“CHARACTERIZATION OF MULTIFUNCTIONAL PROTEINS FROM *Helicobacter pylori*”

Alessandra Scannella

Dottorato in Scienze Biotecnologiche XXIII ciclo
Indirizzo Biotecnologie industriali e molecolari
Università di Napoli Federico II

Dottorato in Scienze Biotecnologiche XXIII ciclo
Indirizzo Biotecnologie industriali e molecolari
Università di Napoli Federico II

"CHARACTERIZATION OF MULTIFUNCTIONAL PROTEINS FROM *Helicobacter pylori*"

Alessandra Scannella

Dottoranda:	Alessandra Scannella
Relatore:	Prof.ssa Adriana Zagari
Coordinatore:	Prof. Giovanni Sannia

To my love, Dario

Table of contents

Abbreviations:

APS	Ammonium persolfate
ATP	Adenosine triphosphate
BSA	Bovine serum albumin
CoA	Coenzyme A
CoaA	Pantothenate Kinase
dNTP	deoxy nucleotide tri-phosphate
DTT	Dithiothreitol
EDTA	Ethylenediaminetetra-acetic acid
Hp	*Helicobacter pylori*
Hsp	Heat shock protein
HPLC	High performance liquid chromatography
kDa	Kilodalton
K_M	Michaelis constant
IPTG	isopropyl-beta-D-thiogalactopyranoside
LB	Luria-Bertani Broth
PanK	Pantothenate kinase
PCR	Polymerase chain reaction
Sa	*Staphilococcus aureus*
SDS-PAGE	sodium dodecyl sulfate polyacrilamide gel electrophoresis
TAE	Tris-acetate EDTA
TEMED	N,N,N,N-tetramethyl ethylene diamine
TEV	tobacco etch virus
Tris	Tris (Hidroxy methyl) amino methane
V_e	Elution Volume
V_0	Bed Volume
Å	Ångström, 1 Å= 0.1 nm

Summary

This thesis reports the cloning, overexpression and characterization of two interesting proteins from *Helicobacter pylori*: Heat Shock protein A (HspA), and CoaX (Type III Pantothenate kinase). *Helicobacter pylori* is a gram negative bacterium that colonises the human stomach. It is present in more than half of the world's population and causes major diseases such as gastritis, peptic ulcers and stomach cancer.
The current treatment against *H. pylori,* consisting of a combination of a proton pump inhibitor and two different antibiotics, is effective but has several drawbacks such as increased risk of antibiotic resistance and no protection against re-infection. Although extensive research is going on to develop a vaccine that would be an attractive alternative or a complement to the current treatment, there is no vaccine available against *H. pylori* today.
Helicobacter pylori produces an unusual GroES homologue protein referred as to HspA (Heat-shock protein A). Besides its classical co-chaperone activity, HspA plays additional roles being involved in nickel binding. It also exhibits an extended subcellular localization, ranging from cytoplasm to bacterial cell surface. In fact, unlike its homologue proteins, HspA is highly immunogenic being also present in the extra cellular media. For this reason it is considered a target for new therapeutic strategies. HspA consists of two domains: an N-terminal domain (domain A, residues 1-90), that is homologous with other GroES bacterial proteins and a C-terminal domain (domain B, residues 91-118), which other GroES-like proteins lack. Domain B is unique to HspA and contains 8 histidines and 4 cysteines which have been suggested to be involved in nickel binding.
This study points on a unique characteristic of HspA among all GroES proteins: a high content of cysteine residues. Cysteine is the less represented residue in all known GroES proteins examined. In this context we have produced and characterized a recombinant HspA and its mutants Cys94Ala and Cys94Ala/Cys111Ala and we have identified the disulphide bridge pattern of the protein. In particular the study has been addressed on the Cys oxidised/reduced state; the disulphide bridge pattern has been assigned by integrating classical biochemical methodologies with mass spectrometry. We found that the cysteines (two from domain A and four from domain B) are engaged in three disulphide bonds between residues Cys51/Cys53, Cys94/Cys111 and Cys95/Cys112. Our results suggest that two of the disulphide bridges, located in the B domain, force the C-terminal domain to adopt a unique closed loop structure that would be optimal for binding to 2 Ni ions as suggested by the different redox environments that the protein experiences inside and outside the bacterial cell.
H.pylori also produces a new Pantothenate Kinase isoform (HpCoaX) encoded by the *CoaX* gene referred as to HP0862, involved in a critical pathway for pathogen survival: that of coenzyme A (CoA) biosynthesis. In particular, pantothenate kinase catalyzes the first step of the universal five-step CoA biosynthetic pathway. HpCoaX enzyme, belonging to the type III class, differs in sequence, structure and enzymatic properties from the previously characterized type I, found essentially in bacteria, and type II forms found in eukarya as well as in some bacteria; its activity is not regulated by CoA and thio-esters. In this context we have cloned the *HpCoaX* gene in pRoEX-HTc vector and

expressed the recombinant protein in *E.coli* BL21(DE3) cells. We have optimized the experimental conditions for purification and stability of the expressed protein. The purified HpCoax was analyzed by gel filtration chromatography and dynamic light scattering to investigate its state of oligomerization. Results show that HpCoaX exists as a homodimer, as predicted by its homology with the other PanK III bacterial proteins. The purified protein was also characterized by mass spectrometry assay. Crystallization trials using a nanodispensator were also carried out in order to obtain crystals suitable for an X-ray diffraction analysis.

H. pylori infections remain a significant global public health problem. Vaccine and antagonist compound development against this infection appears to be a preferable strategy. Actually there is a substantial requirement for new drug targets as alternative strategies for the treatment of *H. pylori* infections, since there is often resistance against traditional antibiotic therapy. Comparison of substrate binding and catalytic sites of PanK-III with those of PanK-II and PanK-I type, reveals drastic differences in the binding modes for both ATP and pantothenate substrates, suggesting that these differences may be exploited in the development of new inhibitors targeting specifically PanK-III type. HpCoaX inhibitors may be a good goal, interfering with the survival mechanism of *H. pylori* within the gastric mucosa. For the all above reasons, both HpCoaX and HspA can be considered to be a good targets for new therapeutic strategies. Our study belongs to a wide research line with the aim of identifying, characterizing and clarifying the role of specific targets, in particular of multifunctional uncharacterized proteins involved in critical pathway for the pathogen survival.

Riassunto

Lo scopo del seguente progetto di dottorato ha riguardato lo studio sperimentale di alcune proteine con caratteristiche multifunzionali del patogeno gram-negativo *Helicobacter pylori*. Il recente sviluppo delle metodiche d'interazione molecolare su larga scala, basate sia su sistemi *in vitro* che *in vivo* ha permesso di evidenziare le mappe d'interazione proteina-proteina relative ad un intero organismo; tali interazioni risultano coinvolte in molteplici processi biologici cellulari. Di particolare interesse risultano quelle mappe d'interazione di organismi altamente patogeni per l'uomo come *Helicobacter pylori* [Ito, T., *et al*. 2001], [Rain G.C., *et al.* 2001]. *Helicobacter pylori* è il principale agente causale dell'infiammazione cronica della mucosa gastrica e dell'ulcera peptica, e rappresenta uno dei principali fattori di rischio per la neoplasia di tipo gastrico [McGowan C.C., *et al.* 1996], [Parsonnet J., *et al* 1994]. Per colonizzare il microambiente acido dell'epitelio gastrico il batterio ha sviluppato delle efficaci strategie di adattamento e colonizzazione. Infatti *H. pylori* è in grado di produrre un elevato assortimento di fattori di virulenza capaci di neutralizzare il pH della mucosa gastrica e quindi di sopravvivere nell'ambiente estremamente acido dello stomaco.
Pertanto lo studio delle proteine coinvolte nell' omeostasi e nella regolazione metallo-dipendente, così come quello relativo alla caratterizzazione dei componenti proteici coinvolti nei principali pathway biosintetici, riveste primaria importanza nel campo delle infezioni ospite-patogeno, poiché la limitazione di metalli è associata all'induzione dell'espressione dei geni di virulenza in molti patogeni batterici. L'obiettivo principale dell' intero progetto di ricerca di dottorato è la caratterizzazione biochimica e strutturale di alcune proteine multifunzionali del patogeno gram-negativo *H. pylori*, il cui genoma risulta essere accuratamente annotato in seguito al completo sequenziamento [Tomb G.C., *et al*. 1997]. L'accessibilità di tale genoma sequenziato ha avuto un profondo impatto su tutti gli aspetti di ricerca che vanno dalle basi biologiche di *H. pylori* all'identificazione di specifici target fino allo sviluppo di nuove strategie terapeutiche volte al controllo e infine alla completa eradicazione di questo importante agente patogeno.
Il potenziale interesse biotecnologico del progetto risiede nella individuazione e caratterizzazione di nuovi target coinvolti nel meccanismo di adattamento e sopravvivenza del patogeno *H. pylori* e indispensabili per il suo metabolismo, allo scopo di effettuare una progettazione razionale di potenziali farmaci (drug design). Le proteine selezionate come oggetto del mio studio sono Heat schok protein A (HspA) e la pantotenato kinasi di tipo III (HpCoaX). In particolare l'HspA essendo una componente del patogeno con caratteristiche immunogeniche rappresenta un target favorito per lo sviluppo di un vaccino. Ciò rappresenterebbe una valida alternativa terapeutica al tradizionale cocktail antibiotico previsto per eradicare l'*Helicobacter* e che purtroppo spesso risulta fallimentare. [Choudhry A. E., *et al*. 2003]
Nell'ambito di questo progetto ci si è avvalsi di metodiche che sfruttano un approccio interdisciplinare integrando discipline biologiche, biotecnologiche, chimiche e biochimiche allo scopo di affrontare problematiche di natura biotecnologica sia base che applicative. In particolare le principali strategie sono basate sull' utilizzo di metodiche di biologia molecolare, come metodiche di clonaggio ed espressione eterologa di proteine ricombinanti, utilizzando specifici vettori di espressione e ceppi modificati di *E. coli*,

tecniche biochimiche per la purificazione delle proteine (Cromatografia per affinità, cromatografia per esclusione molecolare e light scattering dinamico) integrate con indagini di spettrometria di massa e di tipo strutturale (cristallizzazione di proteine).

