following seroepidemiological studies that showed the significantly increasing prevalence of *H.*

pylori infection in abattoir and meat workers increased, as compared with that in subjects not involved in handling animals or animal products (Vaira et al. 1988). These evidences have subsequently been debated, because the increased prevalence in these workers may have resulted from cross-reactivity between *H. pylori* and antibodies to other gastrointestinal microorganisms, such as *Campylobacter jejuni* (Vaira et al. 1988). Historically pigs were the first animal from which at the beginning of 1990s *H. pylori* were detected, from the stomach of a laboratory pig. But all the subsequent studies of this bacterium in pigs failed and the microorganism described the first time was probably *Helicobacter suis* (Grasso et al. 1996). A high prevalence of *H. pylori* was evidenced in Sardinian shepherds (M. Dore et al. 1999), a prevalence significantly higher than that in their family members, who did not have regular contact with sheep and blood donors. Subsequently *H. pylori* was isolated from the milk of sheep, suggesting sheep as ancestral host (M. P. Dore et al. 1999). Houseflies can act as vector to transmit pathogens from feces to food. In 1997 caged houseflies were exposed to freshly grown *H. pylori* o agar plate; it was possible to isolate the bacterium from the external surfaces of the flies and from gut and excreta (Grübel et al. 1997). But the evidence that *H. pylori* couldn't be recovered from houseflies fed human feces infected with the bacterium, suggests that the domestic housefly is neither a vector for transmission nor a reservoir for *H. pylori* (Osato MS, Ayub K, Le HH, Reddy R 1998). Furthermore, *H. pylori* was identified also in pets, but the relationship between pet owner and the prevalence of the infection is not adequately demonstrated. In particular, in 1994 the bacterium was isolated from the stomach of laboratory cats (Handt et al. 1994). These data suggest felines as reservoir of *H. pylori*; but the reliability of this hypothesis is questionable, because these cats were commercially bred and were maintained in isolation. Additionally, *H. pylori* can be isolated also from rhesus monkeys (Handt et al. 1997); but, given the rare contact between humans and monkeys, they can't represent an important reservoir for *H. pylori*. In conclusion, any animal reservoir was confirmed and supported by convincing results; therefore, human stomach remains the unique proven source.

Another potential reservoir for *H. pylori* is the water and objects in contact with it, such as food, vegetables, bathing (Hopkins et al. 1993). Studies employing microbiological techniques have demonstrated that *H. pylori* is able to survive when introduced into water and that it is present in water and other environmental samples all over the world. Epidemiological studies have shown that water source and exposures related to water supply represent risk factors for infection (Bellack et al. 2006). In 1984 Barry Marshall

himself drank a culture of *H. pylori*, to prove water as source of infection and to demonstrate that the bacterium is able to infect a healthy person and cause gastritis. As evidences in favor of water as source of infection, some epidemiological studies were carried out in developing countries. As first example, Peruvian children, whose homes had an external water supply, were found to be three times more likely to be infected with *H. pylori* than children whose homes had an internal water source (Klein et al. 1991). In a second study, in Colombia, acquisition of *H. pylori* in children was evidenced to be associated with swimming more than one time per year in rivers, streams and ponds and drinking stream water (Goodman et al. 1996). Nevertheless, in Southern China, despite the practice to boiling water prior to consumption in the majority of subjects, the prevalence of *H. pylori* infection still remains high (Mitchell et al. 1992). Findings as the last and the unsuccessful attempts to culture *H. pylori* from the water samples lead to a controversial survival of the bacterium in the environment, especially in water. It was supposed that *H. pylori*, when exposed to adverse environmental stimuli (such as also increased oxygen tension, extended incubation, exposure to antibiotics), takes a viable but not-culturable coccoid form (Bode, Mauch, and Malfertheiner 1993). Specifically, the coccoid form of *H. pylori* can be classified in two types: the first one is a degenerative form, pyknotic, not-culturable, as result of cell death of *H. pylori*, in which the cell membrane is disintegrated, but gene material can be detected by PCR in water supplies; the second one is a viable but in any case not-culturable form, potentially virulent (it preserves some virulence factors, such as *cagA* gene, flagellum-mediated motility, adhesion property, urease activity; She et al. 2003), likely able to colonize and induce inflammation in experimental animals (Andersen and Rasmussen 2009). But the possibility of the transformation of the coccoid form in the virulent and infective spiral one in human stomach is not confirmed; therefore, the feasibility of the infection through the coccoid form is debated (Duś et al. 2013) and more probably this shape represents an early stage of bacterial death. In conclusion, therefore, despite extensive studies about potential environmental sources of *H. pylori*, no significant reservoir was identified besides the human stomach, the unique proven source. This evidence is supported by analysis of the genome sequence of *H. pylori*, which shows that this bacterium does not possess the full complement of enzymes required for an exclusive aerobic or anaerobic metabolism (Tomb et al. 1997) and consequently its ability to survive in natural environments appear less likely.

Given the unsuccessful attempts to isolate the bacterium from other reservoirs, the direct person-to-person contact appears to be the most likely route of transmission. Ingestion of bacteria is the main mean of acquisition and it occurs by one or a combination of three modes: oral-oral, gastro-oral or fecal-oral, but the determination of a dominant route is complex. Close personal contact is fundamental for the spread of the bacterium. This evidence explains the acquisition of the bacterium early in childhood, especially between family members (Kivi et al. 2005). The latter is favored by overcrowded conditions in the developing countries, where the infection rate in children is higher than in the developed ones. In familial transmission, the mother plays the key role. As evidenced in a Japanese investigation, mother-to-child transmission is suggested as the most probable route of transmission of *H. pylori*; the fingerprint patterns were found to be identical to those of at least one family member in 76% of the children, with an higher rate of identity in the case of the mother's patterns, compared with those of father (Konno et al. 2008). Furthermore, the prevalence of *H. pylori* infection is significantly increased in family members of children infected by *H. pylori*, compared with that in family members of not infected children (Miyaji et al. 2000). Intra-familial or vertical transmission occurs more commonly, but also horizontal transmission is evidenced, mainly in the case of institutionalized populations. Several studies were carried out about the infection rate in these populations. One of the first was performed in Austrian patients to evaluate the prevalence of *H. pylori* antibodies in mentally and physically handicapped adults living together in a long-term care facility; as could be evidenced, institutionalized patients showed a higher prevalence (75%) of *H. pylori*, compared with controls from the general population (Lambert et al. 1995). Given the location of *H. pylori* infection and the basic requirement of this bacterium for *in vivo* proliferation, ingestion appears to be the most likely means to acquiring the bacterium. *H. pylori* reaches the oral cavity via the oral-oral, gastro-oral or fecal-oral route; it was evidenced that it was possible to isolate and cultivate the microorganism from saliva (Goosen et al. 2002), vomitus (Leung et al. 1999) and feces (Mapstone et al. 1993). But, especially in the case of attempts to culturing *H. pylori* from feces and oral cavity, the results are hard to be clearly evidenced, because of the presence in these sites of the resident microbiota. These bacteria, due to its abundance, tend to grow more rapidly than *H. pylori* and, even if *H. pylori* is present, they will often mask its presence.

A further source for *H. pylori* is the contamination of medical tools for endoscopies. They are common medical procedure used to diagnose and manage gastrointestinal disease.

Because of the complex structure of the endoscope and difficulty in disinfecting it, the possibility of iatrogenic infection in patients following endoscopy is a potential risk factor, not only for *H. pylori* but also for other infectious diseases (Brown 2000). Epidemiological studies of the iatrogenic transmission were carried out; they suggest that rate of iatrogenic infection may approximate to 4 per 1,000 endoscopies (0.4%) when the prevalence of *H. pylori* in the endoscoped population is around 60% (Tytgat 1995). This rate may be higher, reaching 1%, in areas of the world in which improper disinfection techniques are used. Proper cleaning and disinfection of the endoscopic tools reduce the contribution of this source to the spread of *H. pylori* infection.

Despite several investigations, the current knowledge about the dominant route of transmission is still controversial and further epidemiological and microbiological studies are required.

1.4 ADAPTATION AND GASTRIC COLONIZATION

H. pylori infection is implicated in a wide range of disorders of the upper gastrointestinal tract, as well as associated organs. It leads to a chronic and progressive gastric mucosal inflammation and it is responsible of severe disorders, such as gastric inflammation-associated diseases, non-ulcer dyspepsia, peptic and duodenal ulceration, gastric cancer, mucosal-associated lymphoid tissue (MALT) lymphoma. As evidence, *H. pylori* is recognized as the main etiological agent of peptic ulcer. Despite of the high prevalence, most of the infected individuals are asymptomatic and the bacterium inhabits the human stomach for many decades without adverse consequences; fewer than 10% of individuals colonized by *H. pylori* develop peptic ulcer disease, gastric cancer or further associated diseases (Dunn, Cohen, and Blaser 1997). Specific outcomes depend on the interplay among host-, bacterial- and environmental-related factors. The ability of *H. pylori* to survive to acidic conditions and colonize the stomach is essential for the development of virulence and the onset and progress of gastrointestinal disorders. Therefore, the mechanisms of acid resistance and gastric colonization have been greatly investigated.

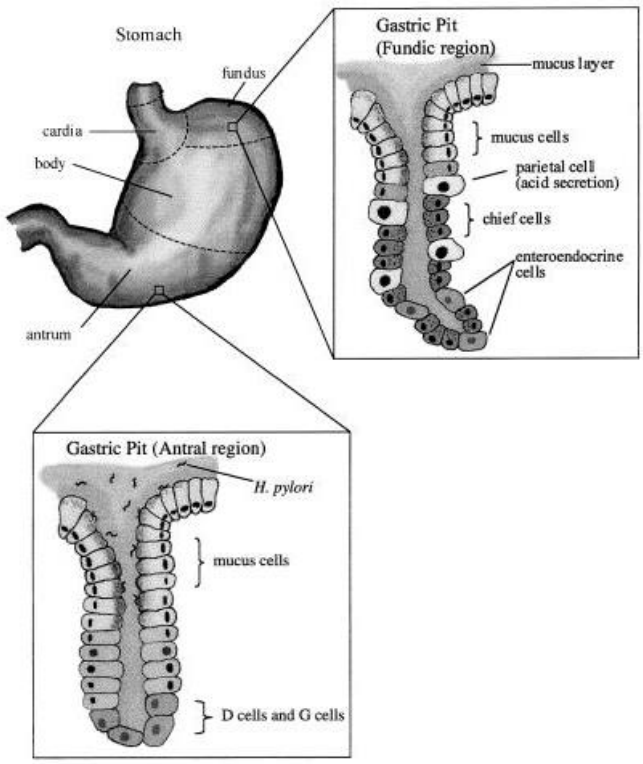

Fig. 1.8 Schematic representation of gastric structure; *H. pylori* colonization is largely confined to the antrum, which lacks acid-secreting parietal cells. Adapted from Mobley et al., 2001.

Following transmission, *H. pylori* faces a new set of challenges in establishing colonization. *H. pylori* exhibits both host tropism, colonizing only primates, and tissue tropism, adhering specifically to gastric mucosal tissue in the stomach or the duodenum, but not to intestinal or squamous-type epithelium. One possible reason for this specific association, between *H. pylori* and gastric epithelium, is that this bacterium takes advantage of specific components of gastric mucus or other factors released by gastric epithelial cells as nutritional sources (Johnson, Gaddy, and Cover 2012). Within the stomach, *H. pylori* can inhabit a range

of different microenvironments. The bacteria are typically more abundant within the gastric antrum, less acid, but they can also be found within the corpus, under conditions of low acid secretion, such as when patients are on long-term acid suppression. Specifically, *H. pylori* mainly adheres to the gastric epithelial lining of the antrum (preferring the mucus-secreting gastric epithelial cells of the upper half of the gastric pits) or stays in the gastric mucus layer; approximately 80% of the bacteria remain in the mucus layer (pH approximately 4.5), with the smaller proportion colonizing the gastric surface (pH close to neutrality). Within the mucus layer, the microorganisms can be found relatively close to the gastric lumen or deep within gastric glands; furthermore, they can be either free-swimming or attached to gastric epithelial cells (Fig. 1.8; Schreiber et al. 2004). Instead, *H. pylori* doesn't adhere to chief cells, parietal cells or endocrine cells in the gastric pit. Moreover, only occasionally *H. pylori* can be internalized by gastric epithelial cells, because of its invasive nature (Dubois and Borén 2007).

The process of bacterial colonization consists of four steps (Mobley, Mendz, and Hazell 2001):

- transmission to a new host;
- bacterial adherence to a specific niche within the host;
- avoidance, subversion or exploitation of host defense mechanisms;
- acquisition of nutrients resulting in successful replication.

Each event is closely related to the others and directed to a successful persistence of the bacterium in its niche. Even in the case of the colonization by *H. pylori*, the mechanisms involved in each event have been extensively investigated. *H. pylori* adhesion to gastric epithelial cells is evidenced to be necessary for the establishment of a successful infection, but many other factors also influence the persistence of *H. pylori* in the gastric mucosa, such as epithelial cell turnover, host immune response, impact of gastric contents. Bacterial factors critical for the colonization of the gastric mucosa include urease, helicoidal shape, flagella, adhesins, δ-glutamyltranspeptidase. Lipopolysaccharide, urease and vacuolating cytotoxin A (VacA) are among the factors that allow *H. pylori* to persist for decades and invoke an intense inflammatory response (McGee and Mobley 2000). Once a chronic infection is established, a considerable amount of bacteria is present in the gastric mucosa, with an infection density estimated between 10^5 and 10^7 *H. pylori* CFU per gram of gastric mucus gel (Atherton et al. 1996).

The bacterial density is higher in the antrum respect to the corpus and plays a key role in the pathogenesis. Higher levels of *H. pylori* density are associated with the development of more severe diseases.

Some bacterial factors involved in *H. Pylori* acid adaptation and colonization are described in the following.

1.4.1 ACID ADAPTATION

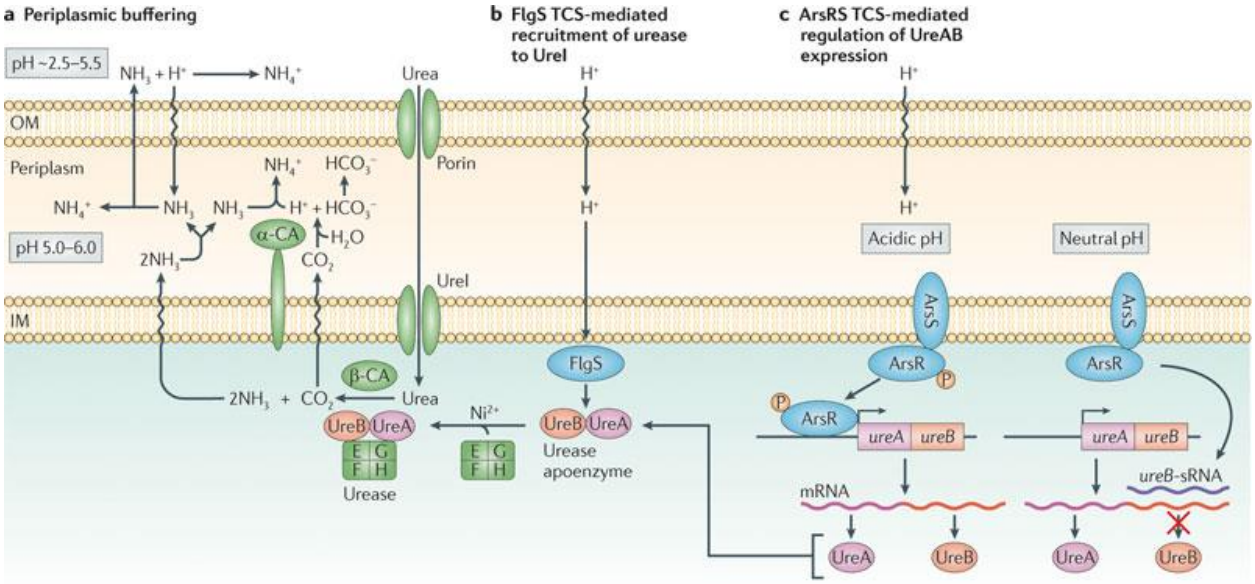

Fig. 1.9 Periplasmic buffering by *H. pylori*. Adapted from Krulwich et al., 2011.

The human stomach is a peculiar ecological niche, characterized by a very acid pH of the gastric juice (1.5-3.5), representing the first line of defense against most gastrointestinal pathogens. It has been estimated that exposure to gastric acid kills more than 99.9% of ingested *Salmonella* species and *Vibrio cholerae* (Gorden and Small 1993). *H. pylori* is the only neutralophile able to survive and to colonize this hostile environment. Since this bacterium is not an acidophile, it has evolved several specialized mechanisms to survive gastric acid. The pathogen has to resist in the gastric lumen (pH ≤2) for a short period, enough to enter into the highly viscous mucosa (average pH ≥5.5), reach the gastric epithelium, find nutrients and multiply (Suerbaum and Josenhans 1999). Some acid-adaptive mechanisms, involved in the gastric juice buffering, include an acid-activated inner membrane urea channel, UreI, a neutral pH-optimum intrabacterial urease, a membrane-anchored periplasmic carbonic anhydrase (Fig. 1.9). This acid acclimation system allows to regulate the pH of the periplasm and of the surrounding liquid in acidic medium at levels where both the cytoplasmic pH and membrane potential are compatible

with survival and growth (Sachs 2005). On the contrary, other neutralophiles cannot maintain a constant intracellular pH and consequently only transit the human stomach. Therefore, *H. pylori* can be classified as an acid tolerant neutralophile, because it grows best at neutral pH, but it is able to survive and grow in acid conditions and to increase its periplasmic pH in the presence of acid by specialized mechanisms (Scott et al. 1998). Urease, also named urea amidohydrolase, was the first colonization factor identified in *H. pylori*. This neutral pH-optimum enzyme is expressed by *H. pylori* at higher levels than any other known bacteria, accounting for as much as 8% of the total protein. It represents one of the most prominent antigenic components, since it induces a strong immunoglobulin response of both IgG and IgA and helps to recruit neutrophils and monocytes in the inflamed mucosa and to activate production of proinflammatory cytokines (Mobley, Island, and Hausinger 1995; Montecucco and Rappuoli 2001). For these reasons it is used as a taxonomic and diagnostic marker for gastritis and peptic ulcer diseases in humans and it has been tested as a vaccine candidate (Suerbaum and Josenhans 1999). Urease-negative mutants retain only 0.4% of the urease activity of the parent strain and, consequently, are unable to colonize animal models, showing that this enzyme is essential for gastric habitation (Eaton et al. 1991). The urease plays a key role regarding the bacterial ability to overcome the acidic conditions of the gastric lumen. It catalyzes the hydrolysis of urea into ammonia and carbamate; the latter compound spontaneously decomposes to yield another molecule of ammonia and carbon dioxide. Ammonia scavenges protons, to mitigate the acidity and increase the pH; carbon dioxide is converted to bicarbonate by a pair of carbonic anhydrases and it buffers the periplasmic space. Therefore, the generation of ammonia and carbon dioxide provides both acid-neutralizing and acid-buffering capacity, enabling *H. pylori* to raise the pH in its microenvironment and periplasm. Besides the neutralization of the intracellular pH, urease activity leads *H. pylori* also to maintain a normal proton motive force (PMF), an electrochemical gradient for protons across the bacterial cell membrane, by adjustment of the transmembrane potential, despite the acid condition, allowing the growth in the gastric environment (Meyer-Rosberg et al. 1996). *H. pylori* survives at a pH range between 4.0 and 8.0 in the absence of urea (Meyer-Rosberg et al. 1996); however, in the presence of urea this microorganism can survive at a pH as low as 2.5. The K_m value of the *H. pylori* urease for urea is 0.8 mmol/L (Dunn et al. 1990); accordingly, *H. pylori* displays a much higher affinity for its substrate than that of ureases produced by other bacterial species. This high affinity allows the utilization of the limited amount of urea

(between 1.7 and 3.4 mM) present in the gastric juice (Dunne 2014). *H. pylori* urease has a cytoplasmic localization, but a significant fraction has been evidenced to be even associated with the bacterial surface. The mechanism whereby urease becomes associated with the surface of the microorganism is unique. This process is defined "altruistic autolysis" and involves release of urease (and other cytoplasmic proteins) by genetically programmed autolysis with subsequent adsorption of the released urease onto the surface of neighboring intact bacteria (Dunn and Phadnis 1998). Therefore, the enzyme is able to maintain a physiological pH value of cytoplasm and periplasm and the membrane potential, and also to develop a protective buffered layer surrounding the bacterial outer membrane. However, the protective role of the surface-associated urease has been challenged because the isolated pure enzyme has a pH optimum of 7.0–8.0 and is rapidly inactivated by short-term exposure to a buffered pH <5, even in the presence of urea (Ha et al. 2001). The urease synthesis and regulation is a complex process, which requires the participation of several gene products. The biosynthesis of urease is controlled by a gene cluster, *ureABIEFGH*, composed by seven genes: an upstream promoter, *ureA* and *ureB*, the two structural subunits of the enzyme, *ureI*, encoding for an urea-specific channel in the cytoplasmic membrane, *ureE*, *ureF*, *ureG* and *ureH*, encoding accessory proteins that interact with the apo-enzyme and deliver nickel ions to the active site in an energy-dependent process (Mobley et al. 1995). All of the genes except *ureI* show considerable homology with urease cluster genes of other bacterial species, including *Klebsiella aerogenes* (Lee et al. 1992), *Proteus mirabilis* (Jones and Mobley 1988), *Yersinia enterocolitica* (de Koning-Ward, Ward, and Robins-Browne 1994). The *H. pylori* urease core enzyme is a huge complex, whose molecular mass has been estimated to be approximately 600 KDa (Evans et al. 1991). The X-ray structure revealed that *H. pylori* urease has a spherical dodecameric assembly, composed of six heterodimers of UreA (26.5 KDa) and UreB (60.3 KDa) subunits (Fig. 1.10; Ha et al. 2001; PDB accession codes: 1E9Z and 1E9Y).

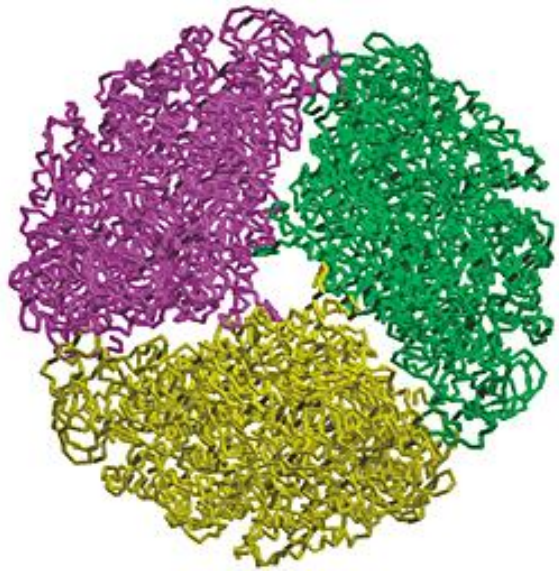

Fig. 1.10 The figure evidences the three-fold symmetry of the dodecameric assembly of UreA/UreB complex and the central huge hallow. Adapted from Nam-Chul Ha et al., 2001.

The supramolecular distribution is organized in four $(UreA/UreB)_3$ units (Fig. 1.11), related by the crystallographic two- and three-fold symmetry axis of a cubic cell. The resulting complex has an outer diameter of the protein shell of approximately 130 Å and a huge central hallow, as observed with transmission electron microscopy (TEM). The urease is originally produced as an immature apo-enzyme and requires Ni^{2+} ions insertion for its catalytic activity. Activation takes place when four chaperone proteins, organized as two complexes, UreE/UreG and UreF/UreH, mediate the Ni^{2+} insertion and assemble the catalytic site of the protein (Mobley et al. 1995). The insertion of two Ni^{2+} ions into each active site is essential for complete activation of the protein, as well as the carbon dioxide, which probably reacts with a Lysine side chain to produce Lysine-carbamate, participating as a ligand in the metal binding, and the GTP hydrolysis, necessary for the Nickel-dependent activation (Ha et al. 2001). Hence urease activity is essential for gastric colonization by *H. pylori* and regulation of urease occurs at several levels: biosynthesis of the structural genes, insertion of Ni^{2+} ions into the apo-enzyme, regulation of urea access to the urease itself (Sachs, Wen, and Scott 2009). The genes encoding UreA and UreB are located in one operon, *ureAB*, while *ureIEFGH* represents another operon (Akada et al. 2000). Recent studies have evidenced that the *ureAB* operon can yield a 2.7 kb full-length transcript that produces a functional enzyme and also a 1.4 kb truncated transcript, the product of which exhibits much lower urease activity and is generated by cleavage of the 3' *ureB* region. Acid pH results in a significant increase in transcription of *ureAB*, while neutral pH increases the truncated transcript. The expression of this smaller transcript was shown to be influenced by pH, the presence of the histidine kinase ArsS and the phosphorylation state of the response regulator ArsR, illustrating the influence that pH has on the expression of an active urease (Wen et al. 2011). ArsR and ArsS belong to a two-component signal transduction system, named ArsRS system, that regulates the transcription of the seven genes of the urease gene cluster and even of other proteins involved in important *H. pylori* pH-sensitive metabolic processes, such as

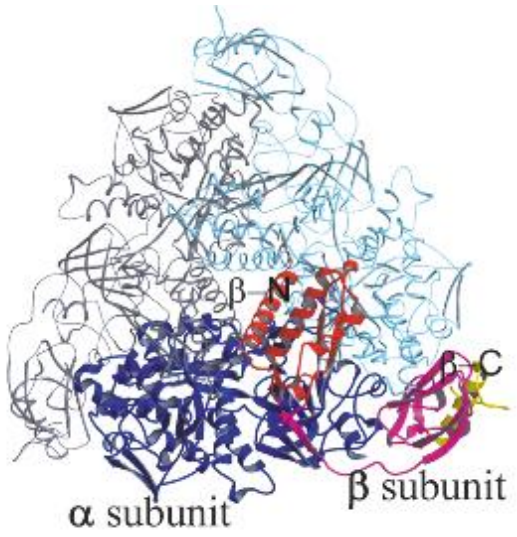

Fig. 1.11 Ribbon diagram of the trimeric unit. One unit is represented with wider ribbons, and the others are differently colored. Each unit consists of the UreA (α-subunit; blue) and two physically distinct domains of the UreB (β-subunit, the N-terminal domain, red, and the C-terminal domain, magenta). Adapted from Nam-Chul Ha et al., 2001.

acid resistance (amidase), acetone metabolism (acetone carboxylase), resistance to oxidative stress (thioredoxin reductase), quorum sensing (enzyme Pfs) (Fig. 1.9; Loh et al. 2010). UreA and UreB are present largely as apo-enzymes; therefore, regulation of Ni^{2+} concentration inside the organism plays a vital role in acid survival. NixA is a specific high-affinity Ni^{2+} transporter with eight transmembrane segments that exists in the inner membrane of *H. pylori* and functions as a monomer (Fulkerson and Mobley 2000). Deletion of *nixA* reduces, but does not prevent, infection of the mouse stomach (Nolan et al. 2002), but many other genes may also be implicated in uptake and storage of this essential cation. HypA and HypB are involved in the insertion of nickel into hydrogenase. Deletion of these nickel assembly genes not only affected hydrogenase activity, but sharply reduced urease activity by 200-fold (Olson, Mehta, and Maier 2001). This finding is attributed to a role of these genes in nickel incorporation into urease. *H. pylori* also expresses *nikR*, which functions as a global regulator of several genes, including *ureA* and *ureB*, in the presence of high Ni^{2+} concentrations (Contreras et al. 2003). Finally, the uptake of urea from the gastric environment is regulated by an acid-activated transmembrane channel, UreI. This transporter is activated by acid pH, when the periplasmic pH falls below approximately 6.2 (Sachs et al. 2003); therefore, it doesn't allow for the transport of urea into the bacterial cell at neutral pH, thus preventing lethal alkalinization of the cytoplasm. At neutral pH little urease activity is measurable in intact organisms, in contrast to lysed bacteria. But as the pH in the medium is reduced from pH 6.5 to 5.5 and down to pH 2.5, carbon dioxide release sharply increases in intact bacteria. The simplest explanation for these observations is that acid pH activates the urea transporter in the inner membrane, increasing urea access to cytoplasmic urease, for a maximal production of ammonia and carbon dioxide. The stable activity down to pH 2.5 also shows that even down to this environmental pH, cytoplasmic pH does not fall (Scott et al. 2000). UreI is a six-transmembrane segment polytopic membrane protein, localized into the inner membrane of the microorganism. UreI shares features of a pH-gated urea channel (non-saturable, voltage independent), having low open probability at neutral pH and high open probability at pH 5.0 and below, with half maximal opening at pH 5.9. It is very selective for urea, with even thiourea being impermeant (Weeks et al. 2000). The *ureI* gene appears to be uniquely expressed by gastric *Helicobacter* species and its absence impairs acid survival. Moreover, *ureI* gene shows homology to the *amiS* gene of the amidase gene cluster (Scott et al. 2000); not coincidentally urea is an amide compound.

Although urease and UreI play a significant role in the acid acclimation processes of *H. pylori*, it would be extremely unlikely that these proteins would be the only means of combating the stress of gastric habitation. Acid-induced gene regulation has been extensively investigated, in absence or presence of urea. Besides urease, several gene products, which might be considered as pH homeostatic genes, are able to produce NH_3, such as acylamide amidohydrolase or amidase (AmiE), formamide amidohydrolase or formamidase (AmiF), two asparaginases (AnsA, AnsB), γ-glutamyltranspeptidase or glutamase (Ggt), glutaminaseaspartate-ammonia lyase or aspartase (AspA). The periplasmic and cytoplasmic carbonic anhydrases are able to produce HCO_3^- as buffer. Also, three hydrogenase expression/formation genes, *hypA*, *hypB* and *hypC*, are up regulated; the first two, as previously described, have a role in nickel uptake and sequestration, essential for adequate urease enzyme activity (Sachs et al. 2009).

Although the NH_3 produced by urease activity is able to neutralize entering protons, the pKa of NH_4^+ is 9.2 and the NH_4^+/NH_3 couple would not effectively buffer the periplasm to a relatively neutral pH. Therefore, NH_3 alone would not account for the finding that the periplasmic pH is relatively constant at 6.1 in the presence of urea. On the other hand, pH 6.1 is the effective pKa of HCO_3^-. Hence, both NH_3 and CO_2 production from intrabacterial urease activity enable acid acclimation by *H. pylori*, the NH_3 to neutralize entering protons and the HCO_3^- to buffer the periplasm (Sachs et al. 2009). *H. pylori* carbonic anhydrases closely cooperate with urease, since the latter produces $2NH_3 + CO_2$ from urea, and these

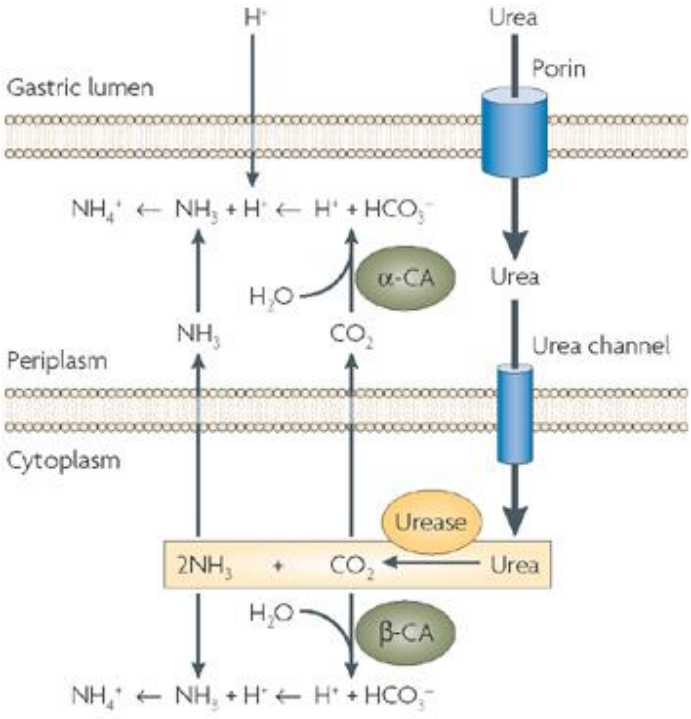

Fig. 1.12 ***H. pylori*** **pH-buffering mechanisms.** The outer membrane contains porins permeable to urea and protons. With acidification, UreI opens and urea moves into the cytoplasm, increasing intrabacterial urease activity. This produces 2NH3 and CO2, gases that readi H. pylori pH-buffering mechanisms. The outer membrane contains porins permeable to urea and protons. With acidification, UreI opens and urea moves into the cytoplasm, increasing intrabacterial urease activity. This produces $2NH_3$ and CO_2, gases that readily exit the cell into the periplasm. Protons entering the cytoplasm are neutralized by NH_3, forming NH_4^+, whereas cytoplasmic β-carbonic anhydrase generates HCO_3^-, which is a stronger buffer at neutral pH than NH_3. Similarly, the NH_3 that effluxes into the periplasm can neutralize entering acidity and the CO_2 due to periplasmic α-carbonic anhydrase activity, producing H^+ and HCO_3^-. A second NH_4^+ is formed along with HCO_3^-, the latter providing buffering in the range of pH 6.1. Adapted from Sachs et al., 2005.

enzymes together establish the main pH-buffering mechanism; therefore, carbonic anhydrases have been investigated in detail. Carbonic anhydrases are metalloenzymes which catalyze the reversible carbon dioxide hydration into bicarbonate and protons. *H. pylori* genome encodes two different classes of carbonic anhydrase with different subcellular localization: a periplasmic enzyme belonging to the α-family and a cytoplasmic enzyme belonging to the β-family (Nishimori et al. 2008). Carbonic anhydrases are involved in pH homeostasis, as well as CO_2 transport or trapping. The mechanism of periplasmic pH-buffering is explained in detail in Fig 1.12. This acid acclimation feature of *H. pylori*, maintaining an intracellular neutral pH while the pH of the environment is acid, is unique to this microorganism and is critical to survival of the pathogen in the stomach. When the periplasmic carbonic anhydrase gene is deleted, many of the properties of the wild type are lost, similar to *ureI* deletion. Furthermore, acid survival is impaired, either with deletion of the enzyme or by addition of acetazolamide, a specific inhibitor for α-carbonic anhydrase, although UreI and urease remain fully functional. There is a 3-log decrease in survival in the knockout mutants or after the addition of acetazolamide at pH 2.0 in the presence of urea (Marcus et al. 2005b). These evidences confirm the crucial role of carbonic anhydrase for the survival of *H. pylori* in the gastric environment.

In addition to urease, *H. pylori* also possesses other NH_3-producing enzymes, including two paralogous amidases, the acylamide amidohydrolase or amidase (AmiE) and the formamide amidohydrolase or formamidase (AmiF). These enzymes are aliphatic amidases that catalyzes the hydrolysis of short-chain amides to produce ammonia and the corresponding organic acid (Skouloubris, Labigne, and De Reuse 2001). Firstly, in 1997, AmiE was identified in *H. pylori* genome, showing high homology to the aminoacid sequence of the aliphatic amidases from other bacterial species, such as *Pseudomonas aeruginosa* or *Rhodococcus* sp. R312 (Skouloubris, Labigne, and De Reuse 1997). Later, analysis of the complete genome sequence has revealed a second aliphatic amidase, belonging to the nitrilase superfamily and named AmiF, whose aminoacid sequence is 34% identical to AmiE. Therefore, AmiE and AmiF are classified as paralogous proteins, presumably generated by a duplication event of an ancestral gene. This evidence suggests some enzymatic specialization: while AmiE hydrolyzes propionamide, acetamide and acrylamide in vitro, AmiF presents an unexpected substrate specificity, since it only hydrolyzes formamide (Skouloubris et al. 2001). Interestingly, *H. pylori* is the first pathogenic microorganism in which two amidases belonging to the aliphatic amidase

family have been found; these enzymes were previously reported only in bacteria present in the environment, like *P. aeuruginosa* (Bury-Moné et al. 2003). The presence of these protein underlines the essential role played by ammonia metabolism in the *H. pylori* survival, as a major nitrogen source, as a compound involved in the acid resistance, and as a cytotoxic molecule. Genes encoding these amidases are up-regulated in response to moderately acid conditions (pH approximately 5.5) encountered at its colonization site, the gastric mucus layer (Bury-Moné et al. 2004). This observation confirms the importance of the buffering role of the ammonia in the protection against the acidity. Furthermore, AmiE and AmiF expression is dependent on the activity of other enzymes involved in the nitrogen metabolism of *H. pylori*, urease and arginase respectively; the production of these amidases is stimulated in *H. pylori* mutants deficient in urease and arginase (Skouloubris et al. 2001). As suggested by these findings, *H. pylori* presents a complex regulatory network that includes several effectors and allows to maintain an appropriate nitrogen balance.

Another enzyme involved in this regulatory network is the above-mentioned arginase (RocF). This protein catalyzes the conversion of L-arginine to L-ornithine and urea, providing endogenously the substrate for the urease. In addition to producing endogenous urea, *H. pylori* also obtains some of its urea exogenously from the host, as urea is present in the gastric mucosa through host arginase activity. Exogenous substrate is imported into *H. pylori* cytoplasm through the UreI transporter, under acid conditions; but this system may be inoperable *in vivo* under neutral pH conditions. Thus, arginase may be important for providing endogenous urea *in vivo* under conditions in which the exogenous one is limited (McGee et al. 2004). RocF in *H. pylori* is metabolically upstream of urease and the latter is known to be required for colonization of animal models by the bacterium. On the contrary, *H. pylori* RocF is required for arginase activity and is crucial for acid protection *in vitro*, but is not essential for in vivo colonization of mice or for urease activity (McGee et al. 1999). *H. pylori* arginase is specifically involved in acid resistance; but it also inhibits host nitric oxide production, evading the immune response and thus contributing to persistent infection. Moreover, it plays a role in inhibiting T cell proliferation. These evidences confirm that *H. pylori* arginase is largely involved in pathogenesis.

Additionally, two asparaginases (AnsA, AnsB), glutaminase (Ggt) and aspartase (AspA), previously mentioned, participate to the ammonia metabolism of *H. pylori*, since they produce NH_3 from aminoacid catabolism. Especially, asparaginase converts L-asparagine to L-aspartic acid and NH_3; glutaminase is responsible for conversion of L-

glutamine to L-glutamic acid and NH_3; finally, aspartase catalyzes the reversible conversion of L-aspartic acid to fumarate and NH_3. Besides the roles in the aminoacid metabolism and in the acid acclimation, asparaginase and glutaminase are also involved in the host cell cytotoxicity and in the suppression of the host immune response (Miller and Maier 2014): asparaginase inhibits either the cell-cycle, particularly the G1 phase, of gastric and non-gastric cells (fibroblasts and epithelial cells) or the normal lymphocyte function at the gastric niche, allowing *H. pylori* to evade the host immune response (Scotti et al. 2010; Shibayama et al. 2011); on the other hand glutaminase induces apoptosis in gastric epithelial cells via a mitochondria-mediated pathway (K.-M. Kim et al. 2007).

Nevertheless, the major ammonia-producing pathway, central to the acid resistance of *H. pylori*, includes urease, amidase, formamidase and arginase. In this complex metabolic network ammonia and also urea could serve as sensors for the regulation of the nitrogen balance, avoiding toxic intracellular accumulation of ammonium and ensuring sufficient nitrogen supply in environmental conditions riches in urea, aminoacids or amides (Skouloubris et al. 2001). Compared to many other bacteria, *H. pylori* possesses relatively few transcriptional regulatory systems; thus it is possible that the few regulatory proteins present control multiple responses and metabolic processes (Alm and Trust 1999). Consistent with this, many genes involved in the *H. pylori* acid resistance, including members of urease genes cluster, amidases and arginase, have been demonstrated to be regulated in the same way by the previously mentioned ArsRS two-component system in response to low pH conditions (Fig. 1.13). It consists of a histidine kinase sensor element ArsS, which phosphorylates itself in response to low pH and transfers the phosphoryl group, and an OmpR-like response regulator ArsR that, by accepting the phosphoryl group, functions as an activator or repressor of gene promoters (Pflock et al. 2006). The influence on the transcription of the pH-sensitive genes by this

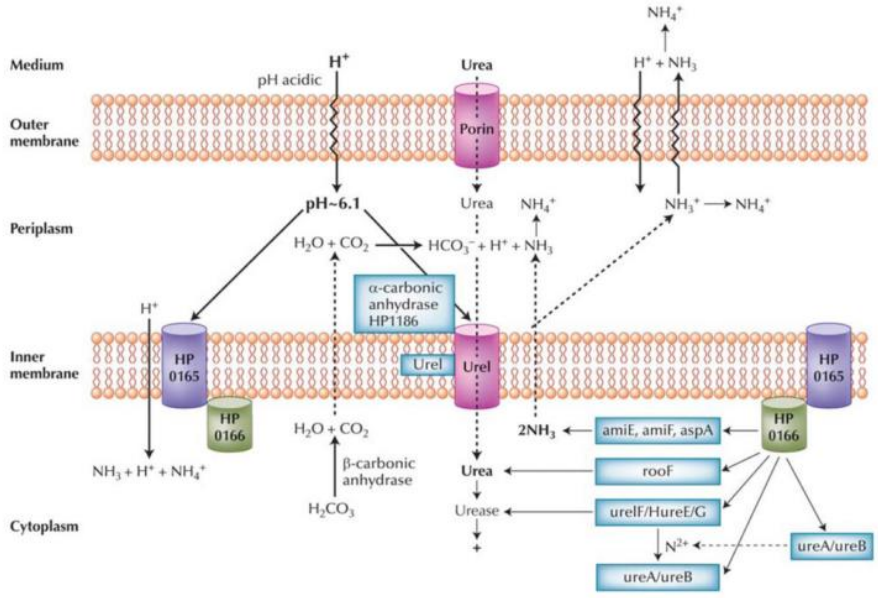

Fig. 1. 13 A model of ArsRS two-component system regulation of periplasmic pH of *H. pylori*. Adapted from Sachs et al., 2009.

two-component system leads to the upregulation of the product in response to low pH. ArsS, in addition to its role in urease gene cluster transcription, is also involved in the recruitment of urease proteins to the inner membrane, precisely to the UreI channel, in order to augment acid acclimation during acute acid exposure (Marcus et al. 2012); this second ArsS function is phosphorylation-independent. Deletion of ArsS renders *H. pylori* unable to colonize the stomach of mouse models, suggesting an essential role of the ArsRS system in the regulation of important virulence properties of the pathogen (Panthel et al. 2003). On the other hand, ArsR has proved to be essential for cell growth (Schär, Sickmann, and Beier 2005). These evidences suggest that this regulatory system and the related target genes are likely essential for the gastric colonization. Besides the genes involved in the acid resistance, several others are regulated by the ArsRS two-component system in a similar mode, including outer membrane proteins, Ni^{2+} storage proteins, detoxifying enzymes involved in the oxidative stress resistance and *H. pylori* specific proteins of unknown function (Pflock et al. 2006). Several of these genes are also under the control of the metal-dependent regulators NikR and Fur. The ferric uptake regulator (Fur) is a well characterized iron-binding transcriptional factor that controls intracellular iron homeostasis *via* concerted expression of iron-uptake and iron-storage genes in response to changes in iron availability (Van Vliet et al. 2002). Binding of ferrous iron to Fur triggers a conformational change that activates the protein for binding to specific DNA sequences named Fur boxes (Dian et al. 2011); it acts as transcriptional repressor that decreases iron uptake in the presence of a high-iron environment. Iron is required as a cofactor by several enzymes and as a catalyst in the electron transport processes. However, iron overload stimulates the formation of reactive oxygen species *via* the Fenton reaction. The resulting species damage DNA, proteins and membrane lipids (Touati 2000). In addition to genes directly related to iron homeostasis, Fur also regulates expression of several enzymes that play a central role in metabolism and energy production, as well as in response to low pH, oxidative stress, salt. Therefore, in *H. pylori* Fur is negatively auto-regulated and is required for the adaptation of the bacterium to two conditions known to oscillate within the gastric mucosa: iron limitation and low pH. Specifically, by DNA array-based transcriptional profiling it has been shown that 16 genes encoding proteins involved in metal metabolism, nitrogen metabolism, motility, cell wall synthesis and cofactor synthesis have displayed iron-dependent Fur-repressed expression; conversely, 16 genes encoding proteins involved in iron storage, respiration, energy metabolism, chemotaxis, and oxygen scavenging have displayed iron-induced

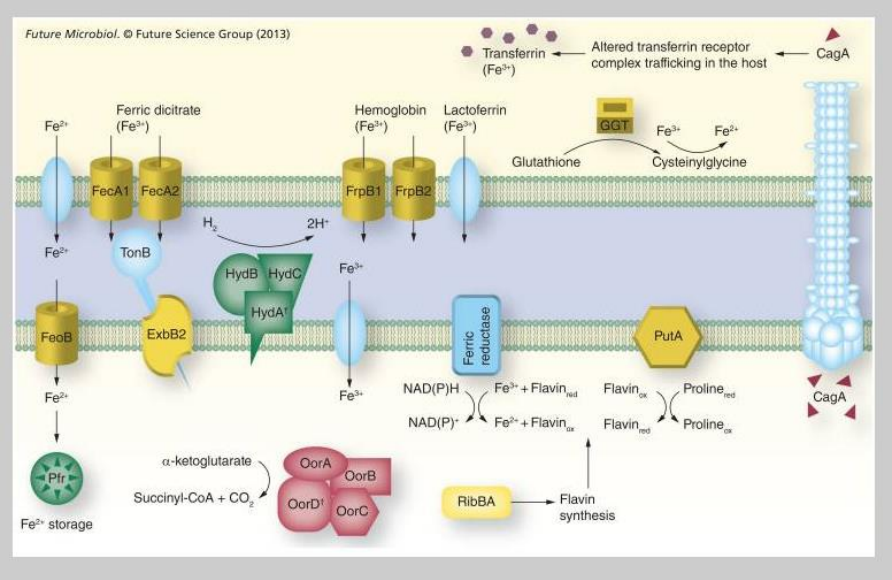

Fig. 1.14 The iron sources available in the stomach and the ferric uptake regulator-mediated responses aimed at maintaining the intracellular iron balance in *H. pylori*. Yellow color indicates that expression is inactivated by iron-bound ferric uptake regulator (Fur). Green color indicates that expression is inactivated by apo-Fur. Red color indicates that expression is activated by iron-bound Fur. Adapted from Pich and Merrell, 2013.

Fur-dependent expression (Ernst et al. 2005). Thus, this iron-sensing protein is actually a global regulator of gene expression in *H. pylori* that contributes significantly to the unique plasticity that is characteristic of this bacterium. Consequently, the Fur regulon includes genes involved in acid acclimation, resistance to oxygen reactive species and nitrogen metabolism. Therefore, collectively, Fur plays a key role in the adaptation of *H. pylori* to the hostile conditions that exist in the stomach (Pich and Merrell 2013). The most widespread group of genes under the control of this metalloregulator is that of the iron uptake systems, including the high-affinity iron transporters FecA1, FecA2, FrpB1 and FeoB (Fig. 1.14; Delany et al. 2001; Danielli et al. 2009). When the intracellular concentration of ferrous iron is low, Fur is unable to bind to the Fur box sequences within the promoters of these genes, transcription is derepressed and the capacity of *H. pylori* to acquire extracellular iron increases drastically. Furthermore Fur activity is also regulated at low pH, because its transcription is repressed in response to low pH (Bury-Moné et al. 2004); this regulation is mediated by the previously mentioned NikR (Van Vliet, Ernst, and Kusters 2004). NikR is a nickel-binding transcription factor that directly controls urease expression and regulates the uptake and storage of nickel, and is also able to regulate the expression of other regulatory proteins, such as Fur. Nickel is a required co-factor for several enzymes, including the fundamental urease; however, because of its potential toxicity, nickel import and homeostasis must be tightly controlled. NikR utilizes allostery and coordination geometry to sense nickel ions and regulate transcription of genes involved in nickel import and processing (Dosanjh and Michel 2006). This regulatory protein acts as both a repressor and an activator within an acid adaptation cascade, which enables the optimal use of the nickel ions under acid conditions, thereby preventing the emergence of toxic nickel concentrations. It has been

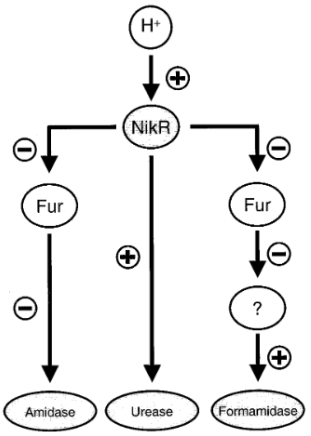

Fig. 1.15 Relationship of NikR, Fur, amidase, formamidase and urease in *H. pylori*. Adapted from Van Vliet et al., 2004.

hypothesized that *H. pylori* uses a repressor cascade to respond to low pH conditions, with NikR initiating the response directly *via* the urease operon and indirectly *via* the members of the Fur regulon. The NikR and Fur proteins generally function as transcriptional repressor proteins through binding of operators in their target promoters, blocking access of RNA polymerase, and subsequent transcription. The regulatory cascade utilizes these two repressors, each only allowing on/off regulation but, by varying the levels of Fur protein in cells, NikR not only regulates expression of its own regulon but also mediates changes in all members of the Fur regulon indirectly *via* regulation of *fur* transcription (Fig. 1.15; Van Vliet, Kuipers, et al. 2004a). This regulatory system demonstrates the close link between metal metabolism and acid resistance in *H. pylori* and may allow the bacterium to mount protective responses to a multitude of different stresses despite a paucity of regulatory systems.

1.4.2 MOTILITY AND CHEMIOTAXIS

Since *H. Pylori* is not an acidophile, the bacterium has evolved several strategies to minimize the exposure to the acid conditions in the stomach lumen by remaining in very close proximity to the surface of the epithelium, where the pH is approximately neutral (Amieva and El-Omar 2008). Therefore, a key factor essential for survival and successful colonization is the bacterial motility. *H. pylori* is considered a “good swimmer”, able to remain within a narrow band of the protective mucus layer that is constantly being secreted by epithelial cells of the stomach. The bacterium moves in the mucus gel *via* a cork-screwing mechanism thanks to its helicoidal shape and to a group of four to eight flagella arranged at one of its poles (Fig. 1.16). The helicoidal shape is crucial for bacterial motility and a prerequisite for successful colonization. The turning helical cell body is thought to interact with large polymers to generate torque that enhances translational movement and reduces circumferential slip, promoting the penetration of the mucus layer. The helicoidal shape improves propulsion efficiency in the form of speed in viscous polymer solutions, such as the gastric mucus gel composed by polymerized mucins, and *H. pylori* has been shown to swim faster at higher viscosities respect to certain rod-shaped bacteria (Magariyama and Kudo 2002). Most bacteria, including *H. pylori*, present

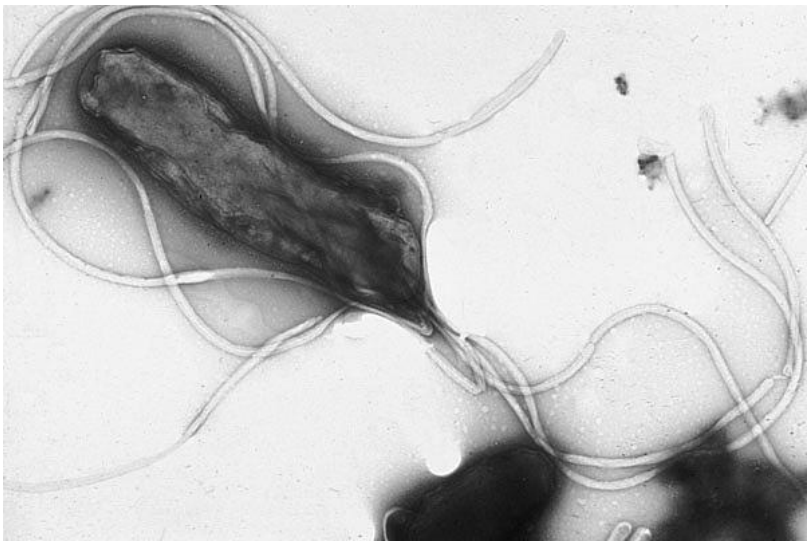

Fig. 1.16 ***H. pylori*** **has unipolar flagella and a spiral shape that enable the bacterium to be motile. Electron micrograph; source: Yutaka Tsutsumi, M.D.**

a peptidoglycan (murein) sacculus, surrounding the cytoplasmic membrane and essential to maintain osmotic stability and cell shape. It consists in a meshwork of glycan strands joined by peptide crosslinks (Sycuro et al. 2010). A number of *H. pylori* genes that induce modifications in peptidoglycan cross-linking of the bacterial cell wall or lead to trimming of peptidoglycan muropeptides have been identified. These include several genes that encode peptidases and determine the helical shape of *H. pylori*, namely, *csd1*, *csd2*, *csd3*/*hdpA*, *ccmA*, *csd4*, *csd5* and *csd6* (Sgouras, Trang, and Yamaoka 2015). Quantitative morphological analyses of multiple-gene deletion mutants have revealed different impairments in shape and motility and each protein uniquely and coordinately contributes to the shape-generating pathway. These findings suggest that the coordinated action of multiple proteins relaxes peptidoglycan crosslinking, enabling helicoidal cell curvature and twist (Sycuro et al. 2010; Sycuro et al. 2012). In addition to the influence of peptidases on bacterial cell shape, recent research has revealed that the shape of *H. pylori* may be dictated by coiled-coil-rich proteins (Ccrps). The bacterium contains four Ccrps (Ccrp58, Ccrp59, Ccrp1143, and Ccrp1142), and all four ccrp deletion mutants significantly exhibit impaired motility, despite unaltered flagella morphology (Sgouras et al. 2015).

Motility is essential for successful *in vivo* colonization by *H. pylori* and is provided by its sheathed unipolar flagella, which are considered as one of the most important colonization and virulence factors of this pathogen, since non-motile mutants lacking flagella are unable to establish persistent infection in animal models (Montecucco and Rappuoli 2001). These propulsive elements allow *H. pylori* to swim in the gastric juice, guided by chemical stimulant, especially urea and bicarbonate ions, till crossing the thick mucus layer protecting the epithelium. In this way the bacterium gets nutrients and avoids to be discharged in the intestinal tract by peristalsis. *H. pylori* possesses two to eight flagella; the filaments extend 3-5 μm from the bacterial surface and exhibit a typical bulb-

like structure at its distal end that represents a dilation of the flagellar sheath (Geis et al. 1989). The sheath itself consists of both proteins and lipopolysaccharide, and is thought to be an extension of the bacterial outer membrane that protects the flagellar filaments from acid in the stomach (Geis et al. 1993).

The *H. pylori* flagella, as those of enteric bacteria, are composed of three structural elements: a basal body, which is embedded in the cell wall and contains the proteins required for rotation and chemotaxis and usually the flagellar export proteins; an external helically shaped filament that works as a propeller when rotated at its base; a hook that serves as a joint between the basal body and the flagellar filament (Fig. 1.17; Chevance and Hughes 2008). In *H. pylori* more than 50 putative proteins are predicted to be involved in expression, secretion and assembly of this complex flagellar apparatus; at least 20 of these proteins constitute the structural components of the basal body, hook and filament (Tomb et al. 1997; Alm et al. 1998). The filament is a copolymer constituted by two subunits, the predominant flagellin FlaA and the minor one FlaB, localized close to the basis of the flagellum (Kostrzynska et al. 1991). Both flagellins have similar molecular mass (approximately 53 KDa), present an acid pI of 5.2 and share considerable aminoacid homology (58% identity), but the respective genes are unlinked in the chromosome (Mobley et al. 2001). They are are necessary for full motility (Josenhans, Labigne, and Suerbaum 1995) and for the establishment of a persistent infection in the gnotobiotic piglet model (Eaton et al. 1996). The hook is composed of FlgE structural subunit (78 KDa); whereas FliD is a hook-associated protein, which is localized at the tip of the flagellar filament and promotes the incorporation of the flagellin

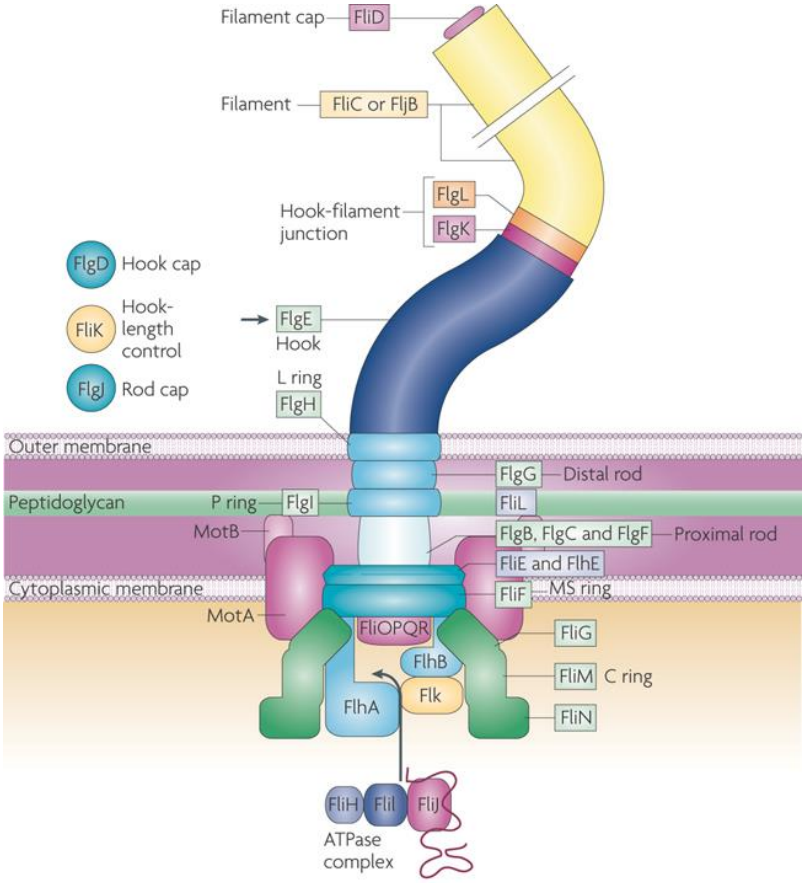

Fig. 1.17 Structural elements of *H. pylori* flagella. Adapted from Chevance and Hughes, 2008.

monomers into the growing flagellar filament (Mobley et al. 2001). Mutants in *flgE*, defective in hook production generated by allele replacement, are non-motile and devoid of flagellar filaments but produce both flagellin subunits (O'Toole, Kostrzynska, and Trust 1994). Mutants in fliD produce truncated flagella and are severely impaired in motility and their ability to colonize the gastric mucosa of mice (Kim et al. 1999). Several other genes involved in the flagellar architecture and function have been identified by sequence homologies studies. The *fliF* gene encodes the subunits of the MS ring, the first complex to be assembled during flagellar morphogenesis (Jones, Homma, and Macnab 1989). This complex is formed on the cytoplasmic side of the inner membrane, where it provides a construction base for the flagellar rod made of the FlgB, FlgC and FlgG proteins and functions as an anchor for the motor switch proteins FliM, FliN and FliG and the motor rotation proteins MotA and MotB. On the other hand, *flgI* and *flgH* genes encode the subunits of the P and L rings (Jones et al. 1989), which anchor to the flagellum in the periplasmic space and to the outer membrane, respectively (Mobley et al. 2001). Additionally, many other putative proteins seem to be implicated in biosynthesis, assembly and function of the flagellar apparatus, but little is known about them (O'Toole, Lane, and Porwollik 2000).

Since most of the flagellar apparatus is localized beyond the cytoplasmic membrane, many of the flagellar proteins have to cross the membrane to reach their final destination. The proteins constituting the P and L rings are secreted via the conventional signal-peptide–dependent Sec pathway, whereas the axial components of the flagellar apparatus, including the structural proteins of filament, hook and rod, which connects the hook to the basal body, are believed to be secreted by a specialized flagellum-specific pathway. This process is constituted of biosynthetic proteins that assemble into a structure at the MS ring of the flagellum that binds the flagellar proteins and transfers them actively into the growing flagellum. The members of this system share homology to components of the widely distributed contact-dependent type III secretion systems (Mobley et al. 2001).

Another ability that plays a crucial role in the gastric colonization by *H. pylori* is the control of the direction of the movement, which is achieved by means of chemotactic response to various chemical stimuli. It has been proved that chemotaxis is responsible of four main events in *H. pylori* mouse infections: establishing infection, achieving high-level infection, maintaining an infection when there are competing *H. pylori* present and colonizing all regions of the stomach (Terry et al. 2005). Most flagellated bacteria perform chemotactic

motility by the recognition of environmental conditions, such as solute concentration, coupled with regulation of swimming behavior (Falke and Hazelbauer 2001); *H. pylori* exhibits attractant chemotactic activity toward various compounds, including the aminoacids glutamine, histidine, lysine and alanine (Abdollahi and Tadjrobehkar 2012), as well as mucin, urea, sodium bicarbonate and sodium chloride (Mizote, Yoshiyama, and Nakazawa 1997), and it is repelled as response to low pH. From the gastric lumen to the apical surface of the gastric mucosa a large pH gradient generated by a different permeability of the mucus to different compound has been postulated (Montecucco and Rappuoli 2001). H. pylori can lose motility rapidly when it encounters acid conditions (Schreiber et al. 2005); therefore, it penetrates the gastric mucus quickly and establish persistent colonization in an area close to the epithelium, where the pH is approximately neutral. *H. pylori* increases its swimming speed when placed within an acid gradient; it can also change its swimming route to favor movement away from the acid environment and toward the gastric epithelium. Like other bacteria, *H. pylori* uses four different chemoreceptors, also named transducer-like proteins (Tlps), to phosphorylate the flagellar rotational response regulator and modulate the flagellar rotational direction. This chemoreceptors (specifically TlpA, TlpB, TlpC and TlpD), located in the bacterial membrane or in the cytoplasm, are molecule that sense the extracellular chemical signals and, following interaction with their respective ligands, initiate a molecular signal transduction cascade, which causes a change in the direction of rotation of the flagellar motors (Sgouras et al. 2015). Positive taxis to arginine and bicarbonate have been observed to be dependent on TlpA function. Whereas negative taxis to acid pH is dependent on the sensor protein TlpB. TlpC has been hypothesized to modulate the TlpB-mediated acid behavior by an as-yet-unknown mechanism. On the other hand, TlpD is a receptor for energy taxis (Cerda et al. 2011). The relevant role of these chemoreceptors has been demonstrated by deletion experiments; an *H. pylori* mutant in the *tlpB* gene can swim but do not move away from acid regions. These *H. pylori* mutant are defective in their ability to colonize mice stomach (Croxen et al. 2006). Therefore, the structure of flagella, the chemotactic movement and its regulation underline the pivotal role of motility and flagella in *H. pylori* colonization.

Additionally, *H. pylori* has been shown to possess the enzymatic ability to disrupt the oligomeric structure of mucins, enabling the pathogen to move freely in the mucus layer, assisted by its highly active flagellum and its ability to down-regulate mucin synthesis (Byrd et al. 2000). A thioredoxin system has been identified to be responsible of the

alteration of the mucus structure; it specifically reduces interchain disulphide bonds of mucins (Windle et al. 2000). This event reduces the gel-forming capabilities of mucins and therefore the viscoelastic properties of mucus aiding movement of the bacterium through mucus.

1.4.3 ADHESION

Approximately only 20% of *H. pylori* bacteria in the stomach adhere to the surface of the gastric epithelial cells, with a majority exhibiting a tropism for intercellular junctions and, occasionally, for deeper intercellular spaces (Fig. 1.18); whereas most of the infecting bacterial cells are found living in the mucus layer (Hessey et al. 1990). *H. pylori* has been shown to disrupt tight junction function, possibly through intraepithelial injection of cytotoxin-associated antigen A (CagA) protein or through the ammonia produced by urease; this change might open the paracellular route for bacterial penetration deep into the mucosa (Amieva et al. 2003). Although *H. pylori* has generally been considered an extracellular pathogen, a number of evidences have shown that a subset of microorganisms is capable of occasionally invading gastric mucosa and entering epithelial cells (Petersen and Krogfelt 2003). *H. pylori* invasion of and survival within epithelial cells is not merely a passive event, but requires active participation of the microorganism. A study on bacterial entry have demonstrated that *H. pylori* enters into gastric epithelial cells through a zipper-like phagocytic mechanism, that requires protein kinase C and phosphatidylinositol 3-kinase (Kwok et al. 2002). Moreover, it has been proposed that VacA toxin promotes bacterial intracellular survival in gastric epithelial cells and contributes to the persistence of infection (Terebiznik et al. 2006); indeed, one possible explanation for the persistence of *H. pylori* infection despite vigorous host immunological defenses and antibiotic therapy is the existence of an intracellular bacterial reservoir. Additionally, electron microscopy studies of gastric biopsy specimens obtained from infected humans have demonstrated

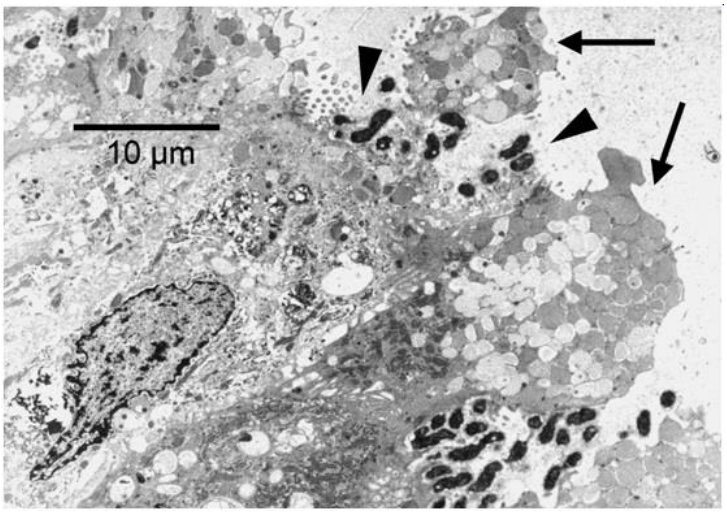

Fig. 1.18 Electron microscopy shows *H. pylori* attached to remaining microvilli (arrowheads). Adapted from Warren, 2010.

extensive areas of adhesion to the host cells and the presence of *H. pylori* within epithelial cells (Wyle et al. 1990). However, significant controversy still remains regarding the ability of *H. pylori* to invade epithelial cells and the mechanisms involved.

H. pylori bacteria which infect mucosal surface share two main goals to achieve a successful colonization: firstly to overcome the mucus barrier, and secondly to interact with the underlying epithelial cells that results in disease (Naughton et al. 2014). *H. pylori* is highly adaptable, as evidenced by the fact that it can occupy a single host for decades, and it is its able to physically interact with various types of gastric host cells and also with host mucins and extracellular matrix proteins using a number of different adhesins displaying a variety of unique receptor specificities (Fig. 1.19; Evans and Evans 2000). Bacterial adhesion involves specialized molecular interactions with the gastric mucosa that may trigger cellular changes including signal transduction cascades and lead to intimate attachment and modification of the cell surface and the underlying cytoskeleton. It is widely accepted that the microorganisms in contact with the epithelial cells cause disease. Moreover, it is considered unlikely that chronic infection with *H. pylori* could occur in the absence of adhesin-host cell interactions. Therefore, adherence of the bacteria to the gastric mucosa is one of the initial steps of *H. pylori* infection and is an important virulence factor; it renders *H. pylori* 100 to 1,000 times more resistant to antibiotics than non-adherent bacteria (Megraud et al. 1991).

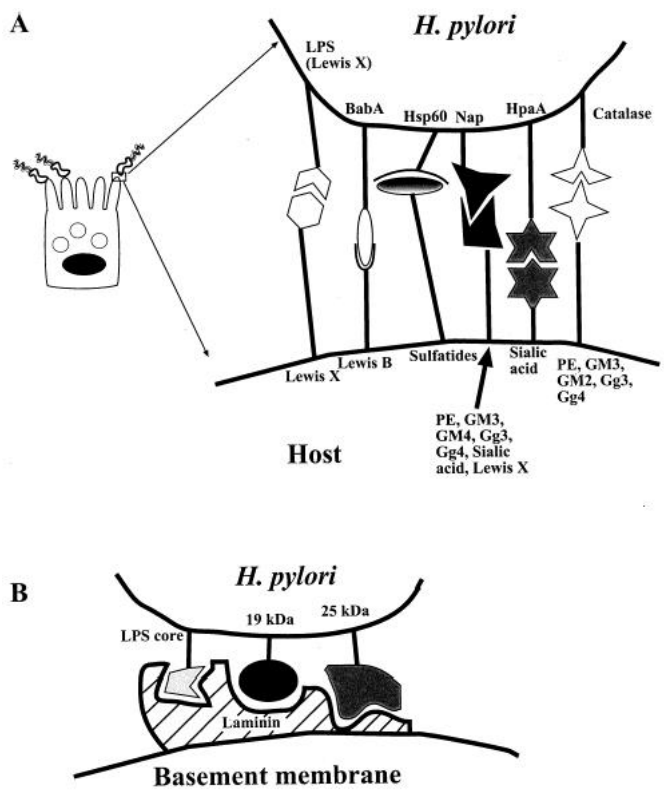

Fig. 1.19 Interaction of *H. pylori* adhesins with host cell receptors. (A) Interaction of *H. pylori* adhesins to gastric epithelial cell receptors. (B) Interaction of *H. pylori* adhesins to the basement membrane protein laminin. Adapted from Mobley, 2001.

More than 30 genes that encode bacterial outer membrane proteins have been identified as possible candidate involved in the adhesion (Odenbreit et al. 2009); several of these have been recognized as adhesins, suggesting multiple and perhaps redundant variable modes of attachment to the gastric cell surface (Amieva and El-Omar 2008). Moreover, different *H. pylori* strains express different adhesins sets, contributing to the strain variability in *H. pylori* and to the adaptation to

changes in mucosal glycosylation at different anatomic locations and over time; adhesin expression can also change within a single strain over time, thus suggesting an evolutional meaning in the redundancy of mechanisms of adhesion. No individual molecule has been shown to be essential for adhesion, further corroborating the multifactorial nature of the adhesin-mediated colonization and the redundancy of adhesive mechanisms. Additionally, to bacterial variability in the adhesin sets, host factors may also influence the adhesion, such as the varying expression of host cell receptors within a single host and the genetic variability in receptor expression in different hosts. Therefore, the cell adhesion mechanism and the role of individual adhesins are very difficult to dissect at the molecular level because of several bacterial and host factors involved.

Particularly, molecular effectors that mediate the interaction between the bacterium and the host cell surface include bacterial proteins, glycoconjugates and lipids. *H. pylori* adheres strongly to the human gastric epithelial cells using fucosylated glycoproteins and sialylated glycolipids as host receptors, highly expressed in the gastrointestinal tract (Boren et al. 1993). The successful adhesion is mediated by several adhesins, including proteins and glycolipids, which play an essential role in the development of the diseases. The primary colonization may occur in the oral cavity, as *H. pylori* has been shown to adhere to MG2 in the human salivary mucin (Andersen 2007). Mucins are a family of high molecular weight oligomeric proteins that are sulfated and heavily glycosylated, containing a wide array of structurally distinct carbohydrate side chains, such as sulfated or non-sulfated sugar moieties, Lewis a, b, X, and/or Y moieties, and sialic acid residues (Carraway and Hull 1991). Adherence to the sialic acid in the mucin seems to be a common feature in most *H. pylori* strains. Several sialic acid-containing glycoconjugates have been shown to exist, such as on the surface of gastric epithelial cells and on human neutrophils; these sialic acid-containing glycoconjugates are up-regulated upon contact of *H. pylori* with the host cell (Miller-Podraza et al. 1999). *H. pylori* possesses at least six adhesins able to bind sialic acid; some of them have been identified, especially HpaA, Nap, SapA, SabA (Fig. 1.19). HpaA has been the first sialic acid-binding adhesion of *H. pylori* to be characterized (Evans et al. 1988). It is a surface-located lipoprotein that binds to sialoconjugates, specifically to the N-acetylneuraminyllactose (Evans et al. 1993), mainly in an α-2,3-specific manner. It has been revealed to be essential to establish a successful colonization in mice (Carlsohn et al. 2006). HpaA has also been observed as a component of the extracellular flagellar sheath (Jones et al. 1997).

The sialic acid-binding adhesin (SabA, also known as HopP because it belongs to the *Helicobacter* outer membrane porin (*hop*) superfamily) is one of the best characterized *H. pylori* adhesins, along with the Lewis b antigen binding adhesin (BabA, also known as HopS). SabA binds to sialylated structures present on gastric mucin and on epithelial cells. This adhesin also binds to sialylated receptors on neutrophils, which leads to non-opsonic activation of the neutrophils, phagocytosis of the bacteria and induction of the oxidative burst response. Furthermore, the adhesin exhibits haemagglutinating activity, binding to gangliosides on erythrocytes in mucosal blood vessels (Dunne 2014). SabA interacts with sialyl-Lewis X antigen (Le^X), an important blood group antigen that is rarely expressed in healthy gastric epithelium but is present in abundance in inflamed and cancerous gastric epithelium (Mahdavi et al. 2002). The SabA-sialyl-Le^X interaction is therefore likely to play a pivotal role in *H. pylori* colonization during chronic infection, almost always associated with chronic active gastritis (Pang et al. 2014). SabA adhesion has also been shown to mediate binding to the structurally related sialyl-Lewis a antigen (Le^a). This glycoconjugate is an established tumor antigen and marker of gastric dysplasia, which may further illustrate *H. pylori* capacity to exploit a full range of host responses to epithelial damage (Mahdavi et al. 2002). SabA-mediated binding of *H. pylori* to sialyl glycoconjugates and sialyl gangliosides requires NeuAcα2–3Gal disaccharide as the minimal binding epitope and is favored by extended and flexible glycan core chains (Aspholm et al. 2006). Furthermore, SabA variants in different *H. pylori* strains exhibit different affinities and specificities for the sialylated glycans, sialyl dimeric Lewis X antigen, sialyl-Lewis a antigen, dialyllactosamine and extracellular matrix protein laminin (Aspholm et al. 2006). Such variable SabA binding specificity might be a host adaptation mechanism by which *H. pylori* rapidly modulates its adhesion properties to achieve optimal colonization and concomitantly evade host immune responses (Pang et al. 2014). BabA is the second best characterized *H. pylori* adhesin, together with SabA, and it belongs to the *hop* superfamily too. Moreover, BabA shares with SabA the structural organization; it contains two domains: a N-terminal extracellular host-binding domain and a C-terminal outer membrane–spanning domain forming an 8 stranded β-barrel structure (Fig. 1.20; Hage et al. 2015). BabA mediates the binding of *H. pylori* to fucosylated structures, including the H type 1 antigen and Lewis b blood group antigen (generating by addition of a fucose residue to the precursor H type 1 by transferase activity), presented by the surface of gastric epithelial cells and the gastric secretory mucins (Ilver et al. 1998). BabA is also centrally involved in *H. pylori* binding to the mucins MUC5AC

and MUC5B, even in non-secreting individuals that either lack an α(1,2)-fucosyltransferase (and are therefore not able to express Le^b in high amounts) or lack Le^b, and thereby acts as an important factor for initial colonization (Van de Bovenkamp et al. 2003). This adhesin is expressed by most disease-causing *H. pylori* strains and its presence on the bacterium correlates with enhanced colonization and virulence (Prinz et al. 2001). A high level of heterogeneity is found in the BabA protein amongst strains, with various polymorphisms being identified, and different levels of Lewis b binding observed.

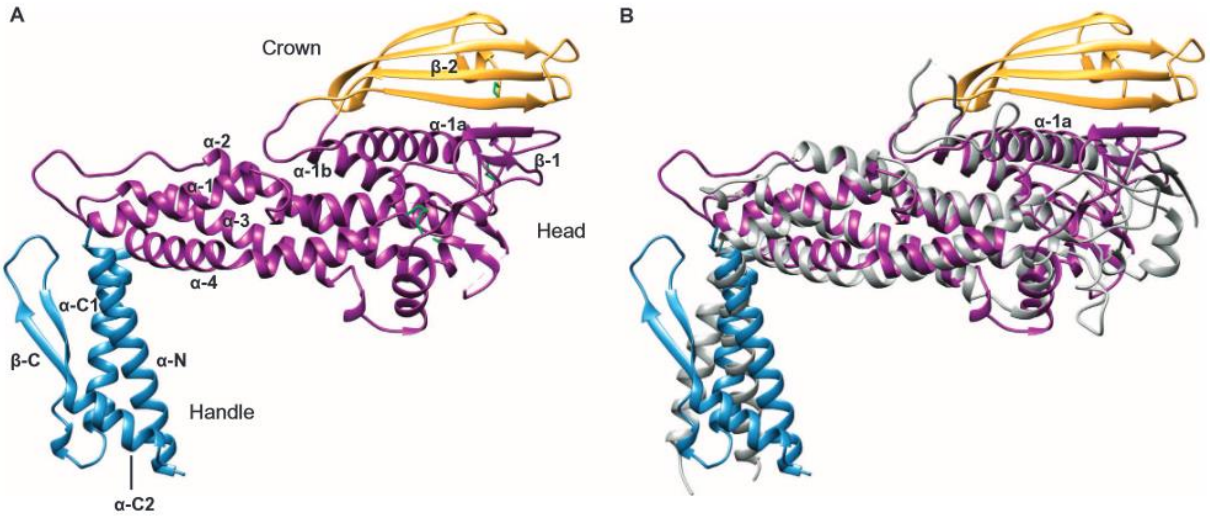

Fig. 1.20 Comparison of the BabA and SabA extracellular domain crystal structures. (A) Crystal structure of the BabA extracellular domain. Indicated are the handle (blue) and head regions (dark magenta) and the crown b-strand unit (gold). The four disulfide bridge are represented as green sticks. (B) Superimposition of the extracellular domains of BabA and SabA (gray). Adapted from Hage, 2015.

Similarly to SabA, BabA also presents a high degree of genetic variability due to phase variation events; their gene includes homopolymeric tracts in the promoter region and dinucleotide repeats in the coding region and phase-variation can occur at these sites *via* slipped-strand mispairing. This high level of sequence flexibility helps *H. pylori* to rapidly adapt to the stomach environment and contributes to differing clinical outcomes among *H. pylori*-infected humans (Hennig et al. 2004; Kao et al. 2012). There is also a high level of allelic variation in *bab* genes. *H. pylori* possesses a closely related gene to *babA*, named *babB*, and both proteins are highly similar in their N- and C-terminus regions, but vary quite significantly in their central region. The paralogous BabB protein does not bind Lewis b antigen, indicating that the central region of the two proteins confers unique functions (Dunne 2014). Despite its specific function is yet unknown, an experimental rhesus macaque model of infection has revealed that BabB, closely related to BabA, could be involved in conversion events where *babA* gene is deleted and *babB* gene is duplicated and replaces the first one (Solnick et al. 2004).

Other adhesins of *H. pylori* include the outer membrane proteins AlpA and AlpB, that are known to mediate the attachment to epithelial cells by defining a macromolecular complex on the bacterial surface with other outer membrane proteins belonging to this family (Odenbreit et al. 1999). AlpA and AlpB are two closely related proteins carried on the same operon in the *H. pylori* genome. The C-terminal portion of both proteins is predicted to form a porin-like β-barrel in the outer membrane, consisting of 14 transmembrane amphipathic beta-strands, similarly to SabA and BabA. Adhesion experiments with defined isogenic mutants indicate that both proteins are necessary for specific adhesion of *H. pylori* to human gastric tissue; loss of these proteins was found to influence the ability of *H. pylori* to colonize the guinea pig stomach (De Jonge et al. 2004) and to bind to gastric tissue (Odenbreit, Faller, and Haas 2002). Virtually all strains express AlpA and AlpB, indicating they have an essential function. Recently, these two proteins have been shown to contribute to the ability of *H. pylori* to bind host laminin (Senkovich et al. 2011). Additionally to its influence on adhesion, the *alpAB* locus has been shown to influence host cell signaling and cytokine production (Odenbreit et al. 2002).

HopZ was identified as a further outer membrane protein of *H. pylori* that plays a role in colonization. It is regulated by a phase-variable CT repeat and two allelic variants of the *hopZ* gene were identified, with a 20 aminoacid region present in only one allele (Kennemann et al. 2012). The role of HopZ in infection has been largely unexplored; HopZ also mediates adherence, since a *hopZ* isogenic mutant has reduced adherence (Peck et al. 1999). There is strong selection in vivo for HopZ expression, as *hopZ* on variants have been recovered from volunteers challenged with a *hopZ* off strain. Transmission of *H. pylori* within families has also been associated with a status change of *hopZ* (Dunne 2014).

Recently, OipA, also named outer inflammatory protein or HopH, has been identified as a phase-variable outer membrane protein characterizing more virulent strains. Its gene is located on the *H. pylori* chromosome approximately 100 kb from the *cag*PAI and its presence is associated with enhanced interleukin-8 secretion, peptic ulcer and increased inflammation (Dabiri et al. 2009). Mutagenesis of *oipA* gene results in reduced bacterial adherence to gastric epithelial cells, but does not alter IL-8 secretion in vitro, probably due to the presence of *cag* Pathogenicity Island (Yamaoka, Kwon, and Graham 2000). OipA may be linked to gastroduodenal diseases owing to its association with other virulence factors or increased bacterial adhesion and colonization. Expression of *oipA* is regulated by the slipped-strand repair mechanism based on the number of CT

dinucleotide repeats in the 5' region of the *oipA* gene in a way such that its switch on would be functional and its switch off would be nonfunctional. This is consistent with the observation that an OipA-positive status has been significantly associated with the presence of duodenal ulceration and gastric cancer, high *H. pylori* density and severe neutrophil infiltration (Yamaoka 2006). OipA on genotype is linked to bacterial virulent determinants, such as functional *vacA*, *babA* and, most strongly, *cagA* genotypes (Dossumbekova et al. 2006).

The biological role of the adhesion process is still controversial. On one hand, adhesive interactions contribute to inflammation, with concomitant expression of sialylated glycans such as sialyl-Le^a and sialyl-Le^X, and seem to be involved in disease progression. In particular, when transgenic mice expressing the Lewis b antigen have been infected with *H. pylori*, the mice have showed increased bacterial attachment, more severe chronic gastritis and parietal cell loss (Guruge et al. 1998). The delivery of the major *H. pylori* virulence factors CagA and VacA are intimately related to adhesion as well, suggesting a role of adhesion in the delivery of toxins (Amieva and El-Omar 2008). Attachment of *H. pylori* to the gastric mucosa activates the type IV secretion system which results in the translocation of CagA protein into the host cells and triggers inflammation (Odenbreit et al. 2000). Some investigators have speculated that inflammation could have an adaptive role, since it promotes the release of nutrients into the gastric lumen. On the other hand, adhesion seems necessary to avoid mechanical clearance, such as liquid flow, peristaltic movements or shedding of the mucous layer, and to allow the bacterium to hold firmly to its environment.

1.5 PATHOGENESIS AND VIRULENCE FACTORS

After stomach colonization, *H. pylori* is able to persist for many decades and invoke an intense inflammatory response, leading to damaged host cells. In most of the time its presence is asymptomatic; but it is sometime associated with an increased risk of developing various gastrointestinal diseases, including non-ulcer dyspepsia, peptic and duodenal ulceration, gastric cancer, mucosal-associated lymphoid tissue (MALT) lymphoma (Kusters, van Vliet, and Kuipers 2006). The ability of *H. pylori* to survive the acid conditions and colonize the stomach is essential for the development of virulence, defined as the degree of pathogenicity of a microorganism; but the latter does not necessarily occur. The pathogenesis of *H. Pylori* infection in humans can be divided into three steps (McGee and Mobley 2000):

- entry to, adherence to, and colonization of the human gastric mucosa;
- avoidance, subversion, or exploitation of the human immune system;
- multiplication, tissue damage, and transmission to a new susceptible host or spread to adjacent tissues.

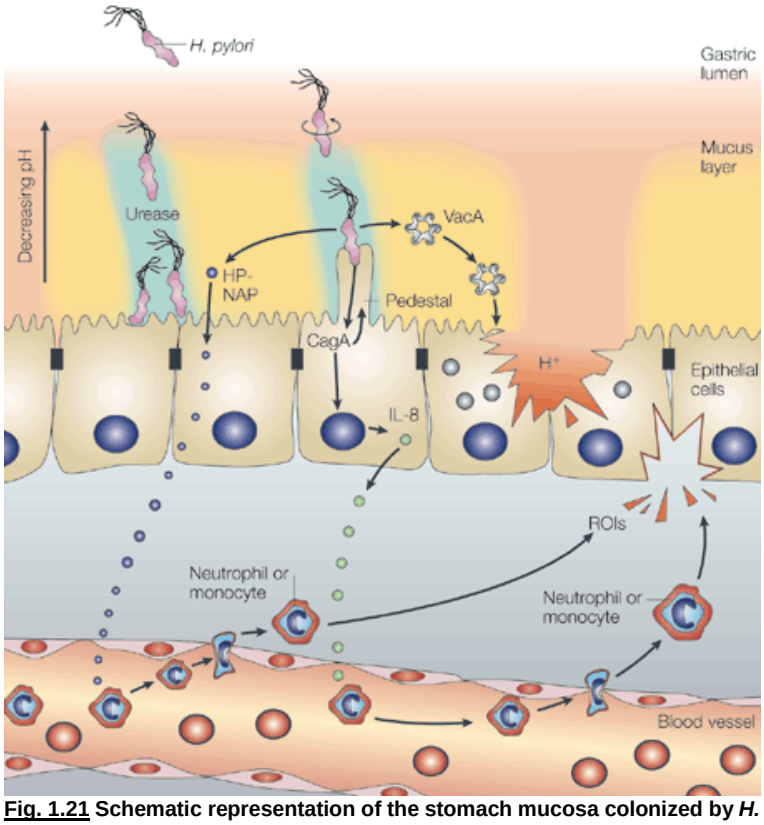

Fig. 1.21 Schematic representation of the stomach mucosa colonized by *H. pylori*, showing the main virulence factors involved in colonization and disease. Adapted from Montecucco and Rappuoli, 2001.

The success of *H. pylori* in causing such a wide spectrum of diseases depends on host, bacterial, and environmental factors. A complex and fascinating balance between these factors takes part in the gastric niche, allowing *H. pylori* to switch between commensalism and pathogenicity. On one hand this network allows the majority of infected individuals to be without any symptom during their entire life; on the other hand it determines the different clinical outcomes (D'Elios and Andersen 2007; Larussa et al. 2015).

The host immune response during *H. pylori* infection plays an important role in the persistence of infection and the pathogenesis. The bacterium is able to escape host defence mechanisms involving both the innate and adaptive immune systems of the host, reducing the inflammatory response in its favor; consequently, it establishes a chronic infection by achieving a delicate balance between inducing immune responses and surviving in the inflammatory condition by using an array of important virulence factors. Both the bacterium and the host adapt to the other in the form of a long-standing dynamic equilibrium. One of the mechanisms employed by *H. pylori* for immune evasion and immune modulation is the ability to adapt to changing environmental conditions, in the gastric epithelium, acidity and nutrient availability, during long-term colonization, made possible by an unusual genetic variability. Particularly, the high prevalence and wide distribution of clusters of strain-specific genes, namely hypervariable plasticity zones, in the genomic sequence throughout several *H. pylori* isolates might favor the emergence of variants after selective pressure and provide a fitness benefit to their hosts for increased persistence (Fischer et al. 2014). *H. pylori* bacteria are also highly competent for uptake of DNA from other *H. pylori* strains. The analysis of genomic sequences shows strong evidence of recombination between strains, to the degree that clonal lineages are largely obscured (Falush et al. 2001). Substantial intragenomic recombination occurs, based largely on the presence of repetitive DNA sequences; their presence allows high-frequency deletion and duplication, including slipped-strand mispairing (Aras et al. 2003). The local selection within an individual stomach also may promote the genetic variability; separate gastric microniches are likely colonized by *H. pylori* subpopulations that have particular attributes to maximize fitness, for example, ligand specificity for local receptors (Ilver et al. 1998).

Following *H. pylori* acquisition, there is a rapid host recognition in the form of both innate and adaptive immune responses; it results in strong specific local and systemic antibody production, cell-mediated immunity as well as an inflammatory infiltrate of neutrophils, lymphocytes, plasma cells, macrophages and eosinophils in the gastric mucosa (Akhiani et al. 2002). Once chronicity is established, the immune stimulation appears remarkably constant, consistent with a model of dynamic equilibrium. The ubiquity and duration of host recognition of *H. pylori* and yet the lifelong colonization by the bacterium demonstrate the effectiveness of bacterial strategies to evade the host immune system. The important first step is to survive without tissue invasion; most of the *H. pylori* bacteria reside in the gastric lumen, beyond the reach of most host immune recognition and effector

mechanisms. However, even in this niche, some *H. pylori* bacteria establish intimate contact with the epithelial surface, some *H. pylori* proteins cross the epithelial barrier, and both innate and adaptive immune systems are activated (Blaser and Atherton 2004).

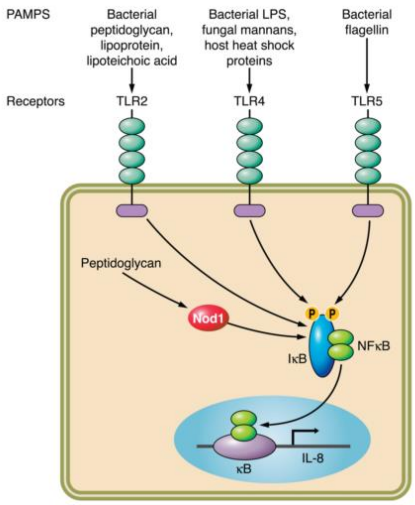

Fig. 1.22 Toll-like receptors and pathogen recognition. Adapted from Peek et al., 2010.

Innate immune recognition involves Toll-like receptors (TLRs), which are present on the surface of gastric epithelial cells and belong to a conserved family of eukaryotic receptors recognizing invariant regions of bacterial molecules termed pathogen-associated molecular patterns (PAMPs). Although the bacterial ligands for TLRs are distinct, the signaling pathways activated by these receptors all result in NF-κB activation and pro-inflammatory cytokine expression TLR (Fig. 1.22). If bacteria invade and penetrate the gastric epithelial barrier, the alternate pathway of complement is activated and invading bacteria encounter macrophages and neutrophils. Since most of *H. pylori* microorganisms are localized within the gastric mucus layer and do not invade gastric tissue, contact between *H. pylori* and phagocytic cells probably occurs infrequently unless there are disruptions in the gastric epithelial barrier (Algood and Cover 2006). Furthermore, *H. pylori* is able to elude the recognition by the TLRs; to avoid the activation of this system the bacterium modulates its surface molecules, including lipopolysaccharide (LPS) and flagellin. TLR5 recognizes bacterial flagella such as those of *Salmonella typhimurium*, but it is not stimulated by *H. pylori* flagella, since *H. pylori* flagellin is not recognized by this receptor. As can be evidenced, *H. pylori* flagellin, especially FlaA, is not released and noninflammatory, since it is unable to induce a pro-inflammatory state by promoting IL-8 secretion (Gewirtz et al. 2004). TLR9 recognizes the largely unmethylated DNA of most bacteria, but the highly methylated *H. pylori* DNA likely minimizes this recognition (Smith 2014). Lipopolysaccharide (LPS) is a surface exposed glycolipid found on the outer membrane of Gram-negative bacteria and it acts as a potent signal for development of an inflammatory response by interaction with TLR4 on gastric epithelial cells and macrophages. It has three distinct units: lipid A, which is responsible for the toxic effects; a core polysaccharide of five sugars linked through ketodeoxyoctulonate to lipid A; and the O-antigen, an outer polysaccharide consisting of up to 25 repeating units of three to five sugars (Lina et al. 2014). In comparison to LPS from *Escherichia coli* or other Gram-negative bacteria, *H. pylori* LPS has approximately

500-fold lower endotoxic activity (Muotiala et al. 1992). The low biological activity of *H. pylori* LPS is attributable to modifications of its lipid A component, especially regarding its phosphorylation pattern and acylation, in order to mask the negatively charged phosphate groups present on the lipid A disaccharide backbone (Cullen et al. 2011). Consequently, *H. pylori* LPS poorly binds to TLR4 and its ability to stimulate macrophage production of pro-inflammatory cytokines, nitric oxide and prostaglandins is decreased, preventing the contribution to innate immune response. Interestingly, *H. pylori* LPS has also been shown to possess antiphagocytic properties *in vitro*, since LPS may mediate direct bacteria/phagocyte interactions and it may also regulate antibacterial activity of the phagocytes (Grebowska et al. 2008). Moreover, *H. pylori* LPS is important not only for the activation of TLR4, but also because the bacterium commonly expresses LPS O-antigens that are structurally related to Lewis blood group antigens found on human cells. Particularly, *H. pylori* O-antigen is composed of N-acetyl-D-glucosamine, L-fucose and D-galactose to form fucosylated oligosaccharide structures that show a clear tendency to mimic the structural motif of human antigens Lewis X (Le^X) and Y (Le^Y) and, thus, are less immunogenic than those of many other enteric bacteria (Aspinall and Monteiro 1996). This similarity in structure between *H. pylori* LPS and Lewis blood group antigens may represent a form of molecular mimicry or immune tolerance that permits *H. pylori* LPS antigens to be shielded from immune recognition because of similarity to “self” antigens (Algood and Cover 2006).

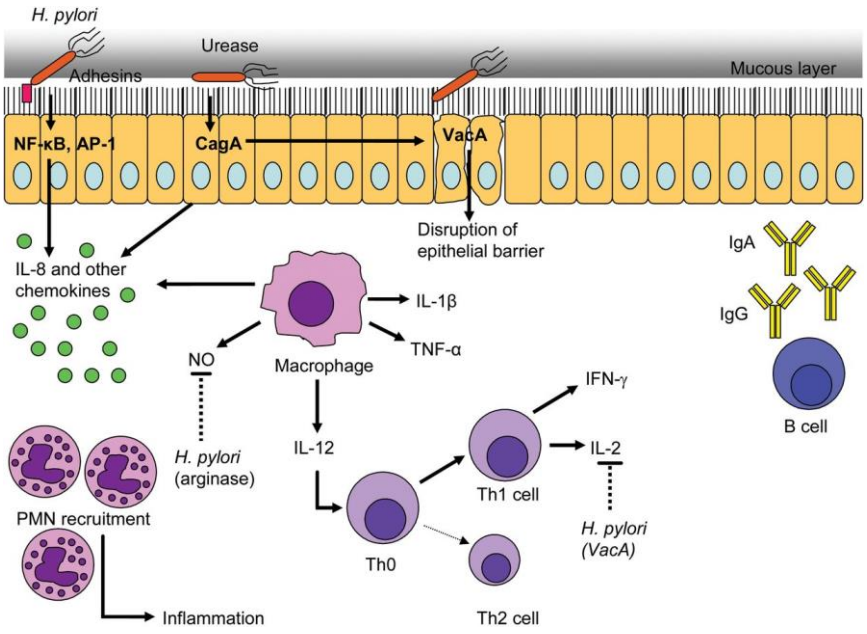

Fig. 1. 23 Inflammatory response to *H. pylori*. Adapted from Portal-Chelay et al., 2006.

H. pylori also activates the adaptive immune system, as indicated by both humoral (B lymphocytes) and cellular (T lymphocytes) recognition of its antigens, although it has evolved to substantially downregulate and avoid acquired immune effectors (Fig. 1.23). The lines between adaptive and innate immunity are frequently blurred by the close interactions between pathways, such that stimulation of antigen-presenting cells (APCs),

such as macrophages and dendritic cells (DCs), leads to activation and recruitment of lymphocytes, particularly $CD4^+$ T cells, and the development of T-helper (Th) cell-specific responses. Classically, differentiation of Th cells involves clonal expansion that is caused by engagement of the T cell receptor. Th cells can differentiate from $CD4^+$ T cells to 2 major functional classes, namely Th1 cells, which produce a set of cytokines that include interferon-γ (IFN-γ) and interleukin IL-2, and Th2 cells, which produce cytokines such as IL-4, IL-5, IL-10, and IL-13. Th1 cells generate cell-mediated immunity, which is important in protection against intracellular microorganisms; whereas Th2 response is associated with humoral immunity and protection against extracellular pathogens (Wilson and Crabtree 2007). Based on the fact that *H. pylori* is non-invasive and that infection is accompanied by an exuberant humoral response, one might predict that a Th2 response would be predominant within *H. pylori* colonized gastric mucosa. Paradoxically, the majority of *H. pylori* antigen-specific T cell clones isolated from infected gastric mucosa produce higher levels of IFN-γ than IL-4, which is reflective of a Th1 response (Bamford et al. 1998). *H. pylori* also stimulates production of IL-12 *in vitro*, a cytokine that promotes Th1 differentiation. These findings raise the hypothesis that an aberrant host response (Th1) to an organism predicted to induce secretory immune responses (Th2) may influence and perpetuate gastric inflammation (Fig. 1.23; Portal-Celhay and Perez-Perez 2006). The reasons for the impaired Th1 immune response could lie in the continuous process of virulence factor elaboration implemented by *H. pylori* over the thousands of years of coexistence with the human host (Larussa et al. 2015). A key role has recently been ascribed to the natural regulatory T cells (Treg); although widely acknowledged to play a role in the maintenance of self-tolerance, recent studies indicate that Treg can be activated and expanded against bacterial antigens. Recent evidence suggests that the activation of regulatory T cells might result in decreased pathological responses and prolonged persistence of infection as a mechanism for the maintenance of pathogen-specific immunological memory, reducing the activation of IFN-γ-producing $CD4^+$ T cells, even at the expense of a higher *H. pylori* load in the gastric mucosa (Fig. 1.24; Raghavan

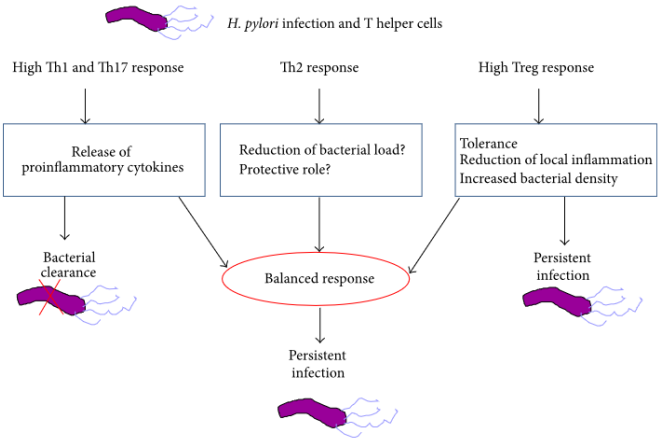

Fig. 1.24 Interplay between *H. pylori*, the effective T helper lymphocytes and the regulatory T cells. Adapted from Larussa et al., 2015.

et al. 2003; Raghavan and Holmgren 2005). *H. pylori* is able to subvert not only the innate, but also the adaptive immune response. Recognition by this immune system requires antigen presentation. Particularly, the proliferation of CD4+ T cells is triggered by recognition of antigenic epitopes bound to a protein receptor, named major histocompatibility complex (MHC) class II, exposed on the surface of antigen-presenting cells (APCs); antigen presentation by APCs plays an essential role in the initiation of adaptive immune responses. *H. pylori* interferes with both uptake and processing of antigens, partially through an effect induced by the toxin VacA. The latter specifically inhibits antigen processing by interfering with late endocytic membrane trafficking by antigen-presenting cells (APCs); this in turn decrease the proliferation of autologous human $CD4^+$ T cells triggered by recognition of antigenic epitopes bound to MHC class II molecules exposed on APC surfaces. Specifically, VacA interferes with the proteolytic processing of antigens and the generation of T cell epitopes loaded on newly synthesized MHC class II molecules (the Ii-dependent pathway of antigen presentation), leaving unaffected generation and presentation of epitopes by class II molecules that recycle through early endosomal compartments (Molinari et al. 1998). Furthermore, the virulence factor VacA acts as an immunomodulatory, since it efficiently inhibits signaling and proliferation of T cells by inducing a G1/S cell cycle arrest, through the interference with the T cell receptor/IL-2 signaling pathway at the level of the Ca^{2+}-calmodulin-dependent phosphatase calcineurin. In this way, VacA avoids the nuclear translocation of nuclear factor of activated T cells (NFAT), the main regulator of the T cell pathway (Gebert et al. 2003). A low-molecular-weight protein, distinct from VacA, has been reported to inhibit proliferation of T cell by blocking cell cycle progression at the G1 phase through G1 cyclin-dependent kinase activity modulation. This suppression factor has been identified as the γ-glutamyltranspeptidase (Ggt); as can be noticed, this enzyme mediates the extracellular cleavage of glutathione, with ROS production and consequently induction of cell cycle arrest in lymphocytes (Schmees et al. 2007). Additionally, among the enzymes involved in nitrogen metabolism, arginase is involved in immune evasion and, specifically, impairs the T cell function. This enzyme is important for urea production by hydrolyzing L-arginine to urea and ornithine. L-arginine is required for T cell activation and function; therefore arginase has been observed to cause a significant decrease in T cell proliferation by depleting L-arginine availability. The results has appeared not to be mediated by apoptosis, but rather correlated with a reduced expression of the chief signal transduction

protein CD3ζ-chain of the T cell receptor (TCR), which is required for the initiation of T cell activation (Zabaleta et al. 2004).

The relative contributions of the different host immune modulation and evasion strategies to *H. pylori* persistence are not precisely established, possibly differing in individual hosts; but the existence of these various mechanisms implies that immune surveillance of the gastric lumen is powerful and that bacterial survival requires its subversion, as fundamental requirement for pathogenesis. A strong correlation exists between gastric infiltration by neutrophils, mucosal damage and development of gastroduodenal diseases in *H. pylori* infection (D'Elios, Montecucco, and de Bernard 2007). Several virulence factors contribute to these outcomes, but the most important and most characterized ones include the Vacuolating cytotoxin A (VacA), the *cag* pathogenicity island (*cag*PAI) and its effector, namely cytotoxin-associated gene A (CagA), and the neutrophil-activating protein (HP-NAP). In recent years, more than hundred different toxins produced by various bacterial species have been characterized. Their cellular effects lead to "intoxication" of the host cell and range from cell death to a variety of non-lethal changes, including permeabilization of membranes, blockade of exocytosis, and alterations of cellular signal transduction, cellular cytoskeletal properties and cellular proliferation (Cover and Blanke 2005). Even *H. pylori* toxins affect a multitude of host cellular pathways and they have also polymorphic nature. These combined evidences can easily contribute to differences in disease outcomes and severity.

1.5.1 VACUOLATING CYTOTOXIN A

The vacuolating cytotoxin A (VacA) is one of the most versatile and most characterized virulence factors produced by *H. pylori*. VacA is a secreted, pore-forming toxin, that was identified when the supernatants of *H. pylori* broth cultures have been found to induce large intracellular vacuoles in cultured mammalian cells (Fig. 1.25; Leunk et al. 1988). VacA is present nearly in all *H. pylori* strains, defining *VacA*+ and *VacA*- populations, and its presence is epidemiologically associated with tissue damage and disease (Tomb et al. 1997). Particularly, VacA triggers intrinsic apoptosis, increases mitogen-activated protein kinases, induces autophagy and cell death, and alters immune cell activity (Cid et al. 2013). The aminoacid sequence of VacA does not show a significant similarity to any other known bacterial or eukaryotic protein. As most of the *H. pylori* factors, VacA presents an important level of polymorphism and is characterized by a considerable

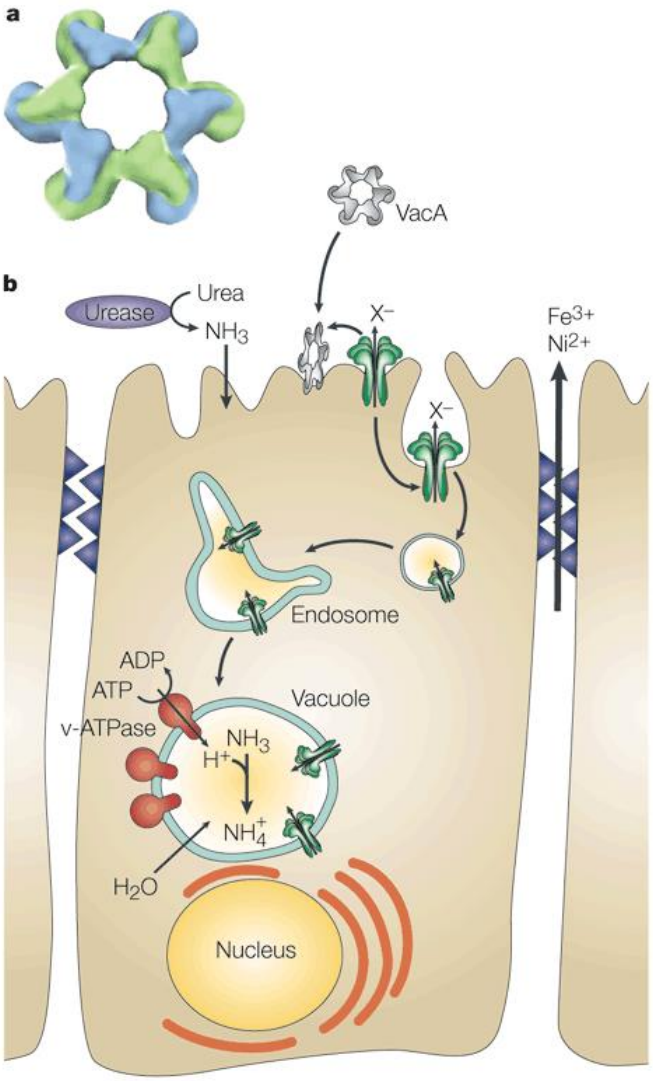

Fig. 1.25 Vacuolating cytotoxin A; a: oligomeric structure of VacA deduced by electron microscopy; b: current model of the cellular alterations induced by VacA cytotxin. Adapted from Montecucco and Rappuoli, 2001.

genetic variability. Genetic variation at this locus could be under strong selection as *H. pylori* adapts to the host immune response, colonizes new human hosts, or inhabits different host environments. Moreover, expression levels, cytotoxicity, disease severity are linked to this sequence variation in different domains of VacA. The toxin is firstly expressed as a 140 kDa proenzyme that is trimmed at both ends during secretion from the bacterial cell. It has been supposed to be secreted by a type Va secretion system (T5aSS) which is SecA-dependent (Kim et al. 2014). The N-terminal domain contains a signal sequence ("s" region of the gene) that shows allelic variability and has been classified into different types, mainly s1 and s2 variants (Fig. 1.26). Strains harboring s1 types of VacA secrete active toxin and are also more highly associated with both ulcers and gastric cancer; instead s2 subtype strains are associated with less inflammation and lower ulcer prevalence (Atherton et al. 1995). The N-terminal signal sequence allows the export of the toxin throughout the bacterial inner membrane and it is hydrolyzed during this process. This product is anchored to the outer membrane by its C-terminal region; it could undergo a subsequent translocation out of the external *H. pylori* envelope and be consequently removed by the mature toxin (Fig. 1.26). The middle region of the VacA gene is classified as the "m" region, which also shows allelic variation, with m1 types having stronger vacuolating activity (Amieva and El-Omar 2008). Among the allowed combinations of these alleles, s1/m1 VacA protein is the most interactive and toxigenic product, while s2/m2 the mildest one. Because nearly all strains are positive for VacA, genotypes of *vacA* are documented as the critical determinant of

pathogenesis, rather than its presence or absence, and s1/m1 genotypes are associated to an increased risk of acquiring peptic and duodenal ulcer. Further allelic variation is attributable to other two region, particularly the intermediate or “i” region, between the “s” region and the “m” region (González-Rivera et al. 2012), and the deletion or “d” region, between the “i” region and the “m” region (Ogiwara et al. 2009). Despite several attempts to relate these subtypes to outcomes or pathogenesis, no truly consistent associations have been found so far.

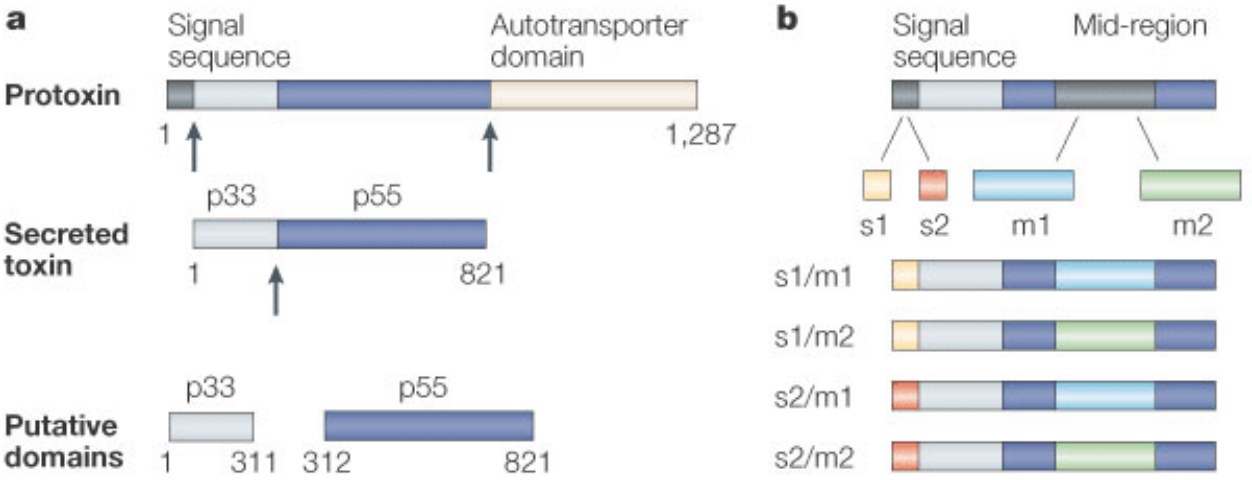

Fig. 1.26 VacA functional domains; a: the amino-terminal signal sequence and carboxy-terminal domain are cleaved from the 140-kDa VacA protoxin to yield an 88-kDa mature toxin that is secreted into the extracellular space via an autotransporter mechanism. A 33-kDa C-terminal beta-barrel domain of VacA is predicted to insert into the outer membrane and form a channel, through which the mature VacA toxin is secreted. Arrows indicate sites of proteolytic cleavage. **b:** there is a high level of diversity among *vacA* alleles from different *H. pylori* strains. Allelic diversity is particularly striking near the 5' terminus of vacA (the s-region) and in the mid-region of the gene (the m-region). Two main families of s-region sequences (s1 and s2) and two main families of m-region sequences (m1 and m2) have been described. Type s1 and s2 VacA proteins differ in the site at which N-terminal signal sequences are cleaved, and consequently, the mature, secreted type s2 VacA toxin contains a 12-amino-acid hydrophilic extension at its N-terminus that is absent from type s1 VacA toxins. VacA toxins that contain this 12-amino-acid hydrophilic extension fail to induce cell vacuolation in vitro. The amino acid sequences of type m1 and m2 VacA proteins are approximately 65% identical within a region comprising 250 aminoacids. Subtypes of vacA alleles (for example, s1a, s1b, s1c, m1a, m1b, m2a and m2b) have been described, and certain subtypes have a geographically restricted distribution. Mosaic forms of vacA are thought to arise via homologous recombination among vacA alleles from different strains. Adapted from Cover and Blanke, 2005.

The mature VacA protein has a molecular mass of 88 kDa, but it can be further nicked into 2 smaller peptides that remain non-covalently associated (Cover et al. 1994). Upon secretion, approximately 50% of the toxin remains associated with the bacterial cell surface, and the rest is released and processed again by unidentified proteases, to give the mature enzyme (Dundon et al. 2000). Interestingly, the surface-associated VacA molecules are still biologically functional and are delivered to the host cells by direct contact between adhered bacteria and the epithelial cell membrane, followed by uptake and intoxication (Ilver et al. 2004). On the other hand, the secreted protein in the supernatant has a strong tendency to oligomerize into water-soluble molecular complexes resembling flowers with six-fold or seven-fold radial symmetry (Figs. 1.25 and 1.27; Lupetti et al. 1996; Reyrat et al. 1999). It is not clear whether these secreted VacA

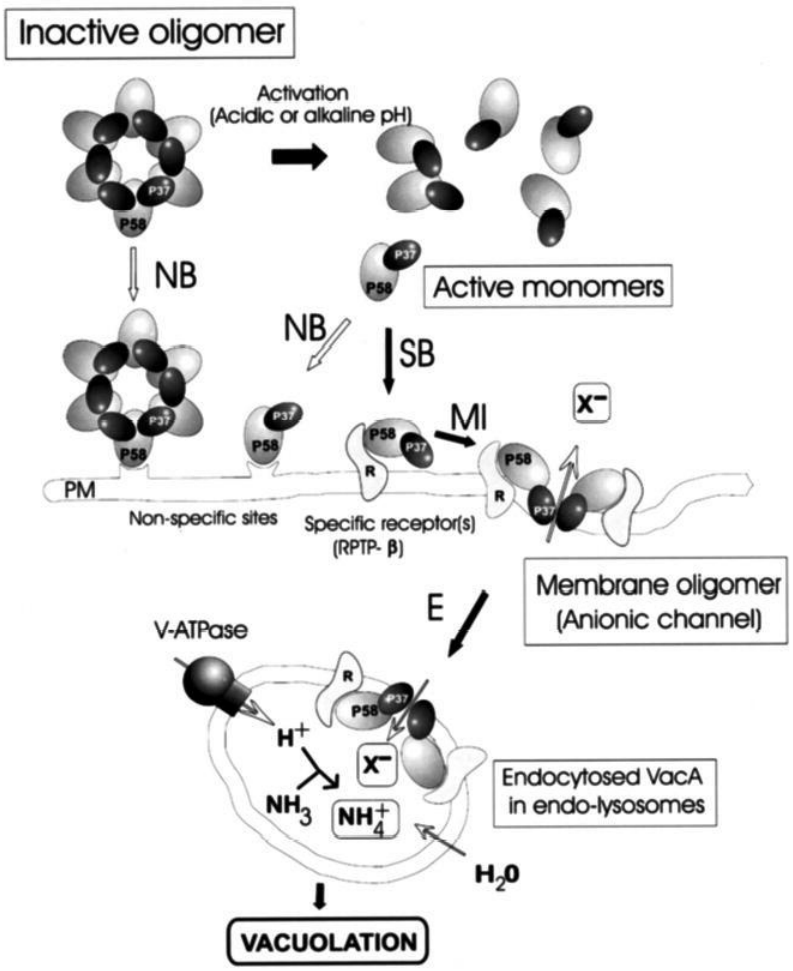

Fig. 1.27 The inactive, soluble oligomer can be disrupted by acidic or alkaline treatment to release active monomeric VacA. Hydrophobic VacA monomers can insert into the plasma membrane, where they associate to form a pore. (NB: non-specific binding; SB: specific binding; MI: membrane interaction; PM: plasma membrane; E: endocytosis; X⁻: anion; R: receptor. Adapted from Mobley, 2001.

complexes are toxigenic because they have little vacuolating activity unless they are dissociated by high or low pH into the monomeric forms that can be taken up by cell membranes. This pH-mediated disassembly is associated with a marked increase in VacA cytotoxic activity. Specifically, an acid pH is necessary for release of the toxin in the environment, dissociation into monomers and activation; but studies *in vitro* have proved that even alkaline pH values could activate VacA toxin. Furthermore, this protein presents unusual structural resistance to extremely acid pH, pepsin digestion at pH 2, and thermal treatments (de Bernard et al. 1995). Following exposure to either acid or alkaline conditions, the oligomeric complexes dissociate in the monomeric components, up to 12 teardrop-shaped subunits, demonstrating that the oligomer consists of 12 subunits of 88 kDa assembled into two interlocked six-membered arrays. VacA monomers are firstly able to interact with cells and their subsequent oligomerization contributes to the ability of the toxin to form anion-selective membrane channels into planar lipid bilayer. Therefore, VacA cytotoxicity requires an initial activation by interaction of the monomeric forms with cells and the flower-like ring oligomer is the conformation it adopts when inserted into membranes, hence its pore-forming properties (Cover, Hanson, and Heuser 1997). Moreover, the VacA monomer of 88 kDa consists of two putative domains and can undergo to limited proteolytic cleavage to yield two fragments, namely p33 and p55, that are located at the N-terminus and the C-terminus, respectively, and remain non-covalently associated (Fig. 1.26). Both the p33 and p55 domains contribute to the binding

and internalization of VacA and that both domains are required for vacuolating cytotoxic activity. Especially, aminoacid sequences located within a hydrophobic region near the N-terminus of the p33 domain are required for the formation of anion-selective membrane channels; on the other hand, p55 domain is the main responsible for VacA binding to host cells, requiring also the contribution of p33 fragment (Torres et al. 2005). Analysis of the X-ray structure of p55 domain (Gangwer et al. 2007; PDB accession code: 2QV3) together with electron microscopy images of the oligomers has led to a model of VacA oligomer organization in which the p33 domain occupies the inner core of the complex and the p55 domain extends outward from the central core, enabling the formation of a stable anionic membrane channel only when both subunits are present (Fig. 1.25; Zeevaart et al. 2009). Near the N-terminus of the p33 domain there is the only strongly hydrophobic region of VacA containing three tandem GXXXG motifs, which are characteristic of transmembrane dimerization sequences. Mutagenesis on several residues within the N-terminal unique hydrophobic region abolishes the ability of VacA to form membrane channels in planar lipid bilayers and also abolishes vacuolating cytotoxic activity. Therefore these results indicate that an intact proline residue and an intact GXXXG motif within the N-terminal hydrophobic region of VacA are essential for membrane channel formation, and they also provide strong evidence that membrane channel formation is required for VacA-induced cell vacuolation (McClain et al. 2003; Cover and Blanke 2005).

The acid-activated VacA binds to the host cells probably by a receptor protein-tyrosine phosphatase β (RPTPβ), which regulates intracellular tyrosine phosphorylation; subsequently the toxin is internalized into the host cell by receptor-mediated endocytosis and it is localized in multiple intracellular sites, including endosomal compartments, the large intracellular vacuoles that form as a consequence of VacA intoxication, and the inner mitochondrial membrane (Garner 1996; Yahiro et al. 1999). Many of the toxigenic effects of VacA, both in epithelial cells and T lymphocytes, can be attributed to the ability of this toxin to be inserted into membranes and to form anion-selective channels (Cover and Blanke 2005). Among its several effects, VacA forms pores in the host cell membranes in order to the release of anions and urea, which are essential for urease enzyme activity (Figs. 1.25 and 1.27). The induction of intracellular vacuoles was the first characterized action of VacA. It is believed that VacA inserts into the membranes of late endosomal vesicles, forms pores with chloride channel activity, alters the composition of anions within endosomes, and subsequently leads to osmotic swelling (Amieva and El-

Omar 2008). Following, VacA induces the formation of large acid vacuoles in the cytoplasm of certain lines of host cells, growing from the perinuclear area to fulfill the entire cytosol and leading eventually to necrosis of the eukaryotic cell (Figura et al. 1989). The formation of VacA-induced vacuoles requires the full activity of the vacuolar-type ATPase proton pump (V-ATPase) and the presence of weak bases, suggesting that vacuoles are derived from the accumulation of membrane-permeable weak bases within acidic compartments followed by water influx and swelling (Cover, Reddy, and Blaser 1993). Alterations in endosomal function in the gastric mucosa could have many effects on the epithelium. A proper function of endosomal trafficking is necessary for antigen presentation process, which is shown to be perturbed by VacA *in vitro*, promoting the host immune evasion and contributing to the persistence of *H. pylori* infection, as previously described (Molinari et al. 1998). Moreover, VacA has been observed to impair the transport of acidic hydrolases to lysosomes in HeLa cells, resulting in release of these enzymes into the extracellular medium and reducing the degradative power of lysosomes (Satin et al. 1997). In addition to its effects on endosomes, VacA also induces host-cell death through apoptosis, through pore formation in mitochondrial membranes and also indirectly through the activation of proapoptotic signaling molecules. A third reported effect of VacA is its ability to cause leakage of ions and small molecules, such as iron, nickel, sugars, and aminoacids, by disrupting the barrier function of tight junctions, without major disruptions in junction integrity. This could be a mechanism by which *H. pylori* acquires nutrients across an intact epithelial barrier (Amieva and El-Omar 2008). VacA plays a further role in pathogenesis interfering with the autophagy process, fundamental to protect against bacterial infection. Prolonged exposure to the toxin disrupts autophagy, by preventing maturation of the autolysosome. It has been observed that VacA-suppressed autophagy facilitates intracellular survival and persistence of the pathogen, since the toxin assists in producing the vacuole where *H. pylori* can survive intracellularly (Raju et al. 2012).

1.5.2 *CAG* PATHOGENICITY ISLAND AND CYTOTOXIN-ASSOCIATED GENE A

The cytotoxin-associated gene *pathogenicity island (*cag*PAI) is widely known as the most important pathogenic factor of* H. pylori *that carries an increased risk for gastric cancer. It consists in a 40-kb region of horizontally acquired DNA inserted into the genome of the more virulent* H. pylori *strains, namely type I strains, and it contains approximately 30*

genes encoding a type IV secretion system (T4SS) apparatus, capable of delivering its

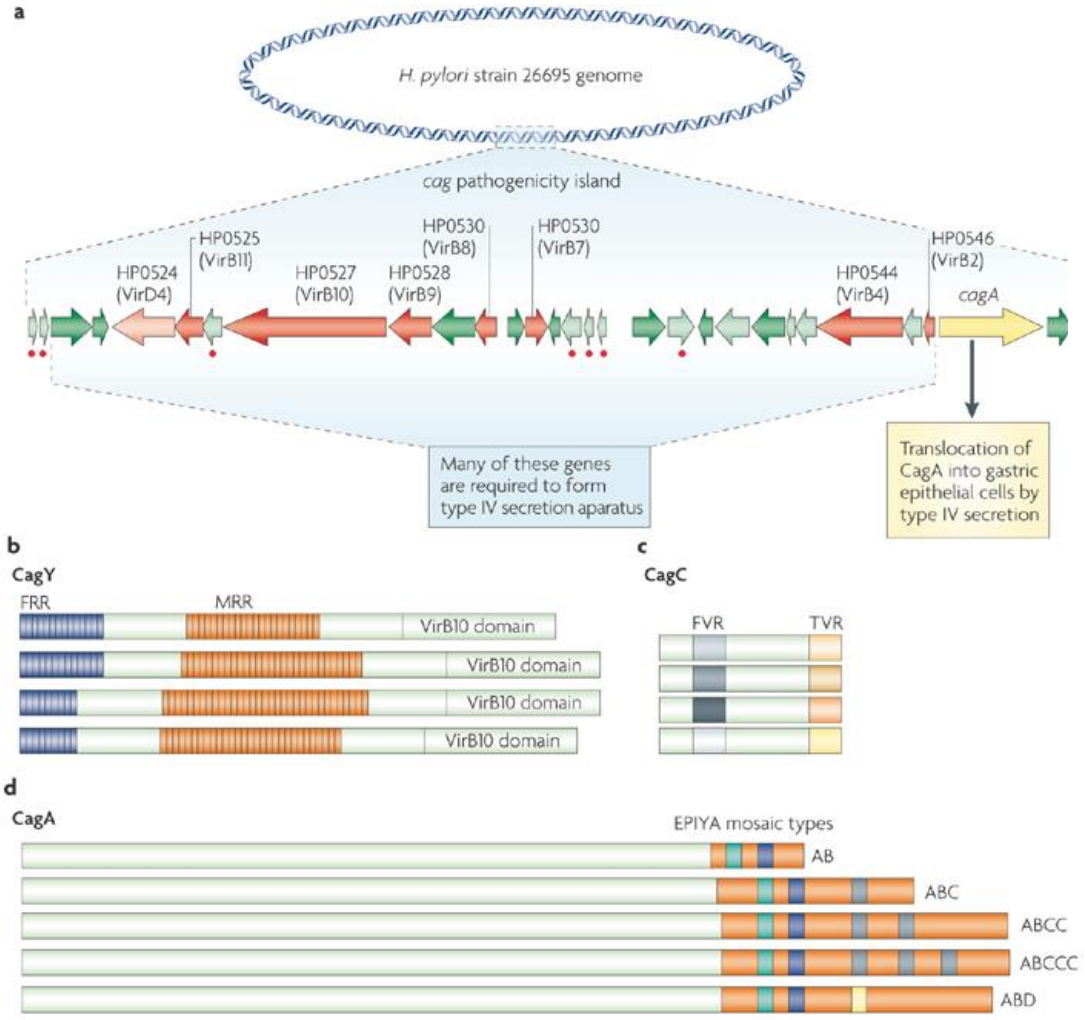

Fig 1.28 ***cag*** **pathogenicity island (*cag*PAI); a:** arrangement of cagPAI genes in *H. pylori* strain 26695; **b:** extensive sequence variation in some Cag proteins; exposure of Cag proteins to the host presumably places them under strong positive selection in vivo.

effector protein cytotoxin-associated gene A (CagA) and a bacterial cell wall component, peptidoglycan, into host cells (Fig. 1.28). The *H. pylori* PAI was originally named *cag* (cytotoxin-associated gene) since the toxin CagA, encoded by the *cag*PAI, has been found to be associated with the vacuolating cytotoxin VacA and their concomitance has been thought correlating (Censini et al. 1996). However, it has been later shown that both factors, VacA and *cag*PAI, are independent of each other. On the genetic level, the *cag*PAI carries from 27 to 31 genes and is flanked by 31-bp direct repeat, which contains the recombination site and corresponds to the last nucleotides of the glutamate racemase gene (*glr*), within which *cag*PAI has been integrated (Fig. 1.29; Censini et al. 1996). The module also forms the core of the left and right ends of an insertion sequence common in *H. pylori*, the IS605 element, whose highest number of copies is statistically related with *cag*PAI-containing strains. Depending on the isolate, different numbers of insertion sequences are associated with *cag*PAI. More often the module is split into two clusters

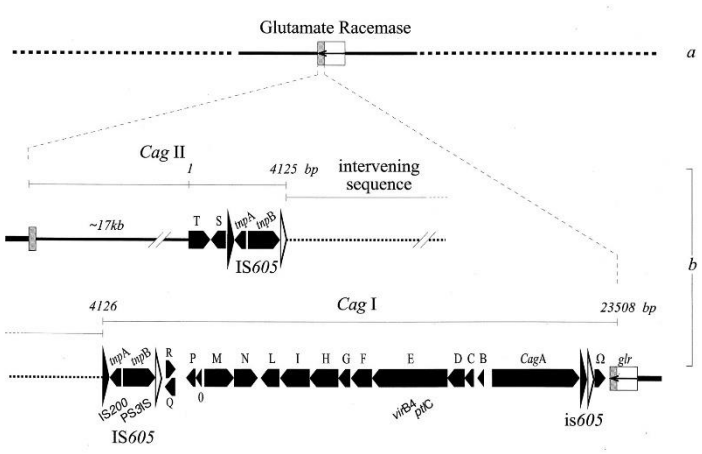

Fig. 1.28 Location and splitting of *cag* pathogenicity island (*cag*PAI). Adapted from Censini et al., 1996.

by a long independent intervening genome sequence, flanked by two IS605 elements; the subregions are named *cagI* and *cagII*, located downstream and upstream this interruption and containing at least 14 and 16 open reading frames, respectively (Fig. 1.29; Bukanov and Berg 1994). Since intervening genome sequence is present between the two insertion sequences, IS605 elements generate *H. pylori* strains with different levels of virulence; especially, strains with many insertions often resemble type II strains (less virulent) more than type I strains (more virulent, *cag*PAI-containing), because the integrity of all the components of the T4SS is needed for translocation of the virulence effector CagA. Additional genetic instability of *cag*PAI results from deletions and inversions. In some cases, the entire module can even be lost completely due to DNA transfer of an empty site from a *cag*PAI-deficient strain into a type I strain and subsequent homologous recombination (D Kersulyte, Chalkauskas, and Berg 1999). *H. pylori cag*PAI is characterized by a different nucleotide composition with respect to the overall bacterial genome, showing a lower G+C content, and is flanked by transposable elements, suggesting an acquisition by horizontal transfer of genetic material (Censini et al. 1996; Tomb et al. 1997). Foreign and rearranged DNA sequences are often responsible for the acquisition of genes involved in virulence. It can be hypothesized that early bacteria communities originating from crop plants, animals or rodent pests rampant in the vicinity of early human societies may have served as donors of some of the virulence cassettes. Such an interspecies gene transfer could be explained partly based on the fact that many constituent genes of the *cag*PAI reveal well-established homologies to the type IV systems of *Agrobacterium tumifaciens* and that *cagA*-like sequences have been reported from some *Aeromonas* isolates, obtained from environmental samples (Ahmed, Tenguria, and Nandanwar 2009). After this initial event of horizontal acquisition, the DNA block has undergone a series of rearrangements in different strains, defining subpopulations. IS605 elements might have played a central role in these recombination events (Censini et al. 1996). The presence or absence of *cag*PAI, classifying *H. pylori* populations in *cag*+ and *cag*- respectively, is one of the most striking differences among *H. pylori* strains,

concurring in making more complex the genetic variability of the pathogen. Diversity within *cag*PAI can also be noted between people belonging to different world's regions: only one-half to two-thirds of the isolates from the Western world carry *cag*PAI; in contrast, nearly all East Asian isolates carry this chromosomal module (Kauser et al. 2004).