Dapprima l'attenzione è stata rivolta prevalentemente allo studio della ***Heat Shock protein A*** **(HspA)** espressa in condizioni di stress cellulare, e coinvolta nell' omeostasi del Nickel [Kansau I., *et al*. 1996], [Ng E.K., *et al*. 1999]. Il nickel è un elemento essenziale per il metabolismo anaerobico dei microrganismi, incluso il patogeno gram-negativo *H. pylori*. L'omeostasi di questo metallo è richiesta per preservare l'equilibrio tra i livelli di Nickel accessibili per le funzioni cellulari batteriche ed i potenziali effetti derivanti da un eccessivo accumulo intracellulare del metallo stesso. Si tratta di un complesso processo che coinvolge l'uptake, l'utilizzazione, l'efflusso ed il deposito degli ioni nickel e che risulta regolato a differenti livelli. Pertanto risultano fortemente attivi all'interno del patogeno quei sistemi deputati alla regolazione dei livelli intracellulari di Ni^{2+} e alla sua detossificazione; *H. pylori* esprime molteplici proteine che risultano coinvolte nell'omeostasi di questo metallo [Robert J. Maier, *et al.* 2007] ed una serie di proteine accessorie tra cui si annovera la *Heat Shock protein A* (HspA) o HP0011, che contiene 6 cisteine impegnate in 3 ponti disolfurici. Lo studio ha condotto alll'assegnazione del corretto pattern di formazione dei ponti disolfurici ed in particolare è stato delucidato il ruolo svolto da due ponti disolfurici vicinali presenti nel dominio B, rilevante per la funzione Ni-dipendente di HspA in *H. pylori*. HspA è una proteina, abbondante in *H. pylori,* conosciuta essere immunogenica. Inaspettatamente presenta una doppia localizzazione: infatti questa proteina citoplasmatica può essere rilevata anche sulla superficie del batterio, attraverso un meccanismo di secrezione ancora non identificato oppure presumibilmente mediante un processo di autolisi programmata [Vanet, *et al*. 1998]. Essa rappresenta un inusuale omologo appartenente alla classe delle GroES, e consta di due regioni: il dominio A (omologo con le altre proteine GroES batteriche) ed il dominio B ricco in cisteine ed istidine. La presenza di tale dominio risulta essere unica e peculiare per *Helicobacter pylori* ed *Helicobacter acinonuchis* e gioca un ruolo essenziale nel recupero e nel rilascio del nickel [Sun X., *et al*. 2008], [Suerbaum S., *et al*. 1994]. L' interesse è stato focalizzato sullo studio di una caratteristica peculiare esibita dall' HspA nell'ambito di tutte le proteine GroES fin'ora conosciute: un elevato contenuto di residui cisteinici. Viceversa, da analisi effettuate mediante tools bioinformatici, la cisteina risulta essere proprio l'aminoacido meno rappresentato all'interno della classe delle proteine GroES. La presenza del dominio B fornisce quindi un ruolo aggiuntivo all'HspA indipendente dalla classica attività di co-chaperone: tale ruolo potrebbe essere mediato da questo elevato contenuto di residui cisteinici, con modalità redox-dipendente. Al fine di indagare sul ruolo giocato da tali residui cisteinici, sono stati prodotti per via ricombinante sia la proteina HspA wild-type che due mutanti a livello di due residui cisteinici nel dominio B: Cys94Ala e Cys94Ala/Cys111Ala. Attraverso la combinanzione sia di metodiche biochimiche, che di spettrometria di massa, è stato determinato lo stato di ossidazione per i residui cisteinici, assegnato il corretto pattern dei ponti disolfurici, in virtù dei quali è stato possibile mettere in evidenza una struttura unica per l'HspA a livello del dominio B. Da questo studio è stato evidenziato che HspA contiene tre ponti disolfurici intramolecolari. I ponti coinvolgono sia il dominio A (Cys51–Cys53) che il dominio B legante il nickel (Cys94–Cys111 e Cys95–Cys112). Ed è proprio in quest'ultimo dominio che il piccolo anello che

si viene a creare tra le quattro Cys rappresenta un forte determinante strutturale fin'ora osservato unicamente per una conotoxina, un piccolo peptide (pdb code 1wct). Per cui in HspA la presenza dei due consecutivi ponti disolfurici, forza la topologia del C-terminale della proteina generando una struttura ''lasso-like'' descritta per la prima volta. Questa struttura potrebbe fornire la conformazione ideale al dominio B per legare gli ioni nickel. Il ruolo di legare il Nickel, potrebbe essere svolto dall'HspA attraverso il sequestro del Ni sia intracellulare che di quello extracellulare, in virtù della presenza della proteina sia a livello subcellulare che all'esterno della membrana batterica. Su questa linea è possibile speculare che la funzione di legare il Nickel potrebbe essere modulata dal differente stato ridotto/ossidato dei residui cisteinici: sulla superficie batterica dove l'ambiente è tipicamente ossidativi la proteina esibisce il pattern dei ponti S-S, mentre nel citoplasma i ponti disolfurici possono esibire una differente stabilità dipendente dalle condizioni quali pH e deplezione del Nickel. In conclusione essendo le cisteine fortemente conservate nei differenti ceppi batterici di *H. pylori*, esse potrebbero assistere il legame/rilascio degli ioni nickel. Pertanto la presenza della peculiare struttura generata dai due ponti disolfurici vicinali nel dominio B è rilevante per la funzione Ni-dipendente di HspA in *H. pylori* [Loguercio S., *et al*. 2008].

In proseguimento di tali studi, sono stati condotti degli esperimenti in presenza di metalli quali il Nickel e il Bismuto, per i quali l' HspA presenta una notevole affinità di legame [Cun*, et al*. 2008]. Si è ipotizzato che la presenza del metallo possa indurre il dominio B dell'HspA a strutturarsi in un folding adatto a legare il metallo. In pratica, dopo opportuna riduzione dei ponti disolfurici (mediante TCEP), e successiva incubazione della proteina con i suddetti metalli, è stato valutato il comportamento della proteina. In seguito all'incubazione in presenza dei due metalli i campioni sono stati sottoposti a esperimenti di cromatografia per esclusione molecolare (Gel filtrazione) per paragonare il differente comportamento della proteina in assenza dei due metalli e senza effetti da parte di un agente riducente, già condotti precedentemente. La proteina così trattata, in entrambe le varianti (in presenza ed in assenza dello ione Nickel e del bismuto) è stata sottoposta ad uno screening di cristallizzazione utilizzando svariati kit a disposizione e mediante l'utilizzo del Nanodispensatore. Sfortunatamente, sia l'HspA-Ni^{+2} che l'HspA-Bi^{+3} non hanno mai formato cristalli evidenziando una intrinseca difficoltà nella cristallizzazione di questa proteina.

Successivamente l'attenzione è stata incentrata sul secondo target proteico oggetto di studio nell'ambito del programma di dottorato. Si tratta di una pantotenato chinasi (PanK) che partecipa al primo di cinque step della via biosintetica che porta alla produzione del coenzima A, un cofattore indispensabile in tutte le forme di vita [Begley T.P., *et al*. 2001]. Il Coenzima A (CoA) rappresenta un cofattore essenziale in tutti i sistemi biologici, in quanto funge da principale fonte di gruppi acili. Inoltre è stato dimostrato che il CoA gioca un ruolo chiave nell'omeostasi dello stato redox provvedendo ad un' importante difesa antiossidante, per un elevato numero di patogeni batterici che, differentemente da *E. coli* sono privi del meccanismo GSH-dipendente. La Pantotenato chinasi (PanK) è un enzima ampiamente conservato nelle differenti specie batteriche (da *E. coli* a *B. subtilis*): catalizza la fosforilazione ATP-dipendente del Pantotenato (vitamina B_5, Pan) con formazione del 4'-fosfopantotenato e svolge un ruolo chiave nella regolazione della biosintesi del CoA. In *E. coli* è stato dimostrato essere

presente in forma omodimerica ed essere regolato nell' attività da feedback negativo a carico del prodotto metabolico. Nonostante l'intero pathway biosintetico del CoA, risulti abbastanza conservato nelle diverse specie batteriche, soltanto alcuni degli enzimi che in *E. coli* convertono il Pan in CoA, risultano conservate in determinati patogeni; infatti è stato identificato recentemente, un nuovo tipo di PanK, codificata dal gene *CoaX*, e che catalizza il primo step della via biosintetica del CoA in patogeni come *B. anthracis* e *H. pylori*. Questa nuova isoforma, denominata Pank III, appartiene alla classe delle chinasi del gruppo 4 (famiglia delle Ribonucleasi H-like) ed esibisce distinte caratteristiche biochimiche rispetto alla classe I e II: in particolare non risente dell' inibizione da feedback negativo ad opera del CoA stesso e suoi tioesteri; non mostra interazione con analoghi del substrato (Pan antimetaboliti); esibisce inoltre caratteristiche cinetiche uniche in quanto possiede bassa specificità per il legame all' ATP [Leisl A., *et al*. 2005], [Hong B.S., *et al*. 2006]. Della pantotenato chinasi da *Helicobacter pylori,* si hanno a disposizione poche conoscenze, data la scarsità di dati in letteratura. Obiettivo dell'attività di ricerca del II anno è stato quello di produrre e caratterizzare la proteina PankIII da *H. pylori* **(HpCoaX**). Sulla base di analisi statistiche (condotte mediante il tool bioinformatico SEED) è stato possibile confrontare la classe delle PanK I e II con la classe III, e selezionare gli organismi patogeni, in cui l'attività pantotenica risulta esercitata esclusivamente dai prodotti del gene CoaX, tra cui *H. pylori.* Da tale analisi derivano importanti implicazioni quali il miglioramento della comprensione dei meccanismi che stanno alla base dell'evoluzione convergente degli enzimi chiave, coinvolti in pathway metabolici essenziali (come la biosintesi del CoA), nonchè l'identificazione di nuovi target terapeutici: queste PanKs III, esibendo infatti distinte caratteristiche, potrebbero rappresentare dei candidati ideali. Attraverso l' applicazione delle principali tecniche di biologia molecolare, Il gene HpCoaX è stato clonato con successo nel vettore di espressione appartenente alla classe pPROEX-HT. Il plasmide di espressione recante la proteina di fusione con l'His_6 –tag all'N-terminale è stato trasformato nel ceppo di espressione competente *E. coli* BL21(DE3)pLysS. Nel secondo anno di attività la proteina è stata così prodotta su larga scala per per via ricombinante ed usata per tutte le caratterizzazioni successive. Analizzando la sequenza amminoacidica della HpCoaX è possibile individuare 8 residui cisteinici che non risultano impegnati nella formazione di alcun ponte disolfurico intramolecolare. Tale dato è stato predetto e successivamente confermato da analisi di spettrometria di massa effettuate dopo il saggio della Vinilpiridina, con denaturazione e successiva alchilazione delle Cys presenti nel campione, in assenza di agenti riducenti. La presenza delle 8 Cys libere comporta la formazione di ponti disolfurici non nativi tra le varie subunità monomeriche con conseguente formazione di diversi stati oligomerici del campione. Questi vengono individuati come un pattern elettroforetico multiplo, con bande a più alto peso molecolare, quando l' analisi SDS-PAGE, è condotta in assenza di β-mercaptoetanolo. Al fine di ottenere una preparazione stabile della proteina, tutti i passaggi di purificazione, dialisi e concentrazione sono stati eseguiti in presenza di 2-3 mM DTT, e per limitare l'ossidazione aerobica dell'agente riducente, il campione è stato sottoposto a flusso di N_2, in tutti gli steps. Allo scopo di valutare lo stato di oligomerizzazione di Hp_0862, la proteina è stata dapprima analizzata mediante cromatografia per gel filtrazione. La Pank III da *H. pylori* mostra un volume di eluizione caratteristico di una proteina di circa 50 kDa (V_e = 15.2 ml/min) e che rappresenta uno stato omodimerico, come risulta dalla predizione per omologia. Il dato è stato

confermato mediante un'analisi molto più accurata di DLS (light scattering dinamico) effettuata sul campione eluito dalla gel filtrazione. La proteina cosi trattata è stata sottoposta ad uno screening di cristallizzazione utilizzando svariati kit di agenti precipitanti, a diverse temperature (20°C e 4°C), avvalendosi del robot Nanodispensatore Hamilton STARLET disponibile presso il nostro dipartimento. Sono state successivamente condotte analisi al fine di ottimizzare mediante la costruzione di una matrice, le condizioni ideali per la cristallizzazione su larga scala: le condizioni di temperatura a 4°C sono risultate essere la condizione essenziale al fine di formare, riprodurre e stabilizzare i cristalli di HpCoaX. La condizione di reservoir di riferimento su cui ci si è basati per riprodurre e sviluppare i cristalli è quella proveniente dal Kit INDEX 2 e corrispondente alla seguente composizione di agente precipitante: 0.1 M Bis-Tris pH 6.5, 0.2 M solfato di ammonio, PEG 400. Su tale condizione si è effettuata la costruzione della matrice di cristallizzazione, variando le concentrazioni e mantenendo costante il valore di pH. I cristallini cosi ottenuti, caratterizzati da forma geometrica regolare, ed opportunamente conservati a 4°C sono stati dapprima testati su SDS-PAGE al fine di valutare l' effettiva formazione di cristalli della proteina di interesse; confermato ciò sono stati di seguito sottoposti a delle analisi preliminari mediante diffrazione ai raggi X. Purtroppo l'elevato disordine all'interno del cristallo e la loro scarsa stabilità, hanno consentito di registrare solo deboli pattern di diffrazione a 5 Å di risoluzione. Nonostante i molteplici esperimenti per migliorare la qualità dei cristalli questi non hanno mai prodotto un pattern di diffrazione ad alta-media risoluzione che ci permettesse di proseguire in tal senso e indirizzare le seguenti analisi verso una risoluzione della struttura tridimensionale. A questo punto sono stati effettuati prove di co-cristallizzazione mediante l'utilizzo sia del substrato (PAN), che del catalizzatore della reazione, un analogo non idrolizzabile dell'ATP (AMPPNP), al fine di ottenere dei cristalli ad elevato potere diffrangente. Nonostante l'aggiunta di tali composti, sia direttamente nella goccia contenete i cristalli di HpCoaX (soaking), che in nuove prove di cristallizzazione in cui i composti venivano miscelati in nuove gocce contenenti la soluzione di proteina, non è stato riscontrato nessun miglioramento nella qualità dei cristalli.

Il potenziale interesse biotecnologico del progetto risiede nell' aver individuato e caratterizzato, nuovi target coinvolti nel meccanismo di adattamento e sopravvivenza del patogeno *H. pylori* ed indispensabili e vitali per il suo metabolismo. Le proteine oggetto del mio studio rappresentano infatti delle molecole di notevole interesse. In particolare la HpCoaX e l' HspA essendo enzimi chiave nel metabolismo del patogeno risultano dei bersagli adatti per la costruzione di specifici inibitori: ciò rappresenterebbe una valida alternativa terapeutica al tradizionale cocktail antibiotico previsto per le infezioni da *Helicobacter* e che purtroppo spesso non rappresenta un trattamento definitivo per la comparsa di casi di recidiva. Dato l'elevato grado di variabilità a livello del genoma del patogeno, infatti, risulta elevato anche il grado di resistenza verso moltissimi antibiotici. Inoltre si pensa che in *Helicobacter pylori* (così come in molti altri organismi patogeni) sia la PanK ad essere responsabile del meccanismo di sviluppo della resistenza a diversi antibiotici. Per cui potrebbe essere interessante sviluppare dei potenziali inibitori della PanK III di *H. pylori*, oppure nuovi composti antimicrobici con modalità d'azione alternativa [Choudhry A. E., *et al*. 2003].

Mediante approcci molecolari, biochimici, spettroscopici sono stati chiariti alcuni meccanismi che governano la funzione delle proteine coinvolte nell'omeostasi redox e nel metabolismo batterico e costituiranno sia la base per lo sviluppo di nuovi farmaci per il blocco dell'infezione, sia la base per lo sviluppo di possibili vaccini. Ci si attende, pertanto, che i risultati ottenuti avranno a lungo termine un elevato impatto nel campo biotecnologico e biomedico, dal momento che le patologie da *H. pylori* colpiscono circa il 50% della popolazione mondiale.

Chapter I:

General Introduction

1.1 History

For a long time, the human stomach was considered to be a sterile organ where no microorganisms could live due to the harsh acidic conditions. This view changed in 1982, when Warren and Marshall were able to culture *Helicobacter pylori* bacteria from gastric biopsies, and found that the bacteria (previously designated *Vibrio rugula* and *Campylobacter pylori*) were present in patients with active chronic gastritis, duodenal ulcer or gastric ulcer. They were later awarded the Nobel Prize for physiology and medicine (in 2005) for their important discovery of *H. pylori* and its role in gastritis and peptic ulcer disease. The possibility that peptic ulcers could be caused by a bacterium, and not by stress, spicy food or other factors was a surprise to the scientific community, and it was difficult for Warren and Marshall to change the prevailing dogma. To convince colleagues and the public, Barry Marshall drank a suspension of bacterium [Marshall B.J., *et al.* 1985]. However, it was not until the early 1990s that it was recognized that *H. pylori* causes peptic ulcer disease. During that period, several researchers confirmed that *H. pylori* eradication cured peptic ulcer [Coghlan *et al.* 1987], [Rauws *et al*. 1990], [Graham *et al.* 1991]. In retrospect, it is interesting to notice that there were many references to the presence of *H. pylori* in gastric mucosa before its culture in 1982. Spiral-shaped bacteria were noted many times in the literature, but their presence was not properly correlated with gastro-duodenal disease [Mobley *et al.* 2001], [Marshall, 2001]. *H. pylori* is, by bacteriological standards, a relative newcomer to medicine. Although its pathogenesis has been studied for the past 20 years, there are reports from as far back as the late 19th century of small, helical bacteria in the stomachs of some patients. One of the first well known reports of gastric helicobacters was done by Bizzozero in Turin in 1893 [Bizzozero G., 1893]. The bacterium was initially named *Campylobacter pyloridis*, then *C. pylori* (after correction to Latin grammar). In 1989, after DNA sequencing and other data had shown that the bacterium did not belong to the *Campylobacter* genus, it was placed in its own genus, *Helicobacter*. The name *pylōri* means “of the pylorus” or pyloric valve (the circular opening leading from the stomach into the duodenum), from the Greek word gatekeeper. Presently, over 20.000 articles have been published on “helicobacter”, not counting the articles under the previous classification of “Campylobacter”. The bacterium has also been shown to play a role in the development of several types of gastric cancers. Indeed, the control of *H. pylori*, may soon be added to the list of high-priority cancer prevention strategies. *H. pylori* seems to have evolved together with humans over tens of thousands of years and arguably has accounted for more deaths from cancer than has any other single carcinogen [Covacci A., *et al*. 1999]. Gastric cancer was the principal cause of worldwide cancer deaths, before being replaced by lung cancer at the end of the twentieth century. It was also the leading cause of cancer death in the United States until about 1940, and its decline in frequency since then probably represents an important and largely unappreciated