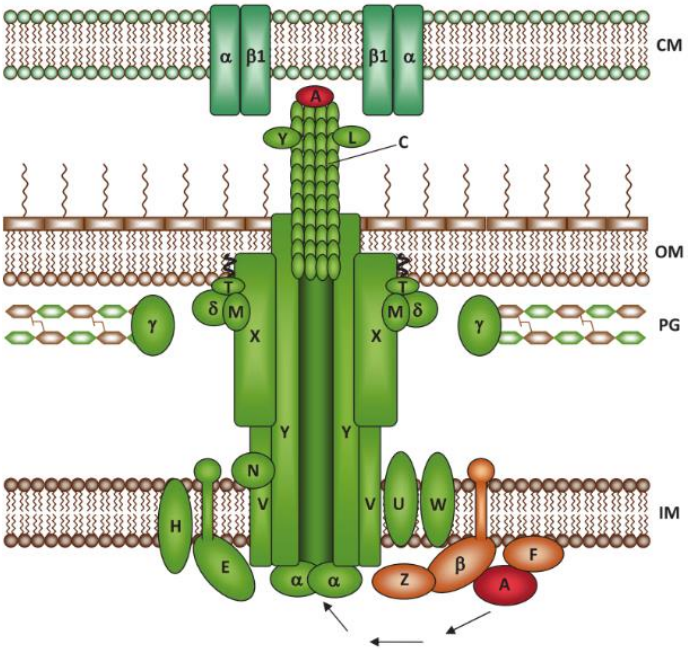

Fig. 1.29 Assembly and interaction model of the Cag type IV secretion system. Adapted from Fischer, 2011.

The pathogenicity island provides *H. pylori* with at least two unique properties: an increased transmission probability and the transformation of what would be an almost commensal into a potential pathogen (Montecucco and Rappuoli 2001). There are approximately 30 open reading frames within the *cag*PAI; one of them encodes the immunodominant antigen CagA, which is localized to the 3' end of the region and does not present any sequence homology to other known proteins, while many of the adjacent genes have significant homology to components of a type IV secretion system (T4SS; Fig. 1.30). Besides the *cag*-T4SS, certain *H. pylori* strains can harbour up to other 3 T4SSs in their genome (*comB*, previously mentioned; *tfs3*; *tfs4*); but the chapter concerns the description of the *cag*-T4SS. This secretion apparatus consists in a macromolecular structure that functions as a minute needle (also named T4SS pilus) for the transfer of bacterial products, both proteins and DNA, from pathogenic bacteria into host cells, and is present in several Gram-negative bacteria ancestrally related to conjugation systems (Cascales and Christie 2003). The best characterized T4SS, namely VirB /D4 machinery, belongs to the plant pathogen *Agrobacterium tumefaciens* and it is regarded as the prototype among that family members. This bacterium injects oncogenic nucleoprotein particles (T-DNA) into host plant cells to induce the formation of a plant tumor or gall. It was thus hypothesized that the *cag*PAI of *H. pylori* could serve as a novel transport system for secretion of virulence factors. T4SS are multicomponent membrane-spanning transport systems ancestrally related to the conjugation processes, which can be responsible for diverse processes

such as DNA transfer, DNA uptake and release, and translocation of proteins that have an effector role in the target cell (Cendron and Zanotti 2011). These transporters are functionally diverse both with respect to delivered substrates (DNA, proteins or DNA-protein complexes) and recipients, which can be either a bacterium of the same or other species, or organisms from a different kingdom like plants, fungi or mammalian cells (Backert and Selbach 2008). T4SS transporters typically are composed of 11 VirB proteins (encoded by *virB1*-*virB11* genes) and the so-called coupling protein, the NTPase VirD4. The proteins comprising this secretion apparatus can be generally categorized into three groups: cytoplasmic or inner membrane proteins, a core complex located in the periplasm or membrane, and a pilus or surface structure that projects beyond the outer membrane (Busler et al. 2006). The VirB proteins from *A. tumefaciens* T4SS can be grouped into these three categories: the core components or putative channel (VirB6-10), the pilus-associated components (VirB2, and possibly VirB3 and VirB5) and the energetic components (the NTPases: VirB4 and VirB11). VirB1 is an enzyme with muraminidase activity possibly enabling localized lysis of murein to achieve T4SS assembling at a given location. In *A. tumefaciens*, signal peptidase-I removes signal peptides from precursors of the main pilus component VirB2 and the minor pilus component VirB5, followed by cyclization of VirB2. Processed VirB2 and VirB5 subsequently associate with the membranes as stabilized by VirB4 and VirB8. Stabilized and properly oriented VirB5 then forms a complex with VirB2, which is a key step in the formation of the T4SS pilus assembly subcomplex (Tegtmeyer, Wessler, and Backert 2013).

Even *H. pylori* T4SS has been characterized in some components; 11 out of the 29 Cag proteins can be ascribed to the secretion machinery itself or have been proposed to represent functional homologues of VirB proteins (Cendron and Zanotti 2011), leading to a T4SS model similar to that of *A. tumefaciens*. Particularly, only a few *cag* genes encode proteins with clear sequence similarities to known T4SS proteins. Obvious similarities exist only for CagE (to VirB4), CagX (to VirB9), CagY (to VirB10), Cagα (to VirB11) and Cagβ (to VirD4). Nevertheless, protein topology predictions and determinations, localization studies and functional studies suggested that Cagγ (VirB1), CagC (VirB2), CagL (VirB5), CagW (VirB6), CagT (VirB7) and CagV (VirB8) are further VirB homologues (Fischer 2011). Moreover, *cag*PAI encode several auxiliary factors unique to the Cag system. Three further gene products are not absolutely necessary, although their absence results in a reduced efficiency of secretion; thus, these proteins, named supporting components, appear to be involved in properly assembling the secretion

apparatus. An additional group of genes has been shown to be required for CagA translocation; therefore, the encoded gene products are termed CagA translocation factors. Finally, several other *cag*PAI gene products do not appear to have a function for the T4SS assembly and their role has not yet been identified (Fischer 2011). *H. pylori* T4SS powering machinery is composed of three cytoplasmic NTPases, Cagα (VirB11), Cagβ (VirD4) and CagE (VirB4 putative homologue), which supply the energy necessary to assemble the apparatus and secrete CagA. Cagα and CagE have been shown to have ATPase activity, while Cagβ corresponds to the coupling protein VirD4, which binds DNA in *A. tumifaciens*. These enzymes couple NTP hydrolysis to conformational changes; these might in turn be coupled to unfolding or transfer of the substrate. The periplasmic core, namely translocation pore, spans the entire periplasmic space and forms channels in the inner and outer membranes. It is subdivided into two layers: the I layer inserting into the inner membrane and the O layer inserting into the outer membrane. The I layer consists of the N-terminal region of CagX and CagY; the O is formed by the C-terminal domains of CagX and CgaY and the full-length CagT. T4SS pilu is generally composed of two proteins, CagC and CagL. The latter, less abundant, decorates the external part of the appendage formed by CagC and plays a role in adhesion, acting as an anchor of the apparatus to the host cell surface. CagF is a chaperone-like protein crucial for CagA translocation. Although T4SS core complexes are able to form autonomously, they are unlikely to do so constitutively and without a positional preference. Accordingly, Cagγ, a lytic transglycosylase responsible of the peptidoglycan-degrading, and CagV, a bitopic inner membrane protein with features similar to a nucleating factor from *A. tumifaciens* T4SS, are essential components assisting in the secretion apparatus assembly (Terradot and Waksman 2011; Fischer 2011).

The only reported effector molecules injected by the *H. pylori* T4SS are peptidoglycan and CagA, although CagA is an effector protein with multiple functions in target cells. A functional activation of the T4SS apparatus, by interaction with a specific host cell receptor, is required to translocate the effector; host cell integrins have been recently shown to directly interact with the CagL protein. The latter contains an Arg-Gly-Asp (RGD) motif which mediates binding of the pilus to integrin $\alpha_5\beta_1$ and is required for injection of CagA (Kwok et al. 2007). After translocation into host cells, CagA becomes rapidly tyrosine phosphorylated ($CagA^{PY}$) in its C-terminally located Glu-Pro-Ile-Tyr-Ala (EPIYA) motifs by a kinase of the host belonging to the Src family, whose members control cytoskeletal processes, cell proliferation and differentiation in normal cells, but are also

key players in carcinogenesis (Selbach et al. 2002). A second conserved motif found to be required for phosphorylation-independent effects is located adjacent to the EPIYA motifs and has been termed microtubule affinity-regulating kinase inhibitor motif because of its binding to this kinase (Neiši et al. 2010). The C-terminal EPIYA motifs are responsible of allelic polymorphism of CagA protein. It has been reported that a number of *cagA* alleles exist and that variation in the C-terminus of the protein is the major difference between the different alleles. Polymorphisms in the C-terminus occur in the EPIYA motifs and typically involve changes in the aminoacid sequences flanking the five-amino-acid repeat (Fig. 1.28). Four different EPIYA regions (A, B, C and D) have been identified and CagA can be categorized based on the alignment of these different motifs on two types: East Asian-type CagA, circulating in Japan, China, and Korea, includes EPIYA-A, EPIYA-B, and usually only one EPIYA-D; while Western-type CagA, prevalent in Europe, America, Australia, and Africa, includes EPIYA-A, EPIYA-B, and one or multiple (up to 5) EPIYA-C motifs (Jones et al. 2009; Sgouras, Trang, and Yamaoka 2015). This evidence explains the size variability of CagA protein (range, 120–145 kDa). Moreover, CagA motifs EPIYA-A and -B have shown to be much weakly susceptible of phosphorylation compared to their variant -C and -D. The phosphorylation level is proportional to the number of EPIYA-C motifs, and thus, increased motifs numbers enhance the pro-inflammatory and carcinogenic potential of the protein, with higher levels of CagA antibody, more severe degrees of atrophy, and reduced survival in a low pH. Thus, both number and types of the motifs, seem to be related to different level of pathogenicity.

After its tyrosine phosphorylation, CagA remains near the plasma membrane; there, it interacts with a number of host proteins, triggering signals that resemble the activation of receptor-tyrosine kinase growth factors, such as c-met. The consequent signal transduction cascade leads to several alterations in the gastric epithelial cells, including activation of Src homology 2 (SH2) domain-containing tyrosine phosphatase 2, alteration in cell structure and motility, alteration in cell scattering and proliferation. The cellular functions of CagA do not always require its tyrosine phosphorylation. These phosphorylation-independent interactions induce activation of β-catenin and NF-κB signalling, leading to disruption of tight and adherens junctions, perturbation of epithelial cell differentiation and polarity, pro-inflammatory and mitogenic responses (Suzuki et al. 2009; Cover and Blaser 2009). As result of its cellular functions, CagA is responsible of evident morphological changes in the host tissues, leading to cell elongation and

formation of filopodia and lamellipodia (“hummingbird” phenotype). According to these various morphogenetic changes that lead to malignant transformation, CagA is defined as a bacterial oncoprotein (Hatakeyama 2004). Furthermore, early reports have indicated that genes encoding components of the T4SS but not CagA itself are required for the induction of pro-inflammatory signalling, including activation of NF-kB, AP-1 and the proto-oncogenes c-Fos and c-Jun. This has suggested that the T4SS might inject factors in addition to CagA (Backert and Selbach 2008). It has been demonstrated that the peptidoglycan is injected by the T4SS into the host cells and induce synthesis of chemokines that recruit neutrophils, such as interleukin-8 (IL-8), and activation of pro-inflammatory signalling pathway, acting through the cytoplasmic receptor Nod1, an intracellular pathogen recognition molecule with specificity for Gram-negative peptidoglycan (Hatakeyama 2004).

1.5.3 NEUTROPHIL-ACTIVATING PROTEIN

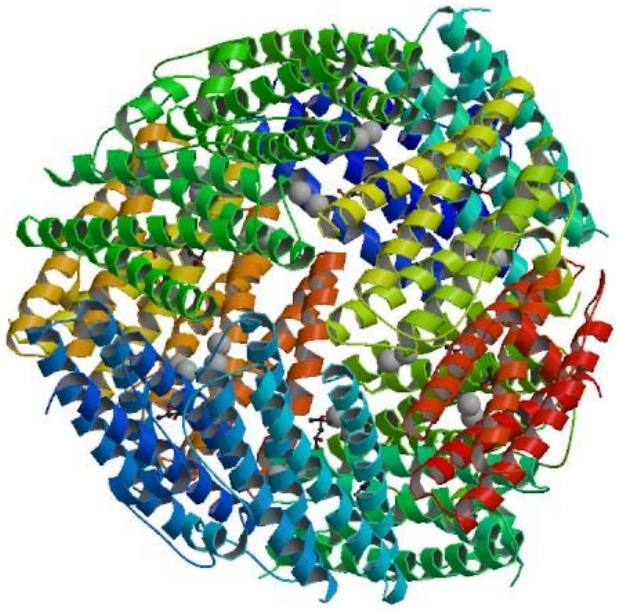

Fig. 1.30 Ribbon representation of Neutrophil activating protein: dodecameric structure (PDB accession code:1JI4)

The *H. pylori* neutrophil-activating protein (HP-NAP) is a highly immunogenic protein, both in humans and mice, responsible for the attraction of neutrophils and monocytes by chemotaxis at the site of *H. pylori* infection; it has been shown to induce neutrophil adhesion to endothelial cells, to increase the adhesion of neutrophils to endothelial cells, to induce migration and activation of human neutrophils and monocytes and to be a potent stimulant of mast cells (Choli-Papadopoulou 2011). The crystallographic structure, which has been determined in our laboratory, reveals that HP-NAP consists in a 150 kDa-oligomer composed by four-helix bundle subunits that oligomerize to form a dodecamer with a central internal iron-containing negative cavity (Fig. 1.31; Zanotti et al. 2002; PDB accession code: 1JI4). HP-NAP shows relevant similarities to some bacterial iron-binding and DNA-protecting

proteins, named Dps protein. These proteins protect bacterial DNA from oxidative radicals generated by the Fenton reaction and also from various other damaging agents; in fact, their expression is induced under stress conditions. The *napA* gene is highly conserved among geographically distinct strains of *H. pylori*, suggesting a precise structurally-dependent function for HP-NAP; as can be evidenced, the few non-conservative differences within the aminoacid sequence do not affect its whole structure. Interestingly, HP-NAP does not bind to DNA fragments like the other Dps proteins, but it is able to storage in its central cavity up to 500 atoms of iron per oligomer (Tonello et al. 1999). An additional peculiarity of this *H. pylori* protein resides in its N-terminal region. Usually Dps proteins contain a positively charged N-terminus that promotes the binding to DNA and its condensation; HP-NAP does not possess a positively charged N-terminus, but it is characterized by a positively charged protein surface, due to the presence of a large number of positively charged residues. This peculiar surface potential charge distribution has been proposed to be responsible for binding and condensing DNA, as well as for activating human leukocytes (Zanotti et al. 2002; Ceci et al. 2007). HP-NAP is the only protein of the Dps family capable of activating human leukocytes. HP-NAP has been found to promote neutrophil adhesion to endothelial cells and to induce neutrophils to produce reactive oxygen radicals. Moreover, HP-NAP has been found to activate the Toll-like receptor 2 (TLR2) and, thus, stimulate either neutrophils or monocytes to increase their expression of interleukin-12 (IL-12), a key cytokine for the differentiation of $CD4^+$ T cells into the Th1 phenotype (Trinchieri 2003). HP-NAP is released in the medium, probably after cell lysis, and binds to the bacterial surface where it can act as an adhesin, mediating binding to mucin or to polymorphonuclear leukocyte sphingomyelin. Therefore, HPNAP is able to cross the epithelium efficiently and to promote rapid neutrophil adhesion in the infection site. Here, the H. pylori protein induces polymorphonuclear leukocytes to synthesize and release several pro-inflammatory cytokines responsible for the inflammation process. Moreover, after crossing epithelial monolayers, HPNAP is also able to activate the underlying mast cells to release tumor necrosis factor-α (TNF-α) and other pro-inflammatory molecules, and to recruit neutrophils and monocytes from the blood, thus amplifying the flogistic process (de Bernard and D'Elios 2010; Choli-Papadopoulou 2011). This inflammation could promote *H. pylori* growth by releasing of nutrients from the degraded epithelium, for the bacterial survival and colonization; but the specific role of HP-NAP inside the bacterium is still controversial.

1.6 *HELICOBACTER PYLORI* AND GASTRODUODENAL DISEASES

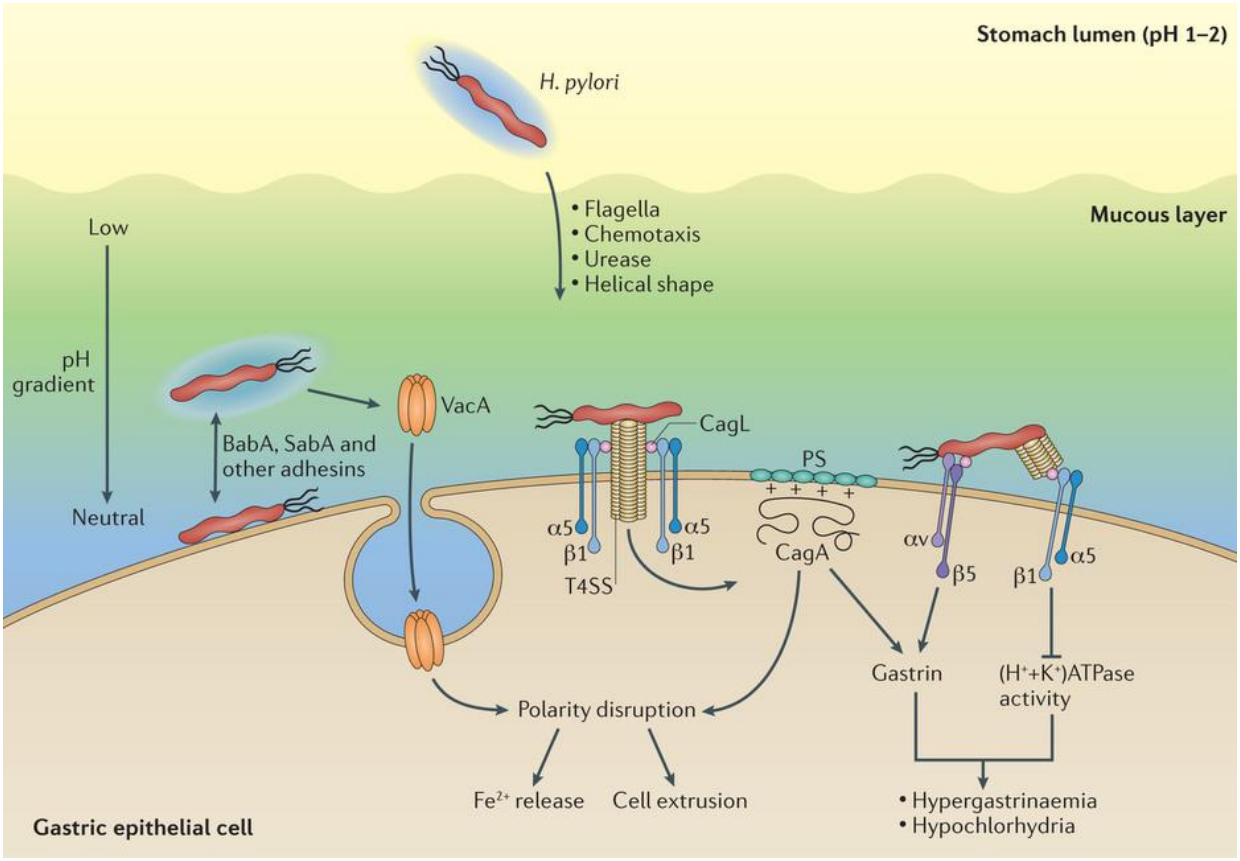

Fig. 1.31 Main *H. pylori* colonization and persistence factors related to gastroduodenal diseases. Adapted from Salama et al., 2013.

H. pylori typically colonizes the human stomach without causing adverse consequences for many decades; but its presence is associate with an increased risk of several diseases, including peptic ulcers, noncardia gastric adenocarcinoma, and gastric mucosa-associated lymphoid tissue (MALT) lymphoma (Cover and Blaser 2009). Interestingly, according to the strong association between *H. pylori* and gastric cancer, the World Health Organization (WHO) classified the pathogen as a Class I Carcinogen (WHO-IARC 1994), and now it is considered the most common etiologic agent of infection-related cancers. Since *H. pylori* is so well adapted to its environment, in most cases the predominant symptom of a prolonged infection is a mild gastritis resulting from the inflammation process linked to nutrients acquisition. One salient factor contributing to the co-existence of *H. pylori* in the human stomach is the localization of the bacterium within the gastric mucus layer, without any substantial invasion of host tissue. Moreover, *H. pylori* is versed in evading immune recognition by the host, by displaying a phase variation and an antigenic variation of its surface components, like outer membrane

proteins (OMPs) and lipopolysaccharide (LPS) antigens, as previously described. Incorporation of a modified form of cholesterol into *H. pylori* membranes and the coating of the bacterium with host molecules, such as plasminogen, might represent additional types of antigenic disguise (Cover and Blaser 2009). Flagellar proteins and many bacterial factors also appear to be designed to reduce inflammation or recognition by the host immune system. Therefore, *H. pylori* induces a robust but specific form of chronic inflammation that is ineffective in clearing the infection, while avoiding forms of inflammation that would eliminate it (Amieva and El-Omar 2008).

A discrete incidence (approximately 10%) of severe gastrointestinal diseases has been correlated with the long-term stomach colonization by *H. pylori*. The various and divergent clinical outcomes deriving from the *H. pylori* infection are dictated by a complex balance between host genetic factors, bacterial virulence determinants, environmental components, such as salt consumption, general alimentary habits, smoking, and living conditions (Fig. 1.32; Sgouras et al. 2015). Moreover, recently the gastrointestinal microbiota, either at the time of exposure or over the course of infection, has been evidenced as another potential determinant in the clinical outcome (Martin and Solnick 2014). The basic process that mediates *H. pylori*-induced damage is gastritis with its associated humoral and cell-mediated immune mechanisms. The extent and distribution of this gastritis ultimately determine the clinical outcome. Three main gastric phenotypes have been identified: “simple or benign gastritis”, “duodenal ulcer”, and “gastric cancer” phenotypes, respectively (Fig. 1.33). Each phenotype is associated with a set of pathophysiologic abnormalities that could explain why a certain outcome occurs (Amieva and El-Omar 2008). Therefore, colonization with *H. pylori* induces a chronic gastritis in all infected individuals. Although the majority of infections are asymptomatic, however, as result of long-term colonization, 10%-15% of infections lead to the development of peptic ulcer disease; especially, *H. pylori* is associated with 95% of duodenal ulcers and 80% of gastric ulcers. Infection with this pathogen is a significant risk factor for the development of gastric cancer and 1%-3% of infected individuals develop the disease. Infection is associated in particular with intestinal-type (approximately 90% of patients) rather than diffuse-type gastric cancers

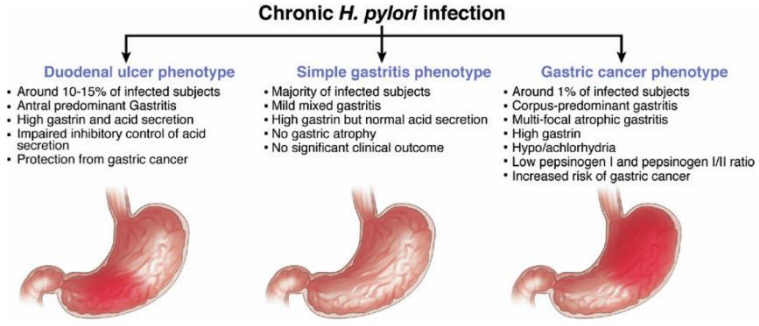

Fig. 1.32 Pathophysiologic and clinical outcomes of chronic *H. pylori* infection. Adapted from Amieva and El-Omar 2008

(approximately 32% of patients). The risk of developing gastric cancer is reduced in patients with duodenal ulcers. The gastric mucosa does not normally contain any mucosa-associated lymphoid tissue (MALT), however the pan-gastric inflammation induced by *H. pylori* infection results in the development of MALT. In < 0.1% of infected individuals this develops into B cell MALT lymphoma; however, at early stages the lymphoma can be cured by eradication of *H. pylori* (Dunne 2014).

First of all, *H. pylori* polymorphism plays a role crucial in planning the consequences of the colonization. Most of polymorphisms associated with various disease risks are associated to genes that encode bacterial factors responsible of the interaction with host cells. *H. pylori* strains isolated from unrelated individuals exhibit a high level of genetic diversity. Nucleotide sequences of conserved genes are 92%–99% identical among different *H. pylori* strains, but several bacterial genes are more highly diverse in sequence; additionally, there is a considerable variation in gene content (Cover and Blaser 2009). One of the most striking differences among *H. pylori* strains is the presence or absence of the *cag* pathogenicity island (*cag*PAI), which also encodes the virulence effector CagA. Several studies have demonstrated that *cag*PAI positive *H. pylori* strains, particularly in Western populations, are associated to a higher risk of peptic ulcer disease, premalignant gastric lesions, and gastric cancer, than the strains lacking *cag*PAI (Basso et al. 2008).

Moreover, as previously mentioned, a higher number of tyrosine phosphorylation EPIYA-C motifs in CagA is correlated with a higher risk of gastric cancer. All *H. pylori* strains contain a *vacA* gene, encoding a further very common virulence factor, the vacuolating cytotoxin A. The gene presents a high polymorphism, with variations at several loci; consequently, the protein effector shows different levels of vacuolating cytotoxin activity. The major variation occurs in the *vacA* signal region (s1 or s2) and the middle region (m1 or m2). Type s1/m1 strains are highly toxigenic and s2/m2 strains are nontoxigenic; while s1/m2 strains are often intermediate. Few s2/m1 strains have been described, suggesting the existence of a strong negative selection toward this allele (Basso et al. 2008). Therefore, both s1/m1 and s1/m2 strains are associated with peptic ulcer disease; additionally, s1/m1 strains are strongly related to gastric cancer risk (Atherton et al. 1997). Similarly, *H. pylori* strains expressing outer membrane proteins (OMPs), such as BabA, SabA, and OipA, are predicted to be highly interactive with the host cells and associated with increased gastric epithelial disease risk, respect to strains lacking these adhesion factors (Israel and Peek 2001; Cover and Blaser 2009). Therefore, concordant with these

evidences, CagA+, s1-VacA+, BabA+ strains are associated with increased gastric mucosal inflammatory cell infiltration and increased gastric epithelial injury, compared to strains that do not express these factors. In addition, the colonization density of CagA+, s1-VacA+, BabA+ strains is typically higher than that of strains that do not express these factors (Cover and Blaser 2009).

The ability of *H. pylori* to cause disease depends not just on bacterial factors, but also on environmental and host factors. A number of host gene polymorphisms have been identified which are thought to increase *H. pylori* colonization and increase susceptibility to disease (Dunne 2014). In effect, genetic polymorphisms directly influence interindividual variation in the magnitude of cytokine response, and this clearly contributes to an individual ultimate clinical outcome (Amieva and El-Omar 2008).

Especially, the IL-1 gene cluster is composed of three genes which encode the pro-inflammatory cytokines IL-1α and lL-1β, as well as their endogenous receptor antagonist IL-1ra. IL-1β is a potent inhibitor of gastric acid secretion in vivo. *H. pylori* infection results in an upregulation of IL-1β, which plays an important role in initiating and amplifying the immune response. Additionally, polymorphisms in IFN-γ are associated with infection by *cag* positive strains and polymorphism in TNF-α are associated with development of peptic ulcer disease. While polymorphisms in IL-10 have been demonstrated to favor the development of intestinal metaplasia and non-cardia gastric cancer. Therefore, polymorphisms which effect cytokine function may explain the highly variable outcomes of *H. pylori* infection (Zambon et al. 2005; Dunne 2014). Moreover, polymorphisms in genes involved in the innate and adaptive immune response play a role in host susceptibility to infection and disease progression. Particularly, IL-2 plays an important role in mediating the T lymphocyte response, including the Th1 phenotype which dominates the immune response to *H. pylori* infection. An increase in IL-2 production as a result of a gene polymorphism has been shown to be associated with a decreased risk of *H. pylori* infection in adults (Queiroz et al. 2009). Furthermore, there is an association between polymorphisms in the gene encoding the membrane-associated mucin MUC1, which result in shorter MUC1 alleles, and *H. pylori*-related gastritis. Short MUC1 alleles have also been linked to the development of gastric adenocarcinoma and intestinal metaplasia in patients with chronic gastritis. Moreover, the secretor status of an individual is also thought to influence their susceptibility to infection. A Fucα1,2-glycan is expressed along the gastro-intestinal tract of individuals with a positive secretor status, which express the secretor-(fucosyl)transferase. Studies involving infected secretor and non-

secretor mice have revealed that the non-secretor mice display reduced adhesion of *H. pylori* to gastric tissue (Magalhaes et al. 2009). A functional polymorphism has been described in TLR4 gene, resulting in alteration in the extracellular domain of the receptor. TLR4 receptor is responsible of recognition and interaction with lipopolysaccharide (LPS) antigen of *H. pylori*. Defective signaling through the mutant TLR4 receptor are associated with an exaggerated and destructive chronic inflammatory phenotype in *H. pylori*-infected subjects. This phenotype is characterized by gastric atrophy and hypochlorhydria, the hallmarks of subsequent increased risk of gastric cancer. Therefore, bacterial factors alone cannot explain why some individuals develop disease upon *H. pylori* infection whereas others remain asymptomatic. The individual-specific synergy between bacterial and host polymorphisms is crucial in determining disease risk and different clinical outcome as result of *H. pylori* colonization (Amieva and El-Omar 2008; Dunne 2014).

1.7 ERADICATION AND POTENTIAL BENEFITS

H. pylori has started to colonize humans so long ago that it is not surprising that its relationship is that of both a commensal bacterium and a pathogen, causing some diseases and possibly protecting against others; a growing body of literature suggests that the eradication of *H. pylori* might also be associated with an increased risk of other diseases (Dorer, Talarico, and Salama 2009).

An absence of *H. pylori* could indicate that an individual was never colonized, or that the microorganism was present earlier in life and subsequently spontaneously eradicated (Cover and Blaser 2009). Natural elimination of the bacterium is a rare event, occurring mainly soon after acquisition in childhood (Xia and Talley 1997). In order to eradicate *H. pylori* from stomach of adults, and even of children, it is necessary a specific treatment. The ideal *H. pylori* eradication therapy should be safe, effective (eradication rate > 90%), simple, and economical (Zhang 2015). The first-line treatment is a triple therapy, consisting of a proton pump inhibitor PPI, clarithromycin and amoxicillin or metronidazole, in populations with less than 15–20% clarithromycin resistance rate (Malfertheiner et al. 2007). The eradication rate of this triple therapy is currently less than 80% in most parts of the world (B. G. Kim et al. 2007); antibiotic resistance is the main reason for this treatment failure. The recourse to antibiotic treatments in the past has dramatically increased the resistance of *H. pylori* and this evidence has always been underestimated. Therefore, these observations lead to strong rationale for development of effective vaccines against *H. pylori*, since they could also allow to prevent the colonization. Over the last years, several approaches have been devised to generate an efficient vaccine against the bacterium. Whole bacterial cell sonicates (first-generation vaccines) and individual *H. pylori* proteins (second-generation vaccines) have been used as antigens to stimulate immunity in the host, but they require adjuvants to elicit effective protection (Ferrero and Labigne 2001). The results of *H. pylori* vaccine development efforts, both preclinical and early clinical, have so far been disappointing. Sterilizing immunity is rarely achieved, even in animal models, and there is no consensus on the delivery route, adjuvants and choice of antigen. The most promising preclinical results have generally been obtained with vaccination strategies that aim to induce protective T cell-mediated immunity rather than humoral immunity, with local gastric T helper 1 (Th1) and Th17 responses being reasonably good correlates of, and prerequisites for, protection

(Salama, Hartung, and Muller 2013). New immunization approaches are going to be explored in the near future, including live vector vaccines, DNA vaccines, microsphere vaccines, ghost vaccines. Briefly, DNA vaccines carry DNA sequences encoding *H. pylori* antigens and they are known because of their safety and efficacy. Microspheres effectively induce humoral and mucosal immunity as well as cell mediated immunity, whereas bacterial ghosts are empty cell envelopes without cytoplasmic contents that retain their cellular morphology with native antigenic structures. Several questions about when and whom to vaccinate will need to be appropriately answered, and a cost-effective vaccine production and delivery strategy will have to be useful for developing countries (Agarwal and Agarwal 2008). Besides the development of new more efficient vaccine against *H. pylori*, the medical research is also focused on finding new pharmacological targets in order to develop new antibiotic treatments, required for a successful eradication of the pathogen. The characterization of new potential pharmacological protein targets is topic of my research project.

Conversely, the complete eradication of *H. pylori* has been suggested to be associated with a potentially increased risk of onset of other disorders. Not surprisingly, most of the potential benefits (as with the costs) come from *cagA*+ strains, which are the most interactive with their human hosts. Firstly, an inverse association has been observed between the presence of *H. pylori* (especially *cagA*+ strains) and disorders such as gastroesophageal reflux disease, Barrett's esophagus and esophageal adenocarcinoma. One potential mechanism for this effect could be that *H. pylori* colonization diminishes gastric acidity; therefore, during reflux episodes, the acidic refluxate might be less damaging to the esophageal epithelium. Another hypothesis is that *H. pylori* alters the expression of multifunctional gastric hormones that have effects on esophageal tissue. Furthermore, the incidence of asthma and related disorders, such as allergies, especially those that appear during childhood, has been observed to arise as result of the absence of *H. pylori*. This inverse association is specific for childhood-onset asthma, and is most pronounced for cagA+ strains. It is possible that the presence of *cagA*+ *H. pylori* strain in the stomach leads to gastric recruitment of T-cell populations, including regulatory T cells, that ultimately affect the activities of T-cells present in other mucosal and cutaneous sites. Recently it has been suggested that *H. pylori* colonization might confer protection against various other infectious diseases. Moreover, *H. pylori* has been proposed to protect against diarrheal diseases, although this relationship has not been consistently observed. Mechanisms for protection might include production of antibacterial peptides by the

pathogen or the host, activating the immune system as an adjuvant, competition for niche, or hypergastrinemia leading to maintenance of gastric acidity throughout childhood. Finally, H. pylori presence is also associated to benefits on metabolism, since it is involved in the regulation of two hormones that control the body weight, namely leptin and ghrelin, partially secreted by the stomach. *H. pylori*-positive persons produce lower amounts of ghrelin than do *H. pylori*-negative persons and the eradication of the bacterium is associated with a subsequent increase in ghrelin production, resulting in long-term metabolic consequences (Rothenbacher et al. 2000; Blaser, Chen, and Reibman 2008; Cover and Blaser 2009). Therefore, if continued studies confirm the findings reported above, then the medical approaches to *H. pylori* infection will need to change.

Chapter II

RESEARCH PROJECT
AND EXPERIMENTAL METHODS

RESEARCH PROJECT AND EXPERIMENTAL METHODS

The main purpose of this research thesis is focused on identification and structural characterization of new potential pharmacological targets of the pathogen *H. pylori*. There is no effective therapy for eradicating *H. pylori* infection. Combination therapies employing one proton pump inhibitor and two or three antibiotics have been used as preferred treatments. However, these therapies have several inherent problems, including the appearance of resistance to the antibiotics used and associated adverse effects, the risk of re-infection and the high cost of antibiotic therapy. Alternative therapeutic approaches are required to control *H. pylori* infection (Ayala et al. 2014). The determination of the three-dimensional structure by x-ray crystallography of *H. pylori* proteins essential for colonization or pathogenesis might allow to develop new potential therapeutic molecules. The strategy adopted consists of the following steps:

- selection of candidates suitable for crystallization;
- amplification of the corresponding gene by PCR from genomic DNA belonging to the G27 strain;
- cloning into expression vectors and eventual insertion of affinity tags to improve the purification steps;
- small scale expression trials in *E. coli* and tests of solubility of the recombinant constructs in order to identify the most effective protocol of expression;
- small scale purification trials in order to define the most fruitful and quality purification protocol;
- characterization in solution;
- crystallization trials adopting various techniques and performing several precipitant solution screening;
- X-ray diffraction tests and optimization of initial crystals;
- X-ray diffraction data collection at synchrotron facilities;
- data processing and structure determination;
- further investigations in order to functionally characterize the candidates.

2.1 SELECTION OF CANDIDATES

H. pylori strain G27, which was originally isolated from an endoscopy patient from Grosseto Hospital (Tuscany, Italy), has been extensively used in *H. pylori* research. It is readily transformable and therefore amenable to gene disruption. In addition, it efficiently delivers the translocated effector protein CagA to cells in culture. The complete genome sequence was determined at the Washington University School of Medicine Genome Sequencing Center. Genomic DNA from an isolate of strain G27 was purified by CsCl gradient centrifugation and used to generate plasmid and fosmid libraries, both of which were subjected to whole shotgun Sanger sequencing (4,609 total reads). The G27 genome consists of a single circular 1,652,983 bp chromosome that is AT rich (61.1%), contains 1,515 open reading frames (ORFs), and is similar in size and composition to the other published *H. pylori* genomes of strains 26695, J99, and HPAG. G27 also contains one 10,032 bp AT-rich (65.2%) plasmid that encodes 11 genes and resembles the plasmid found in strain HPAG. The G27 *cag* pathogenicity island (*cag*PAI) contains a transposon, but this does not disrupt any of the open reading frames and is not predicted to interfere with the type IV secretion system delivery of CagA into host cells. Similar to strains J99 and HPAG but in contrast to 26695, G27 has a single plasticity region, which contains many *H. pylori*-specific genes that are variably present between strains. G27 contains 58 genes that are not found in 26695, J99, or HPAG. The majority of these G27-specific genes are predicted to encode hypothetical proteins (Baltrus et al. 2009).

The criterion for selection of protein candidates is based on two main features; the possible pathogenic role and the predicted solubility have been evaluated. The soluble nature of a protein, in addition to other structural features, is an essential prerequisite for achieving a successful crystallization. Literature searches have been conducted to evaluate metabolic pathways as new potential targets for therapeutic treatment. In this regard, effector proteins involved in *H. pylori* crucial processes, such as acid adaptation, motility, and redox metabolism, have been taken into account. Moreover, G27-specific genes encoding hypothetical proteins have been considered an interesting pharmacological target since they could play a key role in bacterial survival. Taking into account the information from literature, the structural predictions and the functional features, 6 candidates have been selected, of both known and unknown function, as suitable subjects for recombinant production in *E. coli*, purification and following crystallization.

2.2 MOLECULAR CLONING

The first step toward protein production is the cloning of the gene encoding the protein of interest into a proper expression vector. The selected genes have been amplified by PCR from a genomic DNA sample extracted from *H. pylori* strain G27, using specific primers designed individually for each open reading frame. Without exception, the candidates have been cloned as fusion proteins with an affinity tag, which might help to isolate the recombinant fusion protein by affinity chromatography and discard the majority of *E. coli* proteins. The PCR products have been purified by 1% agarose gel and subsequently cloned into a pET vector that simultaneously works as either cloning and expression vector (Fig. 2.1). These pET vectors are optimized for a rapid cloning of His-tagged proteins and are under the control of an inducible promoter (T7 RNA polymerase promoter) with tight regulation for high expression as well as plasmid stability. These vectors are provided in a pre-processed, linearized format for rapid enzyme-free cloning, by homologous recombination. The open reading frames are directionally cloned into the pET vector, in frame with a choice of either an N- or C-terminal 6-His affinity tag. This approach allows to limit as much as possible the time-consuming procedure of traditional cloning methods. In most cases full-length genes have been cloned. For those genes that are predicted to encode a secretion signal sequence, only the portion of the gene encoding the mature protein has been cloned, in order to avoid that the signal sequence made the product more toxigenic to *E. coli*. Prior to proceed with protein expression trials, all the constructs have been confirmed by DNA sequencing.

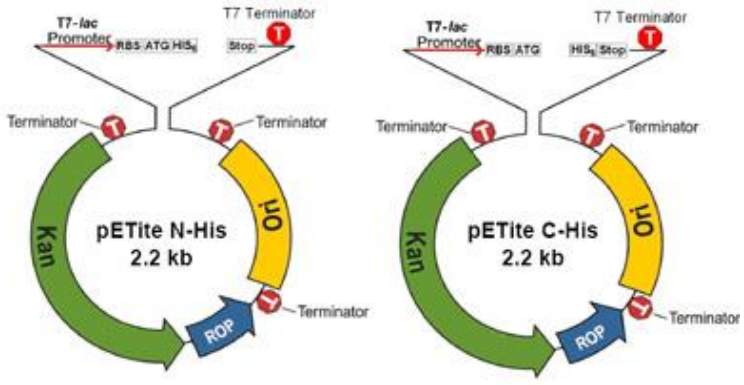

Fig. 2.1 pETite expression vectors (Lucigen), widely used in this research project.

2.3 PROTEIN EXPRESSION IN *E. COLI* AND TEST OF SOLUBILITY

Once the gene of interest has been cloned into the selected expression vector, small scale trials were performed to control and optimize level and quality of protein expression.

Escherichia coli remains the preferred host for heterologous expression of recombinant proteins in order to carry out structural studies. This bacterium is the most characterized host for protein expression and has numerous advantages over other hosts. These advantages include ease of handling, well-known genetics, relatively inexpensive culture medium, availability of vectors and fast high-density cultivation and expression (Makrides 1996; Sørensen and Mortensen 2005a; Sorensen and Mortensen 2005b). Additionally, *E. coli* is a Gram-negative bacterium, like *H. pylori*, and it can be considered the best choice to express *H. pylori* proteins. The sequenced plasmid vectors have been directly transformed into *E. coli* BL21(DE3) cells (Lucigen), which were used for preliminary small scale expression trials. Expression tests have firstly been carried out by growing cultures inoculated with single colonies in small scale using 10-100 ml of Luria Bertani media and taking aliquots of the cultures prior to and following growth induction by adding isopropyl-β-D-thiogalactoside (IPTG). Protein expression has been induced with 0.5-1 mM IPTG and its duration has depended on the growth temperature, specifically 3 h at 37°C, 5 h at 30°C, and overnight at 20°C. Whole-cell lysates have been prepared from cultured cells, by a French pressure cell press or by sonication, and analyzed by Coomassie-stained 15% SDS-polyacrylamide gel electrophoresis (SDS-PAGE) and Western blotting. If the protein expression was too low or not detected, further *E. coli* strains, which can assist in case of problematic target, have been tested. BL21 (DE3) pLysS cells (Invitrogen) contain a chloramphenicol resistant plasmid, pLysS, which carries the gene encoding T7 lysozyme that inhibits the background expression level of target genes under the control of the T7 promoter, but does not interfere with the level of expression achieved following induction by IPTG. C41 (DE3) cells (Lucigen), derived from BL21(DE3), have a mutation that reduces the level of T7 RNA polymerase activity, thereby preventing cell death associated with overexpression of many recombinant toxic proteins. C43 (DE3) cells (Lucigen), derived from C41(DE3), carry at least one additional mutation that provides a greater level of tolerance to toxic proteins. SHuffle® Express Competent *E. coli* cells (New England Biolabs) are engineered to form proteins containing disulfide bonds in the cytoplasm. Lemo21 (DE3) cells (New England Biolabs) offer the host features of BL21(DE3) while also allowing for tunable expression of difficult proteins, achieved by varying the level of lysozyme (*lysY*), the natural inhibitor of T7 RNA polymerase, modulated by adding L-rhamnose to the expression culture. Finally, ArcticExpress (DE3) cells (Agilent) are engineered to address the common bacterial gene expression hurdle of protein insolubility and provide an approach to increasing the yield of soluble protein

produced in *E. coli* at low temperature, since they co-express the cold-adapted chaperonins Cpn10 and Cpn60 from the bacterium *Oleispira antarctica*.

Besides the determination of the expression conditions, the composition of lysis buffers has been evaluated to achieve the highest yield of soluble protein. The protein solubility represents one of the most relevant bottleneck for production of protein for X-ray crystallography. Cultured cells have been resuspended in several lysis buffers, which differ in pH value (at least 2 units different from the pI, to avoid the protein aggregation), salt concentration, and eventual additives, lysed by French pressure cell press or sonication, and finally centrifuged; the supernatant and pellet from the centrifugation have been analyzed by 15% SDS–polyacrylamide gel electrophoresis (SDS-PAGE) and Western blotting to compare solubility of the protein in the various lysis buffers. If most of the protein was found in the pellet, as inclusion bodies, changes in the expression conditions have been made, especially in *E. coli* strain and growth temperature, to decrease the rate of protein translation or enhance protein folding. Despite several trials to improve the expression conditions and the yield of the soluble fraction, the proteins could remain insoluble or behave poorly. Therefore, in case of high expression level of unfortunately insoluble protein, the recovery of the recombinant product has been attempted, by denaturation and subsequent refolding, using solubilizing buffers supplemented with 8 M urea or 6 M Guanidine hydrochloride.

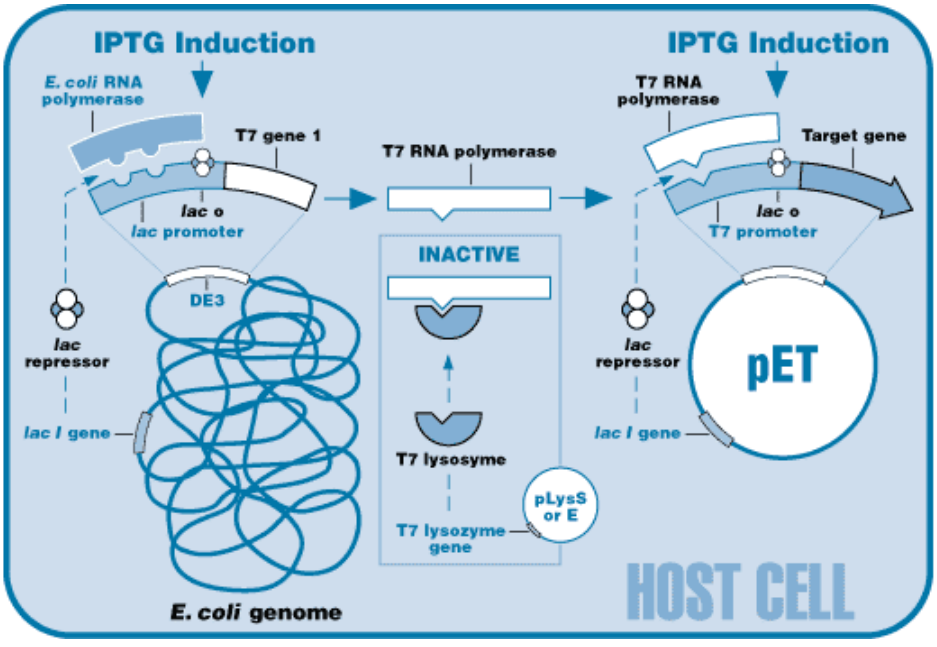

Fig. 2.2 Mechanism of induction of expression by IPTG and T7 promoter. Adapted from Quora web site.

2.4 PROTEIN PURIFICATION AND CHARACTERIZATION

Purification should yield a sample of protein containing only one type of molecule; crystallization step requires samples characterized by high degree of purity and homogeneity. Since every protein have a unique behavior, every purification usually requires constant assessments and modifications based upon the peculiar features of the target. After having developed the expression conditions on small scale, the protein expression has been reproduced on large scale. Purification has generally been approached once a relatively high yield of soluble protein has been achieved. The cultured cells have been resuspended in selected lysis buffer and lysed by French press or sonication. All lysis buffers have also contained protease inhibitors cocktail in order to suppress protein degradation. After the lysis cell debris has been removed by vacuum centrifugation at 18,000 rpm for 30 min at 4 °C. Protein solubility has been checked on 15% SDS–PAGE and Western blotting by loading both the pellet and the soluble fraction after centrifugation. Proteins can be purified by chromatography, on the basis of such characteristics as solubility, size, charge, and binding affinity. The chromatographic separations were generally operated on an ÄKTA FPLC instrument (GE Healthcare). Firstly, an affinity chromatography has been performed for each protein, based on the fused tags, which might help to isolate the recombinant protein and discard the majority of the other *E. coli* proteins. Since all the constructs were 6-His-tagged, the first chromatographic step of purification has been a standard immobilized metal affinity chromatography (IMAC). After extensive washing the proteins have been eluted applying an imidazole gradient (0-500 mM imidazole). If the recombinant fusion protein was not enough pure after affinity chromatography, it has been concentrated by ultrafiltration, with suitable cut-off, and loaded onto a gel-filtration column, chosen according to the molecular weight of the product. Gel-filtration chromatography, also named size-exclusion chromatography, allows to separate a mixture of proteins on the basis of their size. The affinity tag sometimes may negatively affect the protein purification, since it can interfere with folding, oligomerization, protein localization. In case of unsuccessful purification by these leader chromatographic techniques, the preparation of pure proteins has been performed by a classical method, which includes three steps: fractional precipitation with saturated ammonium sulfate solution, ion-exchange chromatography, and finally gel-filtration chromatography. Since the His-tag may affect the crystallization process due to its flexibility and ionic charge, it has been necessary to remove it. In this regard, the

proteins of interest have been cloned into the vectors with a protease restriction site between the affinity tag and the construct. The most efficient one is the TEV protease, from the Tobacco Etch Virus, which shows high specificity, is relatively easy to make in large amounts and cleaves in most case leaving a native N-terminus. Specifically, recombinant 6-His-tagged TEV proteases have been used, enabling an easy separation of the enzyme by means of a new affinity chromatography after the cleavage.
Subsequently, the purified proteins have been characterized in solution by:

- analytical gel-filtration chromatography, to evaluate the homogeneity and eventual oligomeric form of the samples;
- circular dichroism (CD) to ensure the correct folding of the protein product and estimate the content of secondary structures;
- UV/Vis spectroscopy;
- SDS-PAGE and Western blotting;
- functional essays.

2.5 PROTEIN CRYSTALLIZATION

Fig. 2.3 Oryx8 Protein Crystallization Device (Douglas Instruments, UK)

If the protein samples presented high degree of purity and homogeneity, they have been submitted to crystallization trials. The proteins have been concentrated by ultrafiltration to 5-20 mg/ml as starting concentration for crystallization trials. Optimal conditions for crystallization are very difficult to predict, since they depend on a large number of variables, such as pH, precipitant, salt, protein concentration, additives, buffer, temperature, detergents, organic compound, etc. An effective way to search through this large number of parameters that may influence crystal growth is to conduct a sparse matrix search of specific crystallization conditions (Krauss et al. 2013). First crystallization trials have been performed using commercial crystal screening kits, specifically Structure Screen I and II (Molecular Dimension), PACT Suite (Qiagen), PEGs Suite (Qiagen), Crystal Screen I and

II (Hampton Research), JCSG Suite (Molecular Dimension), AmSO4 Suite (Qiagen). Both hanging drop and sitting drop vapor-diffusion methods have been approached; these techniques rely on the exchange of either water and/or some volatile agents between a micro-drop of mother liquor, containing the protein, and much larger reservoir solution. Moreover, microbatch-under-oil method has also been approached; this technique allows to regulate the rate of water evaporation through the layer of oil, leading to a concomitant increase in the concentrations of both protein and precipitant until the nucleation point is reached, as well as preventing complete desiccation of the drops. The crystallization experiments have been performed manually and also using an Oryx8 crystallization robot (Douglas Instruments). Quantity of reservoir solution, drop size, ratio between protein sample and precipitant have been varied in each experiment. Mainly three crystallization temperatures were tested: 4 °C, 12 °C, 20 °C. For initial crystallization condition screenings have been used 96-well crystal plates (Douglas Instruments); conditions that yielded crystals have been optimized with manually prepared condition using 24-wells plates (Hampton Research). Moreover, microseeding and co-crystallization experiments have been performed, in order to optimize the crystal quality and to further investigate the structural features of the protein, respectively. Manual inspection of crystallization experiments has been carried out under an optical microscope; furthermore, the nature and quality of crystals have been evaluated using a fluorescence microscope.

2.6 DATA COLLECTION AND STRUCTURE DETERMINATION

Protein crystals are very often thin and small (microns to millimeters, in a few lucky events) and rich in solvent content, between 30 to 70 %. Therefore, it is often a great advantage to have access to high brilliance synchrotron sources to collect X-ray data with good quality and high resolution. Moreover, the synchrotron radiation is tunable, thus it allows the use of any suitable wavelength in the spectral range selected (usually between 0.9 and 2 Å for proteins). This property is particularly relevant in multiple wavelength anomalous dispersion (MAD) technique and when wavelengths shorter than 1 Å are necessary. Specifically, X-ray diffraction data have been collected at SLS (Villigen, Switzerland) and ESRF (Grenoble, France). Determination of the three-dimensional structure of proteins is achieved by several techniques, such as Multiple wavelength Anomalous Dispersion (MAD), Multiple Isomorphous Replacement (MIR), Single

wavelength Anomalous Dispersion (SAD), Single Isomorphous Replacement (SIR), and Molecular Replacement (MR). Since the solved protein structure presented a structural homologue, molecular replacement technique has been adopted to determine it. This method relies upon the existence of a previously solved protein structure which is homologous to the unknown structure and allows to quite easily solve the phase problem. The goal is to orient and position the determined model, such that it coincides with the position of the unknown protein in the crystal; the model can then provide phase information for the unknown structure.

In the following experimental section, each chapter is dedicated to a protein or a group of proteins; the experimental procedures, briefly described above, are applied to the specific subjects and explained more in detail.

Chapter III

STRUCTURAL CHARACTERIZATION OF α-CARBONIC ANHYDRASE FROM *HELICOBACTER PYLORI*

This Chapter has been adapted from:

Maria Elena Compostella[1], Paola Berto[1], Francesca Vallese[1] and Giuseppe Zanotti[1].
Structure of α-carbonic anhydrase from the human pathogen Helicobacter pylori.
Acta Crystallographica Section F, Structural biology communications (2015); F71, 1005–1011.

[1] Department of Biomedical Sciences, University of Padua, Italy

3.1 INTRODUCTION

Carbonic anhydrases (carbonate hydro-lyase; EC 4.2.1.1) are Zn^{2+}-containing enzymes that catalyze the reversible conversion of carbon dioxide to bicarbonate (Tripp, Smith, and Ferry 2001). This process is extremely rapid, with rate constants in the region of 10^5–10^6 s^{-1} and following a two-step ping-pong mechanism; on the other hand, the reversible hydration occurs rather slowly in the absence of a catalyst (Lindskog and Coleman 1973). Since carbonic anhydrase is ubiquitous in many tissues and plays several important physiological roles in humans, such as bone resorption, gluconeogenesis, production of body fluids, acid–base balance, secretion, signal transduction, oncogenesis, transport of CO_2 and HCO_3^- (Sly and Hu 1995), it is considered to be an attractive pharmacological target. Its inhibitors are clinically used to treat diseases as different as glaucoma, convulsions, osteoporosis and obesity; the enzyme is also a tumour marker (Supuran and Scozzafava 2007; Gilmour 2010). These seemingly disconnected functions are mediated by specific isoforms belonging to the α-carbonic anhydrase family. In humans 16 members of this family have been identified which have distinct tissue-specific expression, kinetic properties, and sensitivity to inhibitors (Becker, Klier, and Deitmer 2014). The mammalian carbonic anhydrase isoforms have been extensively characterized (Lindskog 1997); the catalytic mechanism has been thoroughly described (Boone et al. 2014) and structural studies have led to the deposition of 477 crystal structures in the Protein Data Bank (http://www.pdb.org), either in the native form or with inhibitors or ligands bound. In mammals only one class of carbonic anhydrase is present, the α-carbonic anhydrase, whilst five phylogenetically and structurally different classes of the enzyme, α-carbonic anhydrase, β-carbonic anhydrase, γ-carbonic anhydrase, δ-carbonic anhydrase and ζ-carbonic anhydrase, have been found in bacteria, archaea and diatoms. Very recently, a new carbonic anhydrase family, named η-carbonic anhydrase, has been identified (Del Prete et al. 2014).

Specifically, bacteria encode for enzymes belonging to the α-, β-, and γ-carbonic anhydrase classes; in particular, α-carbonic anhydrases have been detected only in the genomes of Gram-negative bacteria (only in six microorganisms, including *H. pylori*, as well as *Neisseria gonorrhoeae*, *Vibrio cholerae*, *Sulfurihydrogenibium yellowstonense*, *Sulfurihydrogenibium azorense* and *Ralstonia eutropha*), whereas β-carbonic anhydrases and γ-carbonic anhydrases have been identified in both Gram-negative and Gram-positive bacteria (Capasso and Supuran 2015).

These five different carbonic anhydrase classes are distinct in their secondary, tertiary and quaternary structures; specifically, the enzymes belonging to the α-class are usually monomeric, β-carbonic anhydrases are oligomeric with 2–8 monomers, and γ-carbonic anhydrases are homotrimers; whereas δ- and ζ-carbonic anhydrases are less well characterized. Despite the fact that these five classes do not share sequence or structural similarity, their active site structure and catalytic mechanism are common, a clear example of convergent evolution (Liljas and Laurberg 2000; Tripp et al. 2001). A peculiarity that differs these class of enzymes regards the catalytic site; carbonic anhydrases are generally Zn^{2+}-containing metalloenzymes, except for the ζ form, which uses Cd^{2+} ion as alternative metal cofactor (Xu et al. 2008), since these enzyme are diffuse in marine diatoms, environment characterized by a very low concentration of Zn^{2+} ion (Park, Song, and Morel 2007); additionally, γ-carbonic anhydrases contain Fe^{2+} ion in vivo, at least in anaerobic Archaea (Macauley et al. 2009). Since they are often essential for the survival of the organism, bacterial carbonic anhydrases are considered to be potential pharmacological targets for antibacterial drugs (Supuran 2011). Carbonic anhydrases from bacteria and archaea have been less studied structurally than their human counterpart; the crystal structures of carbonic anhydrases of various classes from 15 different bacteria and 2 archaea have been determined to date (Tab. 3.1).

More in detail, despite the structural differences and functional variety among carbonic anhydrases from various organisms, they exhibit a general catalytic mechanism, that has been extensively studied using the human carbonic anhydrase II as a model. The catalysis occurs via a two-step ping-pong mechanism, that consists in the nucleophilic attack of the hydroxide ion coordinated to the Zn^{2+} ion on a carbon dioxide molecule (Lindskog 1997).

$$EZnOH^- + CO_2 \leftrightarrow EZn(OH^-)CO_2 \leftrightarrow EZnHCO_3^- \leftrightarrow EZnH_2O + HCO_3^-$$

$$EZnH_2O \leftrightarrow H^+EZnOH^- + B \leftrightarrow EZnOH^- + BH^+$$

Fig. 3.1 General catalytic mechanism of α-carbonic anhydrase. Adapted from Becker et al., 2014.

In α-carbonic anhydrases, a member of which is examined in this Chapter, the binding of CO_2 in the hydrophobic region adjacent to the Zn^{2+} cofactor promotes the nucleophilic attack by the Zn-OH^- leading to the formation of HCO_3^- which is later displaced by the random diffusion of water in the active site. The transfer of a proton in the second step from the bound water molecule at the Zn^{2+} to an acceptor in the bulk solvent (B) is needed

to regenerate the OH^- ion for a subsequent round of catalysis *via* a proton shuttle histidine residue. The intramolecular proton transport between the Zn^{2+}-bound solvent and the proton shuttle histidine residue occurs *via* intervening water molecules in the active site. From the histidine residue, an intermolecular transfer event delivers the proton to the bulk solvent (Fig. 3.1; Becker, Klier, and Deitmer 2014). This transfer occurs on the order of 10^6 s^{-1} for the human carbonic anhydrase II and is the rate-limiting step of the overall maximum velocity of catalysis; this kinetic properties are also shared by the bacterial α-carbonic anhydrases (Chirica, Elleby, and Lindskog 2001; Russo et al. 2013; Vullo et al. 2013). Additionally, the α-carbonic anhydrases from various organisms share the structural organization of the active site, containing a Zn^{2+} ion coordinated by three histidine residues and a H_2O molecule/OH^- ion. Interestingly, α-carbonic anhydrases can also catalyze the hydrolysis of esters (Steiner and Lindskog 1972) and thioesters (Tanc et al. 2015); whereas no esterase activity was detected so far for enzymes belonging to the other four classes. A peculiar common feature of the bacterial α-carbonic anhydrases known to date is a N-terminal signal sequence indicating a periplasmic or extracellular location and a possible physiological role in CO_2 uptake processes. It can be assumed that the N-terminal sequence of a primordial carbonic anhydrase was changed by introducing a signal peptide and generating a new carbonic anhydrase class, named thereafter α-carbonic anhydrase, localized in a secretory compartment, such as the periplasmic space, in order to supply the metabolic needs also in these compartments (Capasso and Supuran 2015).

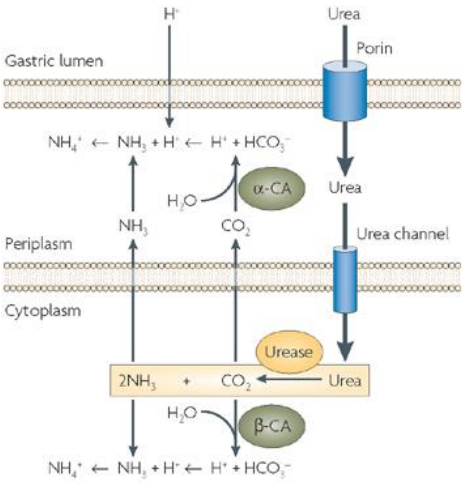

Fig. 3.2 Role of α-carbonic anhydrase and β-carbonic anhydrase in the maintenance of periplasmic pH in *H. pylori*. Adapted from Capasso and Supuran 2015.

H. pylori is a microaerophilic microorganism, which grows under atmospheres with low O_2 partial pressures and, as a capnophile, requires a CO_2-enriched atmosphere (5 to 10% CO_2). Thus, CO_2 appears to be an important component of the physiology of *H. pylori*, and it may also have a role in the pathogenicity of this bacterium (Bury-Moné et al. 2008). Two different carbonic anhydrases are coded by the genome of *H. pylori*: the periplasmic α-carbonic anhydrase and the cytoplasmic β-carbonic anhydrase. In *H. pylori* the role of α-carbonic anhydrase (named HP1186 in strain 26695 and HPG27_1129 in strain G27 examined, respectively) is fundamental in buffering the pH of the periplasm, since when the ammonia and carbon

dioxide produced by urease in the cytoplasm diffuse back, at least partially, into the periplasmic space, this catalyzes the conversion of CO_2 to HCO_3^-. It is reasonable to assume that β-carbonic anhydrase in the cytoplasm plays the same role for the CO_2 molecules that do not freely diffuse out of the inner membrane (Fig. 3.2). The α-carbonic anhydrase expression in *H. pylori* has been shown to be induced under acidic conditions by a two-component (ArsRS) system (Wen et al. 2007). The α-carbonic anhydrase has been shown to be essential for acid acclimatization (Marcus et al. 2005a) and some α-carbonic anhydrase defective *H. pylori* mutants exhibited only reduced stomach colonization in vivo, despite the fact that the enzyme does not seem to have an effect on urease activity in vitro (Bury-Moné et al. 2008). Both carbonic anhydrases from *H. pylori* have been proposed as alternative possible therapeutic targets for the treatment of patients infected by drug-resistant strains of the bacterium, and several inhibitors have been identified (Nishimori, Minakuchi, et al. 2006; Nishimori, Vullo, et al. 2006; Nishimori et al. 2007). Sulfonamides and their isosteres (sulfamates/sulfamides) constitute the main class of carbonic anhydrase inhibitors, which bind to the metal ion in the enzyme active site; an example of sulfonamide inhibitor, specifically acetazolamide (AAZ), bound to the catalytic site of α-carbonic anhydrase from *H. pylori* is reported in Fig. 3.3. Recently the dithiocarbamates, possessing a similar mechanism of action, were reported as a new class of inhibitors. Additionally, other families of compounds possess a distinct mechanism of action: phenols, polyamines, some carboxylates, and sulfocoumarins anchor to the Zn^{2+}-coordinated water molecule; coumarins and five/six-membered lactones are prodrug inhibitors, binding in hydrolyzed form at the entrance of the active site cavity. (Becker et al. 2014). Finally, a strong inhibition has been reported for inorganic and complex anions interacting with Zn^{2+} ion, such as cyanide, cyanate, hydrogen sulfide, divanadate, tellurate, perruthenate, selenocyanide, trithiocarbonate, iminodisulfonate. It has been observed that α-carbonic anhydrase and β-carbonic anhydrase from *H. pylori* are highly inhibited by many primary sulfonamides, including the clinically used acetazolamide, ethoxzolamide, methazolamide, topiramate and sulpiride (Nishimori, Minakuchi, et al. 2006; Nishimori et al. 2007). Furthermore, certain carbonic anhydrase inhibitors, such as acetazolamide and methazolamide, have been observed to inhibit the *H. pylori* growth in cell cultures (Nishimori et al. 2008). In addition, previous studies have shown that treating *H. pylori* with carbonic anhydrase inhibitors drastically reduces the ability of the bacterium to survive within an acid environment, suggesting that carbonic anhydrases are essential for gastric colonization (Bury-Moné et al. 2008).

In this context, the design of selective inhibitors for α-carbonic anhydrase could help to better clarify the effective role played by this enzyme in the control of pH and could eventually open the way to improved classes of therapeutics against the bacterium. The crystallization of α-carbonic anhydrase with acetazolamide in an orthorhombic crystal form has been published (Modak, Revitt-Mills, and Roujeinikova 2013) and its structure complexed with the clinically used compound sulfonamide has recently been published (Fig. ; Modak et al. 2015; PDB accession codes: 4YGF, 4YHA). In this research project the crystal structure of the native form of α-carbonic anhydrase from *H. pylori* has been taken into account and the possible implications for its function have been discussed.

	Type	**Number of structures**	**Ligands or variants**
Bacteria			
Escherichia coli	β, γ	7	Bicarbonate, MSE
Chlostridium difficile	γ	1	
Halothiobcillus neapolitanus	β	1	
Haemophilus influeanzae	β	17	Co^{2+} substituted, bicarbonate, Y181F, V47A, G41A, D44N, W39F, P48S/A49P, W39V/G41A, W39V/G41A/P48S/A49P
Micobacterium tubercolosis	β	6	Thiocyanate ion
Neisseria gonorreae	α	2	
Pseudomonas aeruginosa	β	2	CO_2
Salmonella enterica	β	1	
Streptococcus mutans	?	1	
Sulfurihydrogenibium azorense	α	1	AZM
Sulfurihydrogenibium yellostonense	α	1	AZM
Thermosynecoccus elongatus	γ	3	
Thermovibrio ammonificans	α	3	Sulfanilamide, AZM, B3P
Thiomicrospira crunogena	α	1	Bicarbonate
Vibrio cholerae	β	2	Bicarbonate
Archea			
Methanobacterium thermoautotrophicum	β	1	
Methanosarcina thermophila	γ	12	W19A, W19F, W19N, Y200A, Co substituted, bicarbonate

Tab. 3.1 List of bacteria and archaea whose carbonic anhydrase crystal structure is known. The third column refers to the number of files present in the PDB.

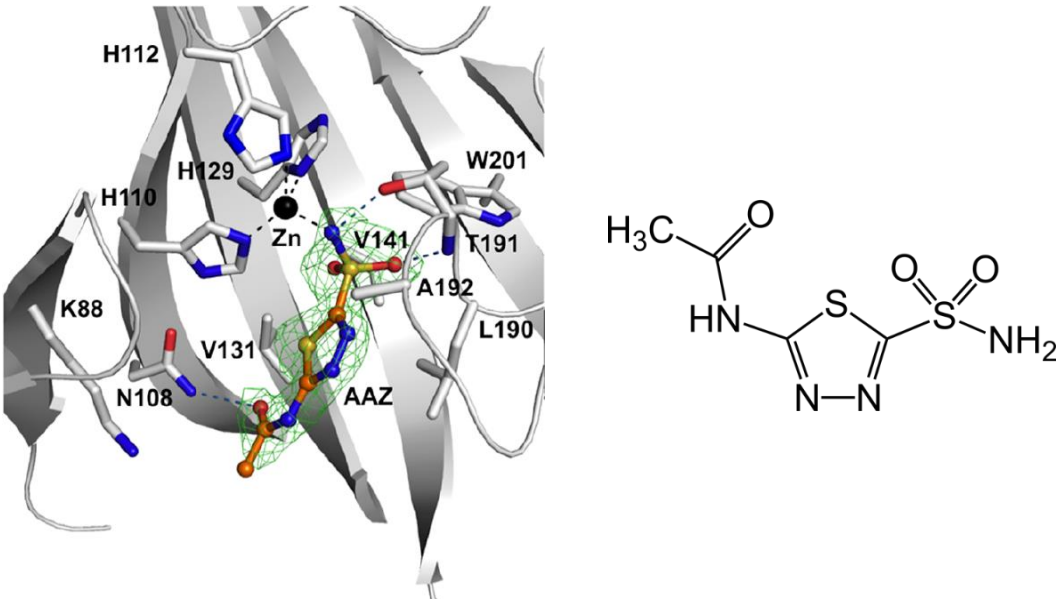

Fig 3.3 Binding of acetazolamide (AAZ) in the active site of α-carbonic anhydrase from *H. pylori*; right: electron density for AAZ bound to α-carbonic anhydrase is shown in green (Adapted from Modak et al., 2015); left: AAZ molecule (N-(5-sulfamoyl-1,3,4-thiadiazol-2-yl)-acetamide), Ki=21 nM (Nishimori, Minakuchi, et al., 2006)

3.2 SEQUENCE ANALYSIS

```
 10         20         30         40         50         60
MKKTFLIALA LTASLIGAEN AKWDYKNKEN GPHRWDKLHK DFEVCKSGKS QSPINIEHYY

         70         80         90        100        110        120
HTQDKADLQF KYAASKPKAV FFTHHTLKAS FEPTNHINYR GHDYVLDNVH FHAPMEFLIN

        130        140        150        160        170        180
NKTRPLSAHF VHKDAKGRLL VLAIGFEEGK ENPNLDPILE GIQKKQNFKE VALDAFLPKS

        190        200        210        220        230        240
INYYHFNGSL TAPPCTEGVA WFVVEEPLEV SAKQLAEIKK RMKNSPNQRP VQPDYNTVII


KRSAETR
```

Property	Value
Aminoacid	247
Cysteines	2
Molecular weight (kDa)	28324.3
Theoretical pI	9.19
Abs_{280nm} (C=1mg/mL; b=1cm)	1.056 (all Cys reduced)
Signal peptide	Yes

Tab. 3.2 Some properties of α-carbonic anhydrase from *H. pylori*

α-carbonic anhydrase from *H. pylori* G27 is composed of 247 aminoacids, including 2 cysteine residues, which will be demonstrated to be involved in an intramolecular disulfide bond formation. As mentioned before, the N-terminal end encodes a 21 aa secretion signal (Fig. 3.5), confirmed by SignalP 4.1 Server prediction (Petersen et al. 2011; http://www.cbs.dtu.dk/services/SignalP). The alignment of the aminoacid sequence of *H. pylori* α-carbonic anhydrase shows how the enzyme shares 39% identity with the α-carbonic anhydrase from *Sulfurihydrogenibium yellowstonense*, 37% identity with the enzyme from *Thermovibrio ammonificans*, and 36% identity with the enzyme from *Neisseria gonorrhoeae* (Fig. 3.4). Moreover, the sequence alignment reveals that the aminoacid residues important for the catalytic mechanism are highly conserved, namely three histidine residues (red labelling in Fig. 3.4).

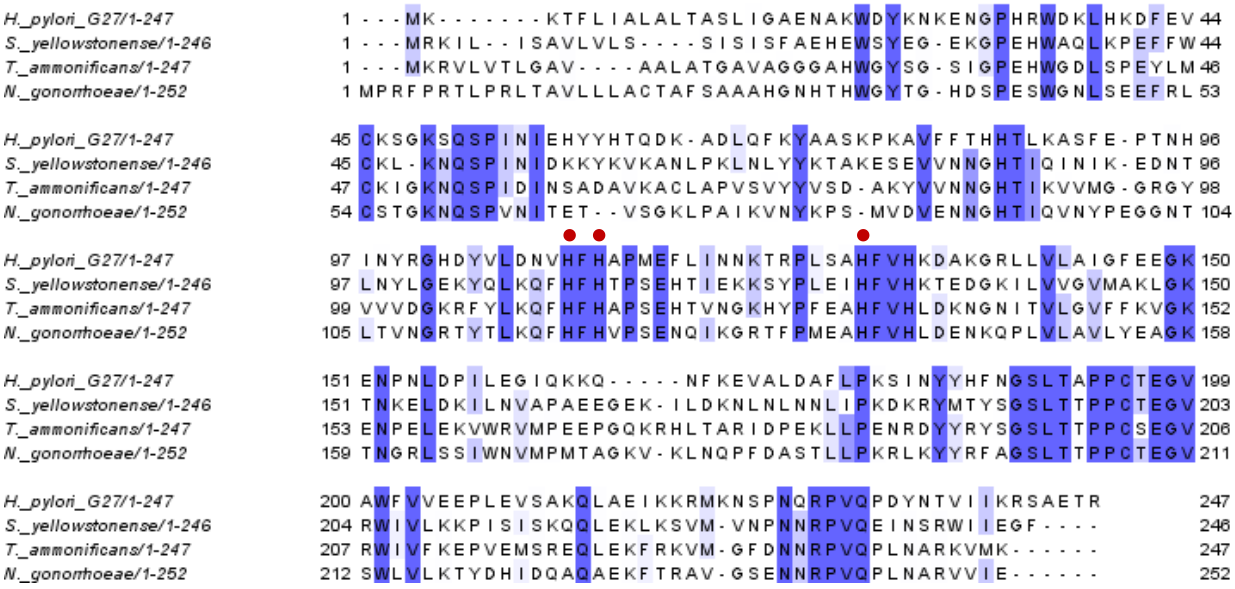

Fig. 3.4 Sequence alignment of α-carbonic anhydrase from *H. pylori* strain G27, *S. yellowstonense*, *T. ammonificans*, and *N. gonorrhoeae*; in shades of blue the identity above 50% is shown; red labelling denotes the catalytic histidine residues (ClustalW and Jalview ver. 2.9.0b2).

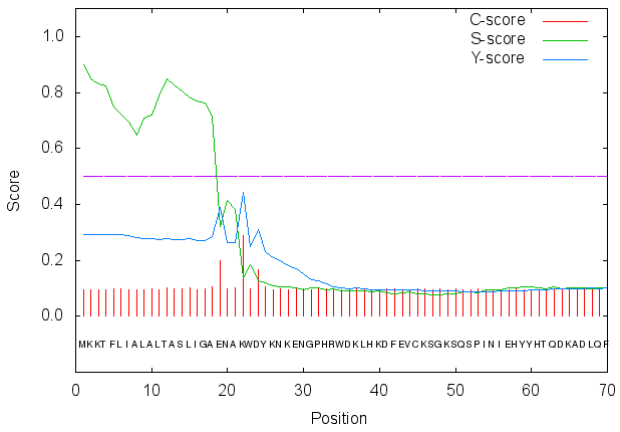

Fig. 3.5 α-carbonic anhydrase from *H. pylori* signal peptide prediction (SignalP 4.1 Server).

3.3 MATERIALS AND METHODS

3.3.1 CLONING, EXPRESSION AND PURIFICATION

The coding sequence for the α-carbonic anhydrase gene (HPG27_1129) was PCR-amplified from genomic *H. pylori* DNA (strain G27) using Phusion High-Fidelity DNA Polymerase (New England Biolabs) and primers 5'-CATCATCACCACCATCACGAAAACCTGTATTTTCAGGGAAAATGGGATTATAAAAATAAAGAA-3' (fw) and 5'-GTGGCGGCCGCTCTATTAGCGGGTCTCAGCTGA-3' (rv). Since the full-length protein is toxic to *Escherichia coli*, the N-terminal export signal sequence was excluded and a N-terminal 6-His-tag and a TEV proteolysis site were included. The following sequence corresponds to the final recombinant protein lacking of the N-terminal 21 aa export signal and added of the N-terminal 6-His-tag flanked by a TEV recognition site.

```
 10         20         30         40         50         60
HHHHHHENLY FQ//GKWDYKNK ENGPHRWDKL HKDFEVCKSG KSQSPINIEH YYHTQDKADL

        70         80         90        100        110        120
QFKYAASKPK AVFFTHHTLK ASFEPTNHIN YRGHDYVLDN VHFHAPMEFL INNKTRPLSA

       130        140        150        160        170        180
HFVHKDAKGR LLVLAIGFEE GKENPNLDPI LEGIQKKQNF KEVALDAFLP KSINYYHFNG

       190        200        210        220        230
SLTAPPCTEG VAWFVVEEPL EVSAKQLAEI KKRMKNSPNQ RPVQPDYNTV IIKRSAETR
```

// corresponds to the cleavage site for TEV protease

Property	Value
Aminoacid	239
Cysteines	2
Molecular weight (kDa)	27840.4
Theoretical pI	8.98
Abs_{280nm} (C=1mg/mL; b=1cm)	1.128 (all Cys reduced)
Signal peptide	No

Tab. 3.3 Some properties of α-carbonic anhydrase referred to the construct lacking of the N-terminal export signal and added of N-terminal 6-His-tag and TEV recognition site

E. coli BL21 (DE3) competent cells (Lucigen), harboring the pETite plasmid encoding for the recombinant construct, were grown in Luria–Bertani medium. Expression was induced by adding 0.5 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG) to the medium and was continued for 5 h at 30 °C with constant shaking. The cells were resuspended in 50 mM Tris–HCl pH 7, 150 mM NaCl, 5 mM imidazole supplemented with protease inhibitors (1 mM phenylmethanesulfonyl fluoride (PMSF), 15 mM aprotinin, 1 mM leupeptin) and lysed using a One Shot Cell disruption system (Constant Systems Ltd). The lysate was centrifuged at 18,000 rpm for 20 min at 4 °C to separate the supernatant from the insoluble fraction. The soluble fraction was loaded onto a 1 mL HisTrap HP Ni–NTA column (GE Healthcare) pre-equilibrated with lysis buffer. The column was extensively washed with buffer A and the protein was eluted using a linear gradient from 350 to 500 mM imidazole. The protein was eluted as a single species (Figg.) and was further purified by buffer exchange using a PD-10 desalting column (GE Healthcare) equilibrated with a buffer consisting of 50 mM Tris–HCl pH 7, 150 mM NaCl. The His-tag was removed by incubation with TEV protease (Sigma–Aldrich) in a 1:100 ratio overnight at 30 °C. The reaction mixture was buffer-exchanged with buffer consisting of 50 mM Tris–HCl pH 8, 500 mM NaCl, 15 mM imidazole, 1% (v/v) glycerol. The cleaved protein was isolated as an unbound sample by loading it onto an Ni Sepharose 6 Fast Flow column (GE Healthcare) equilibrated with buffer. The protein was buffer-exchanged into 30 mM Tris–HCl pH 8, concentrated to 20 mg/ml using a Vivaspin 20 5,000 MWCO centrifugal concentrator (Sartorius) and stored at -20 °C for crystallization trials.

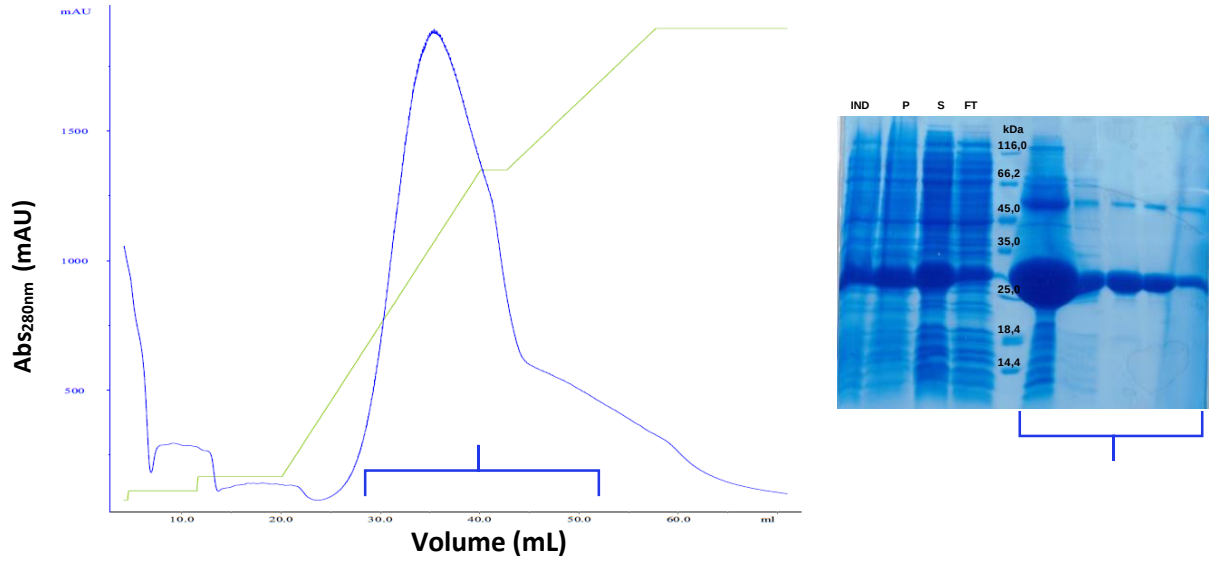

Fig. 3.6 6-His-tagged recombinant α-carbonic anhydrase affinity chromatography; left: absorption profile; right: SDS-PAGE analysis of the fractions collected (IND: induced expression; P: lysed pellet; S: supernatant; FT: flow-throw).

3.3.2 CRYSTALLIZATION

The purified protein was concentrated to 20 mg/ml and used in crystallization tests, which were partially automated using an Oryx8 crystallization robot (Douglas Instruments, UK). The best crystals (Figs. 3.7 and 3.8) were obtained at 4 °C by the sitting-drop vapor-diffusion technique using a solution consisting of 0.2 M sodium nitrate, 0.1 M bis-tris propane pH 8.5, 20% (w/v) PEG 3350 as precipitant (PACT Suite solution n. 89, Qiagen). Moreover, soaking and cocrystallization experiments were performed in order to investigate the molecular binding of inhibitor compounds to the active site; specifically, topiramate and etoxybenzothiazolesulfonamide were selected, since no cocrystal structures of bacterial α-carbonic anhydrase with these compounds have been published. The protein/inhibitor ratio used was 1:20 and crystallization attempts were performed by sitting-drop vapor-diffusion and microbatch-under-oil techniques. Cocrystals suitable to x-ray diffraction measurement have not been obtained yet.

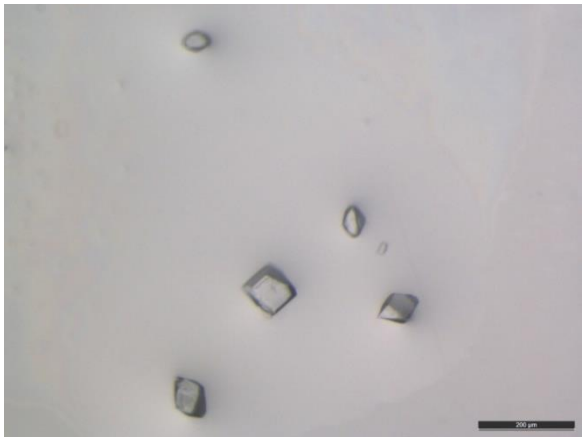

Fig. 3.7 Crystals of recombinant α-carbonic anhydrase from *H. pylori* grown at 20 mg/mL protein concentration, in 0.2 M sodium nitrate, 0.1 M bis-tris propane pH 8.5, 20% (w/v) PEG 3350, at 4 °C.

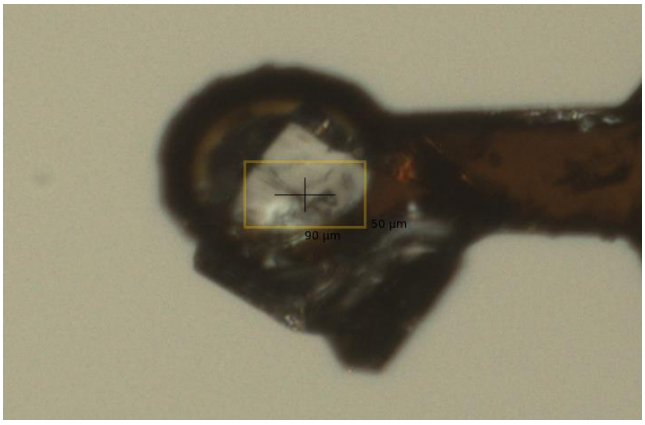

Fig. 3.8 Crystal of recombinant α-carbonic anhydrase from *H. pylori* grown in PACT Suite n. 89 mounted on a loop.

3.3.3 DATA COLLECTION AND PROCESSING

Diffraction data were measured on the PXIII beamline at the SLS synchrotron (Villigen, Switzerland). The crystal was found to belong to the monoclinic space group $P2_1$, with unit cell parameters *a*=44.906, *b*=95.905, c=53.318 Å, β=92.92. Two molecules are present in the asymmetric unit, corresponding to a V_M of 2.03 $Å^3$/Da and an approximate solvent content of 39%. All data sets were indexed and integrated with XDS (Kabsch 2010) and merged and scaled with SCALA (Evans 2006) as contained in the CCP4 crystallographic package (Winn et al. 2011; Tab. 3.4).

Data Collection and processing	
Diffraction source	Beamline PXIII, SLS
Wavelength (Å)	1.00
Temperature (K)	100
Detector	Pilatus 2M
Rotation range per image (°)	0.1
Total rotation range (°)	180
Exposure time per image (s)	0.1
Space group	P 2_1
Cell parameters (Å, °)	*a*=44.906, *b*=95.905, *c*=53.318, β=92.92
Resolution range (Å)	47.961 – 1.517 (1.60-1.517)
Total n. of reflections	229967
N. of unique reflections	69653 (8705)
Completeness (%)	96.7 (85.8)
Multiplicity	3.3 (3.0)
‹*I*/σ(*I*)›	18.5 (3.0)
R_{sym}	0.037 (0.382)
$R_{p.i.m.}$	0.024 (0.256)
Overall B factor from Wilson plot (Å²)	16.97

Tab. 3.4 Statistics on data collection and processing relative to recombinant α-carbonic anhydrase from *H. pylori*; values in parentheses are for the outer shell.

3.3.4 STRUCTURE SOLUTION AND REFINEMENT

The structure was solved by molecular replacement using MOLREP (Vagin and Teplyakov 2010), starting from a model built using the SWISS-MODEL server (Biasini et al. 2014; http://swissmodel.expasy.org/) from PDB accession code 4G7A (Di Fiore et al. 2013). The rebuilding procedure available in the PHENIX package (Adams et al. 2010) was used to rebuild the model, which was subsequently checked and adjusted with Coot (Emsley et al. 2010). Refinement was continued with PHENIX (Adams et al. 2010). The final statistics of the refinement are summarized in Tab. 3.5.

Structure solution and refinement	
Resolution range (Å)	47.96–1.52
Completeness (%)	96.7
σ Cutoff	0
N. of reflections, working set	63791
N. of reflections, test set	3384
Final R_{cryst}	0.1696
Final R_{free}	0.1984
N. of non-H atoms total protein ions others water	 4396 3721 4 8 662
R.m.s. deviations Bond length (Å) Bond angles (°)	 0.007 1.15
Average B factor (Å^2)	22.6
Ramachandran plot Most favored (%) Allowed (%)	 96.0 3.6
Ramachandran outliers (%)	0.4
Rotamer outliers (%)	1.2
Overall score	1.65

Tab. 3.5 Statistics on structure solution and refinement relative to recombinant α-carbonic anhydrase from *H. pylori*.

3.4 RESULTS AND DISCUSSION

3.4.1 OVERALL FOLD OF THE ENZYME

The asymmetric unit of the α-carbonic anhydrase crystal contains two monomers. Their structure is essentially the same (r.m.s.d. of 0.61 Å), with the exception of the long stretch from residues 61 to 68. Each monomer includes 226 residues (22–247; Fig. 3.9; the first 21 amino acids were not included in the gene cloned, since they are predicted to correspond to a signal sequence for export into the periplasmic space of the bacterium). The entire polypeptide chain is very well defined in the electron-density map, with the exception of amino acids 163–167 of monomer B, a stretch that is exposed to the solvent and is far away from the active site and from the contact region with the other monomer (see below). The fold of the *H. pylori* α-carbonic anhydrase monomer corresponds to that of the classical α-carbonic anhydrase, characterized by a central ten-stranded β-sheet

surrounded by three α-helices and by the remainder of the protein chain (Fig. 3.10). One layer of the twisted β-sheet along with another portion of the polypeptide chain defines a conical-shaped cavity hosting a Zn^{2+} ion at the bottom. The latter defines the active site of the enzyme. The structure of the monomer is stabilized by an intramolecular disulfide bond (Cys45–Cys195) that is conserved in most carbonic anhydrase structures(Di Fiore et al. 2013). Since this disulfide bond connects the N-terminus to a loop (residues 192–195) that surrounds the entrance of the catalytic site cavity, it is tempting to speculate that its function is to keep the cavity well opened in order to favour entrance of the substrate.

```
B5Z8I0 1   ---MK-------KTFLIALALTASLIGAENAKWDYKNKENGPHRWDKLHKDFEVCKSGKSQSPINIEHYYHTQDK-ADLQFKYAASKPKAVFFTHHTLKASFE-PTNHINYRGHDYVLDN  108
B2V8E3 1   ---MRKIL--ISAVLVLS----SISISFAEHEWSYEG-EKGPEHWAQLKPEFFWCKL-KNQSPINIDKKYKVKANLPKLNLYYKTAKESEVVNNGHTIQINIK-EDNTLNYLGEKYQLKQ  108
E8T502 1   ---MKRVLVTLGAV----AALATGAVAGGGAHWGYSG-SIGPEHWGDLSPEYLMCKIGKNQSPIDINSADAVKACLAPVSVYYVSD-AKYVVNNGHTIKVVMG-GRGYVVVDGKRFYLKQ  110
Q50940 1   MPRFPRTLPRLTAVLLLACTAFSAAAHGNHTHWGYTG-HDSPESWGNLSEEFRLCSTGKNQSPVNITET--VSGKLPAIKVNYKPS-MVDVENNGHTIQVNYPEGGNTLTVNGRTYTLKQ  116
                  :         .         :        .* *     .*. * .*  ::  *.  *.***::*     ..     :.. *       *  . **::         :   *. : *.:
B5Z8I0 109 VHFHAPMEFLINNKTRPLSAHFVHKDAKGRLLVLAIGFEEGKENPNLDPILEGIQKKQ-----NFKEVALDAFLPKSINYYHFNGSLTAPPCTEGVAWFVVEEPLEVSAKQLAEIKKRMK  223
B2V8E3 109 FHFHTPSEHTIEKKSYPLEIHFVHKTEDGKILVVGVMAKLGKTNKELDKILNVAPAEEGEK-ILDKNLNLNNLIPKDKRYMTYSGSLTTPPCTEGVRWIVLKKPISISKQQLEKLKSVM-  226
E8T502 111 FHFHAPSEHTVNGKHYPFEAHFVHLDKNGNITVLGVFFKVGKENPELEKVWRVMPEEPGQKRHLTARIDPEKLLPENRDYYRYSGSLTTPPCSEGVRWIVFKEPVEMSREQLEKFRKVM-  229
Q50940 117 FHFHVPSENQIKGRTFPMEAHFVHLDENKQPLVLAVLYEAGKTNGRLSSIWNVMPMTAGKV-KLNQPFDASTLLPKRLKYYRFAGSLTTPPCTEGVSWLVLKTYDHIDQAQAEKFTRAV-  234
           .***.* *  :: :  *:. ****   . .  *:.:  : ** * .*. : .              .  . ::*:   *  : ****:***:*** *:*.:    :.  *  ::   :

B5Z8I0 224 NSPNQRPVQPDYNTVIIKRSAETR  247
B2V8E3 227 VNPNNRPVQEINSRWIIEGF----  246
E8T502 230 GFDNNRPVQPLNARKVMK------  247
Q50940 235 GSENNRPVQPLNARVVIE------  252
              *:****      :::
```

Fig. 3.9 Sequence alignment of *H. pylori* α-carbonic anhydrase (B5Z8IO) with α-carbonic anhydrases from *S. yellowstonense* (B2V8E3), *T. ammonificans* (E8T502) and *N. gonorrhoeae* (Q50940). Red, yellow and pale green backgrounds denote the residues involved in the binding of Zn^{2+}, in the conical surface of the active-site cavity and in the areas of contact in the dimer, respectively.

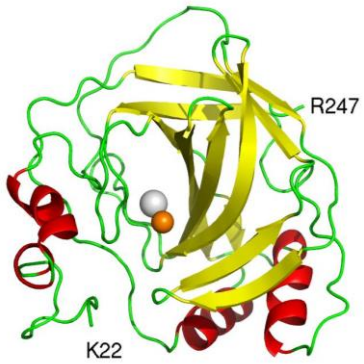

Fig. 3.10 Cartoon view of the monomer of *H. pylori* α-carbonic anhydrase, coloured according to secondary-structure element. The Zn^{2+} and Cl^- ions are shown as orange and yellow spheres, respectively.

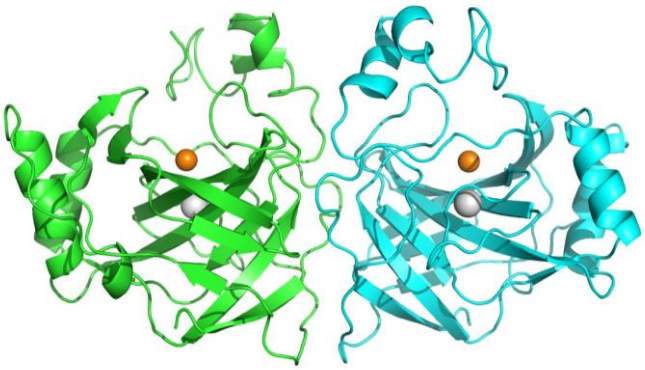

Fig. 3.11 The *H. pylori* α-carbonic anhydrase dimer present in the asymmetric unit of the crystal. A molecular twofold axis runs approximately vertically in the plane of the paper.

3.4.2 PROTEIN DIMERIZATION

The two monomers in the asymmetric unit (Fig. 3.11) are related each other by a rotation axis of 177°. The dimer is stabilized mostly by hydrophilic interactions, in particular by the formation of 18 hydrogen bonds between protein atoms. Other hydrophilic interactions

are mediated by solvent molecules. The surface concealed after formation of the dimer is 1226 Å^2 per monomer, corresponding to about 11% of the total surface of the monomer. This value is relatively low and analysis with PISA (Krissinel and Henrick 2007) assigns quite a low score (0.123) to the dimer formation, suggesting that this interface may play only an auxiliary role in dimer formation. A size-exclusion chromatography experiment (Fig. 3.12) confirmed that the enzyme is present as a monomeric species in solution. Altogether, these data suggest that the dimer observed is the result of crystal packing and that the physiological state of *H. pylori* α-carbonic anhydrase is monomeric, analogously to all mammalian α-carbonic anhydrases. Nevertheless, it is surprising that the dimer observed is the same dimer as found in the α-carbonic anhydrases from the other bacteria *S. yellowstonense* (Di Fiore et al. 2013) and *N. gonorrhoeae* (Huang et al. 1998) and in the dimer of the tetrameric *T. ammonificans* α-carbonic anhydrase (James et al. 2014). This occurs despite the regions involved in the dimerization surface not being particularly conserved. Since the interaction surface between the two monomers is essentially hydrophilic, the equilibrium between monomer and dimer possibly depends on the environmental conditions (ionic strength, pH), which in the case of the *H. pylori* periplasm can change significantly according to the pH of the host stomach. Moreover, the two active sites in the dimer are quite distant and independent, and dimerization does not affect the entrance to the two binding sites.

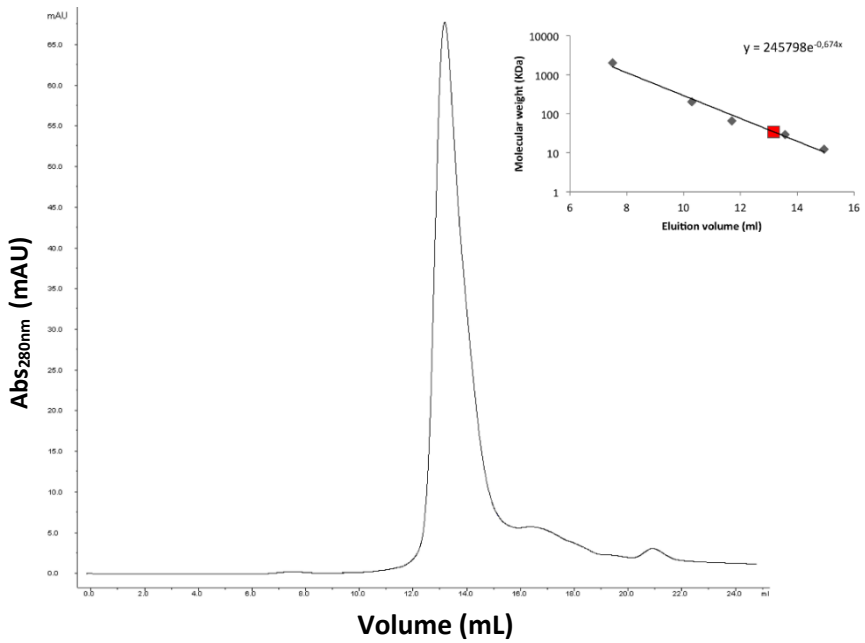

Fig. 3.12 A gel-filtration experiment showing that the apparent molecular mass of *H. pylori* α-carbonic anhydrase (red square) is about 34,000 Da, which is slightly larger than the theoretical calculated mass (26,165.7 Da, referred to the protein removed of 6-His-tag) but significantly smaller than that of a dimer. The small grey square close to it corresponds to bovine α-carbonic anhydrase.

3.4.3 THE ACTIVE SITE

The Zn^{2+} ion present in the active site of *H. pylori* carbonic anhydrase presents a slightly distorted trigonal bipyramidal coordination: the five ligands are the three N atoms of three histidine residues conserved in all other α-carbonic anhydrases for which structures have been determined to date and two solvent molecules (Fig. 3.13). Coordination distances for the Zn^{2+} ion range from 2.14 to 2.19 Å in the two monomers, and from 2.12 to 2.30 Å for the solvent O atoms. The Zn^{2+} coordination in other bacterial α-carbonic anhydrases involves the same three histidines in all cases and varies from trigonal bipyramidal to tetrahedral according to the ligand bound. Another two important residues in the vicinity of the active site, His191 and Glu116, are conserved in several α-carbonic anhydrases. The former interacts with one of the two solvent molecules that coordinate the Zn^{2+} ion, while the latter does not form any significant interactions with active-site residues in the *H. pylori* α-carbonic anhydrase structure, but its negative charge is possibly fundamental to balance the global positive charge of the active site or to orient the substrate (Vullo et al. 2013). A significant difference of *H. pylori* α-carbonic anhydrase with respect to the other bacterial α-carbonic anhydrases is represented by the environment of the active site. In all of them one of the histidines (corresponding to His129 in *H. pylori* α-carbonic anhydrase) interacts with a conserved glutamic acid (position 127 in *H. pylori* α-carbonic anhydrase), the negative charge of which partially neutralizes the positive charge of the histidine–Zn^{2+} complex and possibly stabilizes it. In the structure of *H. pylori* α-carbonic anhydrase, aminoacid 127 is a serine and an atom heavier than a water molecule is present in the position occupied by the glutamate carboxylic group in the other bacterial α-carbonic anhydrases. This atom interacts with $N^{\varepsilon 2}$ of His129 (at a distance of 3.11 Å) and O^{γ} of Ser127 (at a distance of 3.15–3.16 Å). It is reasonable to assume that this atom must be negatively charged to compensate for the absence of the glutamate in this position. An anomalous Fourier difference map with data measured at 1.000 Å wavelength presents a large peak (about 20σ) corresponding to the Zn^{2+} position and a smaller peak (about 5σ) corresponding to this unknown atomic species (Fig. 3.13). This value corresponds to that for the S atoms of methionines and cysteines visible in our map. At this wavelength the *f''* values for Zn^{2+} and S are 2.6 e and 0.265 e, respectively. The only monoatomic anion present in solution is Cl^-, the *f''* value of which is 0.321 e, a value close to that of S. This strongly supports the presence of a Cl^- ion at this position. The latter would compensate for the absence of the negative charge of the glutamate.

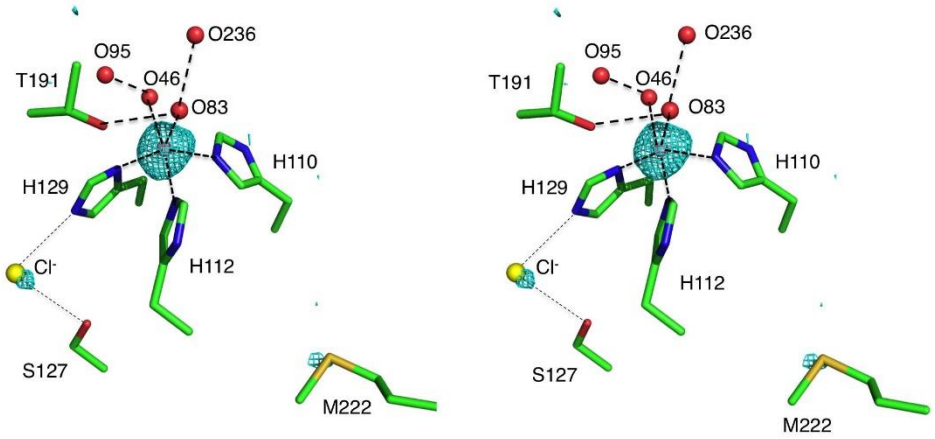

Fig. 3.13 Stereographic view of the enzyme active site, showing the Zn^{2+} ion coordinated by three histidine residues and two solvent molecules. A Cl^- interacting with His129 is shown as a yellow sphere. The cyan grid represents the anomalous-difference map, contoured at 5σ contour level. The position of Met222 side chain is also shown for comparison.

3.4.4 COMPARISON WITH OTHER α-CARBONIC ANHYDRASE STRUCTURES

The overall architecture of *H. pylori* α-carbonic anhydrase is quite similar to that of α-carbonic anhydrases from other bacteria: superposition of the corresponding C^{α} atoms of one monomer with those of the enzyme with the most similar amino-acid sequence, that from *S. yellowstonense* (Di Fiore et al. 2013; PDB accession code: 4G7A), which presents 39% identity to the enzyme of interest, gives an r.m.s.d. of 1.49 Å for 224 residues; on superposition with *T. ammonificans* α-carbonic anhydrase (37% sequence identity; James et al. 2014; PDB accession code: 4COQ) the r.m.s.d. is 1.33 Å for 211 amino acids and with *N. gonorrhoeae* α-carbonic anhydrase (36% identity; Huang et al. 1998; PDB entry 1KOP) the r.m.s.d. is 1.57 Å for 212 residues. Some structural differences among α-carbonic anhydrases from different bacteria are possibly represented by the entrance to the active site and by the surface of the conical cavity that gives access to the active site (De Simone, Alterio, and Supuran 2013). The residues lining the surface of the cavity are highlighted in Fig. 3.9 by a yellow background. In Fig. 3.14a, in which residues are coloured according to their potential charge, it is possible to see that half of the entrance of the active-site cavity is charged, with a prevalence of positive charges, in particular if the histidine is protonated, whilst the other half is neutral and is mostly hydrophobic.

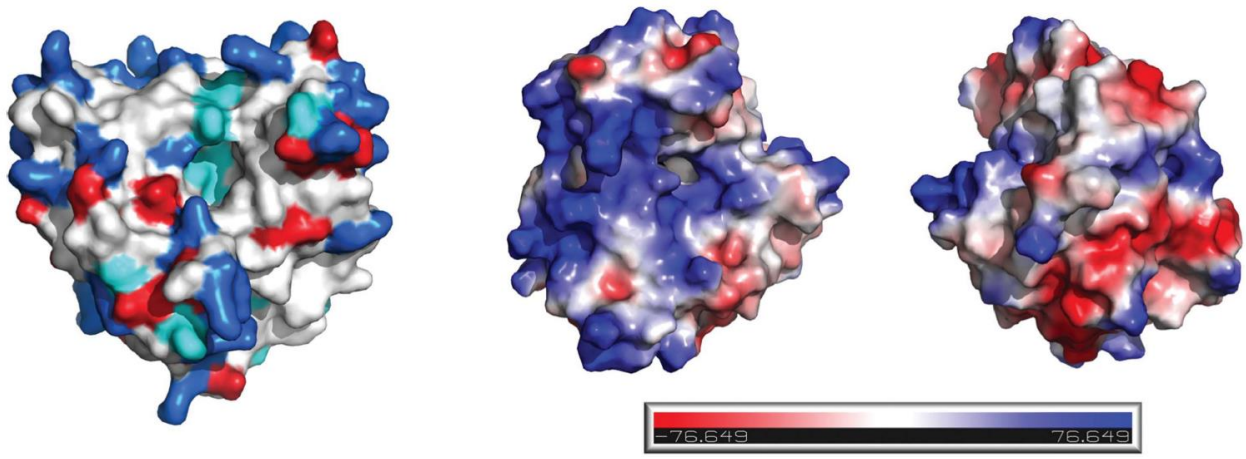

Fig. 3.14 a: Surface of the *H. pylori* α-carbonic anhydrase monomer. Potentially charged residues are coloured blue (Arg and Lys), red (Glu and Asp) and cyan (His). The bottom of the central cavity in the center appears in cyan owing to the presence of the three histidine residues that coordinate the Zn^{2+}. **b: Qualitative electrostatic surface of the *H. pylori* α-carbonic anhydrase monomer.** On the left it is possible to observe the access to the active site, whilst on the right the molecule is rotated 180° in the vertical direction.

3.4.5 LOCALIZATION

H. pylori α-carbonic anhydrase has been detected bound to the inner membrane of the periplasmic space and was not found in the soluble fraction in a Western blot experiment (Marcus et al. 2005a). The crystal structure presented confirms that it is a soluble protein and its surface presents several positively and negatively charged residues (Fig. 3.14b). A qualitative electrostatic potential calculation indicates that positive electrostatic potentials are prevalent on the protein surface, in line with an isoelectric point of 9.1 as estimated from the amino-acid sequence (referred to the protein removed of 6-His-tag). An analysis of the distribution of these charges shows that positive potential is mostly located on the face of the protein containing the opening of the active-site cavity, whilst on the rest of the surface positive and negative potentials are more randomly distributed. It is possible to hypothesize that the positive surface of *H. pylori* α-carbonic anhydrase serves to interact with the negatively charged phospholipids of the membrane. If this is true, then the enzyme would present the active-site entrance roughly oriented towards the membrane in such a way that when CO_2 diffuses from the cytoplasm the enzyme could capture the gas directly as soon as it flows through the membrane.

Chapter IV

CLONING, EXPRESSION AND PURIFICATION OF β-CARBONIC ANHYDRASE FROM *HELICOBACTER PYLORI*

4.1 INTRODUCTION

β-carbonic anhydrase, as well as α-carbonic anhydrase, plays a crucial role in *H. pylori* complex and delicate balance of urea and bicarbonate aimed to survival in the extremely acid conditions of the stomach. Therefore, even this enzyme can be considered as an interesting pharmacological target to develop new potential antibiotic therapies, affecting a critical metabolic network in *H. pylori* survival. Moreover, since the carbonic anhydrases belonging to the β-class are present in pathogenic microorganisms (such as fungi, bacterial and protozoa), and they lack from mammals, in which α-carbonic anhydrase class is physiologically relevant, these enzymes started to be more widely accepted as possible drug targets for developing new antibacterial agents, able thus to overcome the important resistance problems, which are affecting most classes of antibiotics in clinical use (Ceruso et al. 2015).

The β-carbonic anhydrase class is broadly distributed in nature, including archaea, bacteria, algae, plant chloroplasts and invertebrates (Rowlett 2010). Phylogenetic analyses indicate that this enzymatic class is far more diverse in sequence than the other classes, with only five residues (three forming the Zn^{2+} ligands plus an aspartate and an arginine) being completely conserved (Smith et al. 2000). Based on sequence identity, the β-carbonic anhydrase class can be divided into seven clades (A-G), with the plant enzymes forming two clades representing dicotyledonous and monocotyledonous plants. Enzymes within these clades can vary with respect to structure and their response to inhibitors, suggesting different functional mechanisms of action (Smith and Ferry 1999). Moreover, although β-carbonic anhydrases catalyze the same reaction as α-carbonic anhydrases and other forms, important structural differences between these classes exist. To date, the X-ray crystallographic structures of β-carbonic anhydrases deposed in the Protein Data Bank (http://www.pdb.org) are definitely not as many as those of the well-characterized α-carbonic anhydrases; crystal

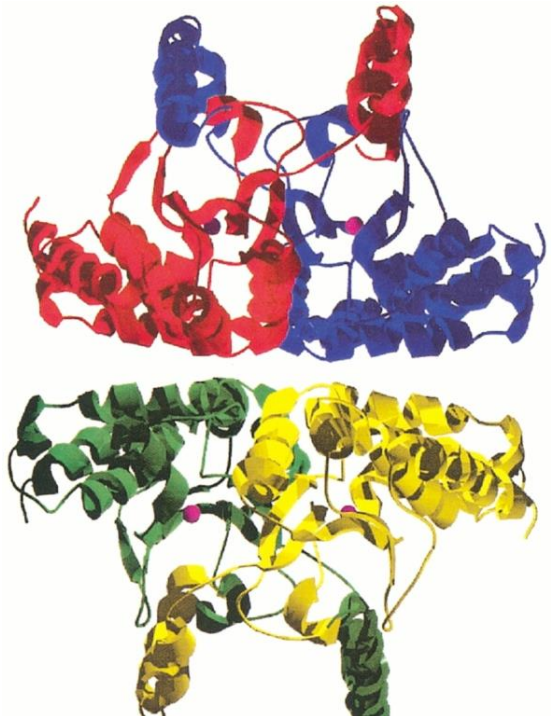

Fig. 4.1 β-carbonic anhydrase from *Escherichia coli*, tetrameric association. Adapted from Cronk et al., 2001.

structures from *Pisum sativum* (Kimber & Pai, 2000; PDB accession code: 1EKJ), the red alga *Porphyridium purpureum* (Mitsuhashi et al., 2000; PDB accession code: 1DDZ), *Methanobacterium thermoautotrophicum* (Strop, Smith, Iverson, Ferry, & Rees, 2001; PDB accession code: 1G5C), *Escherichia coli*, the first bacterium in which β-carbonic anhydrase structure has been determined (Fig. 4.1; Cronk, Endrizzi, Cronk, & Neill, 2001; PDB accession codes: 1I6O, 1I6P; Fig.), *Mycobacterium tuberculosis* (Covarrubias et al., 2005; PDB accession code: 1YM3), *Haemophilus influenzae* (Cronk et al., 2006; PDB accession code: 2A8D), *Halothiobacillus neapolitanus* (Sawaya et al., 2006; PDB accession code: 2FGY), the green alga *Coccomyxa* (Huang et al., 2011; PDB accession codes: 3UCJ, 3UCK, 3UCM, 3UCN, 3UCO), *Sordaria macrospora* (Lehneck et al., 2014; PDB accession codes: 4O1J, 4O1K), *Pseudomonas aeruginosa* (Pinard et al., 2015; PDB accession code: 4RXY) and *Vibrio cholerae* (Ferraroni, Del Prete, Vullo, Capasso, & Supuran, 2015; PDB accession code: 5CXK) have been published. X-ray crystallographic analysis of β-carbonic anhydrases has revealed two distinct subtypes of this enzymatic class based on active-site organization, denoting type I and type II β-carbonic anhydrases. In the type I enzymes, exemplified by β-carbonic anhydrases from *P. sativum*, *M. thermoautotrophicum* and *M. tuberculosis* Rv1284, the active site Zn^{2+} ion is coordinated by one histidine and two cysteine residues (instead of three histidine residues present in α-, γ- and δ-classes), with a fourth coordination site occupied by water or a substrate analogue, the so-called “open conformation”. In contrast, the other subclass of β-carbonic anhydrases, the type II, exemplified by the enzymes from *H. influenzae*, *E. coli*, *P. purpureum* and *M. tuberculosis* Rv3588c, has a unique Zn^{2+}-coordination geometry in which the water molecule is replaced by an aspartate side chain, forming a non-canonical active site, namely the “closed conformation” (Ferraroni et al. 2015). These differences determine the catalytic properties of these enzymes. The type II β-carbonic anhydrases are characterized by little or no CO_2 hydration activity at pH values less than 8.0, compared with the type I enzymes that show catalytic activity at pH values from as low as 6.5 to greater than 9.0. Moreover, type II β-carbonic anhydrases show highly cooperative pH-rate profiles and cooperative inhibition by HCO_3^-, whereas type I are not inhibited by HCO_3^- (Rowlett 2010). Therefore, it has been hypothesized that the closed conformation observed in the structures of type II β-carbonic anhydrases is an allosteric form of the enzyme, the so-called T state, that is an inactive form present at pH values below 8.0. This conformation is stabilized by the presence of HCO_3^-. However, at pH values >8.3 the closed active site is converted to an open one, with an incoming water

molecule replacing the carboxylate moiety of the aspartate residue and thus generating the nucleophile used in the catalytic cycle. Indeed, at this pH value the carboxylate of the aspartate residue makes a strong interaction with the guanidine/guanidinium moiety of a conserved arginine residue present in all β-carbonic anhydrases so far investigated (Ferraroni et al. 2015). Moreover, instead of functioning as obligate monomers like most of α-forms or trimers like γ-forms, β-carbonic anhydrases are found in many oligomerization states; especially, dimeric, tetrameric, hexameric and octameric β-carbonic anhydrase structures have been reported (Syrjänen et al. 2010). Additionally, in contrast to the overall structure of the α-carbonic anhydrases (extended 10-strand twisted β-sheet, flanked by six or more α-helices), the β-enzymes have more compact structures: a β-sheet core composed of four or five strands, and four or more α-helices surrounding this core (Aggarwal et al. 2015). Despite these differences, β-carbonic anhydrase class basically shares the same two-step ping-pong molecular mechanism for reversible hydration of carbon dioxide into bicarbonate, as α-carbonic anhydrase class, even if the mechanism of proton transfer involves different residues between the two enzymatic classes (Fig. 4.2).

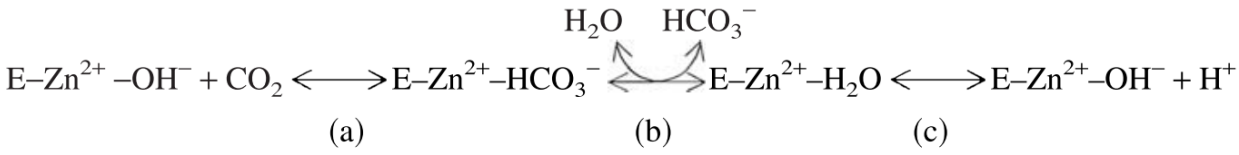

Fig. 4. 2 ***General catalytic mechanism of β-carbonic anhydrase.*** The essential elements of this mechanism are (a) nucleophilic attack of the metal-activated hydroxide ion on carbon dioxide, (b) ligand exchange of the product bicarbonate for a water molecule, and (c) regeneration of the zinc hydroxide form of the enzyme. Adapted from Cronk et al., 2001.

In α-carbonic anhydrases, this function is conducted by a well-ordered network of water molecules and a histidine proton shuttle residue (Aggarwal et al. 2014). In contrast, as previously mentioned, β-class enzymes possess a highly conserved dyad comprising an aspartate residue and an arginine one that seem to be crucial for the catalytic mechanism since mutation of these residues severely reduces the catalytic activity. Particularly, the aspartate makes a hydrogen bond with the Zn^{2+}-coordinated water molecule, activating it for nucleophilic attack of the CO_2 molecule (Fig. 4.3; K. S. Smith, Ingram-Smith, and Ferry 2002).

β-carbonic anhydrases are important accessory enzymes for many CO_2 or HCO_3^--utilizing enzymes, such as RuBisCO in chloroplasts, cyanase in *E. coli*, carboxylases in

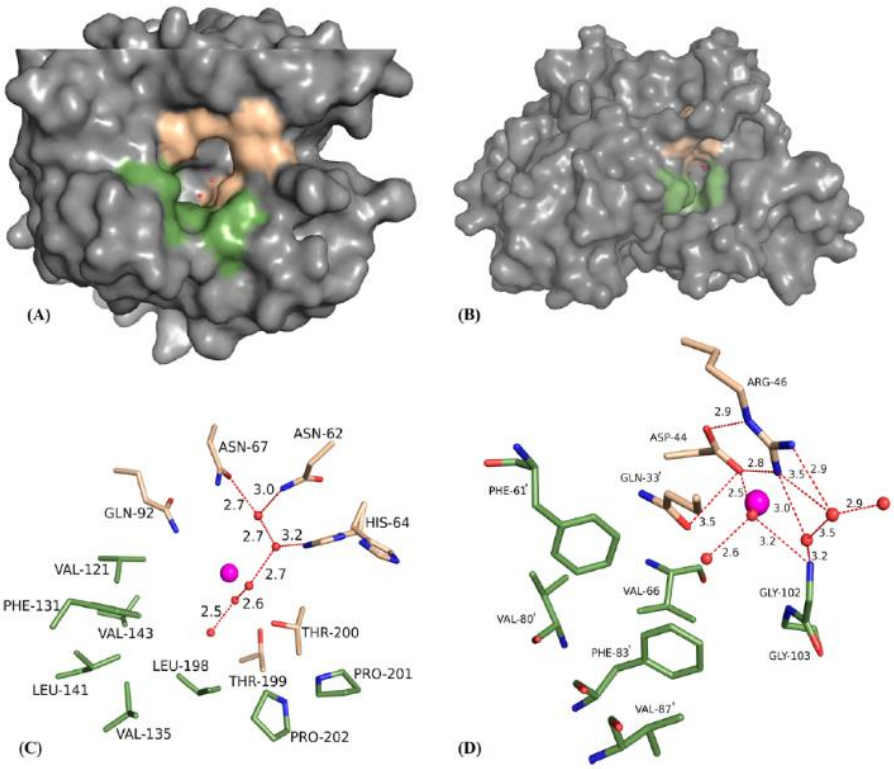

Fig. 4.3 Surface representation of (A) hCA II (human α-carbonic anhydrase II) and (B) psCA3 (β-carbonic anhydrase from *Pseudomonas aeruginosa*). Beige and green regions represent hydrophilic and hydrophobic residues of the active site, respectively. **Stick representations of the active site for (C) hCA II and (D) psCA3.** The active site zinc is depicted as a magenta sphere, ordered waters depicted as red spheres, H-bonds represented by red dashes, and distances given in angstroms. Adapted from Aggarwal et al., 2015.

Corynebacterium glutamicum, and also urease in *H. pylori* (Zolfaghari Emameh et al. 2014). Their importance in prokaryotic biology can be deduced from their widespread presence in metabolically diverse species. The β-carbonic anhydrases play an essential role in facilitating aerobic growth of microbes at low partial pressures of CO_2 by providing endogenous HCO_3^- and are involved in multiple roles such as cyanate degradation, host colonization, host survival, and growth in different organisms (Aggarwal et al. 2015). In plants and photosynthetic bacteria, β-carbonic anhydrases are required for transport and maintenance of CO_2 and HCO_3^- concentrations for carbon fixation and photosynthesis; they might possess different roles, depending on the location of the enzyme and the type of plant. β-carbonic anhydrases are strongly expressed in both roots and green tissues and are located in chloroplasts, cytoplasm and mitochondria with isozyme-specific patterns (Fabre et al. 2007). The highest carbonic anhydrase activity has been found within the chloroplast stroma, but there is also some carbonic anhydrase activity in the cytosol of mesophyll cells. Additionally, β-carbonic anhydrases play a role in photosynthesis by facilitating diffusion into and across the chloroplast, and by catalyzing HCO_3^- dehydration to supply CO_2 for RuBisCO (Zabaleta, Martin, and Braun 2012). The

presence of β-carbonic anhydrases in the animal kingdom has been controversial or ignored due to the paucity and poor quality of the available sequences; but it has been reported that β-class is widespread among invertebrates (Syrjänen et al. 2010). However, the full physiological role of β-carbonic anhydrases in the biosphere, either in prokaryotes or in plants or other organisms, is still to be discovered.

In *H. pylori*, as the α-carbonic anhydrase in the periplasm, the β-carbonic anhydrase (named HP0004 in strain 26695 and HPG27_4 in strain G27 examined, respectively) catalyzes the reversible hydration of the carbon dioxide into bicarbonate in the cytoplasm, cooperating to the urease-dependent response to acidity. Urease activity produces ammonia and carbon dioxide; the latter readily diffuses across the inner membrane into the periplasm, where it becomes substrate of the hydrolysis by the α-carbonic anhydrase (the metabolic mechanism has been described in detail in Chapter I). Carbon dioxide molecules that still remain in the cytoplasm are hydrolyzed by the β-form, thus contributing to buffer the pH of the cytoplasm. Interestingly, it may be observed that the *H. pylori* β-carbonic anhydrase is a catalytically efficient enzyme, possessing an enzymatic activity 3.2 times higher than that of the α-form. The β-carbonic anhydrase from *H. pylori* possesses a catalytic activity higher than that of human carbonic anhydrases III, VA, XII, and XIV among others. Only the human carbonic anhydrases VB and especially II, one of the best catalysts known in nature, show a better activity than *H. pylori* β-enzyme (Nishimori et al. 2007; Morishita et al. 2008). As previously mentioned, deletion or inhibition of the α-carbonic anhydrase drastically reduces acid survival of *H. pylori* even in the presence of urea and impairs the gastric colonization (Marcus et al. 2005b); deletion of both carbonic anhydrases results in a strongly reduced inflammation of the gastric mucosa or even prevents infection of animal models, although some mutants still show carbonic anhydrase activity, probably due to compensatory changes in the activity and/or expression of other enzymes involved in bicarbonate metabolism (Bury-Moné et al. 2008). These findings highlight the crucial importance of these enzymes in *H. pylori* survival and virulence, enrolling them as new attractive pharmacological targets to overcome the current antibiotic resistance problem. As for the α-carbonic anhydrase, also for the β-form many inhibition studies have been reported and several low nanomolar inhibitors have been detected; specifically, sulfamide and sulfamate compounds and complex anions, such as cyanide, cyanate, and hydrogen sulfide, have been observed to effectively inhibit the *in vitro* and *in vivo* growth of the pathogen (Nishimori et al. 2007; Morishita et al. 2008; Nishimori et al. 2010; Maresca et al. 2013).

4.2 SEQUENCE ANALYSIS

```
 10         20         30         40         50         60
MKAFLGALEF QENEYEELKE LYESLKTKQK PHTLFISCVD SRVVPNLITG TQPGELYVIR

         70         80         90        100        110        120
NMGNVIPPKT SHKESLSTIA SIEYAIVHVG VQNLIICGHS DCGACGSIHL ISDETTKAKT

        130        140        150        160        170        180
PYIANWIQFL EPIKEELKNH PQFSNHFAKR SWLTERLNAR LQLNNLLSYD FIQERVMDNE

        190        200        210        220
LKIFGWHYII ETGRIYNYNF ESHFFEPIEE TIKQRISHEN F
```

Property	Value
Aminoacid	221
Cysteines	4
Molecular weight (kDa)	25691.2
Theoretical pI	5.93
Abs_{280nm} (c=1mg/mL; b=1cm)	1.164 (all Cys reduced)
Signal peptide	No

Tab. 4.1 Some properties of β-carbonic anhydrase from *H. pylori*.

β-carbonic anhydrase from *H. pylori* G27 is composed of 221 aminoacids, including 4 cysteine residues, which are predicted to not form disulfide bonds (DISULFIND; http://disulfind.dsi.unifi.it/). Its theoretical isoelectric point is close to 6, in accordance with that of others bacterial β-carbonic anhydrase. The alignment of the aminoacid sequences shows that *H. pylori* β-carbonic anhydrase shares 33% identity with the β-carbonic anhydrase CynT from *Escherichia coli*, whose crystallographic structure is known, 51% identity with the predicted enzyme from the related pathogen *Helicobacter hepaticus*, and 25% identity with the enzyme from *Neisseria gonorrhoeae* (Fig. 4.5), confirming the high phylogenetic diversity of the β-carbonic anhydrase

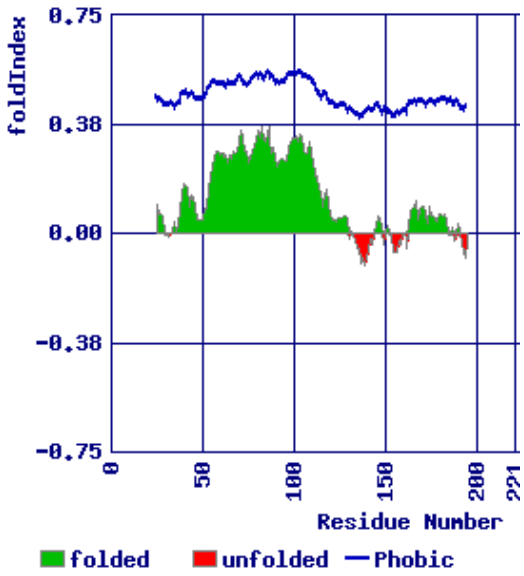

Fig. 4.4 Folding and hydrophobicity prediction of β-carbonic anhydrase from *H. pylori* (FoldIndex©).

class. However, the aminoacid residues important for the catalytic mechanism are highly conserved, namely cysteine, histidine and the dyad aspartate and arginine (dark blue in Fig. 4.5). A folding prediction program, (FoldIndex; http://bip.weizmann.ac.il/fldbin/findex) suggests the presence of an unfolded C-terminal region (Fig. 4.4); on the other hand, the secondary structure expected include 7 α-helices (47%) and 5 β-strands (15%), in accordance to the overall structure of the β-carbonic anhydrases previously mentioned (Phyre[2]; Kelley et al. 2015; http://www.sbg.bio.ic.ac.uk/phyre2/). In order to investigate the structural aspect, a putative structure of the β-carbonic anhydrase from *H. pylori* has been predicted, using the bioinformatics homology modelling tool SWISS-MODEL (Biasini et al. 2014; http://swissmodel.expasy.org/) and β-carbonic anhydrase CynT from *E. coli* as template (PDB accession code: 1I6P); it confirms a core of β-strands surrounded by the α-helices (Fig. 4.6a) and a homo-tetrameric oligomeric status (Fig. 4.6b), like CynT from *E. coli*.

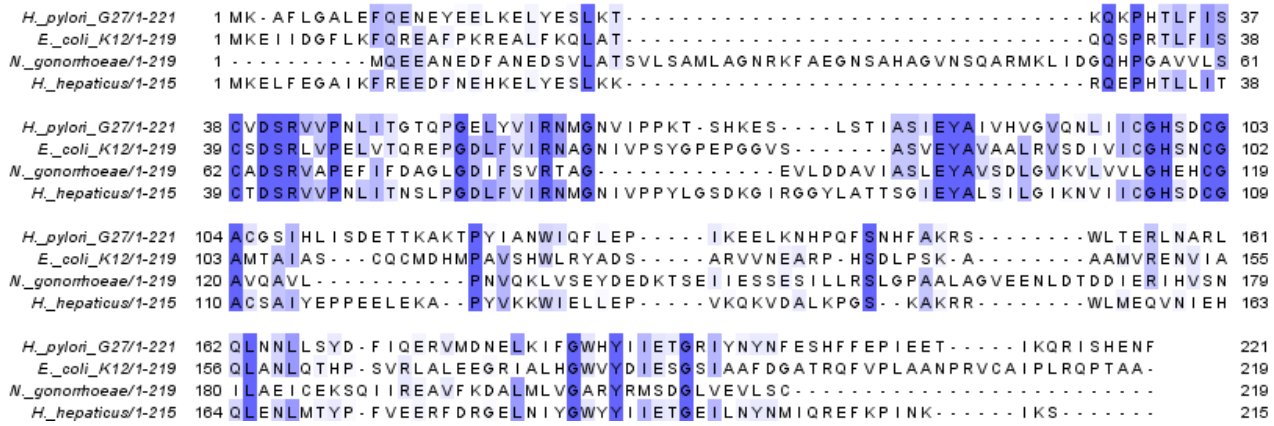

Fig. 4.5 Sequence alignment of β-carbonic anhydrase from *H. pylori* strain G27, *E. coli* strain K12, *N. gonorrhoeae* and *H. hepaticus*; in shades of blue the identity above 50% is shown (ClustalW and Jalview ver. 2.9.0b2).