consequence of the sanitary revolution. Mortality rates for gastric cancer now vary profoundly around the world, with the greatest effect seen in much of Asia and Latin America. Some countries with a high prevalence of chronic *H. pylori* infection, such as South Africa and India, have low rates of gastric cancer and even within high-risk countries, such as China, Colombia, and Costa Rica, rates can vary greatly (Fig. 1.1). The reasons for these findings are not fully understood, but they may in part reflect variations in the distribution of the more virulent strains of *H. pylori* [Bravo LE., *et al.* 2002]. As we learn more about the pathogenesis of *H. pylori*-induced cancer and about the immunologic mechanisms that contribute to chronic infection, the prospects for vaccine-based prevention strategies become increasingly bright. If we can prevent the acquisition of *H. pylori* infection, which in high-risk countries usually occurs in early childhood, we may virtually eliminate gastric cancer as an important cause of cancer death. Eradication of infection with antibiotic therapy seems to offer more promise, and the results of two clinical trials have provided encouraging, although not conclusive, evidence that antibiotic treatment of adults may prevent gastric carcinogenesis. *H. pylori* is notoriously difficult to treat, and the need to use a multidrug regimen over extended periods, as well as the emergence of resistant organisms, has complicated research and, ultimately, implementation programs aimed at eradication. However, the recent report of the effectiveness of a single-day regimen of combined antibiotic therapy in eradicating infection provides some encouragement that mass treatment programs may be feasible in future [Lara L.F., *et al*. 2003]. The current treatment against *H. pylori* infection consists of a combination of two antibiotics and a proton pump inhibitor. The major drawbacks to this therapy are high cost, poor patient compliance and risk of developing antibiotic resistance. Furthermore, such treatment does not protect against re-infection. Extensive research is going on to develop a vaccine against *H. pylori* that will have (a) prophylactic use to prevent infection and/or (b) therapeutic use to eradicate an ongoing infection. A prophylactic vaccine would primarily be useful in young children in high endemic areas, whereas a therapeutic vaccine may be the most relevant one for treatment for those that are already infected (Svennerholm A.M., and Lundgren A., 2006).

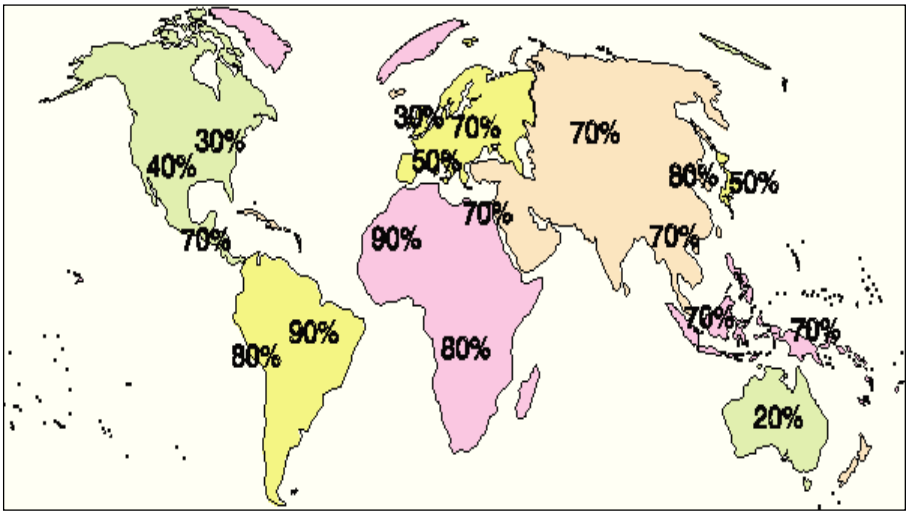

Fig. 1.1: Geographic prevalence of *H. pylori* infection

1.2 Helicobacter pylori

Helicobacter pylori is a spiral-shaped Gram negative flagellated bacterium (Fig. 1.2) that infects the gastric mucosa of more than half of the world's population, making it the most prevalent of all bacterial infections. The prevalence in developing countries can be as high as 80-90%, whereas it is lower in industrialized countries, ranging between 10-50 % (Rothenbacker D., and Brenner H., 2003).

Fig. 1.2: Image of *H. pylori* by electronical microscopy

H. pylori is an important cause of chronic gastritis, peptic ulcer disease and gastric cancer (Goodwin, C.S., 1997). The natural progression of *H. pylori* infection is presented in Fig. 1.3. Infection usually occurs during childhood and causes symptomatic acute gastritis in most patients and persists for decades or life-long; once acquired it, the

infection persists throughout life unless specifically treated. The infection can take multiple courses. Most people infected with *H. pylori* will never develop symptomatic disease. *H. pylori*-induced gastritis, during the decades that follow initial infection, can remain silent or evolve into more-severe diseases: 10-15 % will develop peptic ulcer disease (gastric or duodenal ulcers), approximately 1% will develop gastric adenocarcinoma, and a small group of patients will develop gastric MALT lymphoma (Sauerbaum S., and Josenhans C., 2007). Epidemiological studies first indicated that *H. pylori*-infected subjects have at least a twofold increase in the risk of gastric cancer when compared with uninfected subjects. The strong association between *H. pylori* infection and gastric cancer led the WHO to classify *H. pylori* as a class 1 carcinogen. In industrialised countries, the overall carriage rate of *H pylori* infection in middle-aged adults is 20–50%, compared with 80% or more in many developing countries. The carriage rate of *H. pylori* remains relatively stable, but in the industrialised world these values have substantially decreased over recent decades, probably as a result of improved hygiene and sanitation, especially during childhood, and active elimination of carrier ship via antimicrobial treatment. *H pylori*-associated disorders usually regress or heal completely after successful treatment of *H pylori* infection with antimicrobials. However, the available antimicrobial therapies for *H pylori* infection have many shortcomings e. g, side-effects, the need for combination therapy, and limited efficacy, in particular because of the development of antimicrobial resistance. The continuous increase in the prevalence of antimicrobial resistance in *H pylori*, together with the lack of forthcoming novel treatment options, already negatively affects eradication of *H pylori* infection, and is predicted to lead to serious problems for treatment of *H pylori*-associated disorders in the near future.

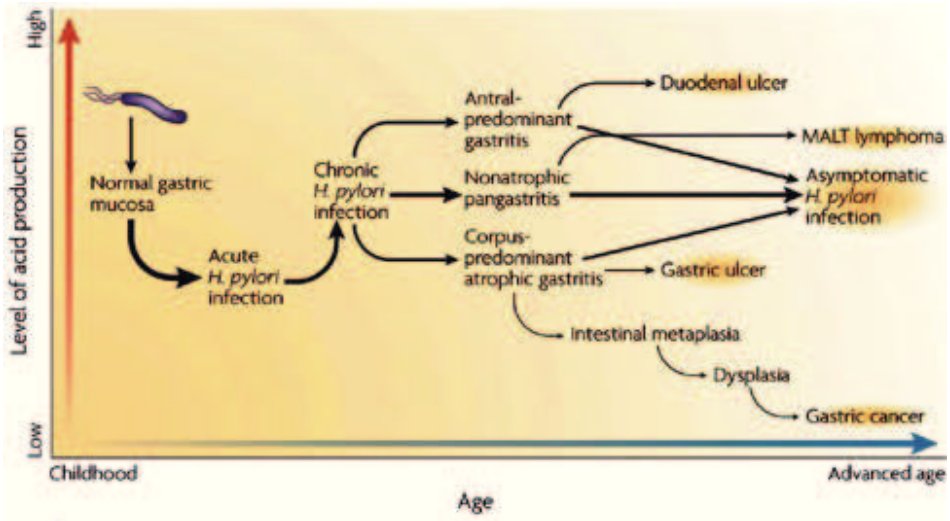

Fig. 1.3: Natural progression of *H. pylori* infection

Acquisition of *H. pylori* usually occurs during childhood. Once acquired and left untreated, the infection persists for life. After the acute phase, most *H. pylori* positive patients develop a chronic gastritis without symptoms. In some patients, more severe manifestations will develop later in life. A normal or high acid secretion predisposes to duodenal ulcers, whereas a low acid secretion predisposes to gastric ulcers and gastric cancer.

The human stomach is a unique ecological niche characterized by very acidic pH a condition lethal for most microbes (Fig. 1.4). *H. pylori* is so well adapted to this unfriendly environment that, after the first infection, which usually occurs early in life, it establishes a life-long chronic infection. The extraordinary success of *H. pylori* in its hostile niche is indicative of very effective adaptation to these conditions. The selection of a niche with no competition and the ability to establish a chronic infection make *H. pylori* one of the most successful human bacterial parasites, which colonizes more than half of the human population. The successful life-lasting colonization of the human stomach by *H. pylori* is achieved through a combination of factors, which address the different challenges presented by the harsh environment. *H. pylori* synthesizes a urease to buffer the pH of its immediate surroundings within the stomach. Its helicoidal shape and the action of flagella allow it to cross the thick layer of mucus lining the stomach.