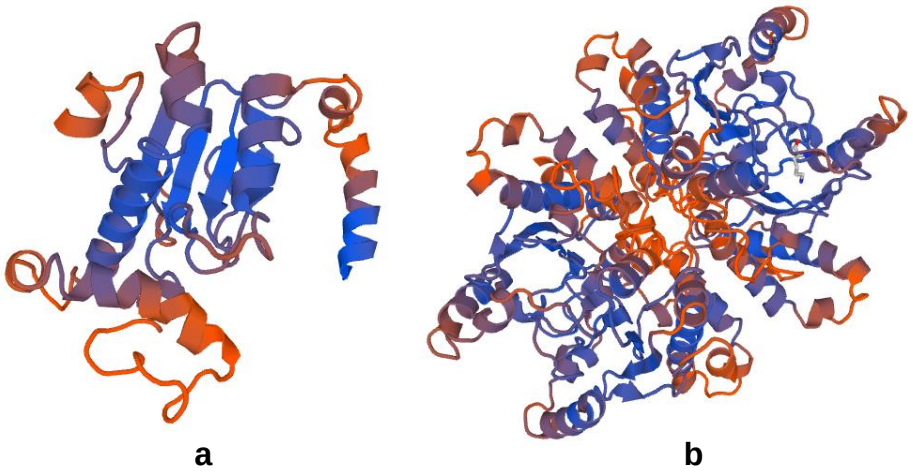

Fig. 4.6 Homology modelling structure of β-carbonic anhydrase from *H. pylori*, monomer (a) and homo-tetramer (b) predicted (SWISS-MODEL; template: β-carbonic anhydrase from *E. coli*, PDB accession code: 1I6P).

4.3 MATERIALS AND METHODS

4.3.1 MOLECULAR CLONING

The coding sequence for the β-carbonic anhydrase gene (HPG27_4) was PCR-amplified from genomic DNA of *H. pylori* strain G27, using Q5® High-Fidelity DNA Polymerase (New England Biolabs) and primers 5'-CATCATCACCACCATCACGAAAACCTGTATTTTCAGGGAAAAGCGTTTTTAGGAGC GTTA-3' (fw) and 5'-GTGGCGGCCGCTCTATTAGAAGTTTTCATGACTTATCCTTTG-3' (rv). The forward primer provided a 18 nt-sequence that recognizes the pETite N-His Kan vector (Lucigen) and encodes a N-terminal 6-His tag, and a TEV proteolysis site. The reverse primer included a Stop anticodon corresponding to the C-terminus. The PCR product, purified by 1% agarose gel, was cloned by thermal shock into the pETite N-His Kan vector (Lucigen) using *E. coli* XL1-Blue competent cells (Agilent), grown overnight on selective LB medium supplemented with 30 µg/ml of kanamycin. The colonies were checked by colony-PCR using EconoTaq DNA Polymerase (Lucigen) and T7 primers. Purified plasmid of positive colonies was double-digested by restriction enzymes NdeI (New England Biolabs) and NotI (New England Biolabs), for 2 h at 37 °C. Positive samples were finally checked by DNA sequencing to evaluate the quality.

The following sequence corresponds to the final recombinant protein with the addition of a N-terminal 6-His tag flanked by a TEV recognition site.

```
 10         20         30         40         50         60
HHHHHHENLY FQ//GKAFLGAL EFQENEYEEL KELYESLKTK QKPHTLFISC VDSRVVPNLI

        70         80         90        100        110        120
TGTQPGELYV IRNMGNVIPP KTSHKESLST IASIEYAIVH VGVQNLIICG HSDCGACGSI

       130        140        150        160        170        180
HLISDETTKA KTPYIANWIQ FLEPIKEELK NHPQFSNHFA KRSWLTERLN ARLQLNNLLS

       190        200        210        220        230
YDFIQERVMD NELKIFGWHY IIETGRIYNY NFESHFFEPI EETIKQRISH ENF
```

// corresponds to the cleavage site for TEV protease

Property	Value
Aminoacid	233
Cysteines	4
Molecular weight (kDa)	27234.8
Theoretical pI	6.14
Abs_{280nm} (C=1mg/mL; b=1cm)	1.153 (all Cys reduced)
Signal peptide	No

Tab. 4.2 Some properties of β-carbonic anhydrase referred to the construct added of N-terminal 6-His-tag and TEV recognition site.

4.3.2 EXPRESSION

The pETite plasmid carrying the gene of interest was transformed into *E. coli* BL21 (DE3) pLysS competent cells (Lucigen), since the basal expression of the recombinant protein resulted to be toxic for cell growth. The *E. coli* BL21 (DE3) pLysS harboring the pETite plasmid were grown in a selective LB medium supplemented with 30 µg/mL of kanamycin. A 2 L culture was grown under mild shaking (180 rpm) at 37 °C until an OD value (optical dispersion, at 600 nm) of about 0.6; then 1 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG) was added to the medium to induce protein expression and the culture was incubated at 20 °C overnight. The cultured cells were harvested and medium eliminated by centrifugation at 5,000 rpm for 30 min at 4 °C. The pellet was resuspended in a selected lysis buffer, containing a protease inhibitor cocktail (1 mM phenylmethanesulfonyl fluoride (PMSF), 15 µM aprotinin, 1 µM leupeptin, 0.5 mM 4-(2-Aminoethyl)-benzenesulfonylfluoride hydrochloride (AEBSF)). The composition of the lysis buffer changed depending on the chromatographic technique adopted for the purification, since various attempts of purification were performed in order to achieve the highest yield of β-carbonic anhydrase. Expression and purification steps were checked on SDS-PAGE and Western blotting, showing good level of expression, but most of the protein of interest in the inclusion bodies. Therefore, the purification of β-carbonic anhydrase was performed both from the soluble and from the insoluble fractions.

4.3.3 PURIFICATION *VIA* AFFINITY CHROMATOGRAPHY

The cultured pellet was resuspended in a lysis buffer containing 50 mM Tris–HCl pH 8, 400 mM NaCl, 5 mM imidazole, supplemented with protease inhibitors and lysed by mechanical pressure using a One Shot Cell disruption system (Constant Systems Ltd) at pressure 1.35 kBar. The lysis procedure was repeated twice to be more effective. The lysed cell suspension was cleared of debris by vacuum centrifugation at 18,000 rpm for 30 min at 4°C. The supernatant was loaded onto a 1 mL His-Trap HP Ni–NTA column (GE Healthcare), previously equilibrated with the lysis buffer, at a flow rate of approximately 1 mL/min. The column was extensively washed with the lysis buffer and subsequently with a buffer supplemented with 4% elution buffer (50 mM Tris–HCl pH 8, 400 mM NaCl, 500 mM imidazole), in order to eliminate most of contaminants. The 6-His-tagged protein was eluted at approximately 50 mM imidazole, by applying a linear gradient from 4% to 100% elution buffer (Fig. 4.7, left). The recombinant β-carbonic anhydrase was eluted at low imidazole concentration, likely because of the low affinity for the Ni-NTA resin of the chromatographic column. Moreover, a generous amount of the recombinant protein was eluted concurrently with the initial washings, as shown in the SDS-PAGE (Fig 4.7, right).

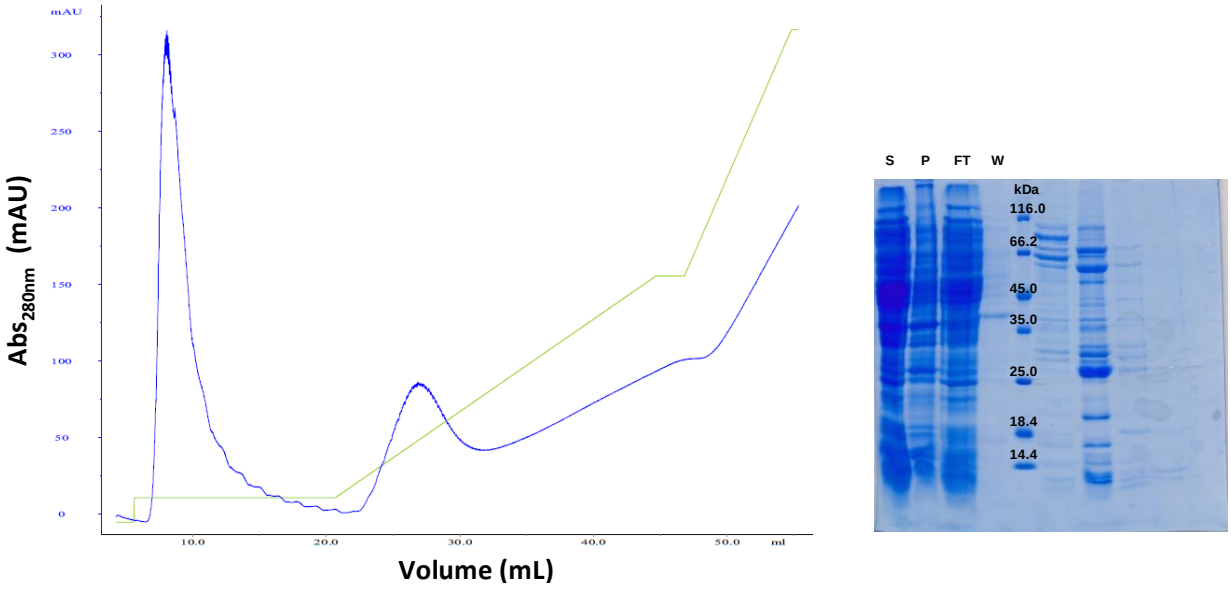

Fig. 4.7 6-His-tagged β-carbonic anhydrase affinity chromatography; left: absorption profile; right: SDS-PAGE analysis of the fractions collected (S: supernatant; P: lysed pellet; FT: flow-throw; W: wash)**.**

The fractions containing the 6-His-tagged β-carbonic anhydrase were pooled and concentrated by ultrafiltration using a Vivaspin 20 5,000 MWCO centrifugal concentrator (Sartorius). Subsequently, the concentrated protein was further purified by gel-filtration chromatography using a Superose 12 10/300 GL column (GE Healthcare), equilibrated with a buffer containing 50 mM Tris-HCl pH 8, 200 mM NaCl. The elution volume (13 mL)

showed a probable dimeric oligomerization of β-carbonic anhydrase in solution (Fig. 4.8, left). Moreover, the protein of interest was occasionally eluted as double species (Fig. 4.8, right), which differs each other for few kDa, therefore the gel-filtration column was not suitable to efficiently separate them.

The purified β-carbonic anhydrase was concentrated by ultrafiltration using a Vivaspin 6 5,000 MWCO centrifugal concentrator (Sartorius) for crystallization trials, despite the limited concentrations reached (4-8.8 mg/mL), verified by UV-Vis absorption spectrum (250-340 nm; NanoDrop 2000 UV-Vis Spectrophotometer). In accordance with that, it was decided not to cleave the 6-His-tag from the recombinant protein because of the risk of losing further protein amount performing this step.

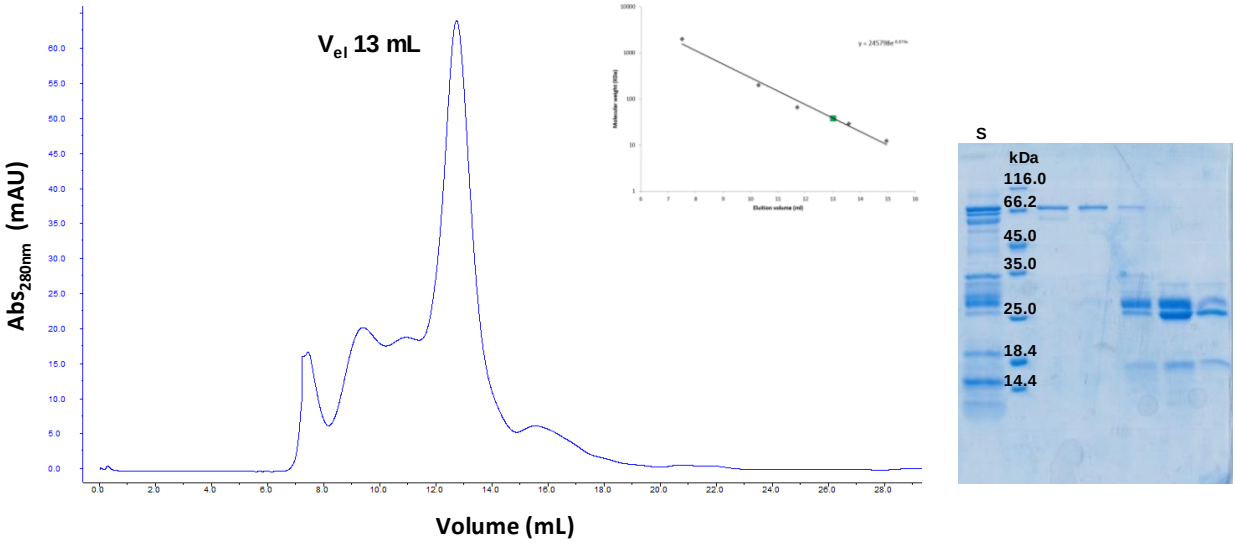

Fig. 4.8 6-His-tagged β-carbonic anhydrase gel-filtration chromatography (Superose 12 10/300 GL column); left: absorption profile; right: SDS-PAGE analysis of the fractions collected, showing the elution as double species (S: sample injected).

4.3.4 WESTERN BLOTTING

Samples from different steps of β-carbonic anhydrase purification were investigated for the presence of the His-tag, to ensure the quality of the protein (Fig. 4.9). After loading of the sample into a SDS-PAGE, the proteins were transferred on a Hybond-ECL nitrocellulose membrane (GE Healthcare) by electrophoresis. After blocking of the membrane using BSA 3% solution in TBS for 1 h and incubation of the Anti-polyHistidine primary antibody (Sigma-Aldrich) at 4 °C overnight, three washing steps using TTBS and a final one with TBS were performed. The peroxidase-conjugated secondary antibody (Promega) was incubated for 1 h at room temperature in TBS and washed for three times

with TTBS and finally with TBS. Subsequently, the membrane was developed using the Pierce ECL Western Blotting Substrate (ThermoFischer) for detection of peroxidase activity, at Image Station 4000 MM PRO (Kodak).

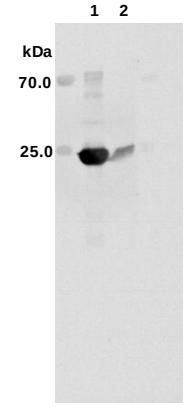

Fig. 4.9 Western blotting analysis of 6-His-tagged β-carbonic anhydrase from affinity chromatography fractions (1, 2: fractions collected)**.**

4.3.5 PURIFICATION *VIA* FRACTIONATED PRECIPITATION AND ION-EXCHANGE CHROMATOGRAPHY

Given the low yield of the affinity chromatography, a classic method of purification was performed. The cultured pellet was resuspended in a buffer containing 50 mM Tris–HCl pH 8, 150 mM NaCl, supplemented with protease inhibitors and lysed by mechanical pressure using a One Shot Cell disruption system (Constant Systems Ltd) at pressure 1.35 kBar. The lysis procedure was repeated twice to be more effective. The lysed cell suspension was cleared of debris by vacuum centrifugation at 18,000 rpm for 30 min at 4°C. A saturated solution of $(NH_4)_2SO_4$ (4.1 M at 25°C) was added to the supernatant till 25% (v/v) and the suspension was cleared by vacuum centrifugation at 16,000 rpm for 20 min at 4°C to separate the precipitated proteins (Fig. 4.11). The clarified supernatant was dialyzed 48 h at 4 °C using a CelluSep H1 5,000 Da cut-off membrane (Biosigma), to remove the excess of $(NH_4)_2SO_4$ and to exchange the buffer with another consisting of 30 mM Tris-HCl pH 8, suitable for the following ion-exchange chromatography step. The dialyzed supernatant was loaded at approximately 1 mL/min onto a HiPrep DEAE FF 16/10 column (GE Healthcare), suitable for a weak anion-exchange, since the recombinant β-carbonic anhydrase is negatively charged at pH 8 (theoretical pI 6.14). The column was equilibrated with the dialysis buffer and, following loading of the supernatant, was extensively washed with 2 CV of the same buffer. The protein was

eluted at 2 mL/min by applying a linear salt gradient in 7 CV from 0 to 70% high salt buffer, consisting of 30 mM Tris-HCl pH 8, 1.5 M NaCl (Figs. 4.10).

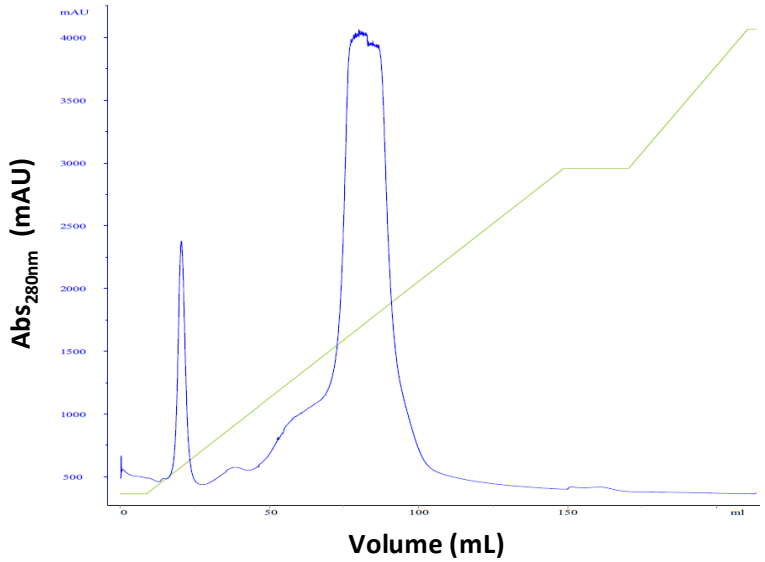

Fig. 4.10 6-His-tagged β-carbonic anhydrase anion-exchange chromatography: absorption profile.

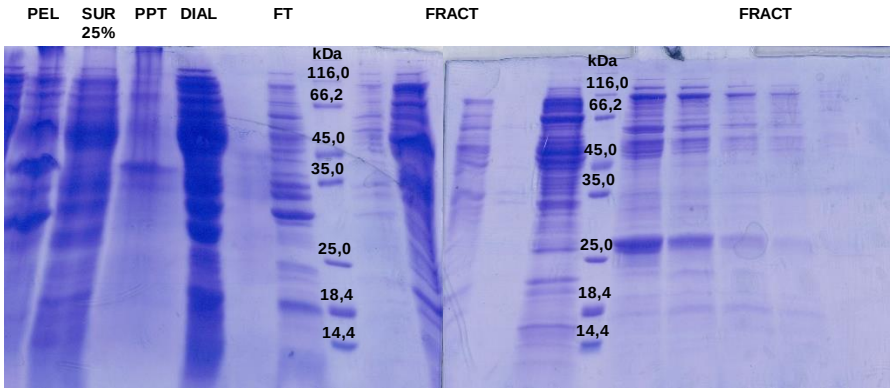

Fig.4.11 SDS-PAGE analysis of $(NH_4)_2SO_4$ fractionated precipitation and anion-exchange chromatography of 6-His-tagged β-carbonic anhydrase (PEL: lysed pellet; SUR 25%: supernatant added of 25% saturated solution of $(NH_4)_2SO_4$; PPT: precipitate; DIAL: dialyzed supernatant; FT: flow-throw; FRACT: fractions collected from anion-exchange chromatography).

The fractions containing the recombinant β-carbonic anhydrase, verified on SDS-PAGE, were pooled and concentrated by ultrafiltration using a Vivaspin 20 5,000 MWCO centrifugal concentrator (Sartorius). A fraction of concentrated protein was further purified by gel-filtration chromatography using a Superdex 200 Increase 10/300 GL column (GE Healthcare), equilibrated with a buffer containing 50 mM Tris pH 8, 200 mM NaCl (Fig. 4.12, left). The elution profile was not clear and didn't allow to separate the protein of interest, because of too many species in the ion-exchange sample. The low final yield of

β-carbonic anhydrase wasn't considered compatible to the several purification steps (Fig. 4.12, right) and the protein not suitable for crystallization trials, also because of the several subsequent handlings. Therefore, since the largest amount of β-carbonic anhydrase was found in the inclusion bodies, the recovery of the recombinant protein from the insoluble fraction was attempted.

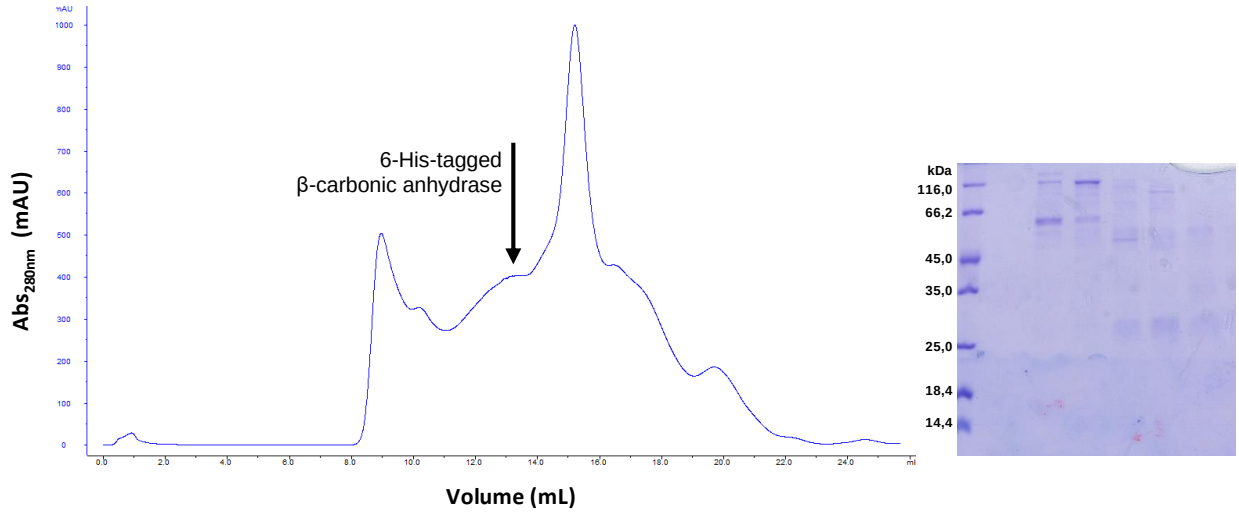

Fig. 4.12 6-His-tagged β-carbonic anhydrase gel-filtration chromatography (Superdex 200 Increase 10/300 GL column); left: absorption profile; right: SDS-PAGE analysis of the fractions collected, showing the low final yield.

4.3.6 PURIFICATION *VIA* ON-COLUMN REFOLDING

The protein purification from the inclusion bodies was performed according to the indications of an on-column chemical refolding protocol, adopting several various washings (Oganesyan, Kim, and Kim 2004).

The cultured pellet was resuspended in a buffer containing 30 mM Tris–HCl pH 8, 150 mM NaCl, supplemented with protease inhibitors and lysed by mechanical pressure using a One Shot Cell disruption system (Constant Systems Ltd) at pressure 1.35 kBar. The lysis procedure was repeated twice to be more effective. The lysed cell suspension was vacuum-centrifuged at 18,000 rpm for 30 min at 4°C to separate supernatant and debris pellet. The latter was resuspended in a denaturing buffer consisting of 30 mM Tris-HCl pH 8, 150 mM NaCl, 8 M urea, supplemented with protease inhibitors, and incubated at 4 °C under vigorous shaking overnight, to solubilize the inclusion bodies. Subsequently, the denatured suspension was cleared by vacuum centrifugation at 18,000 rpm for 30 min at 4°C; the solubilized inclusion bodies, added of 5 mM imidazole, were loaded onto 500 µL of Ni-Sepharose 6 Fast Flow resin (GE Healthcare), previously equilibrated with

the denaturing binding buffer (30 mM Tris-HCl pH 8, 150 mM NaCl, 5 mM imidazole, 8 M urea), and incubated at 4 °C overnight to enhance the batch-absorption. On-column renaturation and purification were performed by several changes of buffers, as listed below:

1) 30 mM Tris-HCl pH 8, 150 mM NaCl, 20 mM imidazole, 8 M urea, 10 mM β-mercaptoethanol
2) 30 mM Tris-HCl pH 8, 150 mM NaCl, 0,1 % Triton X-100, 10 mM β-mercaptoethanol
3) 30 mM Tris-HCl pH 8, 150 mM NaCl, 5 mM β-cyclodextrin, 10 mM β-mercaptoethanol
4) 30 mM Tris-HCl pH 8, 500 mM NaCl, 10 mM β-mercaptoethanol

Every washing step consisted of 10 CV of peculiar buffer, at approximately 1 mL/min. An initial washing at low imidazole concentration was performed to remove nonspecific bindings. The subsequent detergent buffer allowed to ensure a complete protein denaturation. To remove detergent from the protein–detergent complex and to allow the protein refolding, β-cyclodextrin was added to the following washing. A high salt concentration buffer was finally applied to remove remaining impurities and β-cyclodextrin. Refolded β-carbonic anhydrase was eluted with an elution buffer consisting of 30 mM Tris-HCl pH 8, 150 mM NaCl, 100 mM imidazole. As previously mentioned, β-carbonic anhydrase showed a low affinity for the Ni-NTA resin, although the overnight shaking incubation to enhance the binding. Therefore, a consistent amount of the recombinant protein was eluted concurrently with the initial washing at low imidazole, as shown in the SDS-PAGE (Fig. 4.13). The eluted fractions containing the β-carbonic anhydrase were pooled and concentrated by ultrafiltration using a Vivaspin 20 5,000 MWCO centrifugal concentrator (Sartorius), in order to perform some characterization experiments.

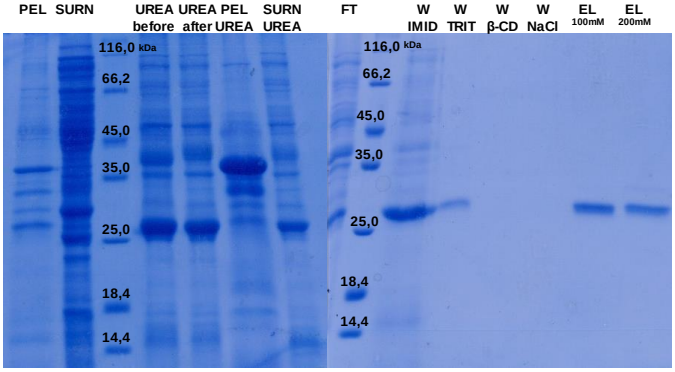

Fig. 4.13 On-column refolding of 6-His-tagged β-carbonic anhydrase by several changes of buffer (PEL: lysed pellet; SURN: supernatant; UREA before: inclusion bodies before incubation in urea 8M; UREA after: inclusion bodies after incubation in urea 8M; PEL UREA: pellet of inclusion bodies suspension cleared by centrifugation; SURN UREA: supernatant of inclusion bodies suspension cleared by centrifugation; FT: flow-throw; W IMID: wash 1; W TRIT: wash 2; W β-CD: wash 3; W NaCl: wash 4; EL 100mM: protein elution at 100 mM imidazole; EL 200mM: protein elution at 200 mM imidazole).

4.3.7 CHARACTERIZATION

Firstly, the concentrated refolded sample was submitted to an analytical gel-filtration chromatography using a Superdex 200 Increase 10/300 GL column (GE Healthcare), equilibrated with a buffer containing 50 mM Tris-HCl pH 8, 200 mM NaCl, to evaluate the oligomerization of the protein in solution. The elution profile demonstrated that the refolded β-carbonic anhydrase was poly-disperse in solution, suggesting that this sample was not suitable for crystallization trials (Fig. 4.14).

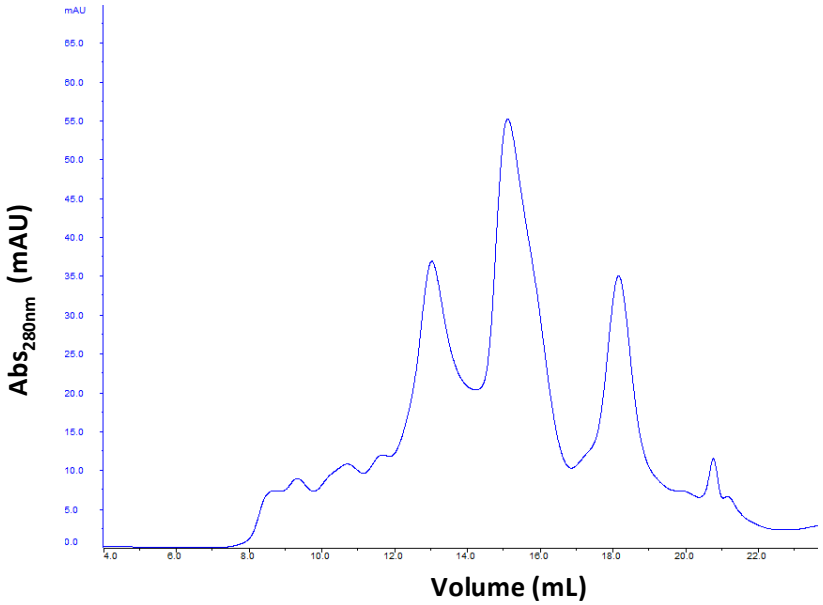

Fig. 4.14 6-His-tagged β-carbonic anhydrase analytical gel-filtration chromatography (Superdex 200 Increase 10/300 GL column): absorption profile, showing the inhomogeneity of the purified protein sample.

In order to evaluate the goodness of the refolding procedure and to estimate the content of secondary structure, a circular dichroism spectrum was collected by a JASCO J-715 Spectropolarimeter (Jasco), in the range 193-250 nm, with a buffer consisting of 30 mM Tris-HCl pH 8 and protein concentration 0.1 mg/mL (Fig. 4.15). Circular dichroism analysis revealed that β-carbonic anhydrase was not fully refolded, since α-helix signal was lacking, despite its predicted prevalence. These evidence confirmed the not suitable quality of this sample for crystallization trials and different approaches of protein refolding need to be investigated to purify β-carbonic anhydrase from inclusion bodies.

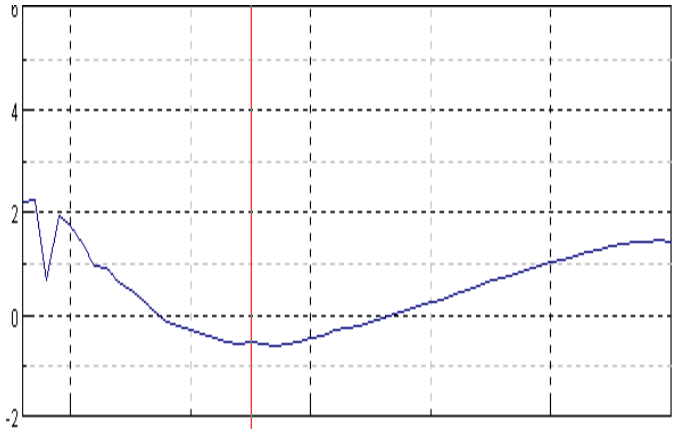

Fig. 4.15 Circular dichroism of refolded 6-His-tagged β-carbonic anhydrase (JASCO J-715 Spectropolarimeter, range 193-250 nm)

4.3.8 CRYSTALLIZATION TRIALS

β-carbonic anhydrase samples from affinity chromatography purification were used for crystallization tests, at different concentration (2-8.8 mg/mL). Crystallization trials with sitting-drop vapor-diffusion technique were carried out using an Oryx8 crystallization robot (Douglas Instruments). Several standard crystal screening kits were explored, including Structure Screen I and II (Molecular Dimensions), Crystal Screen I and II (Hampton Research), PACT Suite (Qiagen), JCSG Suite (Qiagen), PEGs II (Qiagen), AmSO4 (Qiagen). Given the limited protein concentration obtained, crystallization plates were performed mainly at 4 °C, to reduce the protein solubility and facilitate the crystallization process. Nevertheless, several conditions showed precipitation even immediately after drop deposition, most likely owing to the degradation of the purified protein. Microcrystalline precipitates were observed only in few crystallization conditions, mainly characterized by acid pH and PEG precipitant. Several crystallization attempts were performed, but suitable crystals were not obtained.

4.4 RESULTS AND DISCUSSION

The yield of soluble fraction was revealed to be the bottleneck of the purification of the β-carbonic anhydrase. Most likely due to its C-terminal disorder region, the protein of interest was not properly folded by *E. coli* cells, lacking of some abilities of post-translational modifications, and formed inclusion bodies. First attempt of overexpression

was performed in *E. coli* BL21 (DE3) cells, but the basal expression of the recombinant protein was observed to be toxic for cell growth. Therefore, *E. coli* BL21 (DE3) pLysS cells were selected, to inhibit the background expression level of the target protein, and cells were grown slowly, at low temperature, to facilitate the correct folding. Despite these precautions, most of the protein of interest was found in the insoluble fraction, as inclusion bodies. In order to overcome this limitation, overexpression of the recombinant β-carbonic anhydrase was attempted using further engineered *E. coli* strain, including C41 (DE3) cells, C43 (DE3) cells, SHuffle® cells, Lemo21 (DE3) cells, whose peculiarities were previously described. However, despite the high expression level, the yield of soluble protein did not clearly improve (data not shown). ArcticExpress (DE3) cells, not yet tested, might be useful for this purpose, since they provide an approach to increasing the yield of soluble protein produced at low temperature, thanks to two co-expressed chaperonins. A further possible solution could be the subcloning as recombinant protein lacking 15 N-terminal aminoacid residues, as reported in literature (Nishimori, Minakuchi, Kohsaki, Onishi, Takeuchi, Vullo, Scozzafava, and Claudiu T. Supuran 2007); but this region seems to be relevant for the purpose of structural investigations, since it is predicted to be involved in oligomerization.

An additional limiting step was revealed to be the low affinity of the 6-His-tagged β-carbonic anhydrase for the NiNTA resin of the affinity chromatography column. Besides the various chromatographic purifications previously reported, in order to overcome this issue, the target protein was subcloned as C-terminal 6-His-tag fusion protein. But this recombinant product was observed to bind to the NiNTA resin even less effectively, probably owing to the C-terminal disorder region, prone to degradation (data not shown). Moreover, the elution as double species reported in the event of some purifications was possibly due to degradation of this protein domain.

Further investigations are needed to optimize the purification step and the yield of soluble and properly folded protein of interest, in order to explore more crystallization conditions and to characterize the catalytic activity and the peculiarities compared to homologues from other bacterial species.

Chapter V

CLONING AND EXPRESSION TRIALS OF FliK, THE FLAGELLAR HOOK-LENGTH CONTROL PROTEIN FROM *HELICOBACTER PYLORI*

5.1 INTRODUCTION

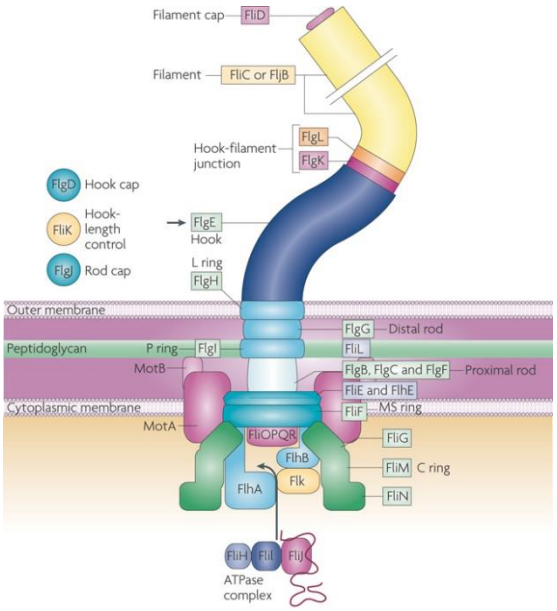

Fig. 5.1 Structural elements of *H. pylori* flagella. Adapted from Chevance and Hughes, 2008.

As previously evidenced, motility is essential for successful stomach colonization by *H. pylori* and is provided by its sheathed unipolar flagella, considered as one of the most important colonization and virulence factors. *H. pylori* flagella, as those of enteric bacteria, are composed of three structural elements: a basal body, which is embedded in the cell wall and contains the proteins required for rotation and chemotaxis and usually the flagellar export proteins; an external helically shaped filament that works as a propeller when rotated at its base; a hook that serves as a joint between the basal body and the flagellar filament (Fig. 5.1; Chevance and Hughes 2008). Regulation of flagellar gene expression and biosynthesis is a complex and well-coordinated process in *H. pylori*; more than 40 proteins are involved in the biosynthesis and operation of flagella and their control by the chemotaxis machinery (Niehus et al. 2004). The assembly of the flagellar structure is a process initiated by the assembly of the basal body on the cell membrane. On the basal body, the flagellar hook forming protein monomers, namely FlgE, are assembled exterior to the cell forming the hook-basal body complex. Reached this step, FliK is responsible for the hook length control, as reported in other Gram-negative bacteria, such as *Salmonella typhimurium* and *Escherichia coli* (Kawagishi et al. 1996). When the flagellar hook substructure reaches its optimal length, sensed by the "checkpoint control" protein FliK, the export of the anti-σ^{28} factor, namely FlgM, is triggered, releasing σ^{28} from a σ^{28}-FlgM complex, which in turn allows the subsequent expression of σ^{28}-dependent genes. Specifically, σ^{28}, as well as σ^{54}, is a RNA polymerase factor responsible for the control of the transcription of genes encoding the flagellar filament subunits, namely FlaA and FlaB (Colland et al. 2001). Therefore, the release of σ^{28} from the complex results in a switch of export substrate specificity from rod/hook type to filament type, consequently initiating filament assembly (Baidya, Bhattacharya, and Chowdhury 2015). Accordingly, FliK acts as a molecular ruler, since it terminates hook

export and assembly and transmits a signal to begin filament export, the final stage in flagellar biosynthesis (Waters, O'Toole, and Ryan 2007).

Moreover, a cooperating effector is required for hook regulation and assembly, namely FlgD, which functions as the hook scaffolding protein and is also considered to be the hook-capping protein and the basal body rod-modification protein. In addition, together with FliK, FlgD regulates the assembly of the hook cap structure to prevent the leakage of hook monomers into the medium and hook monomer polymerization as well playing a role in determination of the correct hook length (Luo et al. 2009). Interestingly, in *S. typhimurium* the N-terminal domain of FliK has been observed to bind to the hook-capping protein FlgD with high affinity (Moriya et al. 2006). Therefore, according to their importance and cooperation in hook regulation and assembly, FliK and FlgD play a critical role in *H. pylori* pathogenesis, becoming the focus of several studies, also performed by this research group (data not published yet).

In *fliK* mutants it has been observed that the hook to filament transition does not occur, and long hooks of unregulated length, named polyhooks, are formed (Muramoto et al. 1998). Moreover, microarray analysis of a *fliK*-null mutant has revealed increased transcription of genes under the control of the σ^{54} sigma factor. This factor has been shown to be responsible for transcription of several flagellar genes, including *flgE* and *flaB*, as previously mentioned. No genes higher in the flagellar hierarchy had altered expression, suggesting specific and localized FliK-dependent feedback on the σ^{54} regulon (Ryan et al. 2005; Douillard et al. 2009). FliK thus appears to be involved in three main processes: hook-length control, export substrate specificity and control of σ^{54} transcriptional activity. Interestingly, it has been reported that contact of *H. pylori* with gastric epithelial cells strongly induces expression of the flagellar hook-length control protein. Host cell contact also up-regulates expression of the *H. pylori* major virulence factor CagA (Kim et al. 2004). It has been demonstrated that FliK is necessary for *cagA* gene upregulation in gastric cell-associated *H. pylori*, since FliK has a role in triggering dissociation of the alternate sigma factor, σ^{28}, from a nonfunctional σ^{28}-FlgM complex, releasing free, functional σ^{28}; consequently, the σ^{28}-RNA polymerase initiates transcription of *cagA* (Baidya et al. 2015).

The flagellar cytoplasmic protein FliK regulates hook elongation by two successive main events: by determining hook length and by stopping the supply of hook protein; but the mechanism remains still controversial. It has been hypothesized that these two distinct roles are assigned to different regions of FliK: the N-terminal domain determines length

and the C-terminal domain switches secretion from the hook protein to the filament protein (Mizuno et al. 2011). Flagellar components are exported through a homologue to a type III secretion system (T3SS) apparatus, as previously mentioned; FlhB is a membrane-bound component of the export apparatus and, particularly, represents the switchable secretion gate. As result of the interaction between the C-terminal domain of FliK and the cytoplasmic domain of FlhB, FliK has been reported to switch substrate specificity of the export apparatus from rod/hook-type to filament-type by causing a conformational change in the cytoplasmic domain of FlhB upon completion of the hook assembly (Minamino et al. 2004). On the other hand, the N-terminal domain of FliK include a secretion signal and is dispensable for length control of the flagellar hook (Hirano et al. 2005). The secreted FliK, thanks to its highly elongated structure, has been hypothesized to interact with the hook cap during hook polymerization, as reported in *S. typhimurium*; thus, when the hook polymerized to a proper length, the FliK C-terminal domain would be in proximity to FlhB to catalyze the secretion-specificity switch, terminate hook polymerization, and initiate filament assembly (Fig. 5.2; Minamino and Pugsley 2005).

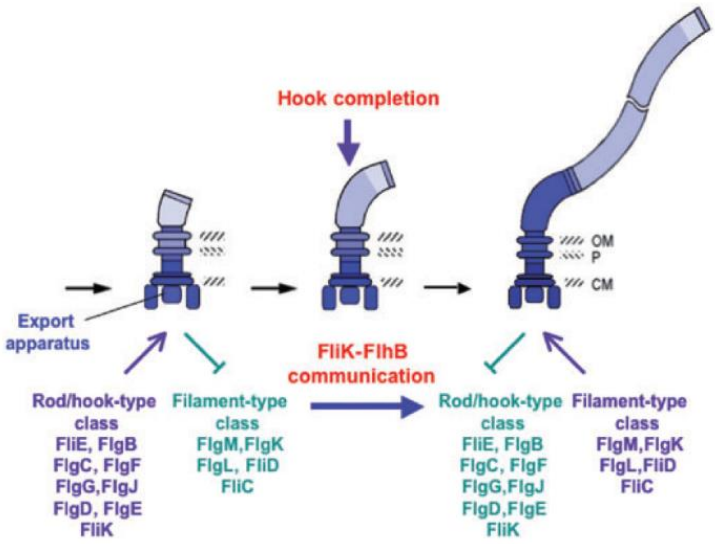

Fig. 5.2 Hypothetical model of substrate specificity switching of the type III flagellar protein secretion apparatus. The flagellar protein secretion apparatus secretes rod-type and hook-type substrates, but does not secrete the filament-type substrates until the hook reaches its mature length. Upon completion of the hook structure, the completion signal is transmitted to FliK and FlhB, shutting off rod/hook-type protein secretion and turning on filament-type protein secretion. Adapted from Minamino and Pugsley, 2005.

Further investigations are needed to comprehend in detail the functional mechanism of this molecular ruler, affecting a proper motility and the colonization ability of *H. pylori*. The information reported up to now suggests FliK, as well as other flagellar proteins, to be an attractive target for pharmacological studies, since abnormalities in the flagella assembly impair the *H. pylori* motility and rate of stomach colonization. Structural investigations of FliK homologues are very limited; to date in the Protein Data Bank (http://www.pdb.org) only a NMR structure of FliK C-terminal domain from *Salmonella enterica* serovar Typhimurium (Fig. 5.4; Mizuno et al. 2011; PDB accession code: 2RRL) has been published. This paucity of structural information is probably due to a largely unstructured nature of FliK, mainly at the N-terminal domain, which contains an unusual sequence of glycine and proline, which

are structure-breaking residues. On the contrary, the C-terminal domain, rich in glutamine, is a more compact and globular region, and conserved among several species (Fig. 5.3). As can be evidenced, FliK aminoacid sequences from various organisms show a low level of identity, even among similar species, such as *S. typhimurium* and *E. coli* (Mizuno et al. 2011); specifically, there are no conserved sequences in the N-terminal domain, whereas conserved sequences are found in the C-terminal domain (Fig. 5.3). A high level of disordered regions has also been detected in *H. pylori* FliK, resulting in problematic investigations and leading to interrupt trials, as described below.

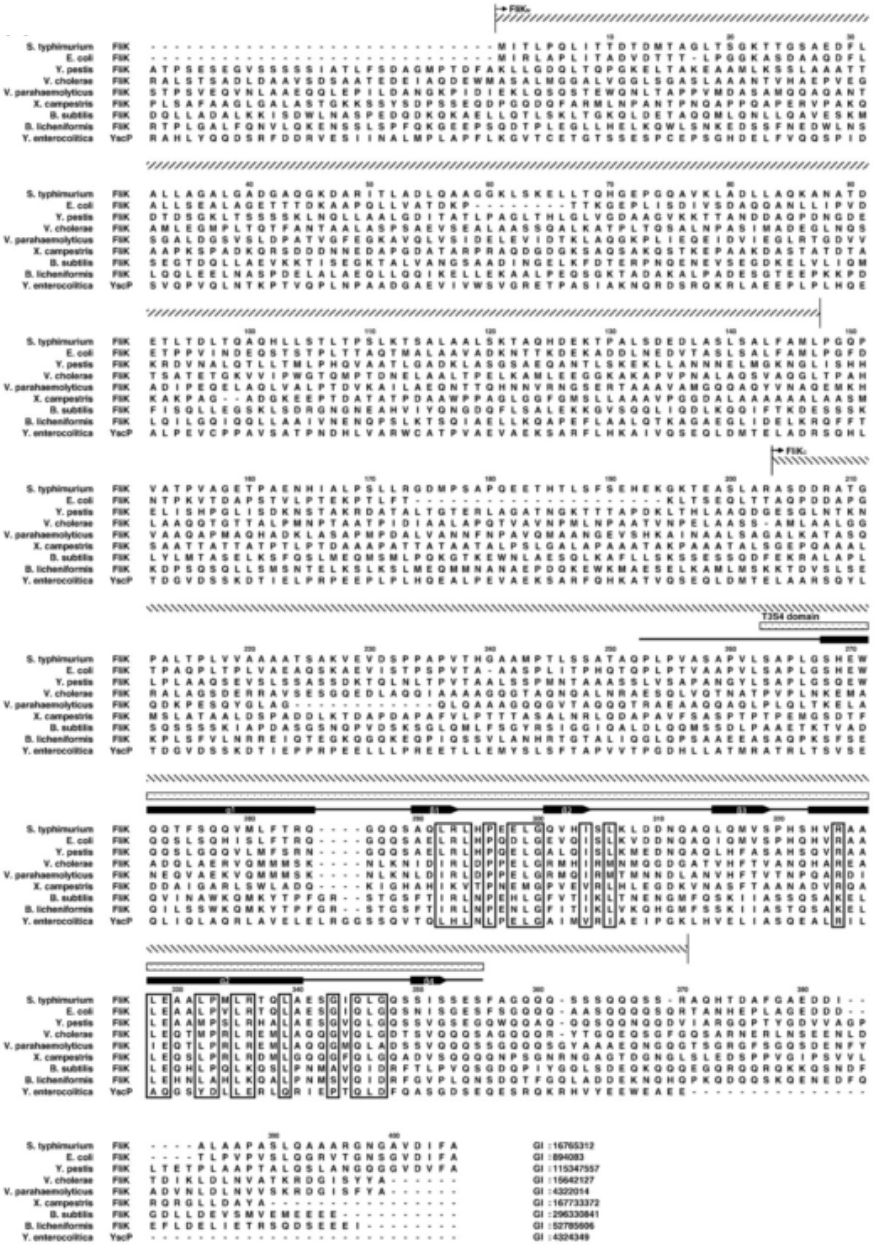

Fig. 5.3 Alignment of the sequences of FliK orthologs from eight bacteria and from YscP, an equivalent protein in the injectisome assembly. The locations of the N-terminal and C-terminal fragments are marked with shaded bars on the alignment; the numbers marked correspond to the positions for the residues in FliK from *S. Typhimurium*. The secondary elements determined by NMR are depicted on the alignment by the filled bars: two α-helices and four β-strands. Adapted from Mizuno et al., 2001.

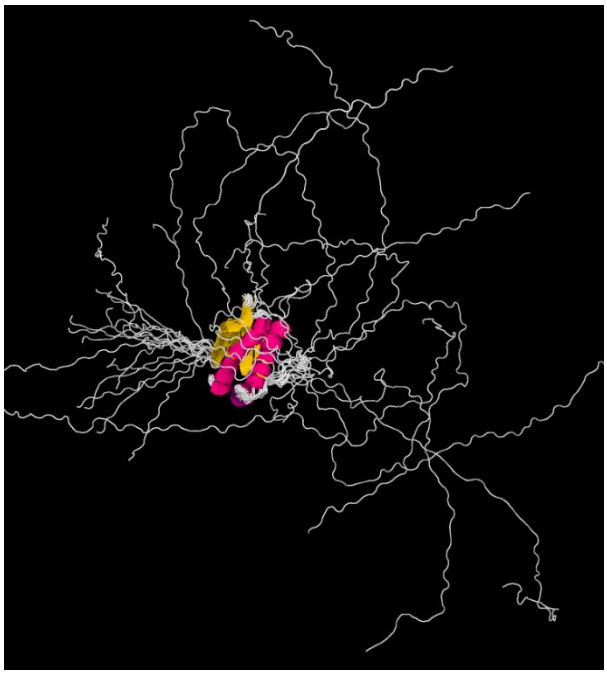

Fig. 5.4 NMR structure of FliK C-terminal domain from *Salmonella enterica serovar Typhimurium* (PDB accession code: 2RRL).

5.2 SEQUENCE ANALYSIS

```
 10         20         30         40         50         60
MPSPINPIHT NASANASTLI NSGAKNKDTK NAPKSASKDF SKILNQKISK DKTAPKENPN

        70         80         90        100        110        120
ALKATPKNTK EDAKVLEKTP TLQPQHAQNP AKDQQAPTLK DLLNHKTTAP HEAQHENHEH

       130        140        150        160        170        180
ETNPKTPNET LNKNEKEPNG VTSNDHQANL TNKNPLTPTN HAIKNPTAPT HNAKEPKTLK

       190        200        210        220        230        240
DIQTLSQKHD LNANNIQAAT IPENKTPLNA SDHLALKTTQ TPINHTLAKN DAKNTANLSS

       250        260        270        280        290        300
VLQSLEKKES HNKEHANPPN NEKKTPPLKE ALQMNAIKRD KTLSKKKSEK TPTKTQTTAP

       310        320        330        340        350        360
SIAPENAPKI PLKTPPLMPL IGANPPNDNP PTPLEKEETT KEASDNKEKT KESSNSAQNA

       370        380        390        400        410        420
QNAQSSDKTS ENKSVTPKET IKHFTQQLKQ EIQEYKPPMS RISMDLFPKE LGKVEVIIQK

       430        440        450        460        470        480
VGKNLKVSVI SHNNSLQTFL DNQQDLKNSL NALGFEGVDL SFSQDSSKEQ PKEQLKEPFK

       490        500        510
EQELTPLKEN ALKSYQENTD HENKETSMQI TLYA
```

Property	Value
Aminoacid	514
Cysteines	0
Molecular weight (kDa)	56868.5
Theoretical pI	9.15
Abs_{280nm} (c=1mg/mL; b=1cm)	0.079
Signal peptide	Yes

Tab. 5.1 Some properties of FliK from *H. pylori*

The flagellar hook-length control protein from *H. pylori* G27 is composed of 514 aminoacids, many more than the other bacterial species considered, and exhibits a basic theoretical isoelectric point, approximately 9, differently from the very acid one of *S. typhimurium* and *E. coli* (approximately 4) and similar to that of the related *Helicobacter hepaticus*. As anticipated, the alignment of the aminoacid sequences confirms that *H. pylori* FliK presents a low level of identity compared to orthologues from similar bacterial species, except for the C-terminal domain; specifically, *H. pylori* FliK shares 29% identity with FliK from *S. typhimurium*, whose C-terminal structure is known, 28% identity with FliK from *E. coli*, and 42% identity with FliK from the related pathogen *H. hepaticus*, and the most conserved residues are located at the C-terminal domain (Fig. 5.6). A folding prediction program (FoldIndex; http://bip.weizmann.ac.il/fldbin/findex) confirms the overall unstructured nature of this flagellar protein (Fig. 5.5), showing the limited folded region located at the C-terminal domain, as predicted; this evidence is also confirmed by a prediction of naturally disordered regions (PONDOR®; http://www.pondr.com), that reveals 75.49% of overall disorder and 12 disordered regions (Fig. 5.7). Additionally, the presence of the N-terminal secretion signal is demonstrated by SignalP 4.1 Server (Fig. 5.8), that is able to predict signal peptide/non-signal peptide based on a combination of several artificial neural

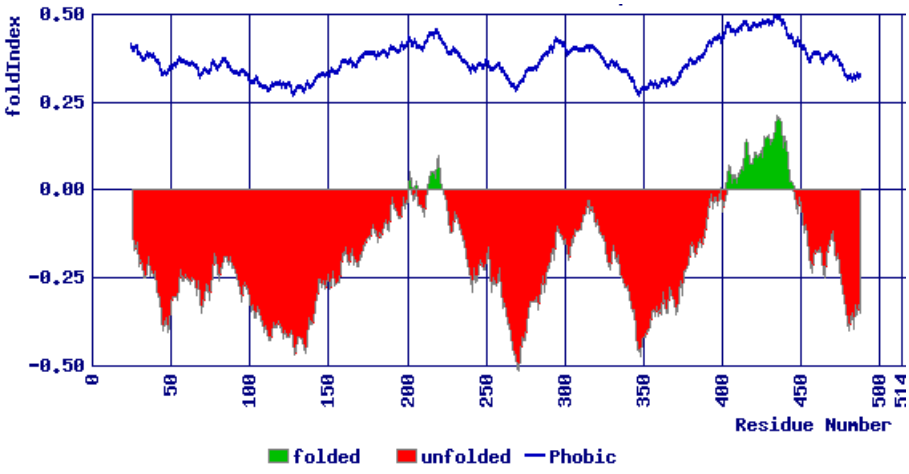

Fig. 5.5 Folding and hydrophobicity prediction of FliK from *H. pylori* (FoldIndex©).

networks (Petersen et al. 2011; http://www.cbs.dtu.dk/services/SignalP). According to the structure prediction by Phyre2 (Kelley et al. 2015; http://www.sbg.bio.ic.ac.uk/phyre2/), the content of secondary structure expected is limited and include 31% α-helices and 7% β-strands. A putative structure of FliK from *H. pylori* (Fig. 5.9) has been predicted, using the SWISS-MODEL (Biasini et al. 2014; http://swissmodel.expasy.org/) and the FliK C-terminal domain from *S. typhimurium* as template (PDB accession code: 2RRL).

Fig. 5.6 Sequence alignment of FliK from *H. pylori* strain G27, *S. typhimurium*, *E. coli* and *H. hepaticus*; in shades of blue the identity above 50% is shown (ClustalW and Jalview ver. 2.9.0b2).

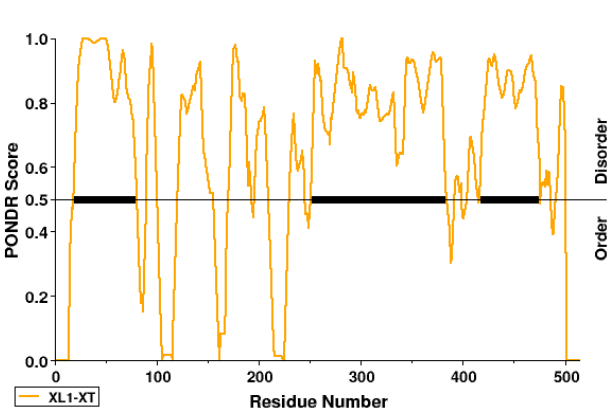

Fig. 5.7 Prediction of naturally disordered regions of FliK from *H. pylori* (PONDOR®; PONDR scores indicate the propensities for folding; the regions having residues with lower PONDR score tend to be part of a folded structure, while the regions having a score over 0.5 are predicted to be disordered).

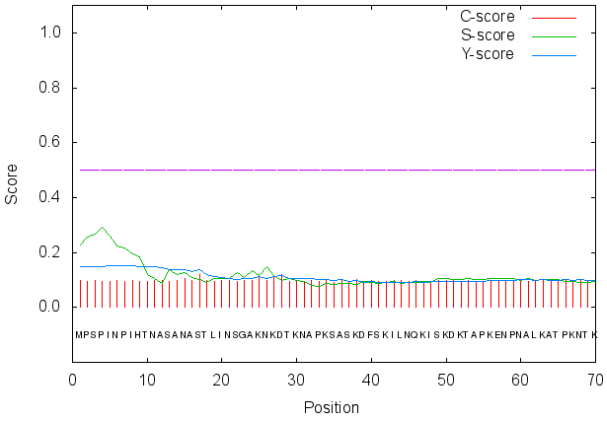

Fig. 5.8 FliK from *H. pylori* secretion signal prediction (SignalP 4.1 Server)

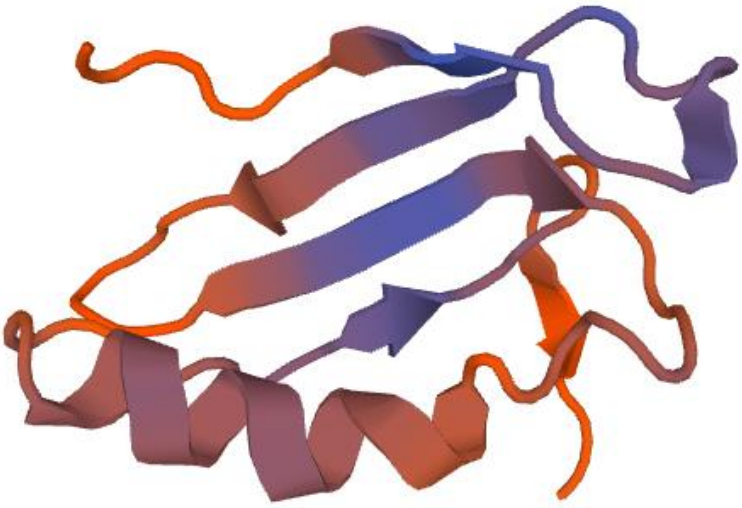

Fig. 5.9 Homology modelling structure of FliK from *H. pylori*, monomer predicted (SWISS-MODEL; template: C-terminal domain of FliK from *S. typhimurium*, PDB accession code: 2RRL).

5.3 MATERIALS AND METHODS

5.3.1 MOLECULAR CLONING

The coding sequence for FliK (HPG27_857) was PCR-amplified from genomic DNA of *H. pylori* strain G27, using Phusion® High-Fidelity DNA Polymerase (New England Biolabs) and primers 5'-GAAGGAGATATACATATGCCATCTCCTATTAATCCCATT-3' (fw) and 5'-GTGATGGTGGTGATGATGCGCATAAAGAGTGATTTGC-3' (rv). The reverse primer provided a 18 nt-sequence that recognizes the pETite C-His Kan vector (Lucigen) and encodes a C-terminal 6-His tag; it did not include a Stop anticodon. The forward primer included a Start codon. It was decided to not introduce a TEV proteolysis site between the 6-His-tag and the C-terminal end of the target protein. The PCR product, purified by 1% agarose gel, was cloned by thermal shock into the pETite C-His Kan vector (Lucigen) using *E. coli* 10G competent cells (Lucigen), grown overnight on selective LB medium supplemented with 30 µg/ml of kanamycin. The colonies were checked by colony-PCR using EconoTaq DNA Polymerase (Lucigen) and T7 primers. Purified plasmid of positive colonies was digested by EcoRV restriction enzyme (New England Biolabs), for 1 h at 37 °C. Positive samples were finally checked by DNA sequencing to evaluate the quality.

The following sequence corresponds to the final recombinant construct with a C-terminal 6-His tag.

```
 10         20         30         40         50         60
MPSPINPIHT NASANASTLI NSGAKNKDTK NAPKSASKDF SKILNQKISK DKTAPKENPN

        70         80         90        100        110        120
ALKATPKNTK EDAKVLEKTP TLQPQHAQNP AKDQQAPTLK DLLNHKTTAP HEAQHENHEH

       130        140        150        160        170        180
ETNPKTPNET LNKNEKEPNG VTSNDHQANL TNKNPLTPTN HAIKNPTAPT HNAKEPKTLK

       190        200        210        220        230        240
DIQTLSQKHD LNANNIQAAT IPENKTPLNA SDHLALKTTQ TPINHTLAKN DAKNTANLSS

       250        260        270        280        290        300
VLQSLEKKES HNKEHANPPN NEKKTPPLKE ALQMNAIKRD KTLSKKKSEK TPTKTQTTAP

       310        320        330        340        350        360
SIAPENAPKI PLKTPPLMPL IGANPPNDNP PTPLEKEETT KEASDNKEKT KESSNSAQNA

       370        380        390        400        410        420
QNAQSSDKTS ENKSVTPKET IKHFTQQLKQ EIQEYKPPMS RISMDLFPKE LGKVEVIIQK

       430        440        450        460        470        480
VGKNLKVSVI SHNNSLQTFL DNQQDLKNSL NALGFEGVDL SFSQDSSKEQ PKEQLKEPFK

       490        500        510        520
EQELTPLKEN ALKSYQENTD HENKETSMQI TLYAHHHHHH
```

Property	Value
Aminoacid	520
Cysteines	0
Molecular weight (kDa)	57691.4
Theoretical pI	9.15
Abs_{280nm} (C=1mg/mL; b=1cm)	0.077
Signal peptide	Yes

Tab. 5.2 Some properties of FliK referred to the recombinant construct added of C-terminal 6-His-tag.

5.3.2 EXPRESSION TRIALS

The pETite plasmid carrying the gene of interest was transformed into *E. coli* BL21 (DE3) competent cells (Lucigen). The *E. coli* BL21 (DE3) harboring the pETite plasmid were grown in small scale in a selective LB medium supplemented with 30 µg/mL of kanamycin, in order to control and optimize level and quality of protein expression. Multiple 50 mL cultures were grown under mild shaking (180 rpm) at 37 °C until an OD value (optical

dispersion, at 600 nm) of about 0.6; then 0.5 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG) was added to the medium to induce protein expression and the cultures were incubated at 37 °C for 3 h, at 30°C overnight, at 20 °C overnight, respectively. The cultured cells were harvested from each small scale culture and medium eliminated by centrifugation at 5,000 rpm for 30 min at 4 °C. The pellet was resuspended in a lysis buffer containing 50 mM Tris-HCl pH 7.2, 150 mM NaCl, supplemented with a protease inhibitor cocktail (1 mM phenylmethanesulfonyl fluoride (PMSF), 15 µM aprotinin, 1 µM leupeptin, 0.5 mM 4-(2-Aminoethyl)-benzenesulfonylfluoride hydrochloride (AEBSF)), and lysed by mechanical pressure using a One Shot Cell disruption system (Constant Systems Ltd) at pressure 1.35 kBar. After vacuum centrifugation at 18,000 rpm for 30 min at 4°C, pellet and supernatant were loaded on SDS-PAGE and Wester blotting, in order to evaluate the propriety of the expression and the solubility of the expressed protein. The recombinant FliK expects a molecular weight of 57691.4 kDa; the expressed protein revealed a lower molecular weight, approximately 45 kDa, as well as a low expression level profile (Fig. 5.10). Moreover, Western blotting analysis to investigate the presence of the His-tag could not detect any species.

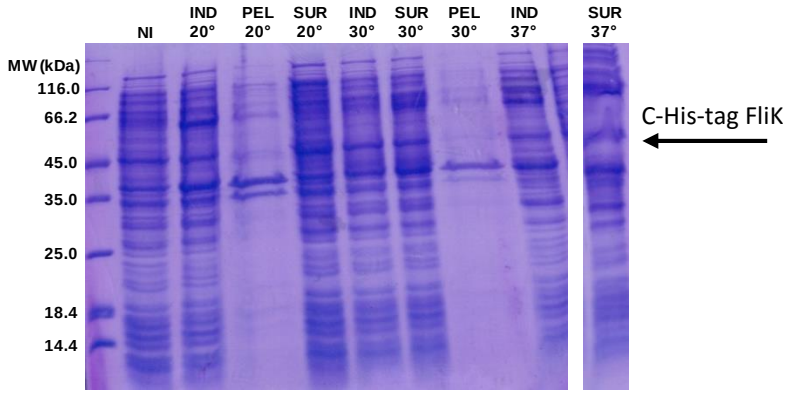

Fig. 5.10 SDS-PAGE analysis of expression and solubilization of 6-His-tagged FliK from *H. pylori*.