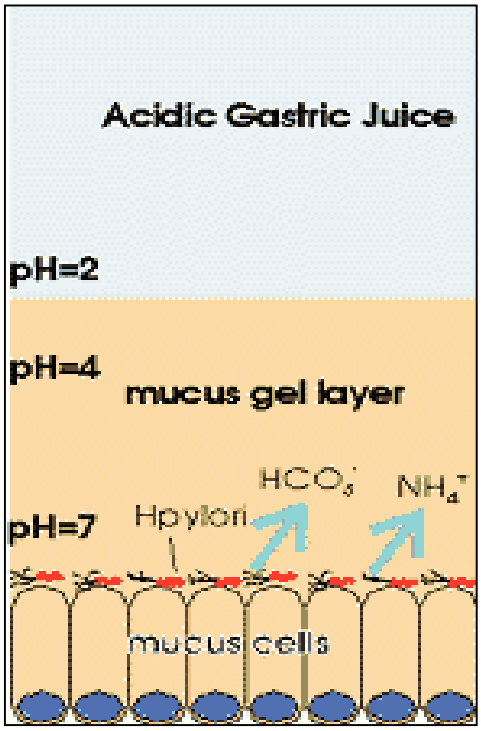

Fig. 1.4: *H. pylori* adhesion to gastric mucosa

H. pylori then binds to Lewis antigens present on host gastric cells, and it secretes factors that attract and stimulate inflammatory cells, as well as the multifunctional toxin VacA. Last, the presence of the *Cag* pathogenicity island, a 40-kb DNA that encodes a type IV secretion system [Blaser M.J., *et al*.1995] seems to be necessary for optimal fitness of the bacterium and the appearance of pathogenic traits. Disease outcome is

complex, because it depends on many factors, including bacterial genotype, host physiology, genotype and dietary habits.

1.2.1 Bacteriology

Helicobacter pylori is a micro-aerophilic, Gram-negative, slow-growing, spiral-shaped and flagellated organism. It is currently classified in the Proteobacteria, a large, diverse division of Gram-negative bacteria which includes, among others, *H. influenzae* and *E. coli*. The bacterium is a member of a rapidly growing genus. New species are being isolated at a fast rate from many vertebrate hosts. Also, other *Helicobacter* species are being isolated from no gastric sites in humans and may be implicated in diseases that previously had no assigned etiologic agent. *H. pylori* is motile via a tuft of polar-sheathed flagella; these structures also carry a terminal bulb, which perhaps makes it more adapted to swimming through mucus. The surface of *H. pylori* is the first point of contact between the pathogen and the host. The biology of this microbe-host interface is sophisticated and complex, and it is made by molecular constituents that account for many of the distinguishing properties of *H. pylori*. Surface proteins of *H. pylori* mediate important pathogen-host interactions that are essential for colonization, adherence, survival, and virulence of this pathogen. Moreover, those proteins, that are particularly accessible to host immune responses and to drug therapy, could be used to control this important human pathogen. On its surface, the lipopolysaccharide (LPS) has unique biological properties and the genes that control addition of the O-side chains can make a phase variation, a mechanism for avoidance of host responses. In addition, it has a unique peptidoglycan structure that differs from other gram-negative bacteria [Appelmelk B.J., *et al.*1996]. Availability of genome sequence of this pathogen has provided a tremendous amount of information that can be useful in a major understanding of virulence mechanism and also in drug target and vaccine target identification. There is unique features of *H. pylori* genomic organisation. A milestone in microbial genomics was set when *H. pylori* became the first bacterial species to have its genome sequenced and compared from two independent isolates. The subsequent comparison has provided the first detailed look at the physical chromosomal organization and has begun to identify a minimal set of common genes that can be used as candidates for therapeutic strategies. There have been 1590 putative open reading frames (ORFs) identified in *H. pylori* 26695 circular chromosome (~1.7 Mbp). Almost 60% of the ORFs have a predicted function, whereas 24% are conserved in other bacterial species but are of unknown function and 17% are *H. pylori* specific with no known homologues in the current databases. These ORFans may encode proteins unique to *H. pylori* and provide selective targets for antibiotic therapy. The sequenced genome revealed a profile of an organism that was fine-tuned for its niche in the gastric mucosa, lacking many of the regulatory features found in the larger *E. coli* genome. (By contrast to the *E. coli* gene sequence, regulatory sequences have been identified in *H. pylori* 3-10 time less). Tight regulation of gene expression is imperative for entero-pathogenic bacteria whether they are continually responding to the harsh acidic environment of the stomach or bile salts in the intestine. The remarkable economy of

regulatory elements in *H. pylori* may reflect the very limited range of environments in the which it survives, suggesting a highly evolved inter-relationship between man and microbe. One intriguing feature of the genomic organization is that ~1% of the genome of *H. pylori* encodes a family of 32 outer-membrane proteins that are well conserved between the J99 and 26695 strains. The OMPs identified provide a tractable subset of the total genome, comprising most of the proteins known to be involved in virulence (for example, those required for adherence to gastric epithelial cells and evasion of the immune system). Of particular interest is the presence of tandem repeat sequences upstream of some OMPs. In other mucosal pathogens increasing or decreasing the numbers of repeats by slipped-strand mispairing and recombination affects transcription of the downstream genes. The extensive use of slipped strand mispairing is a clever form of regulation that allow the organism to present many faces to the host in terms of expression of outer membrane proteins and other surface structures. In this way minor reversible mutations rapidly change the antigen profile of the pathogen, leading to evasion of the host immune system. Numerous restriction-modification systems are present in this species, but they differ between the two genomes analyzed. There are more than twenty homologues associated with DNA restriction and modification systems identified in other bacteria, including type I, type II, and type III systems. The role of these enzymes is unclear at present. It has been suggested that the enzymes are involved in the breakdown of intracellular and/or intercellular DNA or that they are necessary for stimulating the formation of recombinants by DNA fragmentation. The presence of the *cag* pathogenicity island was identified prior to the sequencing of the whole genome, revealing a 40-kb stretch of DNA whose presence correlates with more virulent isolates. Most of the traditional protein secretion systems are encoded in the *H. pylori* genome, including ABC transporters, *sec*-dependent (leader peptide) transport, flagellar assembly (a prototype of the type III secretion system), type IV secretion homologues in the pathogenicity island, and auto-transporters such as VacA [Szabò I., *et al*. 1999]. *H. pylori* populations are extremely diverse at the genomic level. Moreover, a single host can carry several *H. pylori* strains, and isolates within an individual can change over time, as endogenous mutations and/or chromosomal rearrangements or recombination between strains occur. The comparison of *H. pylori* isolated from patients of different ethnical origin and geographical locations indicates that their nucleotide sequences segregate similarly to those of humans. This suggests that this bacterium was already present in the stomach of humans when they left Africa to colonize the world, and co-evolved with them since then. It is possible individuate two groups within the *H. pylori* proteome:

1) Proteins involved in the adaptation of *H. pylori* to the gastric environment:

To survive in the extremely acidic gastric lumen, *H. pylori* utilizes unique mechanisms, the most efficient of which is ammonia production by the potent nickel-containing urease (Fig.1.5). *H. pylori* expresses large amounts of urease, and levels can reach up to 10% of total cellular protein . Urease contains 12 nickel atoms per molecule, and thus *H. pylori* has a relatively high demand for nickel. Urease plays a central role in the pathogenesis of *H. pylori* infection and catalyzes the conversion of urea into carbon dioxide and ammonia. The latter is able to neutralize gastric acid and offer protection to *H. pylori* against the low pH in the stomach. In addition, ammonia may be used as a

nitrogen source supporting growth of *H. pylori*. Urease is regulated by an intricate interplay of different environmental signals such as the concentration of urea and metal ions, or the pH. The activity of urease critically depends on the availability of nickel ions, as a functional urease complex requires 24 Ni^{2+} [Eaton K.A., *et al*. 1991]. *H. pylori* also possesses a NixA nickel specific permease; accessory proteins (UreE and Hyp), required for proper maturation of nickel enzyme; and a histidine and cysteine-rich heat shock protein HspA; Hpn and Hpn-like proteins, which are transcriptionally activated in the presence of nickel by a nickel sensor system. These elements are the major contributors to nickel bioavailability either through sequestering or transporting nickel ion proteins. Their synthesis in cells depends on the Nickel responsive transcriptional regulator NikR, wich together with the Iron dependent regulator Fur was found to be necessary for efficient colonization of the mouse stomach by *H. pylori*. Two other ammonia-producing enzymes, AmiE and AmiF amidases, together with the arginase RocF, which generates urea from arginine, have been implicated in adaptive mechanisms to the low pH environment of the gastric lumen. Finally, the bacterium must deal with reactive oxygen species that are generated by phagocytic cells of the host immune response. In *H. pylori*, genes encoding superoxide dismutase (SOD), catalase, and several putative peroxidases have been identified. *H. pylori* has a potent set of antioxidant defences. Persistent colonization of *H. pylori* in a mouse model appears to require the function of these gene products: *sodB*, *katA* (which encodes for catalase KatA) and *kapA* (KatA-associated protein, that is an important component of the bacterium's resistance to oxidative damage, especially hydrogen peroxide). The expression of *sodB* (which encodes for iron-cofactored SodB protein) is essential for gastric colonization and also required for growth under microaerophilic conditions [Moobley *et al.* 2001].