5.4 RESULTS AND DISCUSSION

These findings are likely ascribed to the high disorder level of *H. pylori* FliK sequence, structural feature shared by orthologous proteins from several organisms. Firstly, *E. coli* competent cells are not able to properly express and fold the recombinant protein, thus it undergoes degradation; these evidences may explain the unexpected molecular weight

on SDS-PAGE, as well as the low expression level. Moreover, the lacking of signal in Western blotting analysis for His-tag could indicate that the C-terminal domain is degraded, despite its more evident structural organization. It has been chosen to locate the 6-His-tag at the C-terminal end just because it is reported to be more compact and folded. To confirm this C-terminal degradation, some purification attempts by affinity chromatography onto a 1 mL His-Trap HP Ni–NTA column (GE Healthcare) were performed, but the purified fractions were not clearly identified, since the corresponding molecular weight is not attributable neither to expected recombinant FliK nor to evidenced expressed protein of 45 kDa (Fig. 5.11). Therefore, the degradation of the 6-His-tagged C-terminal region is supported, since the expressed protein is not able to bind to the Ni-NTA resin of the column. Alternatively, the recombinant FliK undergoes further degradation during the purification and can not be detected. To be sure of the recombinant construct, FliK gene was PCR-amplified and cloned in pETite C-His Kan vector (Lucigen) for a second time. The plasmid carrying the gene of interest was again sequenced, confirming the correctness of the recombinant construct. Therefore, the overall unstructured organization seems to be limiting factor for the purification of *H. pylori* Flik, as well as for the structural characterization of flagellar hook-length control proteins from other organisms. As previously reported, literature data on structural investigations are very limited. Adoptable strategies to overcome the reported problems could be the cloning and purification of single domains, or selecting more sophisticated system of expression, or co-expressing with FlgD demonstrating to bind to FliK during flagellar hook assembly and thus stabilizing the complex.

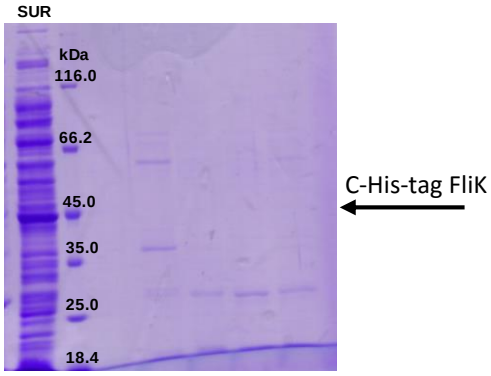

Fig. 5.11 SDS-PAGE analysis of affinity chromatography of 6-His-tagged FliK from *H. pylori*.

Chapter VI

CLONING AND EXPRESSION OF HPG27_1020, A MULTIFUNCTIONAL THIOL: DISULFIDE OXIDOREDUCTASE FROM *HELICOBACTER PYLORI*

6.1 INTRODUCTION

A crucial step in the fast and correct protein-folding pathway is the introduction of disulfide bonds between cysteine residues in a process called oxidative protein folding. Many bacteria use an oxidative protein-folding machinery to assemble proteins that are essential for cell integrity and to produce virulence factors (Heras et al. 2009). Disulfide bonds contribute to the stability and function of many extracytoplasmic, soluble or membrane-bound proteins; the formation of disulfide bonds plays a key role also in bacterial virulence, which often depends on cysteine-rich, extracytoplasmic proteins. In Gram-negative bacteria, the oxidative protein folding takes place in the periplasm and is controlled by proteins from the disulfide bond (Dsb) family. In general, there are two, mostly antagonistic, metabolic pathways acting in the periplasm: an oxidation pathway and an isomerization/reduction pathway, as reported for *Escherichia coli* (Fig. 6.1). The first reaction (catalyzed by DsbA and DsbB) is responsible for the formation of disulfide bonds in the newly synthesized proteins, just after they cross the cytoplasmic membrane. As this process occurs in a non-selective way, a second reaction (driven by DsbC and DsbD) rearranges improperly introduced disulfide bonds (Denoncin and Collet 2012; Bocian-Ostrzycka et al. 2015).

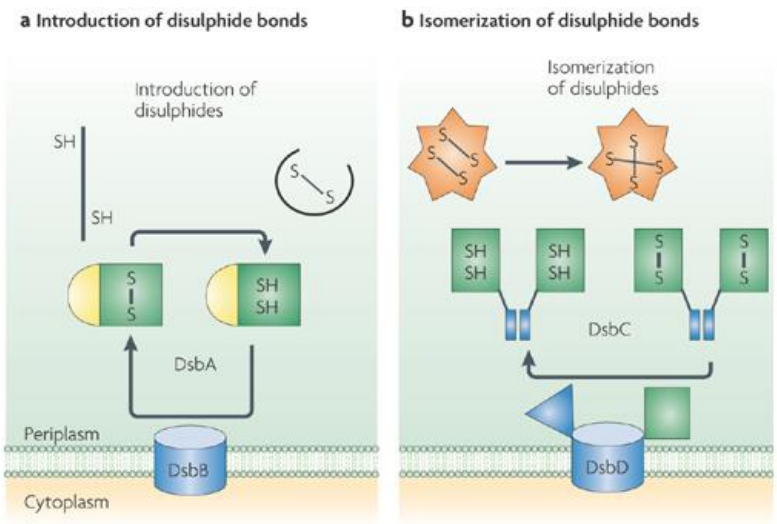

Fig. 6.1 The oxidative apthway and the isomerase pathway in *E. coli*. Adapted from Heras et al., 2009.

In the highly oxidizing environment of the periplasm, there is also a need for selected proteins to be kept in a reduced form. In the assembly of *c*-type cytochromes, essential for electron transfer in a multitude of different cellular processes, energy metabolism, apoptosis (Barker and Ferguson 1999; Martinou, Desagher, and Antonsson 2000; Stevens et al. 2004; Hamel et al. 2009), among the many roles, the cytochrome *c* maturation process requires ligation of heme (iron protoporphirin IX) to reduced thiols of

the Cys-X-X-Cys-His motif of apocytochrome (Bonnard et al. 2010). Biogenesis of cytochrome *c* occurs in the bacterial periplasm, besides the mitochondrial intermembrane space and thylakoid lumen. It consists of the following steps: synthesis and transport across at least one biological membrane of apocytochrome *c* and heme, reduction and maintenance under a reduced form of the heme ferrous iron and the CXXCH sulfhydryls, formation of thioether bonds between the heme and apocytochrome *c* (Hamel et al. 2009). To date, three distinct pathways for cytochrome-*c* biogenesis have been reported, named system I, II, and III, respectively. Specifically, bacteria predominantly use systems I and II, whereas eukaryotes exclusively adopt system III. They comprise two kinds of proteins acting in a coordinated fashion: those involved in heme handling and heme ligation to the apocytochrome, and those contributing to reduction of a disulfide bond of the CXXCH heme-binding motifs (Sanders et al. 2010).

The cytochrome *c* maturation machinery of the system I operating in many Gram-negative bacteria (including *Escherichia coli*), archaea and plant mitochondria, consists of up to ten proteins, named Ccm ABCDEFGHI and DsbD or its shorter version CcdA (Fig. 6.2). The apocytochrome reduction is accomplished by the action of a specific periplasmic thioredoxin-like protein named CcmG (also known as DsbE), as well as of CcmH, and DsbD or CcdA. DsbD/CcdA transfers electrons from cytoplasmic thioredoxin to CcmG. Then CcmG is reoxidized by shuttling its electrons to CcmH, which finally transfers them to apocytochrome *c*. The remaining proteins play a role in heme transport and its ligation to the apocytochrome (Ferguson et al. 2008; Roszczenko et al. 2015).

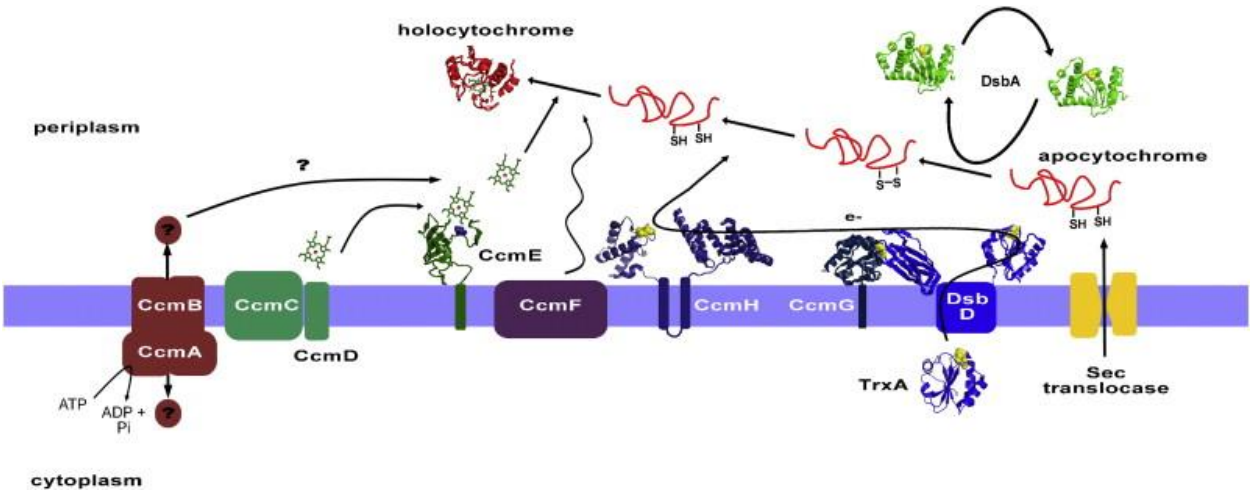

Fig. 6.2 The Ccm (System I) system for *c*-type cytochrome assembly as exemplified by *E. coli*. Adapted from Ferguson et al., 2008.

System II, referred to as the cytochrome *c* synthesis pathway, is simpler and more widespread; it is present in some Gram-negative bacteria (including *H. pylori*), Gram-

positive bacteria (including *Bacillus subtilis*), cyanobacteria and chloroplasts. System II is comprised of four (sometimes three) membrane-bound proteins: ResA (also named CcsX or HelX), ResB (also named CcsB), ResC (also named CcsA) and CcdA, among which ResA is a functional counterpart of CcmG, and CcdA is a functional counterpart of the transmembrane domain of DsbD (Fig. 6.3). CcsA (ResC) and CcsB (ResB) are the components of the cytochrome *c* synthase, whereas CcdA and CcsX (ResA) function in the generation of a reduced heme-binding motif. The CcdA protein is required for reducing CcsX (ResA) and receives its reducing equivalents from thioredoxin in the cytoplasm. In a few bacterial species, such as *Helicobacter*, *Bacteroides* and *Wolinella*, the genes encoding the CcsB and CcsA proteins are naturally fused into one large open reading frame (called *ccsBA*). The integral membrane protein CcsBA acts as a heme exporter and attaches heme to apocytochrome *c* (Simon and Hederstedt 2011; Yoon et al. 2013).

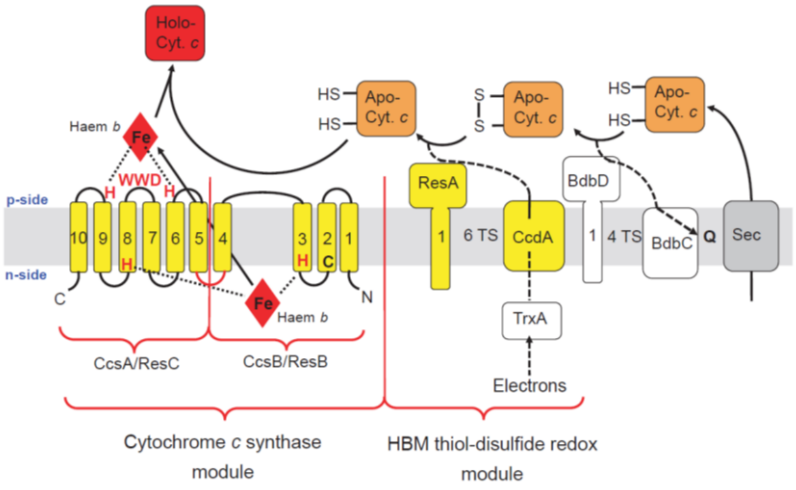

Fig. 6. 3Fig. Overview of the membrane-bound System II machinery for cytochrome *c* biogenesis. Adapted from Simon and Hederstedt, 2011.

Many studies have been carried out to decipher the cooperation among periplasmic Dsb proteins in cytochrome *c* biogenesis, mainly in the maintenance of the proper oxidative status of the CXXC motif of the apocytochrome *c*, showing some peculiarities as regards *H. pylori*. Analysis of its genome nucleotide sequence revealed that *H. pylori* possesses a rather simple respiratory chain consisting of three enzymes: quinol-cytochrome *c* reductase, cytochrome bc_1 complex, cytochrome c_{553} and cb-type cytochrome *c* oxidase (Koyanagi et al. 2000). Little is known about the cytochrome *c* maturation system of this pathogen. In the genome of *H. pylori* 26695 149 proteins containing CXXC motifs characteristic of thiol:disulfide oxidoreductases have been identified to date. Only four of these proteins have a thioredoxin-like (TRX) fold and are periplasm-located, among them HP0377 (HPG27_1020 in strain G27). Additionally, the *H. pylori* genome lacks classical

DsbA and DsbB proteins, members of the Dsb oxidative pathway, as well as classical DsbC and DsbD proteins, members of the Dsb isomerization pathway (Bocian-Ostrzycka et al. 2015). Interestingly, in the genome of *H. pylori* 26695 homologues of the four proteins of system II, namely the cytochrome c maturation system, have been identified. HP0265 is homologous to the CcdA protein. HP0378 contains homologues of two system II components, namely CcsB and CcsA proteins, in a single open reading frame. Finally, HP0377, located directly upstream of HP0378 in a single operon, is homologous to the ResA protein. Specifically, HP0377 shows 15% overall sequence identity compared to ResA of *Bacillus subtilis*, which is required to keep the cysteine residues in the heme-binding CXXCH motif of apocytochrome *c* reduced before heme attachment (Yoon et al. 2013). Although HP0377 was originally described as DsbC homologue, since it was identified as a homologue of the *E. coli* disulfide-bond isomerase DsbC (12% overall sequence identity; Kaakoush et al. 2007), its x-ray crystallographic structure has been determined (Roszczenko et al. 2015; PDB accession codes: 4FYB, 4FYC) and indicates that HP0377 can be characterized as a ResA homologue and, thus, counterpart of CcmG protein. *H. pylori* utilizes the system II pathway for the synthesis of cytochrome *c*. Before covalent attachment of the heme group, the active-site cysteines of apocytochrome c_{553} must be kept reduced by a component of the *H. pylori* system II. HP0377 is a thioredoxin-

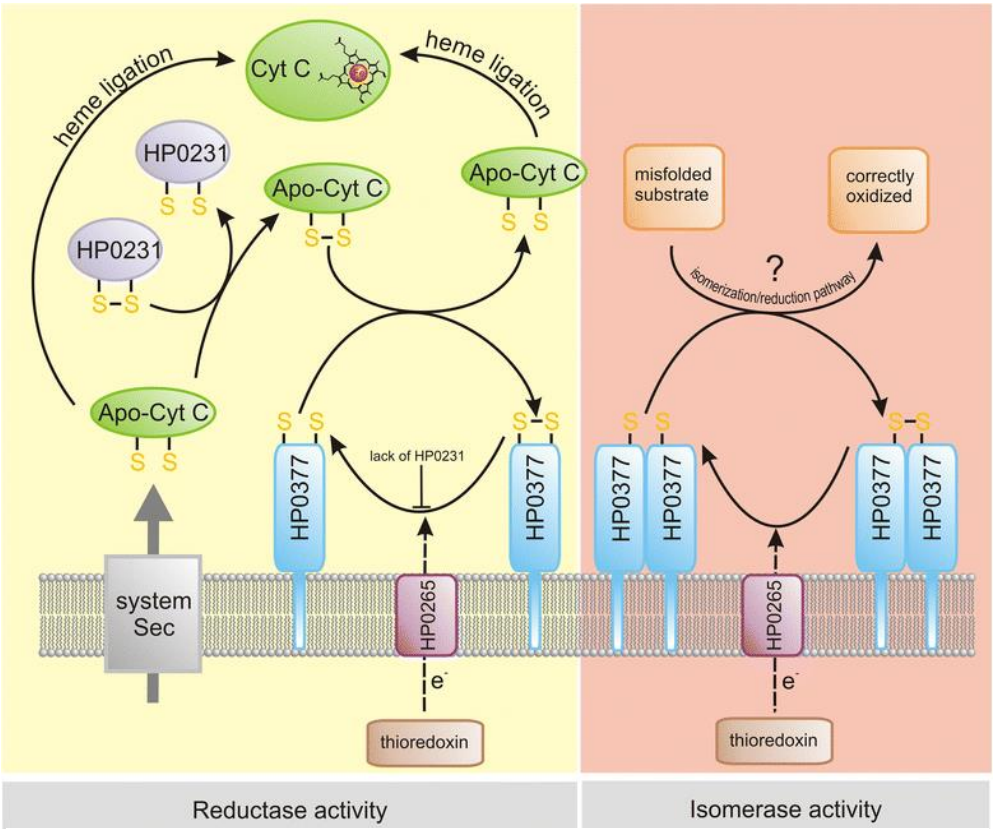

Fig. 6.4 Models representing the role of *H. pylori* HP0377 in cytochrome *c* biogenesis and in oxidized protein folding. Adapted from Roszczenko et al., 2015.

fold protein containing the CSYC motif, which indicates that it functions as a disulfide oxidoreductase (Fig. 6.4; Roszczenko et al. 2015). HP0377 has a classic thioredoxin-fold composed of [β1–(α1a–α1b)–β2]–α2–(β5–β6–α3), with three major modifications. The modifications are the distinct insertion of a two-stranded β-sheet (β3 and β4) between β2 and α2, and extensions at both the N- and C-termini (3_{10}-helix η1 and β0; α4). The CXXC motif and the *cis*-proline loop are conserved active-site features among the redox-active proteins of the thioredoxin-fold and are important in the activity, substrate recognition and stabilization of these proteins. The CXXC motif of HP0377 is located at the N-terminus of the first helix, namely α1, and includes a tyrosine and a serine residues among two

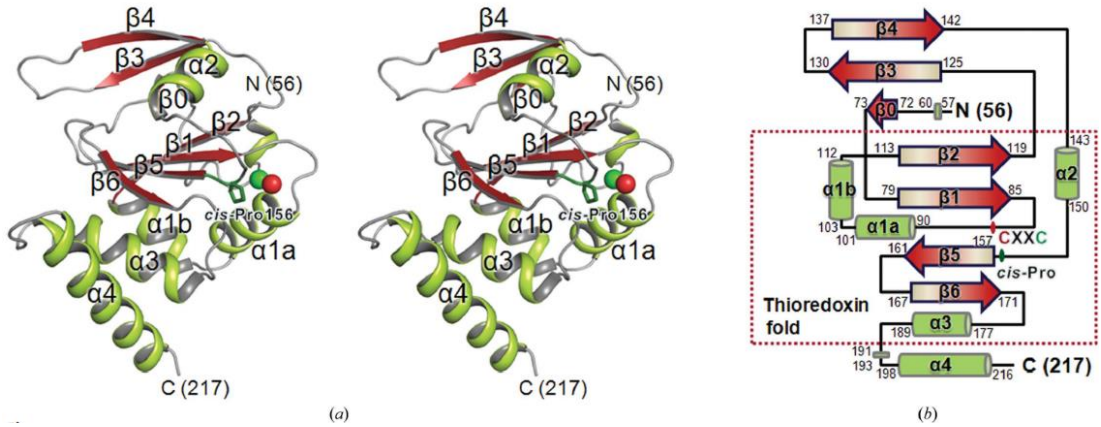

Fig. 6.5 Overall monomer structure and topology of HP0377. (a) Stereoview of an HP0377 monomer in cartoon representation, with the secondary-structure elements labelled. α-helices are depicted in green and β-strands in red. S atoms of Cys89 and Cys92 in the CXXC motif are shown as red and green spheres, respectively; the cis-Pro156 is shown as a green stick model. **(b)** A topology diagram of HP0377 (chain B, oxidized state) is shown with the same colour scheme as in (a). The classic thioredoxin fold is encircled by a red dotted box. Residue numbers for each secondary-structure element are indicated. α-helices, 3_{10}-helices, β-strands and loops are shown as cylinders (green), flat cylinders (green), arrows (graded red) and solid lines, respectively. Adapted from Yoon et al., 2013.

cysteine residues (Fig. 6.5; Yoon et al. 2013). Although there is no evidence that HP0377 is involved in cytochrome *c* assembly *in vivo*, that is the likely case because its resolved structure is similar to that of other CcmG proteins and suggests that HP0377 may function as a reductase towards *H. pylori* apocytochrome c_{553}. Confirming this evidence, in the genome of *H. pylori* 26695 the *hp0377* gene is located between *hemH* (*hp0376*), a gene for heme biosynthesis, and *ccsBA* (*hp0378*), a gene for heme export and the attachment of heme to apocytochrome *c* (Frawley and Kranz 2009; Roszczenko et al. 2015). Moreover, HP0377 has been demonstrated to reduce a putative L,D-transpeptidase (HP0518), which contains a single cysteine residue in the active site, and form a covalent complex *via* the expected disulfide bond, similarly to *E. coli* DsbC, which plays a role as a backup for DsbG in protecting the catalytic cysteine of the YbiS L,D-transpeptidase from oxidation (Yoon et al. 2013). HP0377 and *E. coli* DsbC share a remarkably similar active site, but differ in the oligomerization, since HP0377 exists as a monomer, as well as other

periplasmic thioredoxin-like proteins (*B. subtilis* ResA, *Bradyrhizobium japonicum* CcmG and TlpA, and *E. coli* CcmG), whereas *E. coli* DsbC is a V-shaped dimeric proteins. The structural determination has confirmed that HP0377 possesses a classic thioredoxin-fold with a distinct insertion between α2 and β2 and extensions at both the N- and C-termini (Fig. 6.5). Additionally, in contrast to *B. subtilis* ResA, HP0377 shows no substantial difference between the reduced and partially oxidized structures (Fig. 6.6; Yoon et al. 2013). HP0377 is present in vivo in the reduced form, which is a characteristic feature of thiol oxidoreductases being reductants, and its redox state is influenced by a further thioredoxin-like protein, namely HP0231, creating a redox pair playing a role in introduction of disulfide bonds (Roszczenko et al. 2015). A model of their interaction is shown in Fig. 6.4. The picture evidences how HP0377 seems to be a membrane-anchored lipoprotein, differently than reported by Yoon et al., containing a putative signal sequence with a lipobox that can potentially be processed by signal peptidase II (Paetzel et al. 2002); nevertheless, understanding the role of the HP0377 lipobox requires more investigation. Additionally, the pKa of the N-terminal cysteine of the CXXC motif of HP0377 appears to be similar to those observed for *E. coli* DsbC or DsbA, but not to those determined for most CcmGs. This observation suggests that the activity of HP0377 in *H. pylori* is distinct from that described for the classical CcmGs. Therefore, HP0377, in contrast to most CcmGs that are involved in only the cytochrome *c* biogenesis process, is at least a bifunctional reductase, also showing a Dsb-related isomerization function.

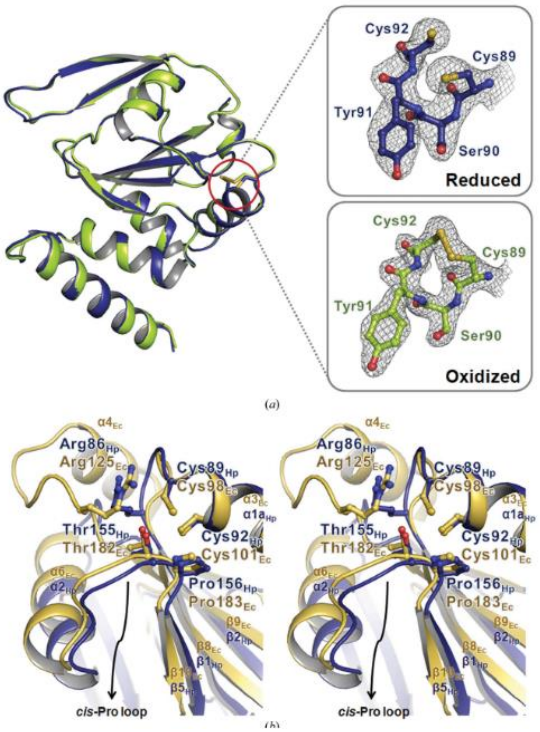

Fig. 6.6 Structural comparisons of reduced HP0377 with partially oxidized HP0377 and reduced *E. coli* DsbC. Adapted from Yoon et al., 2013.

Given the importance of disulfide bond formation to achieve native protein structures and the role of the Dsb proteins in the correct folding of several bacterial virulence factors (Heras et al. 2009), including those from *H. pylori*, this family of proteins represents a possible new drug target. In this regard, HP0377, namely HPG27_1020 in *H. pylori* strain examined, has been considered an attractive target for structural investigations in this research project. To date few crystal structures of bacterial thiol:disulfide oxidoreductase

have been determined, among them ResA from *Bacillus subtilis* (Crow et al. 2004; PDB accession codes: 1ST9, 1SU9), ResA from *Bacillus anthracis* (data not published yet; PDB accession code: 4NMU), DsbC from *Escherichia coli* (McCarthy et al. 2000; PDB accession code: 1EEJ), DsbC from *Salmonella enterica* (Jiao et al. 2013; PDB accession code: 4ILF), DsbC from *Haemophilus influenzae* (Zhang et al. 2004; PDB accession code: 1T3B) and the previously described HP0377 from *H. pylori* 26695 (Yoon et al. 2013; PDB accession codes: 4FYB, 4FYC). Therefore, further structural investigations of these family of proteins could corroborate their use as potential pharmacological target.

6.2 SEQUENCE ANALYSIS

```
 10         20         30         40         50         60
MFSLSYVSKK FLSVLLLISL FLSACKSNNK DKLDENLLSS GSQSSKELND ERDNIDKKSY

        70         80         90        100        110        120
AGLEDVFLDN KSISPNDKYM LLVFGRNGCS YCERFKKDLK NVKELRDYVK EHFSAYYVNI

       130        140        150        160        170        180
SYSKEHDFKV GDKDKNDEKE IKMSTEELAQ IYAIQSTPTI VLSDKTGKTI YELPGYMPST

       190        200        210        220
QFLAVLEFIG DGKYQDAKND EDLTKKLKAY IKYKTNLSKS KSS
```

Property	Value
Aminoacid	223
Cysteines	3
Molecular weight (kDa)	25,552.0
Theoretical pI	8.18
Abs_{280nm} (c=1mg/mL; b=1cm)	0.816 (all Cys reduced)
Signal peptide	Yes

Tab. 6.1 Some properties of HPG27_1020 protein from *H. pylori*.

HP0377 from *H. pylori* G27, namely HPG27_1020, is composed of 223 aminoacids, including 3 cysteine residues, two of which are included in the CXXC motif characteristic of the thioredoxin-fold proteins. The alignment of the aminoacid sequences of HPG27_1020 from *H. pylori* and other bacterial thiol:disulfide oxidoreductases, from *B. subtilis*, *E. coli* and the related *Helicobacter hepaticus*, reports an overall high variability,

but the CXXC motif of the active site is conserved (Fig. 6.7). The secondary structure expected include a core of β-strands (11%) surrounded by α-helices (55%), in accordance to the general structure of the thioredoxin-fold proteins (Phyre^2; Kelley et al. 2015; http://www.sbg.bio.ic.ac.uk/phyre2/). As reported by Yoon et al. 2013, in the case of HP0377 from *H. pylori* 26695, and confirmed by SignalP 4.1 Server prediction (Petersen et al. 2011; http://www.cbs.dtu.dk/services/SignalP), the N-terminal domain includes a 24 aa signal peptide (Fig. 6.8). In order to confirm the thioredoxin-fold, a putative structure of the thiol:disulfide oxidoreductase HPG27_1020 from *H. pylori* G27 has been predicted, using the SWISS-MODEL bioinformatics homology modelling tool (Biasini et al. 2014, http://swissmodel.expasy.org/) and HP0377 from *H. pylori* 26695 as template, which shares 96% sequence identity (PDB accession code: 4FYB); the model has evidenced the typical thioredoxin-fold composed of four-stranded antiparallel β-sheet and three flanking α-helices, added of two two-stranded β-sheet, as reported for the template (Fig. 6.9).

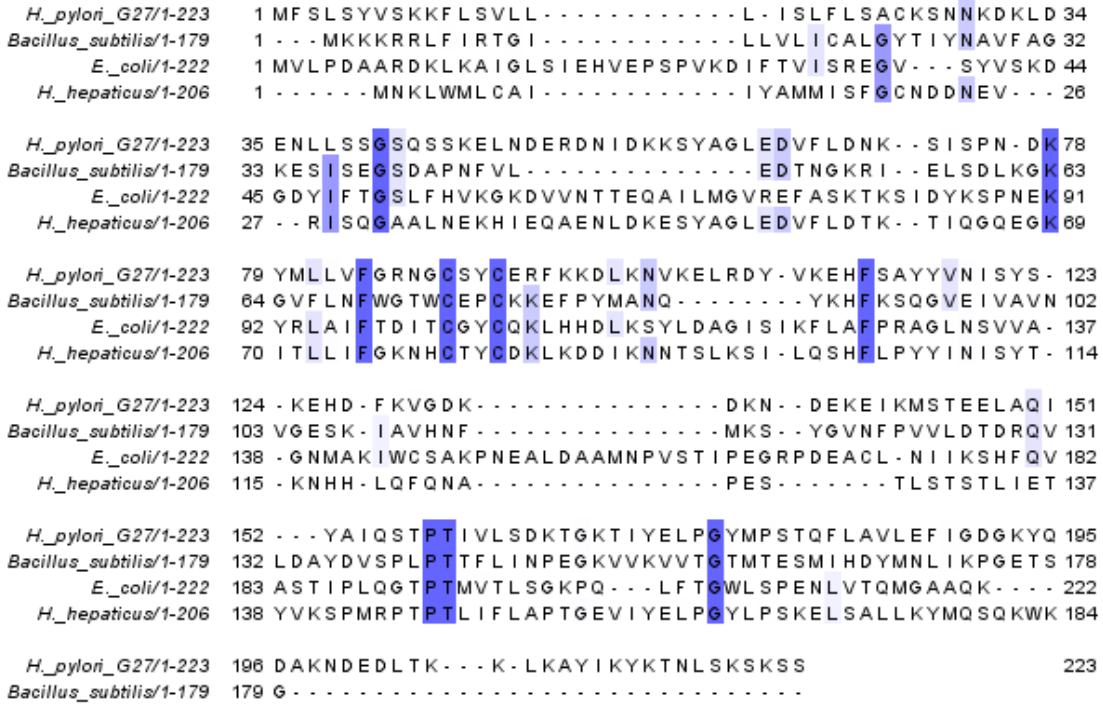

Fig. 6.7 **Sequence alignment of thiol:disulfide oxidoreductases HPG27_1020 from *H. pylori* strain G27, *B. subtilis*, *E. coli* and *H. hepaticus*; in shades of blue the identity above 50% is shown (ClustalW and Jalview ver. 2.9.0b2).**

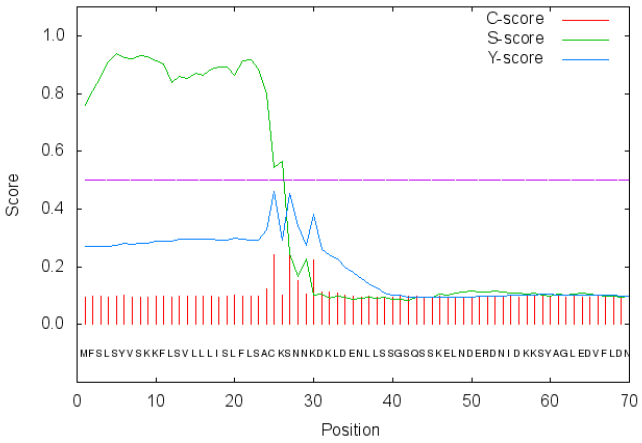

Fig. 6.8 HPG27_1020 from *H. pylori* signal peptide prediction (SignalP 4.1 Server).

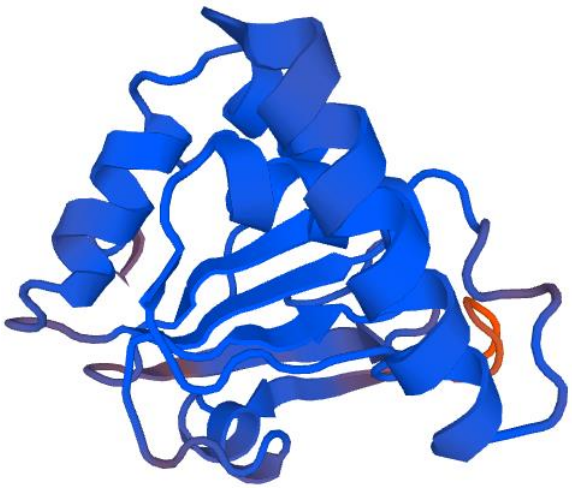

Fig. 6.9 Homology modelling structure of thiol:disulfide oxidoreductases HPG27_1020 from *H. pylori* G27, monomer predicted (SWISS-MODEL; template: thiol:disulfide oxidoreductases HP0377 from *H. pylori* 26695, PDB accession code: 4FYB).

6.3 MATERIALS AND METHODS

6.3.1 MOLECULAR CLONING

The coding sequence for FliK (HPG27_857) was PCR-amplified from genomic DNA of H. pylori strain G27, using Q5® High-Fidelity DNA Polymerase (New England Biolabs) and primers 5'- GAAGGAGATATACATATGTGCAAGTCCAACAATAAAGAC-3' (fw) and 5'-GTGATGGTGGTGATGATGGCTGGATTTGCTTTTAGAAAG-3' (rv). Since the full-length protein is toxic to *E. coli*, the 24 aa N-terminal export signal sequence was excluded and C-terminal 6-His tag was included. The reverse primer provided a 18 nt-sequence that recognizes the pETite C-His Kan vector (Lucigen) and encodes the C-terminal 6-His tag; it did not include a Stop anticodon. The forward primer included a Start codon. It was decided to not introduce a TEV proteolysis site between the 6-His-tag and the C-terminal end of the target protein. The PCR product, purified by 1% agarose gel, was cloned by thermal shock into the pETite C-His Kan vector (Lucigen) using *E. coli* XL1-Blue competent cells (Agilent), grown overnight on selective LB medium supplemented with 30 µg/ml of kanamycin. The colonies were checked by colony-PCR using EconoTaq DNA Polymerase (Lucigen) and T7 primers. Purified plasmid of positive colonies was double-digested by NdeI (New England Biolabs) and NotI (New England Biolabs) restriction

enzyme, for 1 h at 37 °C. Positive samples were finally checked by DNA sequencing to evaluate the quality.

The following sequence corresponds to the final recombinant construct lacking the 24 aa N-terminal export signal and flanked by a C-terminal 6-His tag.

```
 10         20         30         40         50         60
CKSNNKDKLD ENLLSSGSQS SKELNDERDN IDKKSYAGLE DVFLDNKSIS PNDKYMLLVF

        70         80         90        100        110        120
GRNGCSYCER FKKDLKNVKE LRDYVKEHFS AYYVNISYSK EHDFKVGDKD KNDEKEIKMS

       130        140        150        160        170        180
TEELAQIYAI QSTPTIVLSD KTGKTIYELP GYMPSTQFLA VLEFIGDGKY QDAKNDEDLT

       190        200
KKLKAYIKYK TNLSKSKSSH HHHHH
```

Property	Value
Aminoacid	205
Cysteines	3
Molecular weight (kDa)	23,685.5
Theoretical pI	7.23
Abs_{280nm} (C=1mg/mL; b=1cm)	0.818 (all Cys reduced)
Signal peptide	No

Tab. 6.2 Some properties of HPG27_1020 from *H. pylori* referred to the recombinant construct lacking of 24 aa N-terminal export signal and added of C-terminal 6-His-tag.

6.3.2 EXPRESSION

The pETite plasmid carrying the HPG27_1020 gene was transformed into *E. coli* BL21 (DE3) competent cells (Lucigen). The *E. coli* BL21 (DE3) harboring the pETite plasmid were grown in small scale in a selective LB medium supplemented with 30 µg/mL of kanamycin, in order to control and optimize level and quality of protein expression. Multiple 50 mL cultures were grown under mild shaking (180 rpm) at 37 °C until an OD value (optical dispersion, at 600 nm) of about 0.6; then 0.5 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG) was added to the medium to induce protein expression and the cultures were incubated at 37 °C for 3 h, at 30°C overnight, at 20 °C overnight, respectively. The cultured cells were harvested from each small scale culture and medium

eliminated by centrifugation at 5,000 rpm for 30 min at 4 °C. The pellet was resuspended in a lysis buffer containing 50 mM Tris-HCl pH 8.1, 150 mM NaCl, supplemented with a protease inhibitor cocktail (1 mM phenylmethanesulfonyl fluoride (PMSF), 15 µM aprotinin, 1 µM leupeptin, 0.5 mM 4-(2-Aminoethyl)-benzenesulfonylfluoride hydrochloride (AEBSF)), and lysed by mechanical pressure using a One Shot Cell disruption system (Constant Systems Ltd) at pressure 1.35 kBar. After vacuum centrifugation at 18,000 rpm for 30 min at 4°C, pellet and supernatant were loaded on SDS-PAGE and Wester blotting, in order to evaluate the expression level and the solubility of the expressed protein. The highest yield of soluble recombinant protein was obtained under incubation overnight at 30 °C, as reported by SDS-PAGE (Fig. 6.10).

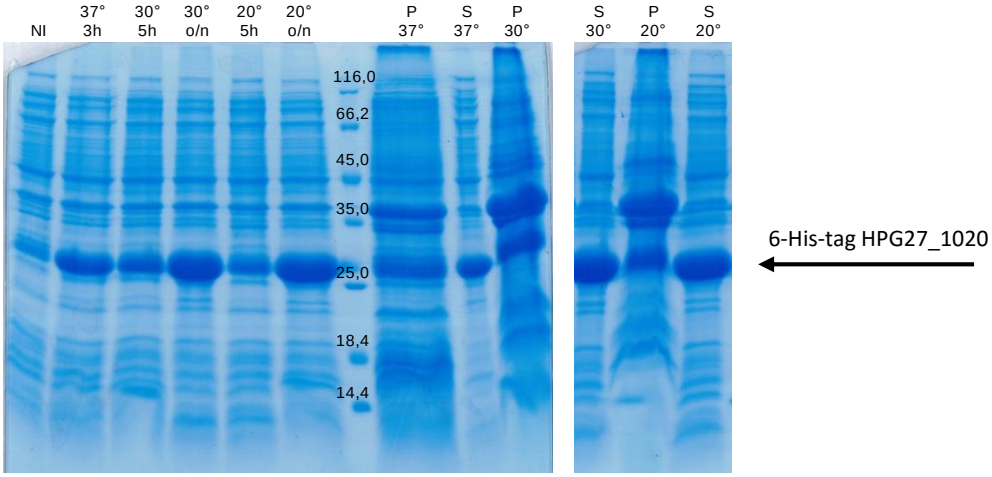

Fig. 6.10 SDS-PAGE analysis of expression and solubilization of 6-His-tag HPG27_1020 from *H. pylori.*

6.3.3 WESTERN BLOTTING

The samples from expression trials of the recombinant HPG27_1020 were investigated for the presence of the His-tag, to ensure the identity of the protein and the yield of soluble fraction. After loading of the sample into a SDS-PAGE, the proteins were transferred on a Hybond-ECL nitrocellulose membrane (GE Healthcare) by electrophoresis. After blocking of the membrane using BSA 3% solution in TBS for 1 h and incubation of the Anti-polyHistidine primary antibody (Sigma-Aldrich) at 4 °C overnight, three washing steps using TTBS and a final one with TBS were performed. The peroxidase-conjugated secondary antibody (Promega) was incubated for 1 h at room temperature in TBS and washed for three times with TTBS and finally with TBS. Subsequently, the membrane was developed using the Pierce ECL Western Blotting Substrate (ThermoFischer) for detection of peroxidase activity, at Image Station 4000 MM PRO (Kodak). Western

blotting analysis confirmed the identity of the recombinant protein and the highest yield of soluble protein under incubation overnight at 30 °C (Fig. 6.11).

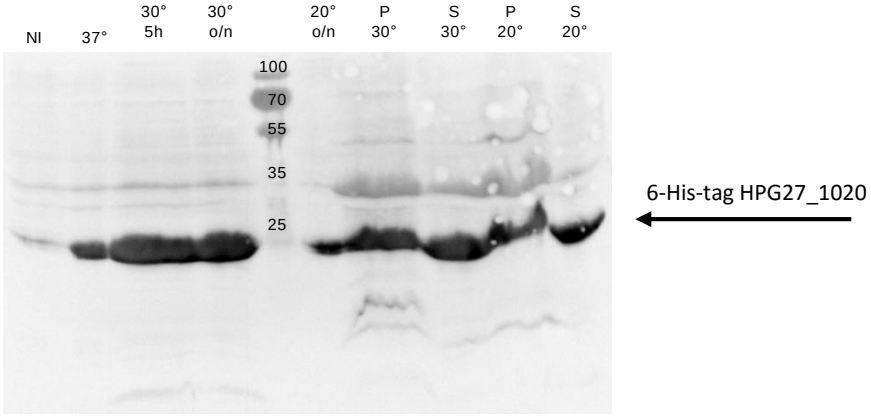

Fig. 6.11 Western blotting analysis of expression and solubilization of 6-His-tag HPG27_1020 from *H. pylori*.

6.4 RESULTS AND DISCUSSION

The thiol:disulfide oxidoreductase HPG27_1020 was previously cloned as full-length protein, but it resulted toxic for the growth of different engineered *E. coli* strains. Therefore, it was decided to exclude the N-terminal signal peptide, as reported in the structural investigation of HP0377 from H. pylori 26695 (Yoon et al. 2013). The recombinant protein lacking 24 aminoacids at the N-terminus demonstrated a high expression level and an optimal solubility; incubation overnight at 30 °C under mild shaking (180 rpm) was identified as the best expression condition for *E. coli* BL21 (DE3) harboring the pETite plasmid carrying the recombinant HPG27_1020 and the soluble fraction of recombinant protein was estimates 60% from SDS-PAGE analysis. On SDS_PAGE analysis the recombinant HPG27_1020 demonstrated a molecular weight lower than expected; but Western blotting analysis confirmed the identity and the molecular weight of the protein of interest. Despite this initial promising results, the investigations were forcedly interrupted since meantime the x-ray structure of the thiol:disulfide oxidoreductase from *H. pylori* 26695, namely HP0377, has been determined and published. The aminoacid sequences from two *H. pylori* strains show a high degree of identity (96%), therefore the investigation has not longer been considered innovative.

Chapter VII

CLONING, EXPRESSION, PURIFICATION AND CRYSTALLIZATION TRIALS OF HYPOTHETICAL PROTEINS FROM *HELICOBACTER PYLORI*

7.1 INTRODUCTION

H. pylori possesses a relatively small genome of less than 1600 genes compared to other Gram-negative bacteria; specifically, the well-characterized strain 26695 contains 1590 open reading frames (ORFs) (Tomb et al. 1997), whereas strain G27, examined in this research project, 1515 (Baltrus et al. 2009), respectively. Interestingly, *H. pylori* possesses several genes that are uncharacterized since: the gene sequences are quite new; the function of genes have not been characterized in any other bacterial systems; sometimes, the protein that is classified into a known protein based on the sequence homology shows some functional ambiguity, which raises questions about the function of the protein produced in *H. pylori* (Park, Son, and Lee 2012). A relevant fraction of these proteins are annotated as “hypothetical proteins”, possibly from 30% to 40% (Fig. 7.1). The function of some of the latter can be hypothesized based on a weak homology with proteins of other organisms, while that of others is completely unknown (Zanotti and Cendron 2014). In order to understand the whole picture of gene functions in *H. pylori* and how this bacterium works as human pathogen, knowledge on the three-dimensional structure of a protein, especially unknown or hypothetical proteins, is frequently useful to elucidate the structure-function relationship of the uncharacterized gene product. Because of its importance as a human pathogen, understanding the mechanism of *H. pylori* colonization and inflammation is the most important part in order to discover effective and specific therapies for *H. pylori*. In this regard, the structural and functional study of *H. pylori* proteins should be the most important point for developing antibiotics and can provide clues to help cope with new antibiotic-resistant bacteria (Kang, Kim, and Lee 2013).

Most of these “hypothetical proteins” are encoded by strain-specific genes and clustered into one locus called hypervariable plasticity zone, which has a lower G+C content, indicating that this region contains fragments acquired by horizontal gene transfer, enriching the pathogen with virulence and survival factors. One of the major challenges is to assign function to and characterize these unclassified proteins, many of which are likely to be novel determinants important in survival and stomach colonization. An initial approach consists in using the genome sequence to perform *in silico* structural and functional prediction. The structure determination can provide detailed insights about the function of these “hypothetical proteins”. In this regard, in this research project it has been considered attractive to include structural studies on these proteins of unknown functions;

specifically, the selected candidates are HPG27_1030 and HPG27_1117, from *H. pylori* strain G27, appointed as “hypothetical proteins”. Their homologues in *H. pylori* 26695, namely HP0367 and HP1173, respectively, were identified by proteomic analysis of the secretome of the pathogen (Bumann et al. 2002); as secreted proteins, they could mediate important pathogen-host interactions and, thus, they represent interesting target for antibiotics and vaccine development.

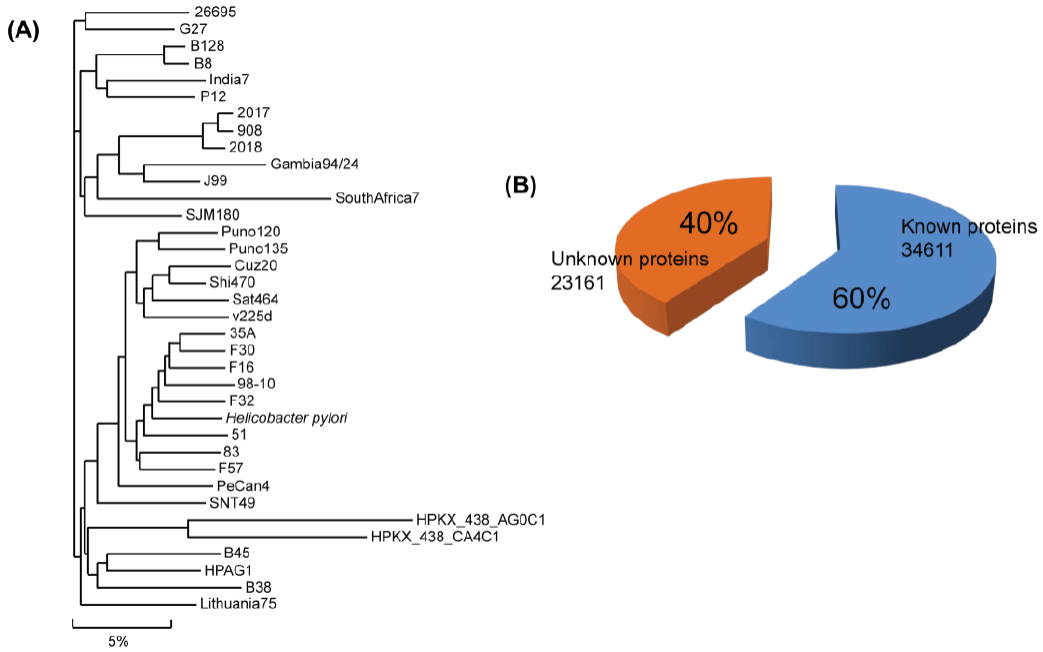

Fig. 7.1 Genome sequence and proteins of *H. pylori*. In the phylogenetic tree, a total of 36 sub-species are branched with a total of about 60,000 genes (A); and among the translated proteins, the biological functions of 40% of the proteins are unidentified (B). Adapted from Park et al., 2012.

7.2 HYPOTHETICAL PROTEIN HPG27_1030

7.2.1 SEQUENCE ANALYSIS

```
 10         20         30         40         50         60
MSKISNNYNP SLMVRDYHTQ RVGSHTKNGE KEENKEIQNL SENDEKIKLA KQAKQDNLAI

         70         80         90        100        110        120
GDLESRLKSL KGMDKDAKEL VGISKAYAHN NEKDRSDFEH FKSRLDKAID SFNQKSGNDG

        130        140        150        160        170        180
LKLPSNIDID DTKALEKFSK SLESEKENIQ NSLHQWKKQL AETNHLNKEY NTLDKTRLNA

        190        200
QKFQDVHDTS KITPSRLQDL LA
```

Property	Value
Aminoacid	202
Cysteines	0
Molecular weight (kDa)	23183.8
Theoretical pI	7.85
Abs_{280nm} (c=1mg/mL; b=1cm)	0.494
Signal peptide	No

Tab. 7.1 Some properties of HPG27_1030 protein from *H. pylori*.

HPG27_1030 from *H. pylori* G27 is a small protein and is composed of 202 aminoacids. There is no evidence of export signals at N- or C- terminal ends, as confirmed by SignalP 4.1 Server prediction (Petersen et al. 2011; http://www.cbs.dtu.dk/services/SignalP). By means of Basic Local Alignment Search Tool (BLAST; http://blast.ncbi.nlm.nih.gov) the aminoacid sequence shows no similarity compared to proteins from organisms different from *Helicobacter* species; it shares 41% and 68% identity with two hypothetical proteins from *Helicobacter cetorum*, respectively, and 80% identity with a hypothetical protein from *Helicobacter acinonychis* (Fig. 7.2). Interestingly, the aminoacid sequence of HPG27_1030 shows complete identity compared to a protein denoted as laminin subunit α2 precursor from a different strain of *H. pylori* (Fig. 7.2). These evidences confirm as this hypothetical protein is a species-specific protein, likely involved in peculiar mechanisms of the pathogen microorganism.

The secondary structure content has been predicted by Phyre2 Server (Kelley et al. 2015; http://www.sbg.bio.ic.ac.uk/phyre2/) and includes 84% α-helices and no presence of β-

structure; therefore, it is plausible to expect a globular structure. To confirm that, a model of HPG27_1030 has been predicted using RaptorX (Källberg et al. 2012; http://raptorx.uchicago.edu/), a web server for protein secondary structure prediction, template-based tertiary structure modeling, alignment quality assessment and sophisticated probabilistic alignment sampling. The model was predicted based on the template 4HPQ_C (transport protein Atg17 from *Lachancea thermotolerans*, chain C) and evidenced an elongated organization, differently from previously announced, and an high solvent accessibility, in accordance with the nature of secreted protein (Fig. 7.3).

Fig. 7.3 Homology modelling structure of HPG27_1030 protein from *H. pylori* G27, monomer predicted (RaptorX web server).

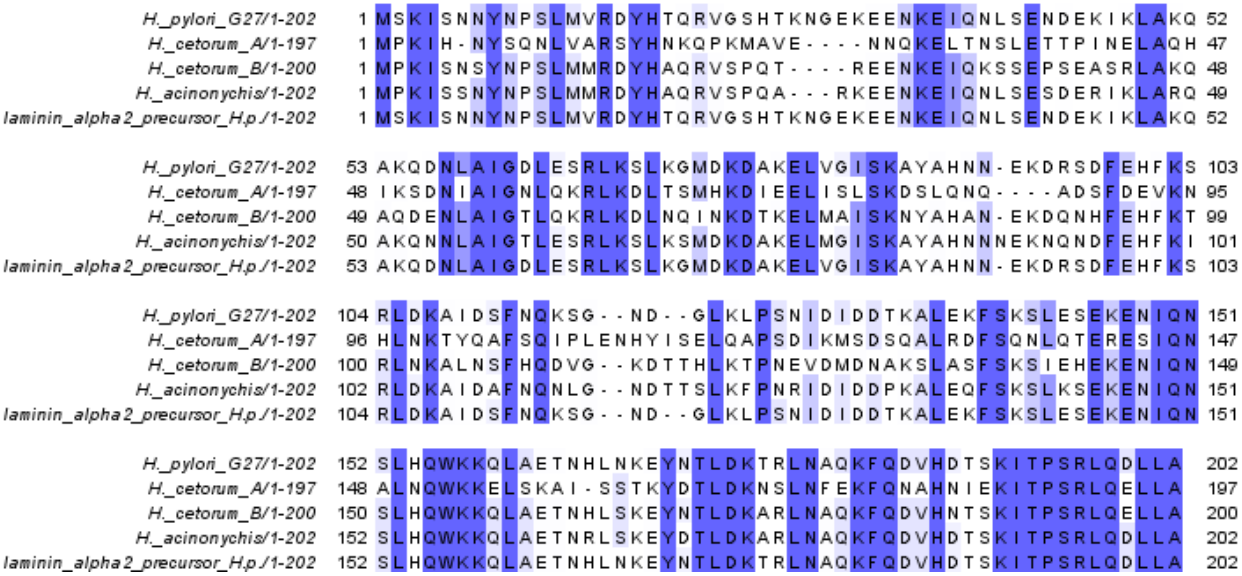

Fig. 7.2 Sequence alignment of HPG27_1030 from *H. pylori* strain G27, hypothetical protein A from *Helicobacter cetorum*, hypothetical protein B from *Helicobacter cetorum*, hypothetical protein from *Helicobacter acinonychis*, and laminin α2 subunit precursor from *Helicobacter pylori*; in shades of blue the identity above 50% is shown (ClustalW and Jalview ver. 2.9.0b2).

7.2.2 MATERIALS AND METHODS

7.2.2.1 MOLECULAR CLONING

The coding sequence for HPG27_1030 protein was PCR-amplified from genomic DNA of *H. pylori* strain G27, using Phusion® High-Fidelity DNA Polymerase (New England

Biolabs) and primers 5' CATCATCACCACCATCACGAAAACCTGTATTTTCAGGGATCTAAGATTTCAAATAAT TATAACC-3' (fw) and 5'-GTGGCGGCCGCTCTATTAAGCGAGCAAGTCTTGCAA-3' (rv). The forward primer provided a 18 nt-sequence that recognizes the pETite N-His Kan vector (Lucigen) and encodes the N-terminal 6-His tag, and a TEV proteolysis site. The reverse primer included a Stop anticodon. The PCR product, purified by 1% agarose gel, was cloned by thermal shock into the pETite N-His Kan vector (Lucigen) using *E. coli* 10G competent cells (Lucigen), grown overnight on selective LB medium supplemented with 30 µg/ml of kanamycin. The colonies were checked by colony-PCR using EconoTaq DNA Polymerase (Lucigen) and T7 primers. Purified plasmid of positive colonies was digested by EcoRV restriction enzyme (New England Biolabs), for 1 h at 37 °C. Positive samples were finally checked by DNA sequencing to evaluate the quality.

The following sequence corresponds to the final recombinant construct with the addition of a N-terminal 6-His tag flanked by a TEV recognition site.

```
 10         20         30         40         50         60
HHHHHHENLY FQ//GSKISNNY NPSLMVRDYH TQRVGSHTKN GEKEENKEIQ NLSENDEKIK

        70         80         90        100        110        120
LAKQAKQDNL AIGDLESRLK SLKGMDKDAK ELVGISKAYA HNNEKDRSDF EHFKSRLDKA

       130        140        150        160        170        180
IDSFNQKSGN DGLKLPSNID IDDTKALEKF SKSLESEKEN IQNSLHQWKK QLAETNHLNK

       190        200        210
EYNTLDKTRL NAQKFQDVHD TSKITPSRLQ DLLA
```

// corresponds to the cleavage site for TEV protease

Property	Value
Aminoacid	214
Cysteines	0
Molecular weight (kDa)	24727.3
Theoretical pI	7.39
Abs_{280nm} (c=1mg/mL; b=1cm)	0.524
Signal peptide	No

Tab. 7.2 Some properties of HPG27_1030 protein from *H. pylori* referred to the recombinant construct with the addition of a N-terminal 6-His-tag flanked by a TEV recognition site.

7.2.2.2 EXPRESSION

The pETite plasmid carrying the HPG27_1030 gene was transformed into *E. coli* BL21 (DE3) competent cells (Lucigen). The *E. coli* BL21 (DE3) harboring the pETite plasmid were grown in a selective LB medium supplemented with 30 µg/mL of kanamycin. A 2 L culture was grown under mild shaking (180 rpm) at 37 °C until an OD value (optical dispersion, at 600 nm) of about 0.6; then 0.5 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG) was added to the medium to induce protein expression and the culture was incubated at 20 °C overnight. The cultured cells were harvested and medium eliminated by centrifugation at 5,000 rpm for 30 min at 4 °C. The pellet was resuspended in a selected lysis buffer consisting of 50 mM MOPS pH 6.8, 150 mM NaCl, 5 mM imidazole, supplemented with a protease inhibitor cocktail (1 mM phenylmethanesulfonyl fluoride (PMSF), 15 µM aprotinin, 1 µM leupeptin, 0.5 mM 4-(2-Aminoethyl)-benzenesulfonylfluoride hydrochloride (AEBSF)), and lysed by mechanical pressure using a One Shot Cell disruption system (Constant Systems Ltd) at pressure 1.35 kBar. The lysis procedure was repeated twice to be more effective. The lysed cell suspension was cleared of debris by vacuum centrifugation at 18,000 rpm for 30 min at 4°C. Pellet and supernatant were loaded on SDS-PAGE and Wester blotting, in order to evaluate the expression level and the solubility of the expressed protein. The recombinant HPG27_1030 protein exhibited an optimal profile of expression and a high yield of soluble fraction, approximately 80%, as estimated from SDS-PAGE analysis and in accordance with its nature of secreted protein (data not shown).

7.2.2.3 PURIFICATION

The HPG27_1030 protein was firstly purified by immobilized metal ion affinity chromatography (IMAC). The cleared supernatant was loaded onto a 1 mL His-Trap HP Ni–NTA column (GE Healthcare), previously equilibrated with the lysis buffer, at a flow rate of approximately 1 mL/min. The column was extensively washed with the lysis buffer and subsequently with a buffer supplemented with 5% elution buffer (50 mM MOPS pH 6.8, 150 mM NaCl, 500 mM imidazole), in order to eliminate unspecifically bound species. The 6-His-tagged protein was eluted at approximately 250 mM imidazole, by applying a linear gradient from 5% to 100% elution buffer. The HPG27_1030 protein elution profile showed a shoulder peak (Fig. 7.3, left) ascribed to degradation processes, as confirmed by SDS-PAGE analysis (Fig. 7.3, right). The 6-His-tagged HPG27_1030 protein ran on

SDS-PAGE with an apparent molecular weight of approximately 26 kDa, in agreement with its calculated molecular mass (24727.3 kDa).

The fractions containing the 6-His-tagged HPG27_1030 showing a profile of lower degradation, corresponding to the peak, were pooled and concentrated by ultrafiltration using a Vivaspin 20 5,000 MWCO centrifugal concentrator (Sartorius). Subsequently, the concentrated sample was further purified by gel-filtration chromatography using a HiLoad 16/60 Superdex 75 preparative grade column (GE Healthcare), equilibrated with a buffer containing 50 mM MOPS pH 6.8, 150 mM NaCl. The recombinant protein was eluted as single peak, with a good profile and a retention volume of 62 mL, corresponding to the molecular weight of the monomer (Fig. 7.4, left).

The fractions collected from the peak at 62 mL volume were checked on SDS-PAGE (Fig. 7.4, right), which still showed a degradation profile, although the protein of interest was the major species, and concentrated by ultrafiltration using a Vivaspin 20 5,000 MWCO centrifugal concentrator (Sartorius), for crystallization trials. After purification approximately 39 mg of whole purified protein were obtained starting from 1 L *E. coli* culture; nevertheless, the concentration of purified sample for crystallization trials was limited, since the protein started to precipitate above 3 mg/mL. The final protein concentration was verified by UV-Vis absorption spectrum (250-340 nm; NanoDrop 2000 UV-Vis Spectrophotometer). Additionally, it was decided to not perform the 6-His-tag removal for initial crystallization test.

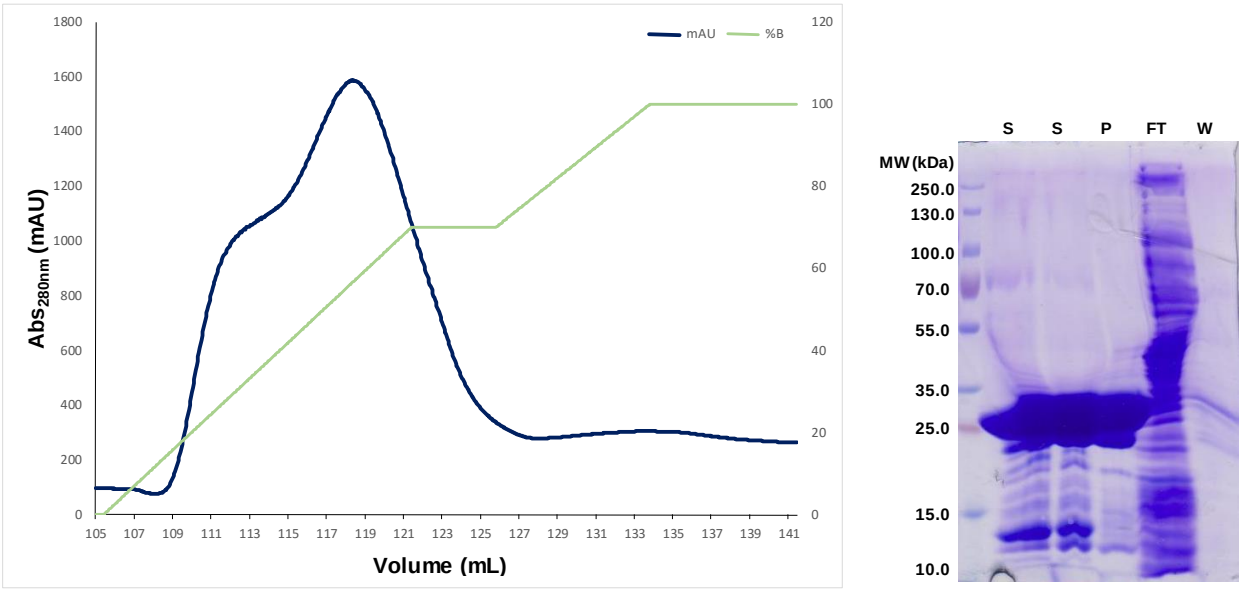

Fig. 7.3 6-His-tagged HPG27_1030 affinity chromatography; left: sbsorption profile; right: SDS-PAGE analysis of the fractions collected (S: shoulder; P: peak; FT: flow-throw; W: wash).

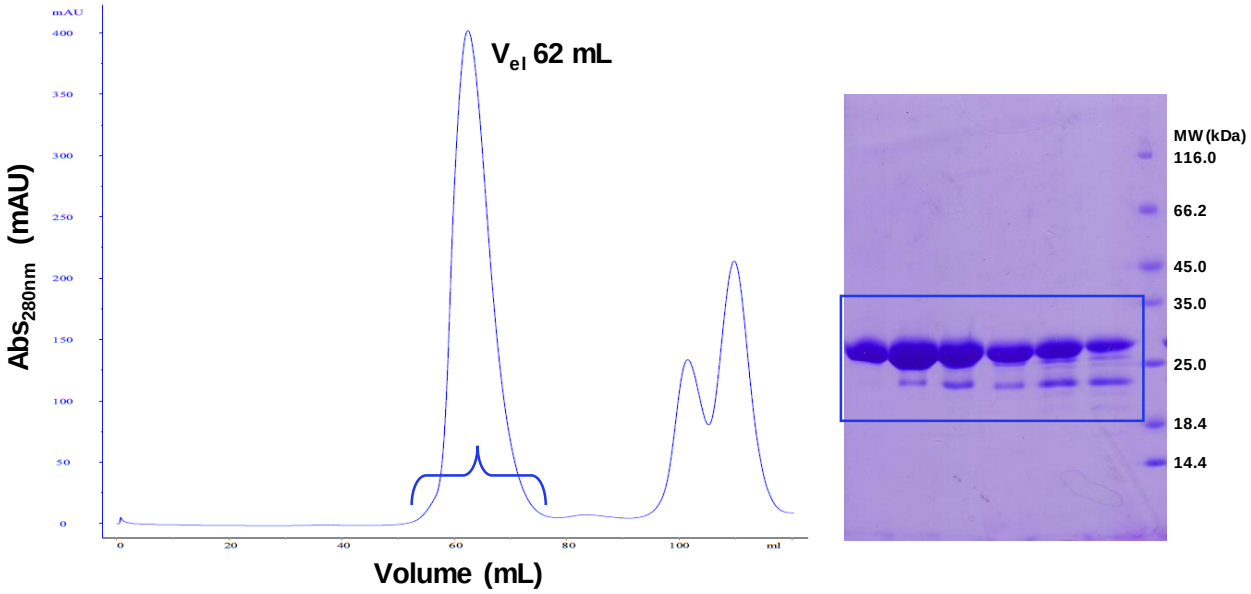

Fig. 7.4 6-His-tagged HPG27_1030 gel-filtration chromatography (HiLoad 16/60 Superdex 75 pg); left: absorption profile; right: SDS-PAGE analysis of the fractions collected from the peak at 62 mL volume.