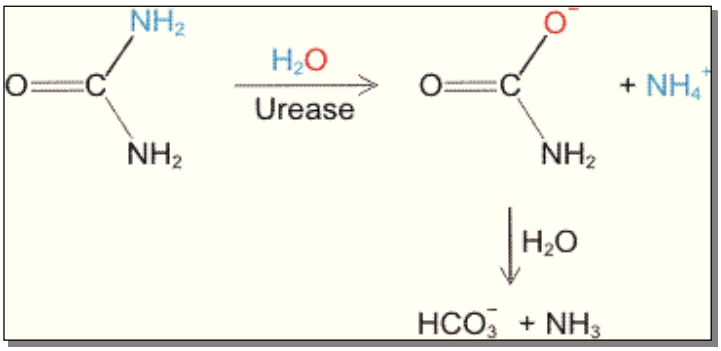

Fig. 1.5: Urease Hidrolysis

2) Main proteins involved in *H. pylori* pathogenicity mechanisms:

Numerous studies have focused on the elucidation of *H. pylori* pathogenicity mechanisms. The *cag*-Pathogenicity Island (PAI), consisting of a group of 31 genes involved in the biogenesis of a type IV secretion system (T4SS) and the translocation of

the immune-dominant CagA antigen, has been shown to contribute to the induction of the pro-inflammatory IL-8 cytokine by host epithelial cells through the activation of the nuclear factor κB (NFκB) pathway, and thus has been recognized as one of the major pro-inflammatory players. T4SSs are ubiquitous secretion machineries formed by at least 12 proteins named VirB1-B11 and VirD4, three of them (VirB4, VirB11, and VirD4) are ATPases believed to power the assembly of the T4SS and drive substrates through it. The vacuolating cytotoxin (VacA) is also one of the most studied virulence factor of *H. pylori*. VacA is thought to interfere with intracellular vesicles trafficking and its role in the persistence of the bacterial infection has been recently proposed [Gebert B., *et al*. 2003]. Adhesion of *H. pylori* to gastric epithelial cells is another important early step in the infection process. One of the most studied adhesins is the blood-group antigen binding adhesin named BabA, encoded by the *bab*A2 gene. Other adhesins have been described such as AlpA-AlpB, SabA or HopZ (Mahdavi *et al*. 2002], [Odenbreit S., *et al*. 2002].

1.2.2 Complete Sequencing of *H. pylori* genome

Prior to the sequencing and annotation of the genomes of *H. pylori* strains 26695 and J99, a large number of studies had elucidated central metabolic pathways, uptake and regulatory systems, responses to various stresses, and virulence factors. Nevertheless, the publication of the genomes has had a marked impact on our knowledge of the bacterium, and the data derived from these sequences have served to confirm experimental results, to provide insights into the biology of the bacterium, to deepen our understanding of its diversity, and to suggest new areas of investigation [Tomb G.C., *et al.* 1997]. *Helicobacter pylori*, strain 26695, has a circular genome of 1,667,867 base pairs and 1,590 predicted coding sequences (Fig. 1.6). Sequence analysis indicates that *H. pylori* has well-developed systems for motility, for scavenging iron, and for DNA restriction and modification. Many putative adhesins, lipoproteins and other outer membrane proteins were identified, underscoring the potential complexity of host–pathogen interaction. Based on the large number of sequence-related genes encoding outer membrane proteins and the presence of homopolymeric tracts and dinucleotide repeats in coding sequences, *H. pylori*, like several other mucosal pathogens, probably uses recombination and slipped-strand mispairing within repeats as mechanisms for antigenic variation and adaptive evolution. Consistent with its restricted niche, *H. pylori* has a few regulatory networks, and a limited metabolic repertoire and biosynthetic capacity. Its survival in acid conditions depends, in part, on its ability to establish a positive inside-membrane potential at low pH. The whole-genome analysis of *H. pylori* gives new insight into its pathogenesis, acid tolerance, antigenic variation and microaerophilic character. The availability of the complete genome sequence will allow further assessment of *H. pylori* genetic diversity. This is an important aspect of *H. pylori* epidemiology as allelic polymorphism within several loci has already been associated with disease outcome. It is predicted that a relatively high percentage of genes involved in central metabolic pathways will be essential for its survival even if their direct equivalents in other organisms are not. The products of these highly diverged essential genes merit further investigation as potential targets for novel, highly specific anti-*H. pylori* agents. The identification of many new putative virulence determinants should allow critical tests of their roles and thus new insight into mechanisms of initial

colonization, persistence of this bacterium during long-term carriage, and the mechanisms by which it promotes various gastroduodenal diseases.

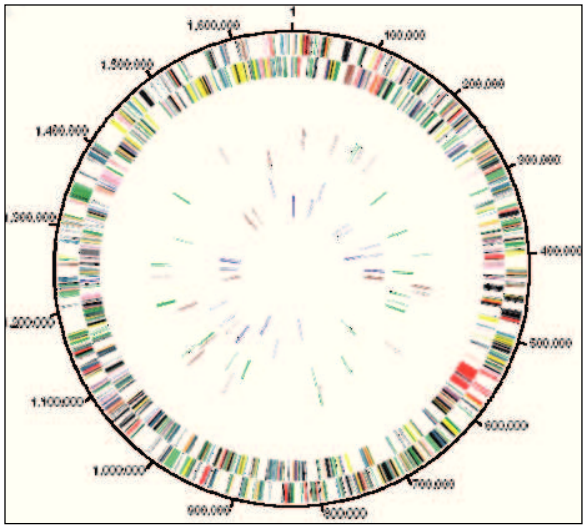

Fig. 1.6: Circular representation of the *H. pylori* 26695 chromosome

1.2.3 Surface Localization of Cytoplasmic Proteins.

The cell surface of *H. pylori* has the unusual property of being able to incorporate proteins such as urease, catalase, HspA, HspB, and superoxide dismutase (SOD), which are found virtually exclusively within the cytoplasm in other bacteria. Cryo-immunolocalization techniques have demonstrated that urease, catalase, and HspB are located strictly within the cytoplasm of freshly subcultured, early log-phase *H. pylori*. However, at the end of the log-phase, these proteins are also surface associated or extracellular. Significant fractions of urease and HspB are also surface associated in vivo. Indirect gold immune-staining of *H. pylori* SOD with a polyclonal antibody directed against the iron-containing SOD of *E. coli* showed a surface localization of the enzyme. Gold particles were distributed on the cell envelope and on the sheath of the flagella. SOD was located on the outer surface of a limited number of bacteria, but the enzyme was cytoplasm in most cells.
In contrast to other *H. pylori* proteins with specific secretion pathways, the mechanism whereby urease, HspA, HspB, catalase, and SOD become associated with the other membrane of *H. pylori* is controversial. Urease, HspA and HspB are large oligomeric proteins that are typically found exclusively in the cytoplasm of bacteria and would not be expected to cross the bacterial outer membrane. *H. pylori* urease, HspA and SOD, and catalase have been shown genetically to lack leader peptides. Alternative transport mechanisms must therefore exist. The ability of released cytoplasmic proteins of *H.*

pylori to bind to the cell surface is considered to be biologically important. Some very elegant work [Phadnis S.H., *et al*.1996], [Vanet A. And Labigne A. 1998] has shown that although *H. pylori* urease, HspA, HspB, and catalase are intrinsic cytoplasmic proteins in log-phase bacteria, these proteins become associated with the outer membrane of *H. pylori* when they are released from the cytoplasm by autolysis of a fraction of the bacteria. They then become adsorbed to the surface of intact bacteria. Although specific export of these proteins cannot be completely excluded, it can be envisaged that the "altruistic autolysis" process aids in protection against environmental stresses and functions in diversion of the immune response of the host (Fig.1.7).

The ability of released cytoplasmic proteins of *H. pylori* to bind to the cell surface is considered to be biologically important. For example, the surface locations of catalase and SOD could enable these enzymes to function more effectively as a defense mechanism against phagocytic attack. In the case of urease, one study indicated that free or extracellular urease is irreversibly inhibited at pH <4.5. The authors concluded that external urease is ineffective as an acid-protective device at the lower pH values in the stomach. However, in another study, bacteria with only cytoplasmic urease showed significantly reduced survival when exposed to acid in the presence of 5 mM urea when compared to the survival of bacteria with both cytoplasmic and extracellular urease. This indicates that cytoplasmic urease activity alone is not enough to enable *H. pylori* to survive in acid. The authors of this study speculate that the association of urease with the outer membrane of *H. pylori* protects urease from inactivation by acid. Development of mutants in which urease is located strictly within the cytoplasm would be helpful to resolve the discrepancies between the two studies. In addition to mediating acid resistance, other roles have been postulated for surface-associated urease that include binding to mucin at acidic pH. The observation that urease, HspA, HspB, and catalase are surface associated helps explain how they can serve as vaccine components in animal trials. However, it is likely that in vivo a subpopulation of *H. pylori* cells do not carry these antigens on the cell surface, and such a subpopulation may therefore evade the protective effect of the vaccine. Characterization of the cell envelope of *H. pylori* has identified a number of important features that distinguish it from other bacterial pathogens. These are the simple structure of its peptidoglycan, its unusual cellular fatty acid and lipid profile, molecular mimicry of Lewis antigens by lipopolysaccharide (LPS) and the presence of cytoplasmic proteins such as urease, Hsps, and SOD on the cell surface.

In addition to these characteristics, polar flagella and a unique repertoire of outer membrane proteins are all likely to contribute to the ability of the organism to colonize the stomach and cause disease. The significance of the outer membrane components, many of which are likely to function as adhesins, is highlighted by the fact that *H. pylori* devotes a significantly high proportion of its coding capacity to them. Many of the vaccine candidates for *H. pylori* are proteins found on the cell surface, underlining the importance of further characterization of these proteins and of elucidating the precise mechanism of interaction between *H. pylori* cell surface and gastric mucosa.