7.2.2.4 CRYSTALLIZATION TRIALS

Purified 6-his-tagged HPG27_1030 protein samples were used for crystallization tests, at a limited concentration (~3 mg/mL). Crystallization trials with sitting-drop vapor-diffusion technique were carried out using an Oryx8 crystallization robot (Douglas Instruments). Several standard crystal screening kits were explored, including Structure Screen I and II (Molecular Dimensions), Crystal Screen I and II (Hampton Research), PACT Suite (Qiagen), JCSG Suite (Qiagen), PEGs II (Qiagen). Hanging-drop vapor-diffusion experiments were also performed, using the previous crystal screening kits. Despite the limited protein concentration, several conditions showed precipitation even immediately after drop deposition, most likely due to the inhomogeneity of the protein sample, prone to degradation. Microcrystalline species were observed only in few crystallization conditions, not suitable for x-ray diffraction measurement. Several crystallization attempts were performed, but suitable crystals were not obtained.

7.3 HYPOTHETICAL PROTEIN HPG27_1117

7.3.1 SEQUENCE ANALYSIS

```
 10         20         30         40         50         60
MSYFYKHCLK FSLVGLLGLL SVQLDARSFV DGDLDIQKFS YEDSLLKKGD PNGVHKVQVR

        70         80         90        100        110        120
DYKGKMQEAE IHSEIRIALK PGVKKEVKKG KIYSAQINDG MCYAFRMLQT GDNTTGLDSK

       130        140        150        160        170        180
EFPKQSREKK GRVITLIGKD EVPYLILETD CQVGDIAKIS LVGNFDGTGF LTEYKFKDAK


PIY
```

Property	Value
Aminoacid	183
Cysteines	3
Molecular weight (kDa)	20643.8
Theoretical pI	8.79
Abs_{280nm} (c=1mg/mL; b=1cm)	0.650 (all Cys reduced)

Tab. 7.3 Some properties of HPG27_1117 protein from *H. pylori*.

HPG27_1117 from *H. pylori* G27 is a small protein composed of 183 aminoacids. By means of Basic Local Alignment Search Tool (BLAST; http://blast.ncbi.nlm.nih.gov) the aminoacid sequence of HPG27_1117 exhibits no similarity compared to proteins from organisms different from *Helicobacter* species; on the contrary, it shares 83% and 51% identity with two hypothetical proteins from *Helicobacter cetorum*, respectively, and 87% identity with a hypothetical protein from *Helicobacter acinonychis*, in the aligned regions (Fig. 7.5). Additionally, the aminoacid sequence of HPG27_1117 shows 31% identity compared to a hypothetical protein from *Hirsutella minnesotensis*, a fungal species (Fig.). As reported for the hypothetical protein HPG27_1030, also HPG27_1117 is most likely a species-specific protein, involved in survival and pathogenicity mechanisms peculiar of the microorganism.

The secondary structure content has been predicted by Phyre2 Server (Kelley et al. 2015; http://www.sbg.bio.ic.ac.uk/phyre2/) and includes 17% α-helices and 48% β-strands, and a limited grade of disordered structure. A model of the hypothetical protein HPG27_1117 has been predicted using RaptorX (Källberg et al. 2012; http://raptorx.uchicago.edu/), a

web server for protein secondary structure prediction, template-based tertiary structure modeling, alignment quality assessment and sophisticated probabilistic alignment sampling. The model was predicted based on the template 2GE8_A (β-mannosidase Btman2A from *Bacteroides thetaiotaomicron*, chain A) and evidenced a prevalence of β-structure, similar to a β-barrel organization, and a N-terminal elongated α-helix region (Fig. 7.6).

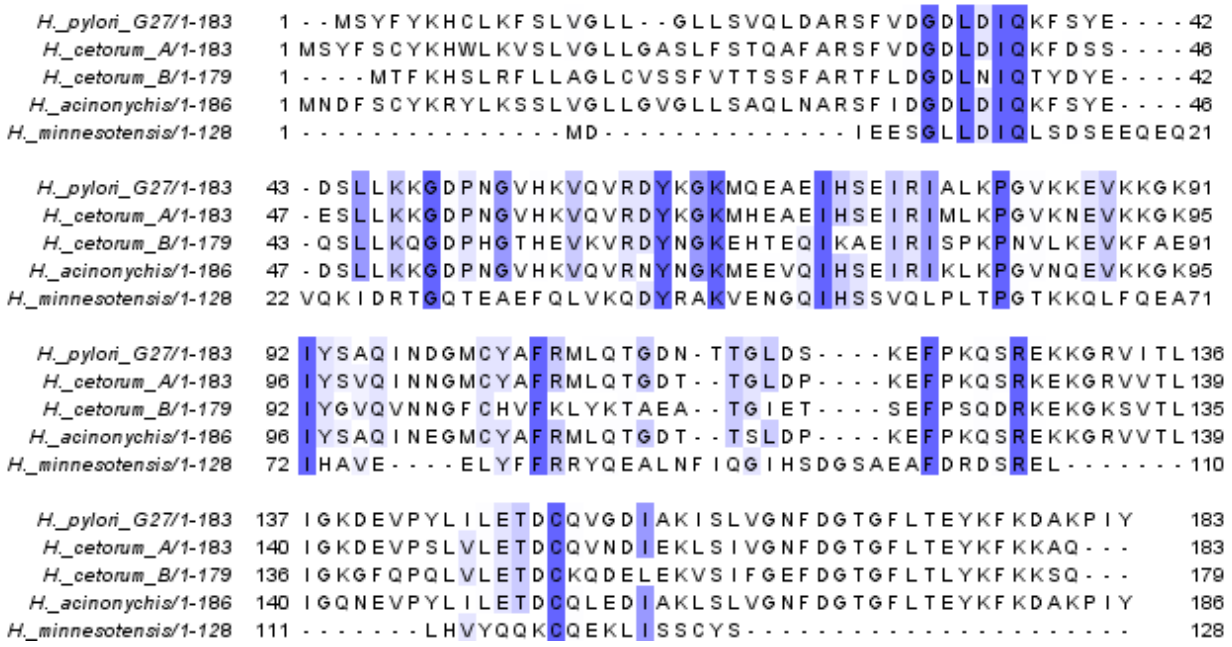

Fig. 7.5 Sequence alignment of HPG27_1117 from *H. pylori* strain G27, hypothetical protein A *from Helicobacter cetorum*, hypothetical protein B *from Helicobacter cetorum*, hypothetical protein from *Helicobacter acinonychis*, and hypothetical protein from *Hirsutella minnesotensis*; in shades of blue the identity above 50% is shown (ClustalW and Jalview ver. 2.9.0b2).

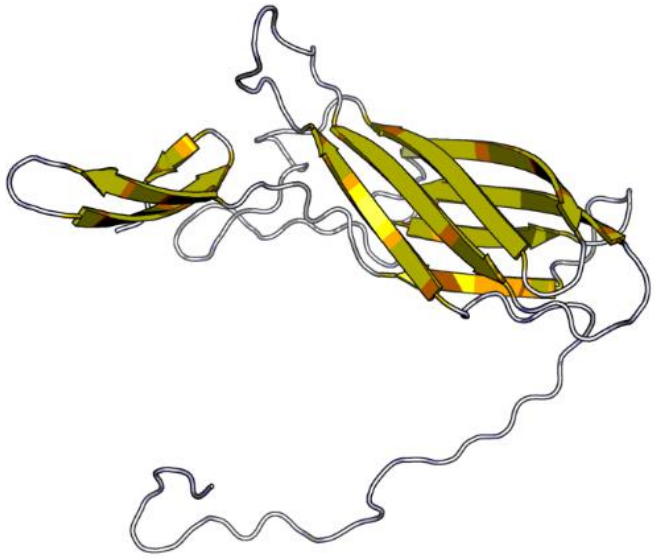

Fig. 7.6 Homology modelling structure of HPG27_1117 protein from *H. pylori* G27 (RaptorX web server).

7.3.2 MATERIALS AND METHODS

7.3.2.1 MOLECULAR CLONING

The coding sequence for HPG27_1117 protein was PCR-amplified from genomic DNA of H. pylori strain G27, using Phusion® High-Fidelity DNA Polymerase (New England Biolabs) and primers 5'-CATCATCACCACCATCACGAAAACCTGTATTTTCAGGGAAGTTATTTTTATAAGCAC TGTTTG-3' (fw) and 5'-GTGGCGGCCGCTCTATTAGTAAATGGGTTTAGCGTCTTT-3' (rv). The forward primer provided a 18 nt-sequence that recognizes the pETite N-His Kan vector (Lucigen) and encodes the N-terminal 6-His tag, and a TEV proteolysis site. The reverse primer included a Stop anticodon. The PCR product, purified by 1% agarose gel, was cloned by thermal shock into the pETite N-His Kan vector (Lucigen) using *E. coli* 10G competent cells (Lucigen), grown overnight on selective LB medium supplemented with 30 µg/ml of kanamycin. The colonies were checked by colony-PCR using EconoTaq DNA Polymerase (Lucigen) and T7 primers. Purified plasmid of positive colonies was digested by EcoRV restriction enzyme (New England Biolabs), for 1 h at 37 °C. Positive samples were finally checked by DNA sequencing to evaluate the quality.

The following sequence corresponds to the final recombinant construct with the addition of a N-terminal 6-His tag flanked by a TEV recognition site.

```
 10         20         30         40         50         60
HHHHHHENLY FQ//GSYFYKHC LKFSLVGLLG LLSVQLDARS FVDGDLDIQK FSYEDSLLKK

        70         80         90        100        110        120
GDPNGVHKVQ VRDYKGKMQE AEIHSEIRIA LKPGVKKEVK KGKIYSAQIN DGMCYAFRML

       130        140        150        160        170        180
QTGDNTTGLD SKEFPKQSRE KKGRVITLIG KDEVPYLILE TDCQVGDIAK ISLVGNFDGT

       190
GFLTEYKFKD AKPIY
```

// corresponds to the cleavage site for TEV protease

Property	Value
Aminoacid	195
Cysteines	3
Molecular weight (kDa)	22187.4
Theoretical pI	8.59
Abs_{280nm} (c=1mg/mL; b=1cm)	0.672 (all Cys reduced)

Tab. 7.4 Some properties of HPG27_1117 protein from *H. pylori* referred to the recombinant construct with the addition of a N-terminal 6-His tag flanked by a TEV recognition site.

7.3.2.2 EXPRESSION

The pETite plasmid carrying the HPG27_1117 gene was transformed into *E. coli* BL21 (DE3) competent cells (Lucigen). The *E. coli* BL21 (DE3) harboring the pETite plasmid were grown in a selective LB medium supplemented with 30 µg/mL of kanamycin. A 2 L culture was grown under mild shaking (180 rpm) at 37 °C until an OD value (optical dispersion, at 600 nm) of about 0.6; then 0.5 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG) was added to the medium to induce protein expression and the culture was incubated at 37 °C for 3 h. The cultured cells were harvested and medium eliminated by centrifugation at 5,000 rpm for 30 min at 4 °C. The pellet was resuspended in a selected lysis buffer consisting of 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 5 mM imidazole, 1 mM DTT, supplemented with a protease inhibitor cocktail (1 mM phenylmethanesulfonyl fluoride (PMSF), 15 µM aprotinin, 1 µM leupeptin, 0.5 mM 4-(2-Aminoethyl)-benzenesulfonylfluoride hydrochloride (AEBSF)), and lysed by mechanical pressure using a One Shot Cell disruption system (Constant Systems Ltd) at pressure 1.35 kBar. The lysis procedure was repeated twice to be more effective. The lysed cell suspension was cleared of debris by vacuum centrifugation at 18,000 rpm for 30 min at 4°C. Pellet and supernatant were loaded on SDS-PAGE and Wester blotting, in order to evaluate the expression level and the solubility of the expressed protein. The recombinant HPG27_1117 protein exhibited a limited expression level; on the contrary, the yield of soluble protein was significant, approximately 80%, as estimated by SDA-PAGE analysis and in accordance with its nature of secreted protein. Additionally, the 6-His-tagged HPG27_1117 protein ran on SDS-PAGE with an apparent molecular weight of approximately 27 kDa, not in full agreement with its calculated molecular mass (22187.4

kDa). To confirm the identity of the protein of interest, a Western blotting analysis was performed to investigate the presence of the 6-His-tag.

7.3.2.3 WESTERN BLOTTING

Samples from expression and solubility test of the recombinant HPG27_1117 protein were investigated for the presence of the His-tag. After loading of the sample into a SDS-PAGE, the proteins were transferred on a Hybond-ECL nitrocellulose membrane (GE Healthcare) by electrophoresis. After blocking of the membrane using BSA 3% solution in TBS for 1 h and incubation of the Anti-polyHistidine primary antibody (Sigma-Aldrich) at 4 °C overnight, three washing steps using TTBS and a final one with TBS were performed. The peroxidase-conjugated secondary antibody (Promega) was incubated for 1 h at room temperature in TBS and washed for three times with TTBS and finally with TBS. Subsequently, the membrane was developed using the Pierce ECL Western Blotting Substrate (ThermoFischer) for detection of peroxidase activity, at Image Station 4000 MM PRO (Kodak). The Western blotting analysis confirmed the identity of the recombinant protein, despite of its apparent molecular weight of approximately 27 kDa (Fig. 7.7).

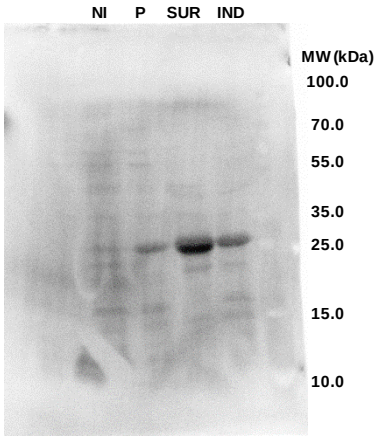

Fig. 7.7 Western blotting analysis of 6-His-tagged HPG27_1117 (NI: not induced *E. coli* culture; P: lysed pellet; SUR: supernatant; IND: induced *E. coli* culture).

7.3.2.4 PURIFICATION

The HPG27_1117 protein was firstly purified by immobilized metal ion affinity chromatography (IMAC). The cleared supernatant was loaded onto a 1 mL His-Trap HP

Ni–NTA column (GE Healthcare), previously equilibrated with the lysis buffer, at a flow rate of approximately 1 mL/min. The column was extensively washed firstly with the lysis buffer, following with a buffer consisting of 50 mM Tris-HCl pH 7.5, 300 mM NaCl, 5 mM imidazole, 1 mM DTT, and finally with lysis buffer supplemented with 5% elution buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 500 mM imidazole, 1 mM DTT), in order to eliminate unspecifically bound species. The 6-His-tagged HPG27_1117 was eluted at approximately 150 mM imidazole, by applying a linear gradient from 5% to 100% elution buffer. The HPG27_1117 protein was eluted as single peak (Fig. 7.8, left); however, the SDS-PAGE analysis of the fractions related to the elution peak showed several species of lower molecular weight, ascribed to degradation processes (Fig. 7.8, right).

The fractions containing the 6-His-tagged HPG27_1117 as clearly predominant species were pooled and concentrated by ultrafiltration using a Vivaspin 20 5,000 MWCO centrifugal concentrator (Sartorius). Subsequently, the concentrated sample was further purified by gel-filtration chromatography using a HiLoad 16/60 Superdex 75 preparative grade column (GE Healthcare), equilibrated with a buffer containing 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1mM DTT. The recombinant protein was eluted as single peak, with a clear profile, although not very intense, and a retention volume of 63 mL, corresponding to the molecular weight of the monomer (Fig. 7.9).

The fractions collected from the peak at 63 mL volume were checked on SDS-PAGE, which showed a limited degradation profile (data not shown), and concentrated by ultrafiltration using a Vivaspin 20 5,000 MWCO centrifugal concentrator (Sartorius), for crystallization trials. After purification a limited amount of whole purified protein, approximately 7 mg, was obtained starting from 1 L *E. coli* culture, owing to the low level of expression. The recombinant protein purified was concentrated till 18 mg/mL, verified by UV-Vis absorption spectrum (250-340 nm; NanoDrop 2000 UV-Vis Spectrophotometer). Additionally, it was decided to not perform the 6-His-tag removal for initial crystallization test.

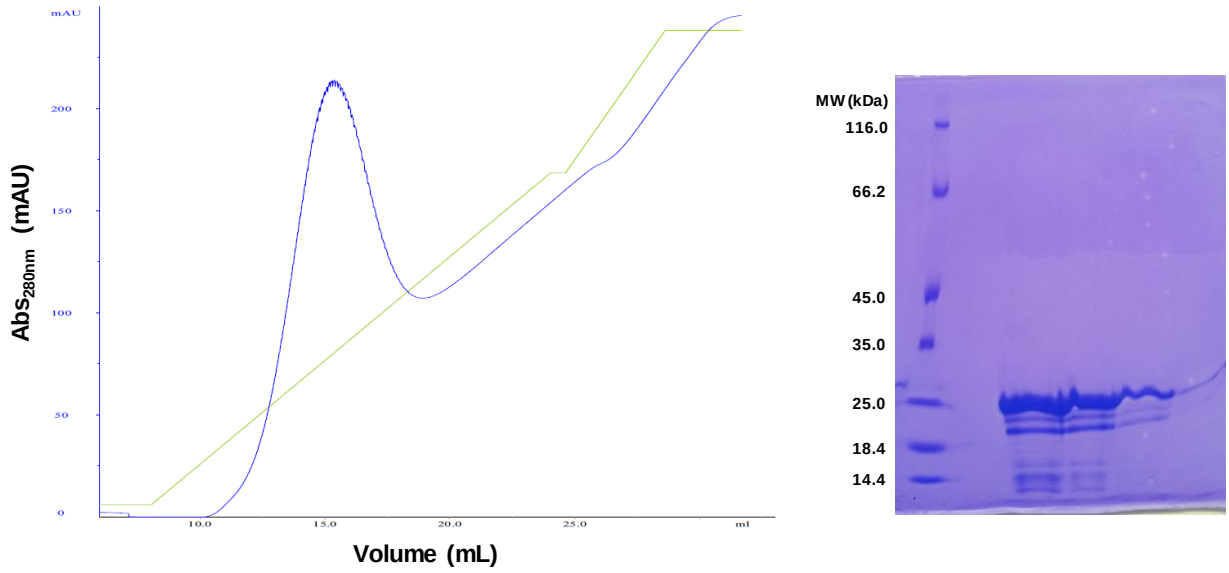

Fig. 7.8 6-His-tagged HPG27_1117 affinity chromatography; left: absorption profile; right: SDS-PAGE analysis of the fractions collected from the peak.

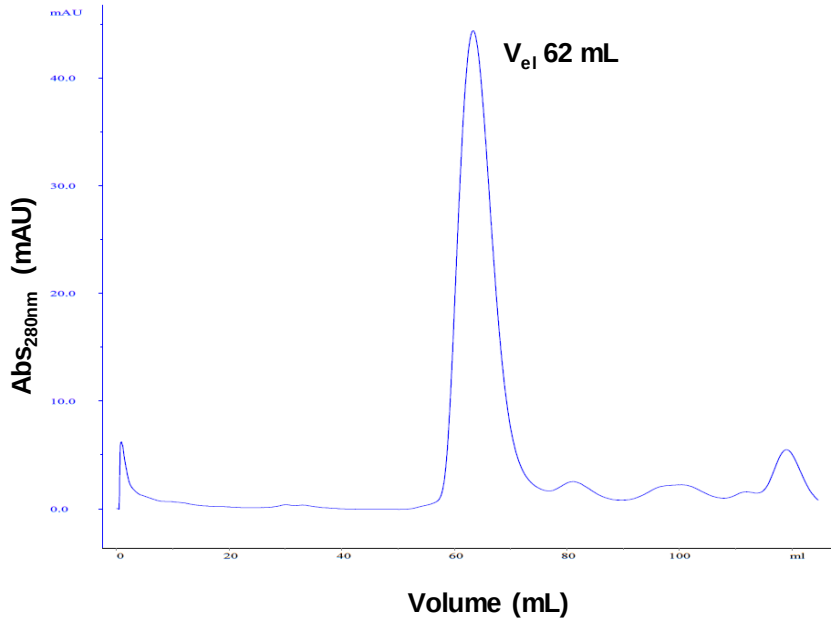

Fig. 7.9 6-His-tagged HPG27_1117 gel-filtration chromatography (HiLoad 16/60 Superdex 75 pg): absorption profile.

7.3.2.5 CRYSTALLIZATION TRIALS

Purified 6-his-tagged HPG27_1117 protein concentrated to 18 mg/mL was used for crystallization tests. Crystallization trials were carried out using the sitting-drop vapor-diffusion technique with an Oryx8 crystallization robot (Douglas Instruments). Several standard crystal screening kits were explored, including Structure Screen I and II

(Molecular Dimensions), Crystal Screen I and II (Hampton Research), PACT Suite (Qiagen), JCSG Suite (Qiagen), PEGs II (Qiagen). Crystallization 96-well plates (Douglas Instrument) were performed at 20 °C and testing various dilutions of the precipitant solutions and drop dimensions. Preliminary crystals were obtained in the following screening conditions: n. 56 from Crystal Screen (1.5 M NaCl, 10% (v/v) EtOH), n. 23 from Structure Screen I (0.2 M $CaCl_2$, 0.1 M Na HEPES pH 7.5, 28% (v/v) PEG 400), n. 27 from Structure Screen I (0.1 M Na Hepes pH 7.5, 1.5 M $LiSO_4$). The crystals obtained were thin and needlelike or exhibited agglomeration (Fig. 7.10). The x-ray diffraction data were measured at synchrotron, but the crystals diffracted at a limited resolution. Further crystallization trials were performed, varying the dilution of the precipitant solutions, the drop size, the protein concentration, and the temperature condition, but suitable crystals were not obtained.

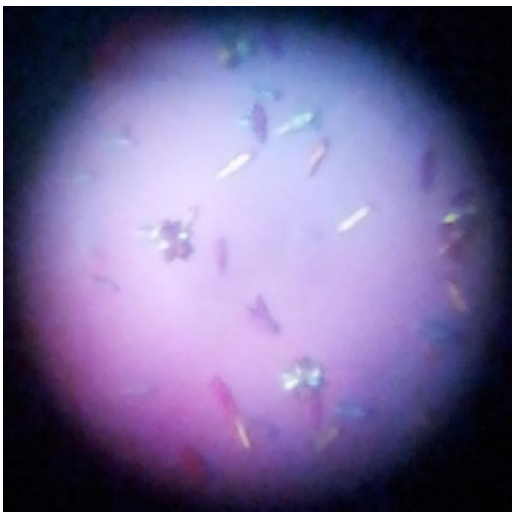

Fig. 7.10 Crystals of 6-His-tag HPG27_1117 grown at 18 mg/mL protein concentration, in 0.2 M $CaCl_2$, 0.1 M Na HEPES pH 7.5, 28% (v/v) PEG 400, at 20 °C.

7.4 RESULTS AND DISCUSSION

Despite their small size, these secreted hypothetical proteins were not demonstrated easy to handle and the lacking information did not support. First of all, as clearly observed, they have in common the tendency to degradation.

HPG27_1030 exhibited an easy and high yielding expression, with the majority of the protein of interest in the soluble fraction (approximately 80%). Despite its propensity to

degradation during purification procedures, the amount of whole purified protein obtained starting from 1 L *E. coli* culture was significant (about 39 mg). Nevertheless, the final concentration of purified protein was limited, since the protein started to aggregate and precipitate above 3 mg/mL concentration. Therefore, to overcome the reported problems of degradation, variations of the buffer composition and the use of additives could enhance the stability of the protein in solution and improve the purification. Increasing the concentration of protease inhibitors and using them over all the purification steps could limit the degradation processes, as well as adding detergents could decrease the aggregation during the final concentration procedure. However, the presence of detergents can affect the crystallization process.

Besides the instability in solution, the second major bottleneck of the purification of HPG27_1117 was the limited expression level. It is possible to assume that the protein of interest is toxic for *E. coli* and prevents its growth. The pETite plasmid carrying the HPG27_1117 gene was transformed in further *E. coli* engineered strains, namely BL21 (DE3) pLysS cells (Invitrogen), C41 (DE3) cells (Lucigen), C43 (DE3) cells (Lucigen), and several overexpression trials were performed, varying the conditions of temperature, duration, isopropyl-β-D-thiogalactopyranoside (IPTG) concentration. Nevertheless, it has not been possible to obtain a significant increase in expression yield in the case of all of the *E. coli* strain used (data not shown). As suggested for HPG27_1030, to overcome the degradation problem, changings in the buffer composition could improve the stability of the protein in solution and, thus, the purification profile, leading to a greater final amount of purified protein. Since the protein of interest contains 3 cysteine residue in its sequence, dithiothreitol (DTT) was added to the buffer composition, to ensure the proper reduced state of the cysteine residues. Moreover, the apparent molecular weight on SDS-PAGE higher than the calculated one could be due to an altered amount of bound sodium dodecyl sulfate (SDS) to the protein, consequent to hydrophobic properties of its surface. SDS is employed in SDS-PAGE to disrupt secondary structure and give all proteins a constant charge/mass ratio, which is assumed to be 1.4 g SDS/1 g protein. This condition is achieved by the aggregation of SDS molecules at hydrophobic protein sites to induce “reconstructive denaturation”. It can be assumed that a lower amount of SDS is bound to HPG27_1117, due to intrinsic properties of the protein; therefore, its mobility would decrease, since the protein presents a lower net negative charge.

Finally, despite several attempts and reasonable protein concentrations, it was not possible to obtain quality crystals. HPG27_1030 showed precipitation in most of the

conditions, most likely due to degradation processes and not homogeneous samples; only few crystallization conditions showed microcrystalline species. On the other hand, it has been possible to obtain preliminary crystal of HPG27_1117; however, despite several attempts to improve their quality, crystals suitable to x-ray diffraction measurement have not been obtained yet, possibly owing to the limited amount of protein sample achieved from each expression and purification event. Improving the stability of these hypothetical proteins could lead to an enhancement of the crystallization possibility.

Chapter VIII

CONCLUSIONS

CONCLUSIONS

This research project, whose results are summarized in this thesis, is focused on the identification and structural characterization of new potential pharmacological targets of the pathogen *H. pylori*. Although its infection rate has globally decreased, as result of gradual improvements in socio-economic status and hygienic practice, the acquisition of *H. pylori* still remains a worldwide emergency. The clinical outcomes, when present, associated to its infection include severe gastroduodenal diseases and the effectiveness of the eradication treatment is often compromised by a widespread antibiotic resistance. Therefore, identification of new pharmacological targets and new treatments is a crucial challenge in managing *H. pylori* infections and related diseases.

This research project provided some additional insights into the overall comprehension about *H. pylori* proteome and the identification of attractive proteins as future subjects aimed to therapy and vaccine development. Colonization factors and putative proteins mediating host-pathogen interactions were considered for the structural characterization, which might allow to develop new potential therapeutic molecules.

Among these subjects, great interest was converged on two of the main players of the intricate network aimed to balance of urea and bicarbonate and to buffer the pH of the microorganism fluids in the extremely acid conditions, namely α-carbonic anhydrase and β-carbonic anhydrase, located in periplasm and cytoplasm, respectively. Both enzymes catalyze the reversible conversion of the carbon dioxide into bicarbonate and cooperate with the crucial enzyme urease; the latter hydrolyzes urea into carbon dioxide and ammonia in the cytoplasm, providing both acid-neutralizing and acid-buffering capacity. Subsequently, carbon dioxide molecules freely diffused out of the inner membrane are substrates of α-carbonic anhydrase; on the contrary, carbon dioxide remaining into cytoplasm is hydrolyzed by β-carbonic anhydrase. From these evidences it is possible to deduce the key role of these enzymes in *H. pylori* survival and colonization, thus making them attractive targets of investigations and promising drug targets. The structural peculiarities of α-carbonic anhydrase have been successfully determined, highlighting the overall structure shared with many other members of the α-carbonic anhydrase family, the elegant expedient to solve the problem of the absence of a negatively charged side chain at position 127 through the use of an anion from the external medium, the localization close to the inner membrane of the bacterium. Taking all these findings

together, the definition of the details of shape and organization of the active site may constitute a basis for the design of inhibitors specific against *H. pylori*. On the contrary, the experimental procedures concerning β-carbonic anhydrase revealed how this enzyme possesses a different handling characteristics compared to α-carbonic anhydrase, confirming how these enzymes present several structural differences. It was not possible to define the molecular details of the three-dimensional structure of β-carbonic anhydrase, since several limits were encountered in its expression and purification running. Consequently, crystallization attempts were sacrificed in the number owing to the limited amount of purified protein and crystals obtained were not adequately suitable to x-ray diffraction measurement.

In this research project further disparate proteins, responsible of various bacterial functions, were subjects of structural investigations, in order to widen the overall comprehension on *H. pylori* and to characterize new pharmacological targets. Specifically, effectors responsible of flagellar motility, oxidative metabolic processes, host-pathogen interactions were considered. From the related investigations only partial findings were achieved, since various issues were encountered. Commonly, these proteins were successfully cloned as 6-His-tag recombinant product; however, improvements can be provided at the level of expression or purification, depending on the subject considered. Crystals of one of these proteins, namely HPG27_1117, were obtained, but not adequately suitable to x-ray diffraction measurement. These partial findings constitute the basis for future improvements of the procedures and structural investigations.

This project thesis contributes to a deeper overall comprehension of the molecular specifications and the physiopathological mechanisms of some features of the human pathogen *H. pylori*; moreover, it sets the experimental basis for future investigations, which may lead to the development of new therapies and vaccines, aimed to limit the diffusion and the related gastric damages of *H. pylori*. Despite the mass of information accumulated to date on *H. pylori*, much work and further researches are still required to achieve a full comprehension of the physiology of the bacterium, the molecular mechanisms, the host-pathogen interactions and related clinical outcomes.

REFERENCES

REFERENCES

Abdollahi, Hamid and Omid Tadjrobehkar. 2012. “The Role of Different Sugars, Amino Acids and Few Other Substances in Chemotaxis Directed Motility of Helicobacter Pylori.” *Iranian Journal of Basic Medical Sciences* 15(3):787–94.

Adams, Paul D. et al. 2010. “PHENIX: A Comprehensive Python-Based System for Macromolecular Structure Solution.” *Acta crystallographica. Section D, Biological crystallography* 66(Pt 2):213–21.

Agarwal, K. and S. Agarwal. 2008. “Helicobacter Pylori Vaccine : From Past to Future.” *Mayo Clinic Proceedings* 83(2):169–75.

Aggarwal, Mayank et al. 2014. “Structural Insight into Activity Enhancement and Inhibition of H64A Carbonic Anhydrase II by Imidazoles.” *IUCrJ* 1(1):129–35.

Aggarwal, Mayank, Teck Khiang Chua, Melissa A. Pinard, Doletha M. Szebenyi, and Robert McKenna. 2015. “Carbon Dioxide ‘Trapped’ in a β-Carbonic Anhydrase.” *Biochemistry* 151016144023009.

Ahmed, Niyaz, Shivendra Tenguria, and Nishant Nandanwar. 2009. “Helicobacter Pylori--a Seasoned Pathogen by Any Other Name.” *Gut pathogens* 1:24.

Akada, Junko K., Mutsunori Shirai, Hiroaki Takeuchi, Masataka Tsuda, and Teruko Nakazawa. 2000. “Identification of the Urease Operon in Helicobacter Pylori and Its Control by mRNA Decay in Response to pH.” *Molecular Microbiology* 36:1071–84.

Akhiani, Ali a et al. 2002. “Protection against Helicobacter Pylori Infection Following Immunization Is IL-12-Dependent and Mediated by Th1 Cells.” *Journal of immunology (Baltimore, Md. : 1950)* 169(12):6977–84.

Algood, H. M. S. and T. L. Cover. 2006. “Helicobacter Pylori Persistence: An Overview of Interactions between H. Pylori and Host Immune Defenses.” *Clinical Microbiology Reviews* 19(4):597–613.

Alm, R. A. et al. 1998. “Genomic-Sequence Comparison of Two Unrelated Isolates of the Human Gastric Pathogen Helicobacter Pylori.” *Nature* 395(2):863–69.

Alm, R. A. and T. J. Trust. 1999. “Analysis of the Genetic Diversity of Helicobacter Pylori: The Tale of Two Genomes.” *Journal of molecular medicine (Berlin, Germany)* 77(12):834–46.

Amieva, Manuel R. et al. 2003. “Disruption of the Epithelial Apical-Junctional Complex by Helicobacter Pylori CagA.” *Science (New York, N.Y.)* 300(5624):1430–34.

Amieva, Manuel R. and Emad M. El-Omar. 2008. “Host-Bacterial Interactions in Helicobacter Pylori Infection.” *Gastroenterology* 134(1):306–23.

Andersen, Leif Percival. 2007. “Colonization and Infection by Helicobacter Pylori in Humans.” *Helicobacter* 12 Suppl 2:12–15.

Andersen, Leif Percival and Lone Rasmussen. 2009. “Helicobacter Pylori - Coccoid Forms and Biofilm Formation.” *FEMS Immunology and Medical Microbiology* 56(2):112–15.

Aras, R. A. 2002. “Helicobacter Pylori Interstrain Restriction-Modification Diversity Prevents Genome Subversion by Chromosomal DNA from Competing Strains.” *Nucleic Acids Research* 30(24):5391–97.

Aras, Rahul a, Josephine Kang, Ariane I. Tschumi, Yasuaki Harasaki, and Martin J. Blaser. 2003. “Extensive Repetitive DNA Facilitates Prokaryotic Genome Plasticity.” *Proceedings of the National*

Academy of Sciences of the United States of America 100(23):13579–84.

Aspholm, Marina et al. 2006. “SabA Is the H. Pylori Hemagglutinin and Is Polymorphic in Binding to Sialylated Glycans.” *PLoS pathogens* 2(10):e110.

Aspinall, G. O. and M. A. Monteiro. 1996. “Lipopolysaccharides of Helicobacter Pylori Strains P466 and MO19: Structures of the O Antigen and Core Oligosaccharide Regions.” *Biochemistry* 35(7):2498–2504.

Atherton, J. C., R. M. Jr Peek, K. T. Tham, T. L. Cover, and M. J. Blaser. 1997. “Clinical and Pathological Importance of Heterogeneity in vacA, the Vacuolating Cytotoxin Gene of Helicobacter Pylori.” *Gastroenterology* 112(1):92–99.

Atherton, J. C., K. T. Tham, R. M. Peek, T. L. Cover, and M. J. Blaser. 1996. “Density of Helicobacter Pylori Infection in Vivo as Assessed by Quantitative Culture and Histology.” *The Journal of infectious diseases* 174(3):552–56.

Atherton, JC et al. 1995. “Mosaicism in Vacuolating Cytotoxin Alleles of Helicobacter Pylori. Association of Specific VacA Types with Cytotoxin Production and Peptic Ulcerataion.” *Journal of Biological Chemistry* 270(30):17771–77.

Ayala, Guadalupe, Wendy Itzel Escobedo-Hinojosa, Carlos Felipe de la Cruz-Herrera, and Irma Romero. 2014. “Exploring Alternative Treatments for Helicobacter Pylori Infection.” *World journal of gastroenterology : WJG* 20(6):1450–69.

Backert, Steffen and Matthias Selbach. 2008. “Role of Type IV Secretion in *Helicobacter Pylori* Pathogenesis.” *Cellular Microbiology* 10(8):1573–81.

Baidya, Amit K., Saurabh Bhattacharya, and Rukhsana Chowdhury. 2015. “Role of the Flagellar Hook-Length Control Protein FliK and σ28 in cagA Expression in Gastric Cell–Adhered Helicobacter Pylori.” *Journal of Infectious Diseases* 211:1779–89.

Baltrus, David a. et al. 2009. “The Complete Genome Sequence of Helicobacter Pylori Strain G27.” *Journal of Bacteriology* 91(1):447–48.

Bamford, K. B. et al. 1998. “Lymphocytes in the Human Gastric Mucosa during Helicobacter Pylori Have a T Helper Cell 1 Phenotype.” *Gastroenterology* 114(3):482–92.

Bardhan, Pradip K. 1997. “Epidemiological Features of Helicobacter Pylori Infection in Developing Countries.” *Clinical Infectious Diseases* 973–78.

Barker, P. D. and S. J. Ferguson. 1999. “Still a Puzzle: Why Is Haem Covalently Attached in c-Type Cytochromes?” *Structure (London, England : 1993)* 7(12):R281–90.

Basso, Daniela et al. 2008. “Clinical Relevance of Helicobacter Pylori cagA and vacA Gene Polymorphisms.” *Gastroenterology* 135(1):91–99.

Becker, Holger M., Michael Klier, and Joachim W. Deitmer. 2014. *Carbonic Anhydrase: Mechanism, Regulation, Links to Disease, and Industrial Applications*.

Bellack, N. R., M. W. Koehoorn, Y. C. MacNab, and M. G. Morshed. 2006. “A Conceptual Model of Water’s Role as a Reservoir in Helicobacter Pylori Transmission: A Review of the Evidence.” *Epidemiology and infection* 134(3):439–49.

de Bernard, M. et al. 1995. “Low pH Activates the Vacuolating Toxin of Helicobacter Pylori, Which Becomes Acid and Pepsin Resistant.” *The Journal of Biological Chemistry* 270(41):23937–40.

de Bernard, Marina and Mario M. D'Elios. 2010. "The Immune Modulating Activity of the Helicobacter Pylori HP-NAP: Friend or Foe?" *Toxicon* 56(7):1186–92.

Biasini, Marco et al. 2014. "SWISS-MODEL: Modelling Protein Tertiary and Quaternary Structure Using Evolutionary Information." *Nucleic acids research* 42(Web Server issue):W252–58.

Björkholm, B. et al. 2001. "Mutation Frequency and Biological Cost of Antibiotic Resistance in Helicobacter Pylori." *Proceedings of the National Academy of Sciences of the United States of America* 98(25):14607–12.

Blaser, Martin J. and John C. Atherton. 2004. "Helicobacter Pylori Persistence: Biology and Disease." *Journal of Clinical Investigation* 113(3):321–33.

Blaser, Martin J., Yu Chen, and Joan Reibman. 2008. "Does Helicobacter Pylori Protect against Asthma and Allergy?" *Gut* 57(5):561–67.

Bocian-Ostrzycka, Katarzyna M., Magdalena J. Grzeszczuk, Lukasz Dziewit, and Elżbieta K. Jagusztyn-Krynicka. 2015. "Diversity of the Epsilonproteobacteria Dsb (disulfide Bond) Systems." *Frontiers in microbiology* 6(June):570.

Bode, G., F. Mauch, and P. Malfertheiner. 1993. "The Coccoid Forms of Helicobacter Pylori. Criteria for Their Viability." *Epidemiology and infection* 111(3):483–90.

Bonnard, Geraldine, Vincent Corvest, Etienne H. Meyer, and Patrice P. Hamel. 2010. "Redox Processes Controlling the Biogenesis of c-Type Cytochromes." *Antioxidants & redox signaling* 13(9):1385–1401.

Boone, Christopher D., Melissa Pinard, Rob McKenna, and David Silverman. 2014. "Catalytic Mechanism of Alpha-Class Carbonic Anhydrases: CO_2 Hydration and Proton Transfer." *Sub-cellular biochemistry* 75:31–52.

Boren, T., P. Falk, K. A. Roth, G. Larson, and S. Normark. 1993. "Attachment of Helicobacter Pylori to Human Gastric Epithelium Mediated by Blood Group Antigens." *Science (New York, N.Y.)* 262(5141):1892–95.

Van de Bovenkamp, Jeroen H. B. et al. 2003. "The MUC5AC Glycoprotein Is the Primary Receptor for Helicobacter Pylori in the Human Stomach." *Helicobacter* 8(5):521–32.

Bridge, Dacie R. and D. Scott Merrell. 2013. "Polymorphism in the Helicobacter Pylori CagA and VacA Toxins and Disease." *Gut Microbes* 4(April):101–17.

Brown, L. M. 2000. "Helicobacter Pylori: Epidemiology and Routes of Transmission." *Epidemiologic Reviews* 22(2):283–97.

Bukanov, N. O. and D. E. Berg. 1994. "Ordered Cosmid Library and High-Resolution Physical-Genetic Map of Helicobacter Pylori Strain NCTC11638." *Molecular microbiology* 11(3):509–23.

Bumann, D. et al. 2002. "Proteome Analysis of Secreted Proteins of the Gastric Pathogen Helicobacter Pylori." *Infect Immun* 70(7):3396–3403.

Bury-Moné, Stéphanie et al. 2003. "Presence of Active Aliphatic Amidases in Helicobacter Species Able to Colonize the Stomach." *Infection and immunity* 71(10):5613–22.

Bury-Moné, Stéphanie et al. 2004. "Responsiveness to Acidity via Metal Ion Regulators Mediates Virulence in the Gastric Pathogen Helicobacter Pylori." *Molecular microbiology* 53(2):623–38.

Bury-Moné, Stéphanie et al. 2008. "Roles of Alpha and Beta Carbonic Anhydrases of Helicobacter Pylori

in the Urease-Dependent Response to Acidity and in Colonization of the Murine Gastric Mucosa." *Infection and immunity* 76(2):497–509.

Busler, V. J. et al. 2006. "Protein-Protein Interactions among Helicobacter Pylori Cag Proteins." *Journal of Bacteriology* 188(13):4787–4800.

Byrd, J. C., C. K. Yunker, Q. S. Xu, L. R. Sternberg, and R. S. Bresalier. 2000. "Inhibition of Gastric Mucin Synthesis by Helicobacter Pylori." *Gastroenterology* 118(6):1072–79.

Capasso, Clemente and Claudiu T. Supuran. 2015. "An Overview of the Alpha-, Beta- and Gamma-Carbonic Anhydrases from Bacteria: Can Bacterial Carbonic Anhydrases Shed New Light on Evolution of Bacteria?" *Journal of enzyme inhibition and medicinal chemistry* 30(2):325–32.

Carlsohn, Elisabet, Johanna Nyström, Ingrid Bölin, Carol L. Nilsson, and Ann-Mari Svennerholm. 2006. "HpaA Is Essential for Helicobacter Pylori Colonization in Mice." *Infection and immunity* 74(2):920–26.

Carraway, K. L. and S. R. Hull. 1991. "Cell Surface Mucin-Type Glycoproteins and Mucin-like Domains." *Glycobiology* 1(2):131–38.

Cascales, E. and P. J. Christie. 2003. "The Versatile Bacterial Type IV Secretion Systems." *Nature reviews. Microbiology* 18(9):1199–1216.

Ceci, Pierpaolo, Laura Mangiarotti, Claudio Rivetti, and Emilia Chiancone. 2007. "The Neutrophil-Activating Dps Protein of Helicobacter Pylori, HP-NAP, Adopts a Mechanism Different from Escherichia Coli Dps to Bind and Condense DNA." *Nucleic acids research* 35(7):2247–56.

Cendron, Laura and Giuseppe Zanotti. 2011. "Structural and Functional Aspects of Unique Type IV Secretory Components in the Helicobacter Pylori Cag-Pathogenicity Island." *FEBS Journal* 278(8):1223–31.

Censini, S. et al. 1996. "Cag , a Pathogenicity Island of Helicobacter Pylori , Encodes Type I-Specific and Disease-Associated Virulence Factors." *Proceedings of the National Academy of Sciences of the United States of America* 93(December):14648–53.

Cerda, Oscar A. et al. 2011. "tlpA Gene Expression Is Required for Arginine and Bicarbonate Chemotaxis in Helicobacter Pylori." *Biological research* 44(3):277–82.

Ceruso, Mariangela et al. 2015. "Inhibition Studies of Bacterial, Fungal and Protozoan β-Class Carbonic Anhydrases with Schiff Bases Incorporating Sulfonamide Moieties." *Bioorganic & Medicinal Chemistry* 23(15):4181–87.

Chevance, Fabienne F. V and Kelly T. Hughes. 2008. "Coordinating Assembly of a Bacterial Macromolecular Machine." *Nat Rev Micro* 6(6):455–65.

Chirica, L. C., B. Elleby, and S. Lindskog. 2001. "Cloning, Expression and Some Properties of Alpha-Carbonic Anhydrase from Helicobacter Pylori." *Biochimica et biophysica acta* 1544(1-2):55–63.

Chiurillo, Miguel Angel et al. 2013. "Genotyping of Helicobacter Pylori Virulence-Associated Genes Shows High Diversity of Strains Infecting Patients in Western Venezuela." *International Journal of Infectious Diseases* 17(9):750–56.

Choli-Papadopoulou, Theodora. 2011. "*Helicobacter Pylori* Neutrophil Activating Protein as Target for New Drugs against *H. Pylori* Inflammation." *World Journal of Gastroenterology* 17(21):2585.

Cid, Trinidad Parra, Miryam Calvino Fernández, Selma Benito Martínez, and Nicola L. Jones. 2013.

"Pathogenesis of *Helicobacter Pylori* Infection." *Helicobacter* 18:12–17.

Colland, F. et al. 2001. "Identification of the Helicobacter Pylori Anti-sigma28 Factor." *Molecular microbiology* 41(2):477–87.

Contreras, Monica, Jean Michel Thiberge, M. a. Mandrand-Berthelot, and Agnès Labigne. 2003. "Characterization of the Roles of NikR, a Nickel-Responsive Pleiotropic Autoregulator of Helicobacter Pylori." *Molecular Microbiology* 49(4):947–63.

Covacci, a, J. L. Telford, G. Del Giudice, J. Parsonnet, and R. Rappuoli. 1999. "Helicobacter Pylori Virulence and Genetic Geography." *Science (New York, N.Y.)* 284(5418):1328–33.

Covacci, Antonello, Stanley Falkow, Douglas E. Berg, and Rino Rappuoli. 1997. "Did the Inheritance of a Pathogenicity Island Modify the Virulence of Helicobacter Priori?" *TRENDS in Microbiology* 5(5):205–8.

Covarrubias, Adrian Suarez et al. 2005. "Structure and Function of Carbonic Anhydrases from Mycobacterium Tuberculosis." *Journal of Biological Chemistry* 280(19):18782–89.

Cover, T. L., L. Y. Reddy, and M. J. Blaser. 1993. "Effects of ATPase Inhibitors on the Response of HeLa Cells to Helicobacter Pylori Vacuolating Toxin." *Infect Immun* 61(4):1427–31.

Cover, T. L., M. K. Tummuru, P. Cao, S. A. Thompson, and M. J. Blaser. 1994. "Divergence of Genetic Sequences for the Vacuolating Cytotoxin among Helicobacter Pylori Strains." *The Journal of biological chemistry* 269(14):10566–73.

Cover, Timothy L. and Steven R. Blanke. 2005. "Helicobacter Pylori VacA, a Paradigm for Toxin Multifunctionality." *Nat Rev Micro* 3(4):320–32.

Cover, Timothy L. and Martin J. Blaser. 2009. "Helicobacter Pylori in Health and Disease." *Gastroenterology* 18(9):1199–1216.

Cover, Timothy L., Phyllis I. Hanson, and John E. Heuser. 1997. "Acid-Induced Dissociation of VacA, the Helicobacter Pylori Vacuolating Cytotoxin, Reveals Its Pattern of Assembly." *Journal of Cell Biology* 138(4):759–69.

Cronk, Jeff D. et al. 2006. "Identification of a Novel Noncatalytic Bicarbonate Binding Site in Eubacterial Beta-Carbonic Anhydrase." *Biochemistry* 45(14):4351–61.

Cronk, Jeff D., James a Endrizzi, Michelle R. Cronk, and Jason W. O. Neill. 2001. "Crystal Structure of E. Coli B-Carbonic Anhydrase, an Enzyme with an Unusual pH-Dependent Activity." *Protein Science* 10:911–22.

Crow, Allister, Richard M. Acheson, Nick E. Le Brun, and Arthur Oubrie. 2004. "Structural Basis of Redox-Coupled Protein Substrate Selection by the Cytochrome c Biosynthesis Protein ResA." *Journal of Biological Chemistry* 279(22):23654–60.

Croxen, Matthew A., Gary Sisson, Roberto Melano, and Paul S. Hoffman. 2006. "The Helicobacter Pylori Chemotaxis Receptor TlpB (HP0103) Is Required for pH Taxis and for Colonization of the Gastric Mucosa." *Journal of bacteriology* 188(7):2656–65.

Cullen, Thomas W. et al. 2011. "Helicobacter Pylori versus the Host: Remodeling of the Bacterial Outer Membrane Is Required for Survival in the Gastric Mucosa." *PLoS Pathogens* 7(12):e1002454.

D'Elios, Mario Milco and Leif P. Andersen. 2007. "Helicobacter Pylori Inflammation, Immunity, and Vaccines." *Helicobacter* 12 Suppl 1:15–19.

D'Elios, Mario Milco, Cesare Montecucco, and Marina de Bernard. 2007. "VacA and HP-NAP, Ying and Yang of Helicobacter Pylori-Associated Gastric Inflammation." *Clinica chimica acta; international journal of clinical chemistry* 381(1):32–38.

Dabiri, Hossein et al. 2009. "Distribution of Helicobacter Pylori cagA, cagE, oipA and vacA in Different Major Ethnic Groups in Tehran, Iran." *Journal Gastroenterology Hepatology* 18(9):1199–1216.

Danielli, Alberto et al. 2009. "Growth Phase and Metal-Dependent Transcriptional Regulation of the fecA Genes in Helicobacter Pylori." *Journal of Bacteriology* 191(11):3717–25.

Delany, Isabel, A. N. A. Beatriz F. Pacheco, Gunther Spohn, and Rino Rappuoli. 2001. "Iron-Dependent Transcription of the frpB Gene of Helicobacter Pylori Is Controlled by the Fur Repressor Protein." *Journal of Bacteriology* 183(16):4932–37.

Denoncin, Katleen and Jean-François Collet. 2012. "Disulfide Bond Formation in the Bacterial Periplasm: Major Achievements and Challenges Ahead." *Antioxidants & Redox Signaling* 19(1):121002061152001.

Dian, Cyril et al. 2011. "The Structure of the Helicobacter Pylori Ferric Uptake Regulator Fur Reveals Three Functional Metal Binding Sites." *Molecular Microbiology* 79(5):1260–75.

Donczew, Rafał, Christoph Weigel, Rudi Lurz, Jolanta Zakrzewska-Czerwinska, and Anna Zawilak-Pawlik. 2012. "Helicobacter Pylori oriC--the First Bipartite Origin of Chromosome Replication in Gram-Negative Bacteria." *Nucleic acids research* 40(19):9647–60.

Dong, Quan-Jiang, Qing Wang, Ying-Nin Xin, Ni Li, and Shi-Ying Xuan. 2009. "Comparative Genomics of Helicobacter Pylori." *World journal of gastroenterology : WJG* 15:3984–91.

Dore, Maria P., Antonia R. Sepulveda, Michael S. Osato, Giuseppe Realdi, and David Y. Graham. 1999. "Helicobacter Pylori in Sheep Milk." *The Lancet* 354(9173):132.

Dore, MariaP. et al. 1999. "High Prevalence of Helicobacter Pylori Infection in Shepherds." *Digestive Diseases and Sciences* 44(6):1161–64.

Dorer, Marion S., Sarah Talarico, and Nina R. Salama. 2009. "Helicobacter Pylori's Unconventional Role in Health and Disease." *PLoS pathogens* 5(10):1–6.

Dosanjh, Nuvjeevan S. and Sarah L. J. Michel. 2006. "Microbial Nickel Metalloregulation: NikRs for Nickel Ions." *Current Opinion in Chemical Biology* 10(2):123–30.

Dossumbekova, Anar et al. 2006. "Helicobacter Pylori HopH (OipA) and Bacterial Pathogenicity: Genetic and Functional Genomic Analysis of hopH Gene Polymorphisms." *The Journal of infectious diseases* 194(10):1346–55.

Douillard, Francois P., Kieran A. Ryan, Jason Hinds, and Paul W. O. Toole. 2009. "Effect of FliK Mutation on the Transcriptional Activity of the σ 54 Sigma Factor RpoN in Helicobacter Pylori." *Microbiology* 155(Pt 6):1901–11.

Dubois, Andre and Thomas Borén. 2007. "Helicobacter Pylori Is Invasive and It May Be a Facultative Intracellular Organism." *Cellular Microbiology* 9(5):1108–16.

Dundon, W. G., D. G. Marshall, C. A. O'Morain, and C. J. Smyth. 2000. "Population Characteristics of Irish Helicobacter Pylori Isolates: A tRNA-Associated Locus." *Irish journal of medical science* 169(2):137–40.

Dunn, B. E., G. P. Campbell, G. I. Perez-Perez, and Blaser M.J. 1990. "Purification and Characterization

of Urease from Helicobacter Pylori." *The Journal of Biological Chemistry* 265(5):9464–69.

Dunn, BE, H. Cohen, and MJ Blaser. 1997. "Helicobacter Pylori." *Clinical Microbiology Reviews* 4:720–41.

Dunn, Bruce E. and Suhas H. Phadnis. 1998. "Structure , Function and Localization of Helicobacter Pyloni Urease." *Yale Journal of Biology and Medicine* 71(1998):63–73.

Dunne, Ciara. 2014. "Factors That Mediate Colonization of the Human Stomach by *Helicobacter Pylori*." *World Journal of Gastroenterology* 20(19):5610.

Duś, Irena et al. 2013. "Role of PCR in Helicobacter Pylori Diagnostics and Research – New Approaches for Study of Coccoid and Spiral Forms of the Bacteria Rola Metody PCR W Diagnostyce I Badaniach Naukowych Helicobacter Pylori – Nowe Możliwości Poznawcze Formy Kokoidalnej I Spir." *Postepy Higieni i Medycyny Doswiadczalnej* 261–68.

Eaton, K. a., C. L. Brooks, D. R. Morgan, and S. Krakowka. 1991. "Essential Role of Urease in Pathogenesis of Gastritis Induced by Helicobacter Pylori in Gnotobiotic Piglets." *Infection and Immunity* 59(7):2470–75.

Eaton, K. A., S. Suerbaum, C. Josenhans, and S. Krakowka. 1996. "Colonization of Gnotobiotic Piglets by Helicobacter Pylori Deficient in Two Flagellin Genes." *Infection and immunity* 64(7):2445–48.

Emsley, P., B. Lohkamp, W. G. Scott, and K. Cowtan. 2010. "Features and Development of Coot." *Acta crystallographica. Section D, Biological crystallography* 66(Pt 4):486–501.

Ernst, Florian D. et al. 2005. "Transcriptional Profiling of Helicobacter Pylori Fur- and Iron-Regulated Gene Expression." *Microbiology (Reading, England)* 151(Pt 2):533–46.

Evans, D. G., D. J. Evans, J. J. Moulds, and D. Y. Graham. 1988. "N-Acetylneuraminyllactose-Binding Fibrillar Hemagglutinin of Campylobacter Pylori: A Putative Colonization Factor Antigen." *Infection and immunity* 56(11):2896–2906.

Evans, D. J. Jr and D. G. Evans. 2000. "Helicobacter Pylori Adhesins: Review and Perspectives." *Helicobacter* 5(4):183–95.

Evans, D., T. Karjalainen, D. Evans, D. Graham, and C. H. Lee. 1993. "Cloning, Nucleotide Sequence, and Expression of an Adhesin Subunit Protein of Helicobacter Pylori." *J.Bacteriol.* 175(3):674–83.

Evans, Doyle J., Dolores G. Evans, Stacy S. Kirkpatrick, and David Y. Graham. 1991. "Characterization of the Helicobacter Pylori Urease and Purification of Its Subunits." *Microbial Pathogenesis* 10(1):15–26.

Evans, Philip. 2006. "Scaling and Assessment of Data Quality." *Acta crystallographica. Section D, Biological crystallography* 62(Pt 1):72–82.

Fabre, Nicolas, Ilja M. Reiter, Noelle Becuwe-Linka, Bernard Genty, and Dominique Rumeau. 2007. "Characterization and Expression Analysis of Genes Encoding Alpha and Beta Carbonic Anhydrases in Arabidopsis." *Plant, cell & environment* 30:617–29.

Falke, J. J. and G. L. Hazelbauer. 2001. "Transmembrane Signaling in Bacterial Chemoreceptors." *Trends in biochemical sciences* 26(4):257–65.

Falush, D. et al. 2001. "Recombination and Mutation during Long-Term Gastric Colonization by Helicobacter Pylori: Estimates of Clock Rates, Recombination Size, and Minimal Age." *Proceedings of the National Academy of Sciences of the United States of America* 98(26):15056–61.

Falush, Daniel et al. 2003. “Traces of Human Migrations in Helicobacter Pylori Populations.” *Science (New York, N.Y.)* 299(5612):1582–85.

Farzi, Nastaran, Tannaz Malekian, Masoud Alebouyeh, Farzam Vaziri, and Mohammad Reza Zali. 2015. “Genotype Diversity and Quasispecies Development of Helicobacter Pylori in a Single Host.” *Japanese Journal of Infectious Diseases* 68(3):176–80.

Ferguson, Stuart J., Julie M. Stevens, James W. A. Allen, and Ian B. Robertson. 2008. “Cytochrome c Assembly: A Tale of Ever Increasing Variation and Mystery?” *Biochimica et Biophysica Acta - Bioenergetics* 1777:980–84.

Fernandez-Gonzalez, Esther and Steffen Backert. 2014. “DNA Transfer in the Gastric Pathogen Helicobacter Pylori.” *Journal of Gastroenterology* 49(4):594–604.

Ferraroni, Marta, Sonia Del Prete, Daniela Vullo, Clemente Capasso, and Claudiu T. Supuran. 2015. “Crystal Structure and Kinetic Studies of a Tetrameric Type II β-Carbonic Anhydrase from the Pathogenic Bacterium *Vibrio Cholerae*.” *Acta Crystallographica Section D Biological Crystallography* 71(12):2449–56.

Ferrero, R. L. and A. Labigne. 2001. “Helicobacter Pylori Vaccine Development in the Post-Genomic Era : Can In Silico Translate to In Vivo ?” *Scandinavian Journal of Immunology* 443–48.

Figura, N. et al. 1989. “Cytotoxin Production by Campylobacter Pylori Strains Isolated from Patients with Peptic Ulcers and from Patients with Chronic Gastritis Only.” *Journal of Clinical Microbiology* 225–26.

Di Fiore, Anna et al. 2013. “X-Ray Structure of the First `extremo-Alpha-Carbonic Anhydrase’, a Dimeric Enzyme from the Thermophilic Bacterium Sulfurihydrogenibium Yellowstonense YO3AOP1.” *Acta crystallographica. Section D, Biological crystallography* 69(Pt 6):1150–59.

Fischer, Wolfgang. 2011. “Assembly and Molecular Mode of Action of the Helicobacter Pylori Cag Type IV Secretion Apparatus.” *FEBS Journal* 278(8):1203–12.

Fischer, Wolfgang et al. 2014. “A Comprehensive Analysis of Helicobacter Pylori Plasticity Zones Reveals That They Are Integrating Conjugative Elements with Intermediate Integration Specificity.” *BMC genomics* 15(1):310.

Fischer, Wolfgang, Dirk Hofreuter, and Rainer Haas. 2001. “Natural Transformation, Recombination, and Repair.” edited by H. L. T. Mobley, G. L. Mendz, and S. L. Hazell. Washington (DC).

Fraser, A. G., R. Scragg, P. Metcalf, S. McCullough, and N. J. Yeates. 1996. “Prevalence of Helicobacter Pylori Infection in Different Ethnic Groups in New Zealand Children and Adults.” *Australian and New Zealand journal of medicine* 26(5):646–51.

Frawley, Elaine R. and Robert G. Kranz. 2009. “CcsBA Is a Cytochrome c Synthetase That Also Functions in Heme Transport.” *Proceedings of the National Academy of Sciences of the United States of America* 106(25):10201–6.

Frenck, Robert W. and John Clemens. 2003. “Helicobacter in the Developing World.” *Microbes and Infection* 5(8):705–13.

Fulkerson, John F. and Harry L. T. Mobley. 2000. “Membrane Topology of the nixA Nickel Transporter of Helicobacter Priori: Two Nickel Transport-Specific Motifs within Transmembrane Helices II and III.” *Journal of Bacteriology* 182(6):1722–30.

Gangwer, Kelly A. et al. 2007. “Crystal Structure of the Helicobacter Pylori Vacuolating Toxin p55 Domain.” *Proceedings of the National Academy of Sciences of the United States of America* 104(41):16293–98.

Garner, Juli a. 1996. “Binding and Internalization of the Helicobacter Pylori Vacuolating Cytotoxin by Epithelial Cells . Binding and Internalization of the Helicobacter Pylori Vacuolating Cytotoxin by Epithelial Cells.” *Microbiology* 64(10):4197–4203.

Gebert, Bettina, Wolfgang Fischer, Evelyn Weiss, Reinhard Hoffmann, and Rainer Haas. 2003. “Helicobacter Pylori Vacuolating Cytotoxin Inhibits T Lymphocyte Activation.” *Science* 301(5636):1099–1102.

Geis, Gabriele, Hermann Leying, Sebastian Suerbaum, U. W. E. Mai, and Wolfgang Opferkuch. 1989. “Ultrastructure and Chemical Analysis of Campylobacter Pylori Flagella.” *Journal of clinical microbiology* 27(3):436–41.

Geis, Gabriele, Sebastian Suerbaum, B. Forsthoff, H. Leying, and W. Opferkuch. 1993. “Ultrastructure and Biochemical Studies of the Flagellar Sheath of Helicobacter Pylori.” *Journal of Medical Microbiology* 38(5):371–77.

Gewirtz, Andrew T. et al. 2004. “Helicobacter Pylori Flagellin Evades Toll-like Receptor 5-Mediated Innate Immunity.” *The Journal of infectious diseases* 189(10):1914–20.

Gilmour, K. M. 2010. “Perspectives on Carbonic Anhydrase.” *Comparative biochemistry and physiology. Part A, Molecular & integrative physiology* 157(3):193–97.

Goh, Khean Lee. 1997. “Prevalence of and Risk Factors for Helicobacter Pylori Infection in a Multi-Racial Dyspeptic Malaysian Population Undergoing Endoscopy.” *Journal of Gastroenterology and Hepatology* 12(6):S29–35.

González-Rivera, Christian, Holly M. Scott Algood, Jana N. Radin, Mark S. McClain, and Timothy L. Cover. 2012. “The Intermediate Region of Helicobacter Pylori Vaca Is a Determinant of Toxin Potency in a Jurkat T Cell Assay.” *Infection and Immunity* 80(8):2578–88.

Goodman, K. J. et al. 1996. “Helicobacter Pylori Infection in the Colombian Andes: A Population-Based Study of Transmission Pathways.” *American journal of epidemiology* 144(3):290–99.

Goosen, C. et al. 2002. “Evaluation of a Novel Heminested PCR Assay Based on the Phosphoglucosamine Mutase Gene for Detection of Helicobacter Pylori in Saliva and Dental Plaque Evaluation of a Novel Heminested PCR Assay Based on the Phosphoglucosamine Mutase Gene for Detection of.” *Journal of clinical microbiology* 40(1):205–9.

Gorden, J. E. D. and P. L. C. Small. 1993. “Acid Resistance in Enteric Bacteria.” *Survival* 61(1):364–67.

Grasso, G. M., G. Ripabelli, M. L. Sammarco, A. Ruberto, and G. Iannitto. 1996. “Prevalence of Helicobacter-like Organisms in Porcine Gastric Mucosa: A Study of Swine Slaughtered in Italy.” *Comparative Immunology, Microbiology and Infectious Diseases* 19(3):213–17.

Grebowska, Aneta et al. 2008. “Anti-Phagocytic Activity of Helicobacter Pylori Lipopolysaccharide (LPS)--Possible Modulation of the Innate Immune Response to These Bacteria.” *Polish journal of microbiology / Polskie Towarzystwo Mikrobiologow = The Polish Society of Microbiologists* 57(3):185–92.

Grübel, Peter et al. 1997. “Vector Potential of Houseflies (Musca Domestica) for Helicobacter Pylori.”

Journal of Clinical Microbiology 35(6):1300–1303.

Gunaletchumy, Selva Perumal et al. 2014. “Helicobacter Pylori Genetic Diversity and Gastro-Duodenal Diseases in Malaysia.” *Scientific Reports* 4:7431.

Guruge, J. L. et al. 1998. “Epithelial Attachment Alters the Outcome of Helicobacter Pylori Infection.” *Proceedings of the National Academy of Sciences of the United States of America* 95(7):3925–30.

Ha, Nam-chul et al. 2001. “Supramolecular Assembly and Acid Resistance of Helicobacter Pylori Urease.” *Nature structural Biology* 8(6):505–9.