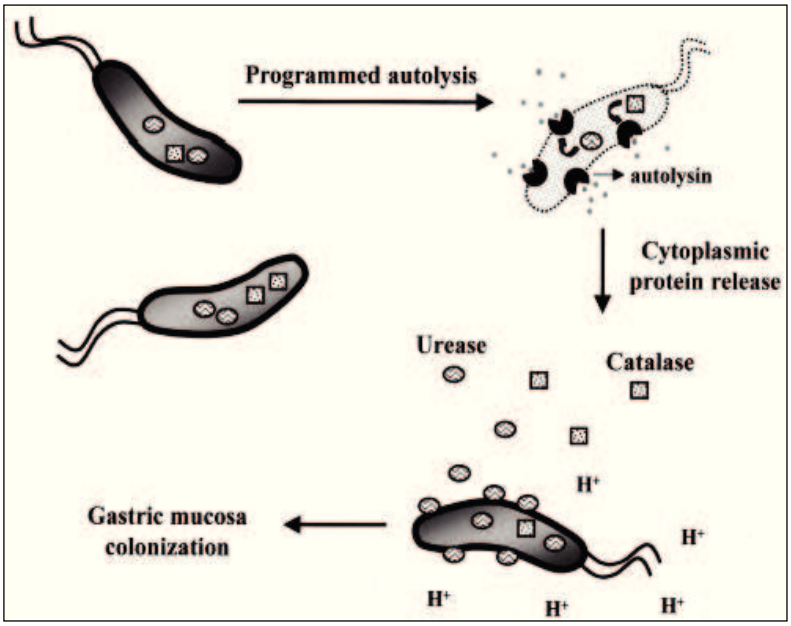

Fig 1.7: Autolysis of *H. pylori:*

Hypotetical model that explains how the bacterium is able to answer against external signals, activating a genetic program of autolysis. This mechanism is responsible for cytoplasmatic protein's releasing, such as urease, catalase and Heat shock proteins (Hsp).

1.3 *H. pylori* Heat shock protein A (HspA)

1.3.1 GroEL and GroES chaperones

Heat shock protein A is a unusual GroES homologue. In general, these proteins are involved in the intracellular folding and assembly of various polypeptide chains into oligomeric complexes [Horwich A.L., *et al*. 2007]. Heat shock proteins (Hsp) of the GroEL/GroES class are a highly conserved group of proteins found in all domains of life. Despite their designation, they are expressed at a low level and play a vital role at all temperatures in maintaining normal cell function. Increased synthesis of these proteins occurs in response to many environmental stresses, including temperature changes (hence their name), oxygen limitation, glucose limitation and iron deprivation for bacteria, but also, as far as eukaryotic cells are concerned, inflammation, irradiation, malignant transformation and T-lymphocyte activation with mitogens or lymphokines. These proteins are involved in the intracellular folding and assembly of various other

polypeptide chains into oligomeric complexes, without becoming part of the mature products.
The bacterial Hsp60 chaperonin (GroEL homologue) act together with the Hsp10 protein (GroES homologue). Both form homo-oligomers of two stacked rings, each ring having a seven-fold symmetry. The present view of the mechanism by which they act suggests that the Hsp60 alone interacts with the unfolded polypeptide, subsequently the Hsp10 binds to the Hsp60-polypeptide complex and plays a role in the release of the polypeptide by modulating the ATPase activity of the Hsp60 moiety. The Hsp10 and Hsp60 proteins are commonly encoded by a single bi-cistronic operon that leads to the coordinated expression of the two genes. The thermo-inducible nature of the heat-shock genes depends upon the bacteria. Two major mechanisms of regulation at the transcriptional level have been identified:

I. In most bacteria the genes are induced by activating transcription from promoters specifically recognized by an RNA polymerase containing the Sigma 32 factor (σ32). Therefore, the level of expression of the *hsp* genes is directly proportional to the amount of the σ32 factor present in the cell, encoded by the *rpoH* gene.

II. A second mechanism involves the presence of a conserved motif forming a hairpin-loop structure, TTAGCACTC-N9-GAGTGCTAA, located between a vegetative promoter sequence and the start site of the structural gene. This motif is known as CIRCE (controlling inverted repeats of chaperone expression).

Interestingly, none of the two regulation systems mentioned above have been identified in the *H. pylori hsp* gene cluster. It seems that the Hsp10 and Hsp60 proteins have vital functions for *H. pylori,* since the attempts to construct *H. pylori* mutants in which either the *hsp60* or the *hsp10* gene has been disrupted have been unsuccessful.

1.3.2 *H. pylori* HspA:

The transition metal nickel plays a central role in the human gastric pathogen *Helicobacter pylori* because it is required for two enzymes indispensable for colonization, the nickel metalloenzyme urease and [NiFe] hydrogenase. To sustain nickel availability for these metalloenzymes while providing protection from the metal's harmful effects, *H. pylori* is equipped with several specific nickel-binding proteins, an assortment of factors in order to adapt itself to the extremely acidic environment of the stomach. Among these, *H. pylori* possesses a particular chaperone, HspA, that is a homolog of the highly conserved and essential bacterial heat shock protein GroES. No other gene encoding a GroES homolog is found in the genome of *H. pylori*. It shows an extended subcellular localization, ranging from cytoplasm to the bacterial cell surface Indeed, *H. pylori* HspA consists of 118 amino-acids divided in two domains: an N-terminal domain (domain A,

residues 1-90), which is homologous with other GroES bacterial proteins, and a C-terminal domain (domain B, residues 91-118), which other GroES-like proteins lack and which contains eight His and four Cys residues among 27 amino acids (Fig. 1.8). This His/Cys-rich C-terminal extension is unique and was demonstrated to bind nickel *in vitro* [Kansau I., *et al.*1996].
Whilst *H. pylori* GroEL-like (Heat-Shock protein B, HspB) was shown to be very similar to other bacterial GroEL-homologues, *H. pylori* HspA appeared to be unique in sequence and structure. The domain B sequence is restricted to *H. pylori* and the closely related *Helicobacter acinonychis* species but is absent from all other available sequenced *Helicobacter* species. This domain also contains an HX_4DH motif that is considered to be a nickel-binding signature sequence in the nickel-cobalt (NiCoT) transporter family. Besides its usual co-chaperone activity, HspA plays two additional roles, being involved in nickel binding and urease activity, it facilitates nickel acquisition by donating it to appropriate proteins in a Ni(II)-deficient environment and carries out detoxification *via* sequestration of nickel excess.
This striking C-terminal domain of HspA was found only in *H. pylori* 26695, *H. pylori* J99 and *H. acinonychis* genomes.

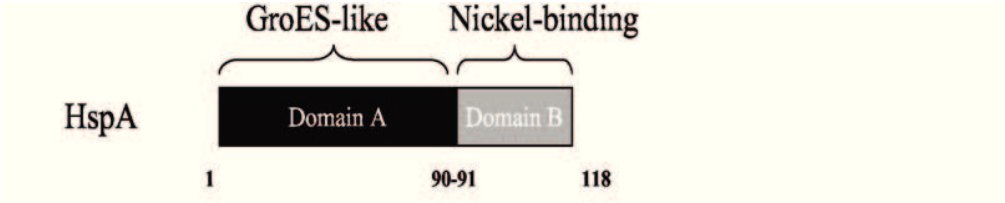

Fig 1.8: Schematic representation of the GroES-like domain A and the nickel-binding domain B of *H. pylori* HspA.

A very unique characteristic of HspA among all GroES proteins is a very high content of cysteine residues, although these amino acids are the less represented residues in all GroES proteins examined so far. Indeed the protein sequence contains 6 cysteines (two in the A domain and four in the B domain) whereas cysteine is the less represented residue in all GroES proteins examined so far. It can be deduced that cysteine is not an essential residue for chaperonins, and hence HspA has to play additional roles mediated by its unusually high number of Cys residues (Fig. 1.9).
The B domain is highly charged and it exhibits a high and specific affinity for nickel ions in comparison with its affinity for other divalent cations (copper, zinc, cobalt). It has two distinct Ni^{2+}-binding sites: a high-affinity site with Kd ~ 2.8 µM and a lower-affinity site, which binds Ni^{2+} at concentrations above 30 µM. The presence of a C-terminal histidine-rich domain associated with nickel binding has already been described in other bacteria, for UreE, a nickel carrier protein that is expressed by most of the urease gene clusters as one of the accessory proteins required for urease metallocentre assembly (i.e.

activation of the apo-enzyme). Curiously, none of the *H. pylori* accessory proteins, including the *H. pylori* UreE homologue, exhibits such a nickel-binding motif. This led to speculate that there might be a link between HspA and the nickel-containing urease enzyme. So far, however, all attempts to visualize a complex consisting of the polypeptides UreA, UreB, HspA and/or HspB have been unsuccessful. Interestingly, evidence for a possible interaction of HspA with the *H. pylori* urease holo-enzyme was provided by a genetic approach. In fact, the co-expression in *E. coli* of the urease gene cluster and of the HspA gene leads to a fourfold increase of urease activity when compared to the expression of the urease gene cluster alone.
A recent study [Cun S., *et al.* 2008] found that HspA binds Bi^{3+} with high affinity via coordination to cysteine residues. Therefore bismuth is a commonly recommended metallodrug for the treatment of *H. pylori* infections. This study showed, that HspA reversibly binds two Ni(II) ions/monomer with moderate binding affinity, whereas two Bi(III) ions bind to the protein incomparably stronger inducing the changes in quaternary structure (from native heptamer to a dimer). The binding is irreversible at physiological pH, which clearly indicates that bismuth may interfere with the biological functions of HspA. It is probably the C-terminal His- and Cys rich domain which is beneficial in nickel resistance but harmful to cell growth when exposed to bismuth. Therefore, it seemed interesting to compare the interactions of Ni(II) and Bi(III) with the C-terminal domain, and also to have a closer look at the interactions of these metals with His- and Cys-rich peptides in general.

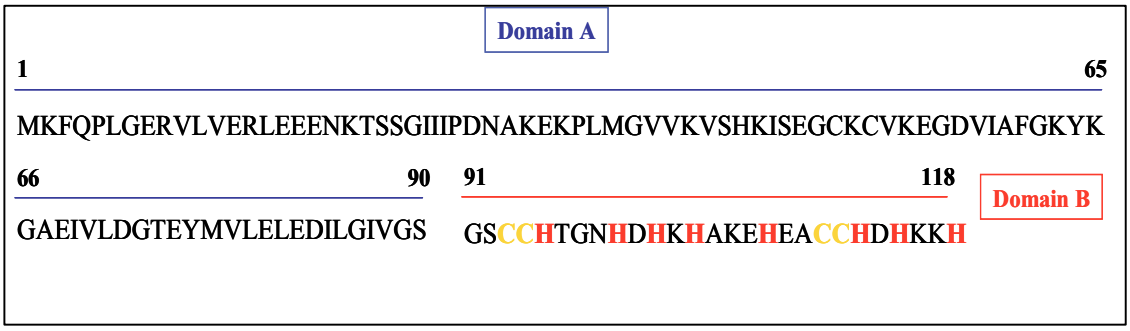

Fig. 1.9: Amino acid sequence of HspA: in the be domain we can observe 8 His residues evidenced in red, and 4 cys residues evidenced in yellow.