Hage, Naim et al. 2015. “Structural Basis of Lewis B Antigen Binding by the Helicobacter Pylori Adhesin BabA.” *Science advances* (August):1–9.

Hamel, Patrice, Vincent Corvest, Philippe Giege, and Geraldine Bonnard. 2009. “Biochemical Requirements for the Maturation of Mitochondrial c-Type Cytochromes.” *Biochimica et biophysica acta* 1793(1):125–38.

Handt, L. K. et al. 1994. “Helicobacter Pylori Isolated from the Domestic Cat: Public Health Implications.” *Infection and immunity* 62(6):2367–74.

Handt, L. K. et al. 1997. “Diagnosis of Helicobacter Pylori Infection in a Colony of Rhesus Monkeys (Macaca Mulatta).” *Journal of clinical microbiology* 35(1):165–68.

Hatakeyama, Masanori. 2004. “Oncogenic Mechanisms of the Helicobacter Pylori CagA Protein.” *Nature reviews. Cancer* 4(9):688–94.

He, S. et al. 2015. “The IS200/IS605 Family and ‘Peel and Paste’ Single-Strand Transposition Mechanism.” *Microbiology spectrum* 3(4).

Hennig, Ewa E., Ray Mernaugh, Jennifer Edl, Ping Cao, and Timothy L. Cover. 2004. “Heterogeneity among Helicobacter Pylori Strains in Expression of the Outer Membrane Protein BabA Heterogeneity among Helicobacter Pylori Strains in Expression of the Outer Membrane Protein BabA.” *Infection and immunity* 72(6):3429–35.

Heras, Begona et al. 2009. “DSB Proteins and Bacterial Pathogenicity.” *Nature reviews. Microbiology* 7(3):215–25.

Hessey, S. J. et al. 1990. “Bacterial Adhesion and Disease Activity in Helicobacter Associated Chronic Gastritis.” *Gut* 31(2):134–38.

Hirano, Takanori, Satoshi Shibata, Kouhei Ohnishi, Tomomi Tani, and Shin-Ichi Aizawa. 2005. “N-Terminal Signal Region of FliK Is Dispensable for Length Control of the Flagellar Hook.” *Molecular Microbiology* 56(2):346–60.

Hopkins, R. J. et al. 1993. “Seroprevalence of Helicobacter Pylori in Chile: Vegetables May Serve as One Route of Transmission.” *The Journal of infectious diseases* 168(1):222–26.

Huang, S. et al. 1998. “Crystal Structure of Carbonic Anhydrase from Neisseria Gonorrhoeae and Its Complex with the Inhibitor Acetazolamide.” *Journal of molecular biology* 283(1):301–10.

Huang, Shenghua et al. 2011. “Structural Studies of β-Carbonic Anhydrase from the Green Alga Coccomyxa: Inhibitor Complexes with Anions and Acetazolamide.” *PLoS ONE* 6(12):e28458.

Ilver, D. et al. 1998. “Helicobacter Pylori Adhesin Binding Fucosylated Histo-Blood Group Antigens Revealed by Retagging.” *Science* 279(5349):373–77.

Ilver, Dag, Silvia Barone, David Mercati, Pietro Lupetti, and John L. Telford. 2004. “Helicobacter Pylori

Toxin VacA Is Transferred to Host Cells via a Novel Contact-Dependent Mechanism." *Cellular Microbiology* 6(2):167–74.

Israel, I. D. a and R. M. Peek. 2001. "Review Article : Pathogenesis of Helicobacter Pylori-Induced Gastric Inflammation." *Alimentary pharmacology & therapeutics* (15):1271–90.

James, Paul et al. 2014. "The Structure of a Tetrameric Alpha-Carbonic Anhydrase from Thermovibrio Ammonificans Reveals a Core Formed around Intermolecular Disulfides That Contribute to Its Thermostability." *Acta crystallographica. Section D, Biological crystallography* 70(Pt 10):2607–18.

Jiao, Li et al. 2013. "Crystal Structure of the Periplasmic Disulfide-Bond Isomerase DsbC from Salmonella Enterica Serovar Typhimurium and the Mechanistic Implications." *Journal of structural biology* 183(1):1–10.

Johnson, Elizabeth M., Jennifer a Gaddy, and Timothy L. Cover. 2012. "Alterations in Helicobacter Pylori Triggered by Contact with Gastric Epithelial Cells." *Frontiers in cellular and infection microbiology* 2(February):17.

Jones, A. C. et al. 1997. "A Flagellar Sheath Protein of Helicobacter Pylori Is Identical to HpaA, a Putative N-Acetylneuraminyllactose-Binding Hemagglutinin, but Is Not an Adhesin for AGS Cells." *Journal of bacteriology* 179(17):5643–47.

Jones, B. D. and H. L. T. Mobley. 1988. "Proteus-Mirabilis Urease - Genetic Organization, Regulation, and Expression of Structural Genes." *Journal of Bacteriology* 170(8):3342–49.

Jones, C. J., M. Homma, and R. M. Macnab. 1989. "L-, P-, and M-Ring Proteins of the Flagellar Basal Body of Salmonella Typhimurium: Gene Sequences and Deduced Protein Sequences." *Journal of bacteriology* 171(7):3890–3900.

Jones, K. R. et al. 2009. "Polymorphism in the CagA EPIYA Motif Impacts Development of Gastric Cancer." *Journal of Clinical Microbiology* 47(4):959–68.

De Jonge, Ramon et al. 2004. "Role of the Helicobacter Pylori Outer-Membrane Proteins AlpA and AlpB in Colonization of the Guinea Pig Stomach." *Journal of Medical Microbiology* 53(5):375–79.

Josenhans, C., A. Labigne, and S. Suerbaum. 1995. "Comparative Ultrastructural and Functional Studies of Helicobacter Pylori and Helicobacter Mustelae Flagellin Mutants: Both Flagellin Subunits, FlaA and FlaB, Are Necessary for Full Motility in Helicobacter Species." *Journal of bacteriology* 177(11):3010–20.

Kaakoush, Nadeem O., Torsten Sterzenbach, William G. Miller, Sebastian Suerbaum, and George Louis Mendz. 2007. "Identification of Disulfide Reductases in Campylobacterales: A Bioinformatics Investigation." *Antonie van Leeuwenhoek* 92(4):429–41.

Kabsch, Wolfgang. 2010. "XDS." *Acta crystallographica. Section D, Biological crystallography* 66(Pt 2):125–32.

Källberg, Morten et al. 2012. "Template-Based Protein Structure Modeling Using the RaptorX Web Server." *Nat. Protocols* 7(8):1511–22.

Kang, Su-Jin, Do-Hee Kim, and Bong-Jin Lee. 2013. "NMR Study on Small Proteins from Helicobacter Pylori for Antibiotic Target Discovery: A Review." *Molecules* 18(11):13410–24.

Kao, Cheng-Yen, Shew-Meei Sheu, Bor-Shyang Sheu, and Jiunn-Jong Wu. 2012. "Length of Thymidine Homopolymeric Repeats Modulates Promoter Activity of sabA in Helicobacter Pylori." *Helicobacter*

17(3):203–9.

Karnholz, Arno et al. 2006. “Functional and Topological Characterization of Novel Components of the comB DNA Transformation Competence System in Helicobacter Pylori.” *Journal of Bacteriology* 188(3):882–93.

Kauser, Farhana et al. 2004. “The Cag Pathogenicity Island of Helicobacter Pylori Is Disrupted in the Majority of Patient Isolates from Different Human Populations.” *Journal of clinical microbiology* 42(11):5302–8.

Kawagishi, I., M. Homma, a W. Williams, and R. M. Macnab. 1996. “Characterization of the Flagellar Hook Length Control Protein fliK of Salmonella Typhimurium and Escherichia Coli.” *Journal of bacteriology* 178(10):2954–59.

Kawahara, T. et al. 2001. “Helicobacter Pylori Lipopolysaccharide from Type I, but Not Type II Strains, Stimulates Apoptosis of Cultured Gastric Mucosal Cells.” *The journal of medical investigation : JMI* 48(3-4):167–74.

Kelley, Lawrence A., Stefans Mezulis, Christopher M. Yates, Mark N. Wass, and Michael J. E. Sternberg. 2015. “The Phyre2 Web Portal for Protein Modeling, Prediction and Analysis.” *Nat. Protocols* 10(6):845–58.

Kennemann, Lynn et al. 2012. “In Vivo Sequence Variation in HopZ, a Phase-Variable Outer Membrane Protein of Helicobacter Pylori.” *Infection and immunity* 80(12):4364–73.

Kersulyte, D., H. Chalkauskas, and D. E. Berg. 1999. “Emergence of Recombinant Strains of Helicobacter Pylori during Human Infection.” *Molecular microbiology* 31(1):31–43.

Kersulyte, Dangeruta et al. 2004. “Sequence Organization and Insertion Specificity of the Novel Chimeric IS Hp609 Transposable Element of Helicobacter Pylori Sequence Organization and Insertion Specificity of the Novel Chimeric ISHp609 Transposable Element of Helicobacter Pylori †.” *Journal of Bacteriology of Bacteriology* 186(22):7521–28.

Kersulyte, Dangeruta, Natalia S. Akopyants, Sandra W. Clifton, Bruce A. Roe, and Douglas E. Berg. 1998. “Novel Sequence Organization and Insertion Specificity of IS605 and IS606: Chimaeric Transposable Elements of Helicobacter Pylori.” *Gene* 223(1-2):175–86.

Kersulyte, Dangeruta, Henrikas Chalkauskas, and Douglas Berg. 1999. “Emergence of Recombinant Strains of Helicobacter Pylori during Human Infection.” *Molecular Microbiology* 31(1):31–43.

Khalifa, Mohammed Mahdy, Radwa Raed Sharaf, and Ramy Karam Aziz. 2010. “Helicobacter Pylori: A Poor Man’s Gut Pathogen?” *Gut pathogens* 2(1):2.

Kidd, M., J. C. Atherton, A. J. Lastovica, and J. A. Louw. 2001. “Clustering of South African Helicobacter Pylori Isolates from Peptic Ulcer Disease Patients Is Demonstrated by Repetitive Extragenic Palindromic-PCR Fingerprinting.” *J Clin Microbiol* 39(5):1833–39.

Kim, Byeong Gwan et al. 2007. “Comparison of 7-Day and 14-Day Proton Pump Inhibitor-Containing Triple Therapy for Helicobacter Pylori Eradication: Neither Treatment Duration Provides Acceptable Eradication Rate in Korea.” *Helicobacter* 12(1):31–35.

Kim, J. S., J. H. Chang, S. I. Chung, and J. S. Yum. 1999. “Molecular Cloning and Characterization of the Helicobacter Pylori fliD Gene, an Essential Factor in Flagellar Structure and Motility.” *Journal of bacteriology* 181(22):6969–76.

Kim, Kyung-Mi et al. 2007. “γ-Glutamyltranspeptidase of Helicobacter Pylori Induces Mitochondria-Mediated Apoptosis in AGS Cells.” *Biochemical and Biophysical Research Communications* 355(2):562–67.

Kim, Nayoung et al. 2004. “Genes of Helicobacter Pylori Regulated by Attachment to AGS Cells.” *Infection and immunity* 72(4):2358–68.

Kim, Sa-Hyun et al. 2014. “Cyanidin 3-O-Glucoside Reduces Helicobacter Pylori VacA-Induced Cell Death of Gastric KATO III Cells through Inhibition of the SecA Pathway.” *International journal of medical sciences* 11(7):742–47.

Kimber, M. S. and E. F. Pai. 2000. “The Active Site Architecture of Pisum Sativum Beta-Carbonic Anhydrase Is a Mirror Image of that of Alpha-Carbonic Anhydrases.” *EMBO Journal* 19(7):1407–18.

Kivi, M., a L. V Johansson, M. Reilly, and Y. Tindberg. 2005. “Helicobacter Pylori Status in Family Members as Risk Factors for Infection in Children.” *Epidemiology and infection* 133:645–52.

Klein, P. D. et al. 1991. “Originally Published as Volume 1, Issue 8756Water Source as Risk Factor for Helicobacter Pylori Infection in Peruvian Children.” *The Lancet* 337(8756):1503–6.

De Koning-Ward, T. F., A. .. C. Ward, and R. M. Robins-Browne. 1994. “Characterisation of the Urease-Encoding Gene Complex of Yersinia Enterocolitica.” *Gene* 145(1):25–32.

Konno, Mutsuko et al. 2008. “Predominance of Mother-to-Child Transmission of Helicobacter Pylori Infection Detected by Random Amplified Polymorphic DNA Fingerprinting Analysis in Japanese Families.” *The Pediatric infectious disease journal* 27(11):999–1003.

Konturek, J. W. 2003. “Discovery by Jaworski of Helicobacter Pylori and Its Pathogenetic Role in Peptic Ulcer, Gastritis and Gastric Cancer.” *Journal of Physiology and Pharmacology* 54(SUPPL. 3):23–41.

Kostrzynska, M., J. D. Betts, J. W. Austin, and T. J. Trust. 1991. “Identification, Characterization, and Spatial Localization of Two Flagellin Species in Helicobacter Pylori Flagella.” *Journal of Bacteriology* 173(3):937–46.

Koyanagi, S., K. Nagata, T. Tamura, S. Tsukita, and N. Sone. 2000. “Purification and Characterization of Cytochrome c-553 from Helicobacter Pylori.” *Journal of biochemistry* 128(3):371–75.

Kraft, Christian and Sebastian Suerbaum. 2005. “Mutation and Recombination in Helicobacter Pylori: Mechanisms and Role in Generating Strain Diversity.” *International journal of medical microbiology : IJMM* 295(5):299–305.

Krauss, Irene Russo, Antonello Merlino, Alessandro Vergara, and Filomena Sica. 2013. “An Overview of Biological Macromolecule Crystallization.” *International Journal of Molecular Sciences* 14(6):11643–91.

Krissinel, Evgeny and Kim Henrick. 2007. “Inference of Macromolecular Assemblies from Crystalline State.” *Journal of molecular biology* 372(3):774–97.

Krokan, H. E., R. Standal, and G. Slupphaug. 1997. “DNA Glycosylases in the Base Excision Repair of DNA.” *Biochem J* 325 (Pt 1:1–16.

Kuipers, Ernst J. et al. 2000. “Quasispecies Development of Helicobacter Pylori Observed in Paired Isolates Obtained Years Apart from the Same Host.” *The Journal of Infectious Diseases* 181(1):273–82.

Kusters, J. G., A. H. M. van Vliet, and E. J. Kuipers. 2006. “Pathogenesis of Helicobacter Pylori Infection.”

Clinical Microbiology Reviews 19(3):449–90.

Kuzminov, A. 1999. “Recombinational Repair of DNA Damage in Escherichia Coli and Bacteriophage Lambda.” *Microbiology and molecular biology reviews : MMBR* 63(4):751–813, table of contents.

Kwok, Terry et al. 2007. “Helicobacter Exploits Integrin for Type IV Secretion and Kinase Activation.” *Nature* 449(7164):862–66.

Kwok, Terry, Steffen Backert, Heinz Schwarz, Jürgen Berger, and Thomas F. Meyer. 2002. “Specific Entry of Helicobacter Pylori into Cultured Gastric Epithelial Cells via a Zipper-Like Mechanism Specific Entry of Helicobacter Pylori into Cultured Gastric Epithelial Cells via a Zipper-Like Mechanism.” *Infection and immunity* 70(4):2108–20.

Lambert, J. R. et al. 1995. “High Prevalence of Helicobacter Pylori Antibodies in an Institutionalized Population: Evidence for Person-to-Person Transmission.” *The American journal of gastroenterology* 90(12):2167–71.

Lara-Ramírez, Edgar Eduardo et al. 2011. “New Implications on Genomic Adaptation Derived from the Helicobacter Pylori Genome Comparison.” *PloS one* 6(2):e17300.

Larussa, Tiziana, Isabella Leone, Evelina Suraci, Maria Imeneo, and Francesco Luzza. 2015. “Helicobacter Pylori and T Helper Cells : Mechanisms of Immune Escape and Tolerance.” *Journalof Immunology Research* 2015.

Lee, Mann Hyung, Scott B. Mulrooney, Michael J. Renner, Yves Markowicz, and Robert P. Hausinger. 1992. “Klebsiella Aerogenes Urease Gene Cluster: Sequence of ureD and Demonstration That Four Accessory Genes (ureD, ureE, ureF, and ureG) Are Involved in Nickel Metallocenter Biosynthesis.” *Journal of bacteriology* 174(13):4324–30.

Lehneck, Ronny et al. 2014. “Crystal Structures of Two Tetrameric β-Carbonic Anhydrases from the Filamentous Ascomycete Sordaria Macrospora.” *FEBS Journal* 281:1759–72.

Leung, Wai-Keung et al. 1999. “Isolation of Helicobacter Pylori from Vomitus in Children and Its Implication in Gastro-Oral Transmission.” *Am J Gastroenterol* 94(10):2881–84.

Leunk, R. D., P. T. Johnson, B. C. David, W. G. Kraft, and D. R. Morgan. 1988. “Cytotoxic Activity in Broth-Culture Filtrates of Campylobacter Pylori.” *Journal of medical microbiology* 26(2):93–99.

Liljas, A. and M. Laurberg. 2000. “A Wheel Invented Three Times. The Molecular Structures of the Three Carbonic Anhydrases.” *EMBO reports* 1(1):16–17.

Lina, Taslima T. et al. 2014. “Immune Evasion Strategies Used by Helicobacter Pylori.” *World Journal of Gastroenterology* 20(36):12753.

Lindskog, S. 1997. “Structure and Mechanism of Carbonic Anhydrase.” *Pharmacology & therapeutics* 74(1):1–20.

Lindskog, S. and J. E. Coleman. 1973. “The Catalytic Mechanism of Carbonic Anhydrase.” *Proceedings of the National Academy of Sciences of the United States of America* 70(9):2505–8.

Loh, John T., Shobhana S. Gupta, David B. Friedman, Andrzej M. Krezel, and Timothy L. Cover. 2010. “Analysis of Protein Expression Regulated by the Helicobacter Pylori ArsRS Two-Component Signal Transduction System.” *Journal of bacteriology* 192(8):2034–43.

Luo, Miao, Siqiang Niu, Yibing Yin, Ailong Huang, and Deqiang Wang. 2009. “Cloning, Purification, Crystallization and Preliminary X-Ray Studies of Flagellar Hook Scaffolding Protein FlgD from

Pseudomonas Aeruginosa PAO1." *Acta Crystallographica Section F: Structural Biology and Crystallization Communications* 65(8):795–97.

Lupetti, P. et al. 1996. "Oligomeric and Subunit Structure of the Helicobacter Pylori Vacuolating Cytotoxin." *Journal of Cell Biology* 1994(4):801–7.

Macauley, Sheridan R. et al. 2009. "The Archetype Gamma-Class Carbonic Anhydrase (Cam) Contains Iron When Synthesized in Vivo." *Biochemistry* 48(5):817–19.

Magalhaes, Ana et al. 2009. "Fut2-Null Mice Display an Altered Glycosylation Profile and Impaired BabA-Mediated Helicobacter Pylori Adhesion to Gastric Mucosa." *Glycobiology* 19(12):1525–36.

Magariyama, Yukio and Seishi Kudo. 2002. "A Mathematical Explanation of an Increase in Bacterial Swimming Speed with Viscosity in Linear-Polymer Solutions." *Biophysical Journal* 83(2):733–39.

Mahdavi, Jafar et al. 2002. "Helicobacter Pylori SabA Adhesin in Persistent Infection and Chronic Inflammation." *Science* 297(5581):573–78.

Makrides, SC. 1996. "Strategies for Achieving High-Level Expression of Genes in Escherichia Coli." *Microbiol. Rev.* 60(3):512–38.

Malaty, H. M., L. Engstrand, N. L. Pedersen, and D. Y. Graham. 1994. "Helicobacter Pylori Infection: Genetic and Environmental Influences. A Study of Twins." *Annals of internal medicine* 120(12):982–86.

Malfertheiner, P. et al. 2007. "Current Concepts in the Management of Helicobacter Pylori Infection : The Maastricht III Consensus Report." *Gut* 56(6):772–81.

Malnick, Stephen David Howard, Ehud Melzer, Malka Attali, Gabriel Duek, and Jacob Yahav. 2014. "Helicobacter Pylori: Friend or Foe?" *World journal of gastroenterology : WJG* 20(27):8979–85.

Mapstone, N. P. et al. 1993. "PCR Identification of Helicobacter Pylori in Faeces from Gastritis Patients." *The Lancet* 341(8842):447.

Marcus, Elizabeth a, George Sachs, Yi Wen, Jing Feng, and David R. Scott. 2012. "Role of the Helicobacter Pylori Sensor Kinase ArsS in Protein Trafficking and Acid Acclimation." *Journal of bacteriology* 194(20):5545–51.

Marcus, Elizabeth A., Amiel P. Moshfegh, George Sachs, and David R. Scott. 2005a. "The Periplasmic Alpha-Carbonic Anhydrase Activity of Helicobacter Pylori Is Essential for Acid Acclimation." *Journal of bacteriology* 187(2):729–38.

Marcus, Elizabeth A., Amiel P. Moshfegh, George Sachs, and David R. Scott. 2005b. "The Periplasmic α-Carbonic Anhydrase Activity of Helicobacter Pylori Is Essential for Acid Acclimation." *Journal of Bacteriology* 187(2):729–38.

Maresca, Alfonso, Daniela Vullo, Andrea Scozzafava, and Claudiu T. Supuran. 2013. "Inhibition of the Alpha- and Beta-Carbonic Anhydrases from the Gastric Pathogen Helycobacter Pylori with Anions." *Journal of enzyme inhibition and medicinal chemistry* 28(2):388–91.

Maria, Ruth et al. 2011. "Determination of Strain of Helicobacter Pylori and of Polymorphism in the Interleukin-8 Gene in Patients with Stomach Cancer." *Arquivos de Gastroeneterologia Journal* (1):46–51.

Marshall, BarryJ and J. Robi. Warren. 1984. "Unidentified Curved Bacilli in the Stomach of Patients with Gastritis and Peptic Ulceration." *The Lancet* 323(8390):1311–15.

Marsin, Stéphanie, Aurélie Mathieu, Thierry Kortulewski, Raphaël Guérois, and J. Pablo Radicella. 2008. "Unveiling Novel RecO Distant Orthologues Involved in Homologous Recombination." *PLoS genetics* 4(8):e1000146.

De Martel, Catherine and Julie Parsonnet. 2006. "Helicobacter Pylori Infection and Gender: A Meta-Analysis of Population-Based Prevalence Surveys." *Digestive Diseases and Sciences* 51(12):2292–2301.

Martin, Miriam E. and Jay V. Solnick. 2014. "The Gastric Microbial Community, Helicobacter Pylori Colonization, and Disease." *Gut Microbes* 5(3):345–50.

Martinou, J. C., S. Desagher, and B. Antonsson. 2000. "Cytochrome c Release from Mitochondria: All or Nothing." *Nature cell biology* 2(3):E41–43.

McCarthy, A. A. et al. 2000. "Crystal Structure of the Protein Disulfide Bond Isomerase, DsbC, from Escherichia Coli." *Nature structural biology* 7(3):196–99.

McClain, Mark S. et al. 2003. "Essential Role of a GXXXG Motif for Membrane Channel Formation by Helicobacter Pylori Vacuolating Toxin." *Journal of Biological Chemistry* 278(14):12101–8.

McClain, Mark S., Carrie L. Shaffer, Dawn a Israel, Richard M. Peek, and Timothy L. Cover. 2009. "Genome Sequence Analysis of Helicobacter Pylori Strains Associated with Gastric Ulceration and Gastric Cancer." *BMC genomics* 10:3.

McGee, D. J. and H. L. Mobley. 2000. "Pathogenesis of Helicobacter Pylori Infection." *Current opinion in gastroenterology* 16(1):24–31.

McGee, David J. et al. 2004. "Purification and Characterization of Helicobacter Pylori Arginase, RocF: Unique Features among the Arginase Superfamily." *European journal of biochemistry / FEBS* 271(10):1952–62.

McGee, David J., Fiona J. Radcliff, George L. Mendz, Richard L. Ferrero, and Harry L. T. Mobley. 1999. "Helicobacter Pylori rocF Is Required for Arginase Activity and Acid Protection in Vitro but Is Not Essential for Colonization of Mice or for Urease Activity." *Journal of Bacteriology* 181(23):7314–22.

Mégraud, F. and N. Broutet. 2000. "Review Article: Have We Found the Source of Helicobacter Pylori?" *Alimentary pharmacology & therapeutics* 14 Suppl 3:7–12.

Megraud, F., Trimoulet pascale, H. Lamouliatte, and L. Boyanova. 1991. "Bactericidal Effect of Amoxicillin on Helicobacter Pylori in an in Vitro Model Using Epithelial Cells." *Antimicrobial agents and chemotherapy* 35(5):869–72.

Meyer-Rosberg, K., David R. Scott, David Rex, Klaus Melchers, and George Sachs. 1996. "The Effect of Environmental pH on the Proton Motive Force of Helicobacter Pylori." *Gastroenterology* 111(5):886–900.

Miller, Erica F. and Robert J. Maier. 2014. "Ammonium Metabolism Enzymes Aid Helicobacter Pylori Acid Resistance." *Journal of bacteriology* 196(17):3074–81.

Miller-Podraza, H. et al. 1999. "Helicobacter Pylori and Neutrophils: Sialic Acid-Dependent Binding to Various Isolated Glycoconjugates." *Infection and immunity* 67(12):6309–13.

Minamino, Tohru et al. 2004. "Domain Organization and Function of Salmonella FliK, a Flagellar Hook-Length Control Protein." *Journal of Molecular Biology* 341(2):491–502.

Minamino, Tohru and Anthony P. Pugsley. 2005. "Measure for Measure in the Control of Type III

Secretion Hook and Needle Length." *Molecular Microbiology* 56(2):303–8.

Mitchell, H. M. et al. 1992. "Epidemiology of Helicobacter Pylori in Southern China: Identification of Early Childhood as the Critical Period for Acquisition." *The Journal of infectious diseases* 166(1):149–53.

Mitchell, H. M. 1999. "The Epidemiology of Helicobacter Pylori." *Current topics in microbiology and immunology* 241:11–30.

Mitsuhashi, Satoshi et al. 2000. "X-Ray Structure of Beta Carbonic Anhydrase from the Red Alga,Porphyridium Purpureum, Reveals a Novel Catalytic Site for CO2 Hydration." *Journal of Biological Chemistry* 275(8):5521–26.

Miyaji, H. et al. 2000. "Helicobacter Pylori Infection Occurs via Close Contact with Infected Individuals in Early Childhood." *Journal of gastroenterology and hepatology* 15(3):257–62.

Mizote, T., H. Yoshiyama, and T. Nakazawa. 1997. "Urease-Independent Chemotactic Responses of Helicobacter Pylori to Urea, Urease Inhibitors, and Sodium Bicarbonate." *Infection and immunity* 65(4):1519–21.

Mizuno, Shino, Hirokazu Amida, Naohiro Kobayashi, Shin-Ichi Aizawa, and Shin-ichi Tate. 2011. "The NMR Structure of FliK, the Trigger for the Switch of Substrate Specificity in the Flagellar Type III Secretion Apparatus." *Journal of Molecular Biology* 409(4):558–73.

Mobley, H. L., M. D. Island, and R. P. Hausinger. 1995. "Molecular Biology of Microbial Ureases." *Microbiological reviews* 59(3):451–80.

Mobley, Harry L. T., George L. Mendz, and Stuart L. Hazell, eds. 2001. *Helicobacter Pylori: Physiology and Genetics.* Washington (DC).

Modak, Joyanta K., Yu C. Liu, Mayra A. Machuca, Claudiu T. Supuran, and Anna Roujeinikova. 2015. "Structural Basis for the Inhibition of Helicobacter Pylori Alpha-Carbonic Anhydrase by Sulfonamides." *PloS one* 10(5):e0127149.

Modak, Joyanta K., Sarah A. Revitt-Mills, and Anna Roujeinikova. 2013. "Cloning, Purification and Preliminary Crystallographic Analysis of the Complex of Helicobacter Pylori Alpha-Carbonic Anhydrase with Acetazolamide." *Acta crystallographica. Section F, Structural biology and crystallization communications* 69(Pt 11):1252–55.

Molinari, M. et al. 1998. "Selective Inhibition of Ii-Dependent Antigen Presentation by Helicobacter Pylori Toxin VacA." *The Journal of experimental medicine* 187(1):135–40.

Momtaz, Hassan, Hossein Dabiri, Negar Souod, and Mohsen Gholami. 2014. "Study of Helicobacter Pylori Genotype Status in Cows, Sheep, Goats and Human Beings." *BMC gastroenterology* 14(1):61.

Montecucco, Cesare and Rino Rappuoli. 2001. "Living Dangerously: How Helicobacter Pylori Survives in the Human Stomach." *Nat Rev Mol Cell Biol* 2(6):457–66.

Morishita, Saori et al. 2008. "Cloning, Polymorphism, and Inhibition of Beta-Carbonic Anhydrase of Helicobacter Pylori." *Journal of gastroenterology* 43(11):849–57.

Moriya, Nao, Tohru Minamino, Kelly T. Hughes, Robert M. Macnab, and Keiichi Namba. 2006. "The Type III Flagellar Export Specificity Switch Is Dependent on FliK Ruler and a Molecular Clock." *Journal of molecular biology* 359(2):466–77.

Muotiala, A., I. M. Helander, L. Pyhala, T. U. Kosunen, and A. P. Moran. 1992. "Low Biological Activity of

Helicobacter Pylori Lipopolysaccharide." *Infection and Immunity* 60(4):1714–16.

Muramoto, K., S. Makishima, S. I. Aizawa, and R. M. Macnab. 1998. "Effect of Cellular Level of FliK on Flagellar Hook and Filament Assembly in Salmonella Typhimurium." *Journal of molecular biology* 277(4):871–82.

Naughton, J., G. Duggan, B. Bourke, and M. Clyne. 2014. "Interaction of Microbes with Mucus and Mucins." *Gut Microbes* 5(1):48–52.

Neiši, Dragana, Marshall C. Miller, Zachary T. Quinkert, Markus Stein, and Brian T. Chait. 2010. "Helicobacter Pylori CagA Inhibits PAR1/MARK Family Kinases by Mimicking Host Substrates." *Nature Structural & Molecular Biology* 17(1):130–32.

Niehus, Eike et al. 2004. "Genome-Wide Analysis of Transcriptional Hierarchy and Feedback Regulation in the Flagellar System of Helicobacter Pylori." *Molecular microbiology* 52(4):947–61.

Nishimori, Isao, Daniela Vullo, et al. 2006. "Carbonic Anhydrase Inhibitors: Cloning and Sulfonamide Inhibition Studies of a Carboxyterminal Truncated Alpha-Carbonic Anhydrase from Helicobacter Pylori." *Bioorganic & medicinal chemistry letters* 16(8):2182–88.

Nishimori, Isao, Tomoko Minakuchi, et al. 2006. "Carbonic Anhydrase Inhibitors: DNA Cloning and Inhibition Studies of the Alpha-Carbonic Anhydrase from Helicobacter Pylori, a New Target for Developing Sulfonamide and Sulfamate Gastric Drugs." *Journal of medicinal chemistry* 49(6):2117–26.

Nishimori, Isao, Tomoko Minakuchi, Takuhiro Kohsaki, Saburo Onishi, Hiroaki Takeuchi, Daniela Vullo, Andrea Scozzafava, and Claudiu T. Supuran. 2007. "Carbonic Anhydrase Inhibitors: The Beta-Carbonic Anhydrase from Helicobacter Pylori Is a New Target for Sulfonamide and Sulfamate Inhibitors." *Bioorganic & medicinal chemistry letters* 17(13):3585–94.

Nishimori, Isao, Tomoko Minakuchi, Takuhiro Kohsaki, Saburo Onishi, Hiroaki Takeuchi, Daniela Vullo, Andrea Scozzafava, and Claudiu T. Supuran. 2007. "Carbonic Anhydrase Inhibitors: The β-Carbonic Anhydrase from Helicobacter Pylori Is a New Target for Sulfonamide and Sulfamate Inhibitors." *Bioorganic and Medicinal Chemistry Letters* 17:3585–94.

Nishimori, Isao et al. 2010. "The β-Carbonic Anhydrases from Mycobacterium Tuberculosis as Drug Targets." *Current Pharmaceutical Design* 3300–3309.

Nishimori, Isao, Saburo Onishi, Hiroaki Takeuchi, and Claudiu T. Supuran. 2008. "The Alpha and Beta Classes Carbonic Anhydrases from Helicobacter Pylori as Novel Drug Targets." *Current Pharmaceutical Design* 14(7):622–30.

Nitharwal, Ram Gopal, Vijay Verma, Santanu Dasgupta, and Suman Kumar Dhar. 2011. "Helicobacter Pylori Chromosomal DNA Replication: Current Status and Future Perspectives." *FEBS Letters* 585(1):7–17.

Nolan, K. J. et al. 2002. "In Vivo Behavior of a Helicobacter Pylori SS1 nixA Mutant with Reduced Urease Activity." *Infection and Immunity* 70(2):685–91.

Nowak, Martin A. 1992. "What Is a Quasispecies?" *Trends in ecology & evolution*.

O'Toole, P. W., M. Kostrzynska, and T. J. Trust. 1994. "Non-Motile Mutants of Helicobacter Pylori and Helicobacter Mustelae Defective in Flagellar Hook Production." *Molecular microbiology* 14(4):691–703.

O'Toole, Paul W., Michael C. Lane, and Steffen Porwollik. 2000. "Helicobacter Pylori Motility." *Microbes and Infection* 2(10):1207–14.

Odenbreit, S. et al. 2000. "Translocation of Helicobacter Pylori CagA into Gastric Epithelial Cells by Type IV Secretion." *Science (New York, N.Y.)* 287(5457):1497–1500.

Odenbreit, Stefan et al. 2009. "Outer Membrane Protein Expression Profile in Helicobacter Pylori Clinical Isolates." *Infection and immunity* 77(9):3782–90.

Odenbreit, Stefan, Gerhard Faller, and Rainer Haas. 2002. "Role of the alpAB Proteins and Lipopolysaccharide in Adhesion of Helicobacter Pylori to Human Gastric Tissue." *International journal of medical microbiology : IJMM* 292(3-4):247–56.

Odenbreit, Stefan, Markus Till, Dirk Hofreuter, Gerhard Faller, and Rainer Haas. 1999. "Genetic and Functional Characterization of the *alpAB* Gene Locus Essential for the Adhesion of *Helicobacter Pylori* to Human Gastric Tissue." *Molecular Microbiology* 31(5):1537–48.

Oganesyan, Natalia, Sung-hou Kim, and Rosalind Kim. 2004. "On-Column Chemical Refolding of Proteins." *PharmaGenomics* (September).

Ogiwara, Hiroaki et al. 2009. "Role of Deletion Located between the Intermediate and Middle Regions of the Helicobacter Pylori vacA Gene in Cases of Gastroduodenal Diseases." *Journal of clinical microbiology* 47(11):3493–3500.

Olson, J. W., N. S. Mehta, and R. J. Maier. 2001. "Requirement of Nickel Metabolism Proteins HypA and HypB for Full Activity of Both Hydrogenase and Urease in Helicobacter Pylori." *Molecular Microbiology* 39:176–82.

Orillard, Emilie, J. Pablo Radicella, and Stéphanie Marsin. 2011. "Biochemical and Cellular Characterization of Helicobacter Pylori RecA, a Protein with High-Level Constitutive Expression." *Journal of Bacteriology* 193(23):6490–97.

Orodovsky Mark, Rppuoli Rino, and Covacci Antonello. 1996. "Cag , a Pathogenicity Island of Helicobacter Pylori , Encodes Type I-Specific and Disease-Associated Virulence Factors." *Proceedings of the National Academy of Sciences of the United States of America* 93(December):14648–53.

Osato MS, Ayub K, Le HH, Reddy R, Graham DY. 1998. "House Flies Are Unlike Reservoir of Vector for Helicobacter Pylori." *J Clin Microbiol* 36(9):2786–88.

Paetzel, Mark, Andrew Karla, Natalie C. J. Strynadka, and Ross E. Dalbey. 2002. "Signal Peptidases." *Chemical reviews* 102(12):4549–80.

Pang, Siew Siew et al. 2014. "The Three-Dimensional Structure of the Extracellular Adhesion Domain of the Sialic Acid-Binding Adhesin SabA from Helicobacter Pylori." *The Journal of biological chemistry* 289(10):6332–40.

Panthel, K., P. Dietz, R. Haas, and D. Beier. 2003. "Two-Component Systems of Helicobacter Pylori Contribute to Virulence in a Mouse Infection Model." *Infect Immun* 71(9):5381–85.

Park, Haewon, Bongkeun Song, and Francois M. M. Morel. 2007. "Diversity of the Cadmium-Containing Carbonic Anhydrase in Marine Diatoms and Natural Waters." *Environmental microbiology* 9(2):403–13.

Park, Sung Jean, Woo Sung Son, and Bong-Jin Lee. 2012. "Structural Analysis of Hypothetical Proteins

from Helicobacter Pylori: An Approach to Estimate Functions of Unknown or Hypothetical Proteins." *International journal of molecular sciences* 13(6):7109–37.

Peck, B., M. Ortkamp, K. D. Diehl, E. Hundt, and B. Knapp. 1999. "Conservation, Localization and Expression of HopZ, a Protein Involved in Adhesion of Helicobacter Pylori." *Nucleic acids research* 27(16):3325–33.

Petersen, Andreas Munk and Karen Angeliki Krogfelt. 2003. "Helicobacter Pylori: An Invading Microorganism? A Review." *FEMS Immunology and Medical Microbiology* 36(3):117–26.

Petersen, Thomas Nordahl, Soren Brunak, Gunnar von Heijne, and Henrik Nielsen. 2011. "SignalP 4.0: Discriminating Signal Peptides from Transmembrane Regions." *Nature methods* 8(10):785–86.

Pflock, Michael et al. 2006. "Characterization of the ArsRS Regulon of Helicobacter Pylori, Involved in Acid Adaptation." *Journal of Bacteriology* 188(10):3449–62.

Pich, Oscar Q. and Douglas Scott Merrell. 2013. "The Ferric Uptake Regulator of Helicobacter Pylori: A Critical Player in the Battle for Iron and Colonization of the Stomach." *Future microbiology* 18(9):1199–1216.

Pinard, Melissa A. et al. 2015. "Structure and Inhibition Studies of a Type II Beta-Carbonic Anhydrase psCA3 from Pseudomonas Aeruginosa." *Bioorganic & medicinal chemistry* 23(15):4831–38.

Portal-Celhay, Cynthia and Guillermo I. Perez-Perez. 2006. "Immune Responses to Helicobacter Pylori Colonization: Mechanisms and Clinical Outcomes." *Clinical science (London, England : 1979)* 110(3):305–14.

Pounder, R. E. and D. Ng. 1995. "The Prevalence of Helicobacter Pylori Infection in Different Countries." *Alimentary pharmacology & therapeutics* 9 Suppl 2:33–39.

Del Prete, Sonia et al. 2014. "Discovery of a New Family of Carbonic Anhydrases in the Malaria Pathogen Plasmodium Falciparum--the Eta-Carbonic Anhydrases." *Bioorganic & medicinal chemistry letters* 24(18):4389–96.

Prinz, Christian et al. 2001. "Key Importance of the Helicobacter Pylori Adherence Factor Blood Group Antigen Binding Adhesin during Chronic Gastric Inflammation 1." *Cancer Research* 1903–9.

Queiroz, Dulciene M. M. et al. 2009. "IL2-330G Polymorphic Allele Is Associated with Decreased Risk of Helicobacter Pylori Infection in Adulthood." *Microbes and infection / Institut Pasteur* 11(12):980–87.

Raghavan, S., M. Fredriksson, A. M. Svennerholm, J. Holmgren, and E. Suri-payer. 2003. "Absence of CD4+CD25+ Regulatory T Cells Is Associated with a Loss of Regulation Leading to Increased Pathology in Helicobacter Pylori-Infected Mice." *Clinical & Experimental Immunology* 132(3):393–400.

Raghavan, Sukanya and Jan Holmgren. 2005. "CD4+CD25+ Suppressor T Cells Regulate Pathogen Induced Inflammation and Disease." *FEMS immunology and medical microbiology* 44(2):121–27.

Raju, D. et al. 2012. "Vacuolating Cytotoxin and Variants in Atg16L1 That Disrupt Autophagy Promote Helicobacter Pylori Infection in Humans." *Gastroenterology* 18(9):1199–1216.

De Reuse, Hilde and Stefan Bereswill. 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 Immunology and Medical Microbiology* 50(2):165–76.

Reyrat, J. M. et al. 1999. "3D Imaging of the 58 kDa Cell Binding Subunit of the Helicobacter Pylori

Cytotoxin." *Journal of molecular biology* 290(2):459–70.

Robinson, Karen, Michael F. Loughlin, Rebecca Potter, and Peter J. Jenks. 2005. "Host Adaptation and Immune Modulation Are Mediated by Homologous Recombination in Helicobacter Pylori." *The Journal of infectious diseases* 191(4):579–87.

Ronning, Donald R. et al. 2005. "Active Site Sharing and Subterminal Hairpin Recognition in a New Class of DNA Transposases." *Molecular Cell* 20(1):143–54.

Roszczenko, Paula et al. 2015. "Helicobacter Pylori HP0377, a Member of the Dsb Family, Is an Untypical Multifunctional CcmG That Cooperates with Dimeric Thioldisulfide Oxidase HP0231." *BMC Microbiology* 15(1):135.

Rothenbacher, D., M. J. Blaser, G. Bode, and H. Brenner. 2000. "Inverse Relationship between Gastric Colonization of Helicobacter Pylori and Diarrheal Illnesses in Children: Results of a Population-Based Cross-Sectional Study." *The Journal of infectious diseases* 182(5):1446–49.

Rothenbacher, Dietrich and Hermann Brenner. 2003. "Burden of Helicobacter Pylori and H. Pylori-Related Diseases in Developed Countries: Recent Developments and Future Implications." *Microbes and Infection* 5(8):693–703.

Rowlett, Roger S. 2010. "Structure and Catalytic Mechanism of the Beta-Carbonic Anhydrases." *Biochimica et biophysica acta* 1804(2):362–73.

Russo, Maria Elena et al. 2013. "Kinetic Study of a Novel Thermo-Stable Alpha-Carbonic Anhydrase for Biomimetic CO_2 Capture." *Enzyme and microbial technology* 53(4):271–77.

Ryan, Kieran A., Najma Karim, Mulugeta Worku, Charles W. Penn, and Paul W. O'Toole. 2005. "Hook-Filament Transition Is Controlled by a FliK Functional Homolog Encoded by the Gene HP0906." *Journal of bacteriology* 187(16):5742–50.

Sachs, G. 2005. "Acid Acclimation by Helicobacter Pylori." *Physiology* 20(6):429–38.

Sachs, George, David L. Weeks, Klaus Melchers, and David R. Scott. 2003. "The Gastric Biology of Helicobacter Pylori." *Annual review of physiology* 65:349–69.

Sachs, George, Yi Wen, and David R. Scott. 2009. "Gastric Infection by Helicobacter Pylori." *Current Gastroenterology reports* 11(6):455–61.

Salama, N. et al. 2000. "A Whole-Genome Microarray Reveals Genetic Diversity among Helicobacter Pylori Strains." *Proceedings of the National Academy of Sciences of the United States of America* 97(26):14668–73.

Salama, Nina R., Mara L. Hartung, and Anne Muller. 2013. "Life in the Human Stomach: Persistence Strategies of the Bacterial Pathogen Helicobacter Pylori." *Nat Rev Micro* 11(6):385–99.

Salih, Barik A. 2009. "Helicobacter Pylori Infection in Developing Countries: The Burden for How Long?" *Saudi Journal of Gastroenterology : Official Journal of the Saudi Gastroenterology Association* 15(3):201–7.

Sanders, Carsten, Serdar Turkarslan, Dong-Woo Lee, and Fevzi Daldal. 2010. "Cytochrome c Biogenesis: The Ccm System." *Trends in microbiology* 18(6):266–74.

Satin, B. et al. 1997. "Effect of Helicobacter Pylori Vacuolating Toxin on Maturation and Extracellular Release of Procathepsin D and on Epidermal Growth Factor Degradation." *Journal of Biological Chemistry* 272(40):25022–28.

Saunders, N. J., J. F. Peden, D. W. Hood, and E. R. Moxon. 1998. “Simple Sequence Repeats in the Helicobacter Pylori Genome.” *Molecular Microbiology* 27(6):1091–98.

Sawaya, Michael R. et al. 2006. “The Structure of β-Carbonic Anhydrase from the Carboxysomal Shell Reveals a Distinct Subclass with One Active Site for the Price of Two.” *Journal of Biological Chemistry* 281(11):7546–55.

Schär, Jennifer, Albert Sickmann, and Dagmar Beier. 2005. “Phosphorylation-Independent Activity of Atypical Response Regulators of Helicobacter Pylori.” *Journal of Bacteriology* 187(9):3100–3109.

Schmees, C. et al. 2007. “Inhibition of T-Cell Proliferation by Helicobacter Pylori Gamma-Glutamyl Transpeptidase.” *Gastroenterology* 132(5):1820–33.

Schreiber, S. et al. 2004. “The Spatial Orientation of Helicobacter Pylori in the Gastric Mucus.” *Proceedings of the National Academy of Sciences of the United States of America* 101(14):5024–29.

Schreiber, Sören et al. 2005. “Rapid Loss of Motility of Helicobacter Pylori in the Gastric Lumen In Vivo Rapid Loss of Motility of Helicobacter Pylori in the Gastric Lumen In Vivo So.” *Infection and immunity* 73(3):1–7.

Scott, D., D. Weeks, K. Melchers, and G. Sachs. 1998. “The Life and Death of Helicobacter Pylori.” *Gut 1998;43(Suppl 1):S56–60.* 43(supplem(suppl 1):S56–60.

Scott, David R. et al. 2000. “Expression of the Helicobacter Pylori ureI Gene Is Required for Acidic pH Activation of Cytoplasmic Urease.” *Infection and Immunity* 68(2):470–77.

Scotti, Claudia et al. 2010. “Cell-Cycle Inhibition by Helicobacter Pylori L-Asparaginase.” *PloS one* 5(11):e13892.

Selbach, Matthias, Stefan Moese, Christof R. Hauck, Thomas F. Meyer, and Steffen Backert. 2002. “Src Is the Kinase of the Helicobacter Pylori CagA Protein in Vitro and in Vivo.” *The Journal of biological chemistry* 277(9):6775–78.

Senkovich, Olga a. et al. 2011. “Helicobacter Pylori AlpA and AlpB Bind Host Laminin and Influence Gastric Inflammation in Gerbils.” *Infection and Immunity* 79(8):3106–16.

Sgouras, Dionyssios N., Tran Thi Huyen Trang, and Yoshio Yamaoka. 2015. “Pathogenesis of *Helicobacter Pylori* Infection.” *Helicobacter* 20:8–16.

She, Fei-fei, Jian-yin Lin, Jun-yan Liu, Cheng Huang, and Dong-hui Su. 2003. “Virulence of Water-Induced Coccoid Helicobacter Pylori and Its Experimental Infection in Mice.” *World Journal of Gastroenterology* 9(3):516–20.

Shibayama, Keigo, Hiroaki Takeuchi, Jun-Ichi Wachino, Shigetarou Mori, and Yoshichika Arakawa. 2011. “Biochemical and Pathophysiological Characterization of Helicobacter Pylori Asparaginase.” *Microbiology and immunology* 55(6):408–17.

Simon, Jörg and Lars Hederstedt. 2011. “Composition and Function of Cytochrome c Biogenesis System II.” *The FEBS journal* 278(22):4179–88.

De Simone, Giuseppina, Vincenzo Alterio, and Claudiu T. Supuran. 2013. “Exploiting the Hydrophobic and Hydrophilic Binding Sites for Designing Carbonic Anhydrase Inhibitors.” *Expert opinion on drug discovery* 8(7):793–810.

Sipponen, P. 1995. “Helicobacter Pylori: A Cohort Phenomenon.” *The American journal of surgical*

pathology 19 Suppl 1:S30–36.

Skouloubris, S., a Labigne, and H. De Reuse. 1997. “Identification and Characterization of an Aliphatic Amidase in Helicobacter Pylori.” *Molecular microbiology* 25(5):989–98.

Skouloubris, Stéphane, Agnès Labigne, and Hilde De Reuse. 2001. “The AmiE Aliphatic Amidase and AmiF Formamidase of Helicobacter Pylori: Natural Evolution of Two Enzyme Paralogues.” *Molecular Microbiology* 40(3):596–609.

Sly, W. S. and P. Y. Hu. 1995. “Human Carbonic Anhydrases and Carbonic Anhydrase Deficiencies.” *Annual review of biochemistry* 64:375–401.

Smith, K. S. and J. G. Ferry. 1999. “A Plant-Type (beta-Class) Carbonic Anhydrase in the Thermophilic Methanoarchaeon Methanobacterium Thermoautotrophicum.” *Journal of bacteriology* 181(20):6247–53.

Smith, K. S., C. Ingram-Smith, and J. G. Ferry. 2002. “Roles of the Conserved Aspartate and Arginine in the Catalytic Mechanism of an Archaeal -Class Carbonic Anhydrase.” *Journal of Bacteriology* 184(15):4240–45.

Smith, Kerry S., Nathaniel J. Cosper, Christina Stalhandske, and Robert A. Scott. 2000. “Structural and Kinetic Characterization of an Archaeal β-Class Carbonic Anhydrase.” *Journal of Bacteriology* 182(23):6605–13.

Smith, S. I., S. Miehlke, K. S. Oyedeji, A. A. Arigbabu, and A. O. Coker. 2002. “Fingerprinting of Nigerian Helicobacter Pylori Isolates by Plasmid Profile and PCR.” *Journal of basic microbiology* 42(1):45–53.

Smith, Sinéad M. 2014. “Role of Toll-like Receptors in Helicobacter Pylori Infection and Immunity.” *World journal of gastrointestinal pathophysiology* 5(3):133–46.

Solnick, Jay V, Lori M. Hansen, Nina R. Salama, Jenni K. Boonjakuakul, and Michael Syvanen. 2004. “Modification of Helicobacter Pylori Outer Membrane Protein Expression during Experimental Infection of Rhesus Macaques.” *Proceedings of the National Academy of Sciences of the United States of America* 101(7):2106–11.

Sorensen, Hans Peter and Kim Kusk Mortensen. 2005. “Advanced Genetic Strategies for Recombinant Protein Expression in Escherichia Coli.” *Journal of biotechnology* 115(2):113–28.

Sørensen, Hans Peter and Kim Kusk Mortensen. 2005. “Soluble Expression of Recombinant Proteins in the Cytoplasm of Escherichia Coli.” *Microbial cell factories* 4(Figure 1):1.

Steiner, H. and S. Lindskog. 1972. “Effects of High Concentrations of Salt on the Esterase Activity of Human Carbonic Anhydrase.” *FEBS letters* 24(1):85–88.

Stevens, Julie M., Oliver Daltrop, James W. A. Allen, and Stuart J. Ferguson. 2004. “C-Type Cytochrome Formation: Chemical and Biological Enigmas.” *Accounts of chemical research* 37(12):999–1007.

Strop, Pavel, Kerry S. Smith, Tina M. Iverson, James G. Ferry, and Douglas C. Rees. 2001. “Crystal Structure of the ‘Cab’-Type β Class Carbonic Anhydrase from the Archaeon Methanobacterium Thermoautotrophicum.” *Journal of Biological Chemistry* 276(13):10299–305.

Su, B. et al. 1998. “Type I Helicobacter Pylori Shows Lewis(b)-Independent Adherence to Gastric Cells Requiring de Novo Protein Synthesis in Both Host and Bacteria.” *Journal of Infectious Diseases* 178(5):1379–90.

Suerbaum, S. 2000. “Genetic Variability within Helicobacter Pylori.” *International journal of medical microbiology : IJMM* 290(2):175–81.

Suerbaum, S. and C. Josenhans. 1999. “Virulence Factors of Helicobacter Pylori: Implications for Vaccine Development.” *Molecular medicine today* 5(1):32–39.

Suerbaum, Sebastian and Mark Achtman. 2004. “Helicobacter Pylori: Recombination, Population Structure and Human Migrations.” *International journal of medical microbiology : IJMM* 294(2-3):133–39.

Suerbaum, Sebastian and Christine Josenhans. 2007. “Helicobacter Pylori Evolution and Phenotypic Diversification in a Changing Host.” *Nature reviews. Microbiology* 5(6):441–52.

Supuran, Claudiu T. 2011. “Carbonic Anhydrase Inhibitors and Activators for Novel Therapeutic Applications.” *Future medicinal chemistry* 3(9):1165–80.

Supuran, Claudiu T. and Andrea Scozzafava. 2007. “Carbonic Anhydrases as Targets for Medicinal Chemistry.” *Bioorganic & medicinal chemistry* 15(13):4336–50.

Suzuki, Masato et al. 2009. “Helicobacter Pylori CagA Phosphorylation-Independent Function in Epithelial Proliferation and Inflammation.” *Cell host & microbe* 5(1):23–34.

Sycuro, Laura K. et al. 2010. “Peptidoglycan Crosslinking Relaxation Promotes Helicobacter Pylori’s Helical Shape and Stomach Colonization.” *Cell* 141(5):822–33.

Sycuro, Laura K. et al. 2012. “Multiple Peptidoglycan Modification Networks Modulate Helicobacter Pylori’s Cell Shape, Motility, and Colonization Potential.” *PLoS Pathogens* 8(3):e1002603.

Syrjänen, Leo et al. 2010. “Characterization of the First Beta-Class Carbonic Anhydrase from an Arthropod (Drosophila Melanogaster) and Phylogenetic Analysis of Beta-Class Carbonic Anhydrases in Invertebrates.” *BMC biochemistry* 11:28.

Tanc, Muhammet, Fabrizio Carta, Andrea Scozzafava, and Claudiu T. Supuran. 2015. “Alpha-Carbonic Anhydrases Possess Thioesterase Activity.” *ACS medicinal chemistry letters* 6(3):292–95.

Taylor, Diane E. 1999. “Helicobacter Pylori and Its Genome: Lessons from the Treasure Map.” 13(3):218–23.

Tegtmeyer, N., S. Wessler, and S. Backert. 2013. “Role of the Cag Pathogenicity Island Encoded Type IV Secretion System in Helicobacter Pylori Pathogenesis.” *FEBS Journal* 18(9):1199–1216.

Terebiznik, M. R. et al. 2006. “Helicobacter Pylori VacA Toxin Promotes Bacterial Intracellular Survival in Gastric Epithelial Cells.” *Infection and Immunity* 74(12):6599–6614.

Terradot, Laurent and Gabriel Waksman. 2011. “Architecture of the Helicobacter Pylori Cag-Type IV Secretion System.” *FEBS Journal* 278(8):1213–22.

Terry, Karianne, Susan M. Williams, Lynn Connolly, and Karen M. Ottemann. 2005. “Chemotaxis Plays Multiple Roles during Helicobacter Pylori Animal Infection Chemotaxis Plays Multiple Roles during Helicobacter Pylori Animal Infection.” *Infection and Immunity* 73(2):803–11.

Tomb, J. F. et al. 1997. “The Complete Genome Sequence of the Gastric Pathogen Helicobacter Pylori.” *Nature* 388(6642):539–47.

Tonello, F. et al. 1999. “The Helicobacter Pylori Neutrophil-Activating Protein Is an Iron-Binding Protein with Dodecameric Structure.” *Molecular microbiology* 34(2):238–46.

Torres, Victor J., Susan E. Ivie, Mark S. McClain, and Timothy L. Cover. 2005. “Functional Properties of

the p33 and p55 Domains of the Helicobacter Pylori Vacuolating Cytotoxin." *Journal of Biological Chemistry* 280(22):21107–14.

Touati, Danièle. 2000. "Iron and Oxidative Stress in Bacteria." *Archives of Biochemistry and Biophysics* 373(1):1–6.

Trinchieri, Giorgio. 2003. "Interleukin-12 and the Regulation of Innate Resistance and Adaptive Immunity." *Nature reviews. Immunology* 3(2):133–46.

Tripp, Brian C., Kerry Smith, and James G. Ferry. 2001. "Carbonic Anhydrase: New Insights for an Ancient Enzyme." *Journal of Biological Chemistry* 276(52):48615–18.

Tytgat, G. N. 1995. "Endoscopic Transmission of Helicobacter Pylori." *Alimentary pharmacology & therapeutics* 9 Suppl 2:105–10.

Vagin, Alexei and Alexei Teplyakov. 2010. "Molecular Replacement with MOLREP." *Acta crystallographica. Section D, Biological crystallography* 66(Pt 1):22–25.

Vaira, D. et al. 1988. "CAMPYLOBACTER PYLORI IN ABATTOIR WORKERS: IS IT A ZOONOSIS?" *The Lancet* 332(8613):725–26.

Van Vliet, Arnoud H. M. et al. 2002. "The Role of the Ferric Uptake Regulator (Fur) in Regulation of Helicobacter Pylori Iron Uptake." *Helicobacter* 7(4):237–44.

Van Vliet, Arnoud H. M., Florian D. Ernst, and Johannes G. Kusters. 2004. "NikR-Mediated Regulation of Helicobacter Pylori Acid Adaptation." *Trends in Microbiology* 12(11):489–94.

Van Vliet, Arnoud H. M., Ernst J. Kuipers, Jeroen Stoof, Sophie W. Poppelaars, and Johannes G. Kusters. 2004. "Acid-Responsive Gene Induction of Ammonia-Producing Enzymes in Helicobacter Pylori Is Mediated via a Metal-Responsive Repressor Cascade." *Infection and Immunity* 72(2):766–73.

Vullo, Daniela et al. 2013. "The Alpha-Carbonic Anhydrase from the Thermophilic Bacterium Sulfurihydrogenibium Yellowstonense YO3AOP1 Is Highly Susceptible to Inhibition by Sulfonamides." *Bioorganic & medicinal chemistry* 21(6):1534–38.

Wang, Ge, Praveen Alamuri, M. Zafri Humayun, Diane E. Taylor, and Robert J. Maier. 2005. "The Helicobacter Pylori MutS Protein Confers Protection from Oxidative DNA Damage." *Molecular microbiology* 58(1):166–76.

Wang, Ge, F. Lo Leja, and Robert J. Maier. 2011. "NIH Public Access." *DNA Repair* 18(9):1199–1216.

Waters, Richard C., Paul W. O'Toole, and Kieran a Ryan. 2007. "The FliK Protein and Flagellar Hook-Length Control." *Protein science : a publication of the Protein Society* 16(5):769–80.

Weeks, D. L., S. Eskandari, D. R. Scott, and G. Sachs. 2000. "A H+-Gated Urea Channel: The Link between Helicobacter Pylori Urease and Gastric Colonization." *Science (New York, N.Y.)* 287(5452):482–85.

Wen, Yi, Jing Feng, David R. Scott, Elizabeth A. Marcus, and George Sachs. 2007. "The HP0165-HP0166 Two-Component System (ArsRS) Regulates Acid-Induced Expression of HP1186 Alpha-Carbonic Anhydrase in Helicobacter Pylori by Activating the pH-Dependent Promoter." *Journal of bacteriology* 189(6):2426–34.

Wen, Yi, Jing Feng, David R. Scott, Elizabeth a. Marcus, and George Sachs. 2011. "A Cis-Encoded Antisense Small RNA Regulated by the HP0165-HP0166 Two-Component System Controls

Expression of ureB in Helicobacter Pylori." *Journal of Bacteriology* 193(1):40–51.

WHO-IARC. 1994. "Schistosomes, Liver Flukes and Helicobacter Pylori. IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. Lyon, 7-14 June 1994." *IARC monographs on the evaluation of carcinogenic risks to humans / World Health Organization, International Agency for Research on Cancer* 61:1–241.

Willey, Joanne M., Linda. Sherwood, Christopher J. Woolverton, Lansing M. Prescott, and Joanne M. Willey. 2011. *Prescott's Microbiology*. New York: McGraw-Hill.

Wilson, Keith T. and Jean E. Crabtree. 2007. "Immunology of Helicobacter Pylori: Insights Into the Failure of the Immune Response and Perspectives on Vaccine Studies." *Gastroenterology* 133(1):288–308.

Windle, H. J., A. Fox, D. Ni Eidhin, and D. Kelleher. 2000. "The Thioredoxin System of Helicobacter Pylori." *The Journal of biological chemistry* 275(7):5081–89.

Winn, Martyn D. et al. 2011. "Overview of the CCP4 Suite and Current Developments." *Acta crystallographica. Section D, Biological crystallography* 67(Pt 4):235–42.

Wyle, F. A., A. Tarnawski, D. Schulman, and W. Dabros. 1990. "Evidence for Gastric Mucosal Cell Invasion by C. Pylori: An Ultrastructural Study." *Journal of clinical gastroenterology* 12 Suppl 1:S92–98.

Xia, H. H. and N. J. Talley. 1997. "Natural Acquisition and Spontaneous Elimination of Helicobacter Pylori Infection: Clinical Implications." *The American journal of gastroenterology* 92(10):1780–87.

Xu, Yan, Liang Feng, Philip D. Jeffrey, Yigong Shi, and Francois M. M. Morel. 2008. "Structure and Metal Exchange in the Cadmium Carbonic Anhydrase of Marine Diatoms." *Nature* 452(7183):56–61.

Yahiro, K. et al. 1999. "Activation of Helicobacter Pylori VacA Toxin by Alkaline or Acid Conditions Increases Its Binding to a 250-kDa Receptor Protein-Tyrosine Phosphatase." *Journal of Biological Chemistry* 274(51):36693–99.

Yamaoka, Y. 2006. "Helicobacter Pylori Outer Membrane Proteins and Gastroduodenal Disease." *Gut* 55(6):775–81.

Yamaoka, Y. 2009. "Helicobacter Pylori Typing as a Tool for Tracking Human Migration." *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases* 15(9):829–34.

Yamaoka, Y., D. H. Kwon, and D. Y. Graham. 2000. "A M(r) 34,000 Proinflammatory Outer Membrane Protein (oipA) of Helicobacter Pylori." *Proceedings of the National Academy of Sciences of the United States of America* 97(13):7533–38.

Yoon, Ji Young et al. 2013. "Structural and Functional Characterization of HP0377, a Thioredoxin-Fold Protein from *Helicobacter Pylori*." *Acta Crystallographica Section D Biological Crystallography* 69(5):735–46.

Zabaleta, Eduardo, M. Victoria Martin, and Hans-Peter Braun. 2012. "A Basal Carbon Concentrating Mechanism in Plants?" *Plant science : an international journal of experimental plant biology* 187:97–104.

Zabaleta, Jovanny et al. 2004. "Helicobacter Pylori Arginase Inhibits T Cell Proliferation and Reduces the Expression of the TCR Zeta-Chain (CD3zeta)." *Journal of immunology (Baltimore, Md. : 1950)* 173(1):586–93.

Zambon, Carlo-F. et al. 2005. “Pro- and Anti-Inflammatory Cytokines Gene Polymorphisms and Helicobacter Pylori Infection: Interactions Influence Outcome.” *Cytokine* 29(4):141–52.

Zanotti, Giuseppe et al. 2002. “Structure of the Neutrophil-Activating Protein from Helicobacter Pylori.” *Journal of molecular biology* 323(1):125–30.

Zanotti, Giuseppe and Laura Cendron. 2014. “Structural and Functional Aspects of the Helicobacter Pylori Secretome.” *World Journal of Gastroenterology* 20(6):1402.

Zawilak, a et al. 2001. “Identification of a Putative Chromosomal Replication Origin from Helicobacter Pylori and Its Interaction with the Initiator Protein DnaA.” *Nucleic acids research* 29(11):2251–59.

Zeevaart, Jacob G. et al. 2009. “Structural Analysis of the Oligomeric States of Helicobacter Pylori VacA Toxin.” *Journal of Molecular Biology* 130(29):9492–99.

Zhang, Man, Arthur F. Monzingo, Laura Segatori, George Georgiou, and Jon D. Robertus. 2004. “Structure of DsbC from Haemophilus Influenzae.” *Acta crystallographica. Section D, Biological crystallography* 60(Pt 9):1512–18.

Zhang, Mei. 2015. “High Antibiotic Resistance Rate: A Difficult Issue for *Helicobacter Pylori* Eradication Treatment.” *World Journal of Gastroenterology* 21(48):13432.

Zolfaghari Emameh, Reza, Harlan Barker, Martti E. E. Tolvanen, Csaba Ortutay, and Seppo Parkkila. 2014. “Bioinformatic Analysis of Beta Carbonic Anhydrase Sequences from Protozoans and Metazoans.” *Parasites & vectors* 7(1):38.