HspA exhibits another important feature: it is known to be immunogenic. The immune response to this protein in *H. pylori*-positive adults increases with age. Among many clinical isolates tested, the A domain is highly conserved, whereas the B domain encompasses two variant type sequences that differ from each other by the simultaneous substitution of seven amino acids. Interestingly, none of the substitutions affects the hystidine and cysteine residues. The different domains of HspA elicit distinct host immunological responses, and the A domain was found to be the immune-dominant domain. In 1996, Labigne *et al.* demonstrated that both HspA and HspB confer

protective immunity against mucosal infection in mice. Since then, the development of HspA-based vaccines against *H. pylori* became an active research field. *H. pylori* HspA is particularly appealing as a vaccine component because it has a unique structure, with the C terminal domain which is absent from other known heat shock homologues, including those of eukaryotic organisms. Moreover, the capacity of this domain to bind nickel ions facilitates large-scale purification of the polypeptide. Lastly, this antigen is also expressed by all the isolates and is highly conserved at the amino acid level.
It is worth noting that HspA is the only nickel chaperon which was found to be surface-associated among the several nickel-binding and accessory proteins facilitating Ni-enzyme maturation in *H. Pylori* [Cun S., *et al.* 2008].
Several mechanisms have been proposed for the HspA extracellular release, including specific secretion pathways, autolysis, and membrane vesicle formation, but a conclusive explanation is still missing. Therefore, the protein experience different pH and redox environments when it moves from the bacterial cell to the extracellular space.
In conclusion, HspA plays a critical role as a specialized nickel chaperone . Because nickel ions play a crucial role during *H. pylori* colonization but seem not to be utilized by humans, the nickel metabolism of *H. pylori* is an attractive therapeutic target, particularly in the age of increasing microbial antibiotic resistance. Our study thus provides key insights for the development of alternative antimicrobials.

1.4 Coenzyme A, central to metabolism

Research involving Coenzyme A has always been popular due to its biochemical centrality. Coenzyme A (CoA) is a ubiquitous and essential cofactor in all living organisms [Leonardi R., *et al*. 2005]. The German biochemist Fritz Lipmann discovered CoA in 1945. He was the first to show that a coenzyme was required to facilitate biological acetylation reactions. In 1953, Lipmann was awarded the Nobel Prize in physiology and medicine for his pioneering work in elucidating the role of this important coenzyme [Baddiley J., *et al*. 1953]. It functions as an acyl group carrier and acyl activating group in a number of central metabolic transformations, including the tricarboxylic acid cycle and fatty acid metabolism. It has been estimated that CoA and its thioester are involved in over 100 different reactions in intermediary metabolism of microorganisms. As a result of CoA's ubiquitous nature and its role as a cofactor in metabolism, its levels must be stringently regulated. In addition to this, the CoA biosynthetic pathway is an energetically expansive pathway, so it makes sense that the pathway itself be regulated as not to waste cellular energy. The biosynthetic pathway from pantothenate to CoA is essential in both prokaryotes and eukaryotes. CoA is produced through a series of five enzymatic reactions from pantothenate or vitamin B_5. All the genes coding for the enzymes that catalyze the reactions in the biosynthetic pathway are known.
The biosynthesis of CoA can be divided into two parts in bacteria. First, pantothenate is synthesized, where after the universal biosynthesis of CoA from pantothenate occurs. The second part of the pathway is present in most organisms, and the best characterized CoA pathway is those of *E. coli*. Pantothenate is one of the B complex of vitamins (vitamin B_5) and it plays a fundamental role in all organisms. Animals and some microbes lack the capacity to synthesize pantothenate and are totally dependent on the

uptake of pantothenate in their diets. However, most bacteria, plants and fungi are capable of synthesizing panthotenate. As a result of pantothenate being found vitually everywhere in biology it was designated pantothenate, which is derived from Greek "*pantothen*" meaning "from everywhere" .

1.4.1 The Pantothenate Biosynthetic Pathway in *E. coli*: an overview

The biosynthesis of CoA begins with the decarboxylation of aspartate to give β-alanine. Pantoic acid is formed by the hydroxymethylation of a α-ketoisolverate followed by reduction of ketopantoate (Fig. 1.9). Pantoic acid and β-alanine are then condensed to generate pantothenic acid. Bacteria divert amino acids and intermediates from central metabolism to produce pantothenate [Song W.J*., et al.* 1992]. All the enzymes which are involved in this biosynthetic pathway are also important as research candidates. The CoA and its thioesters production must be tightly controlled to prevent metabolic activity running away. CoA levels are regulated in one of five ways:

- The compartmentalization of CoA
- Feedback regulation by Pantothenate kinase
- Secondary regulation by adenyltransferase
- Rgulation of CoA levels by gene expression
- Regulation by degradation

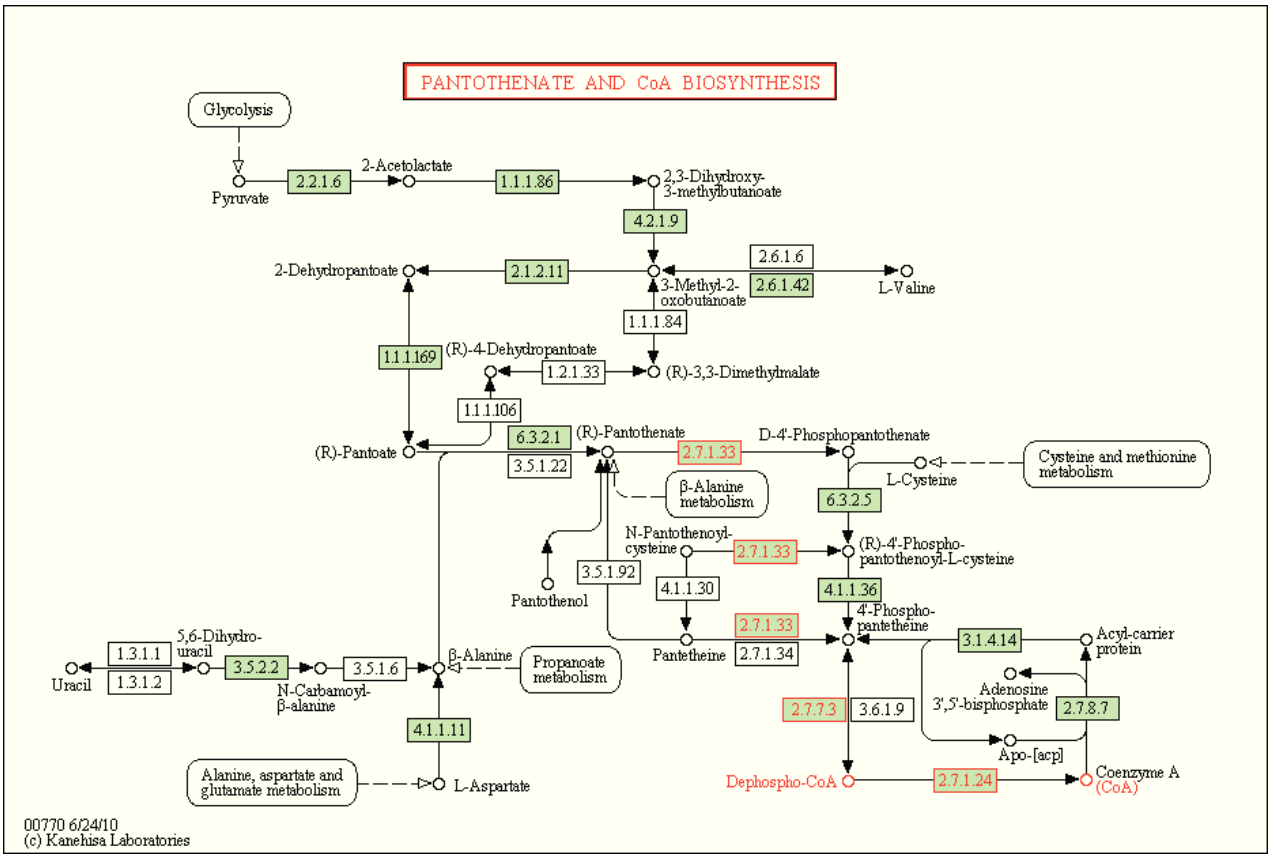

Fig. 1.9: A schematic illustration of the complete biosynthetic pathway of CoA in *E. coli*

1.4.1.1 Phosphorylation of Pantothenate

Pantothenate Kinase, also known as PanK or CoaA, catalyses the ATP-dependent phosphorylation of pantothenate to form phosphopantothenate (Fig. 1.10). This is the first committed step in the biosynthesis of CoA, since none of the phoshorylated intermediates formed in the subsequent reactions can enter the cell. It catalyzes the transfer of the terminal phosphate from ATP to pantothenate, forming phosphopantothenate. Pantothenate kinase is encoded by the *CoaA* gene. Because of CoA's metabolic centrality, the enzyme(s) that regulated its production is of paramount importance. Pantothenate kinase is the key regulatory point in the control of CoA levels in the cell. It is subject to feedback inhibition by CoA itself and to a lesser extent by CoA thioesters.

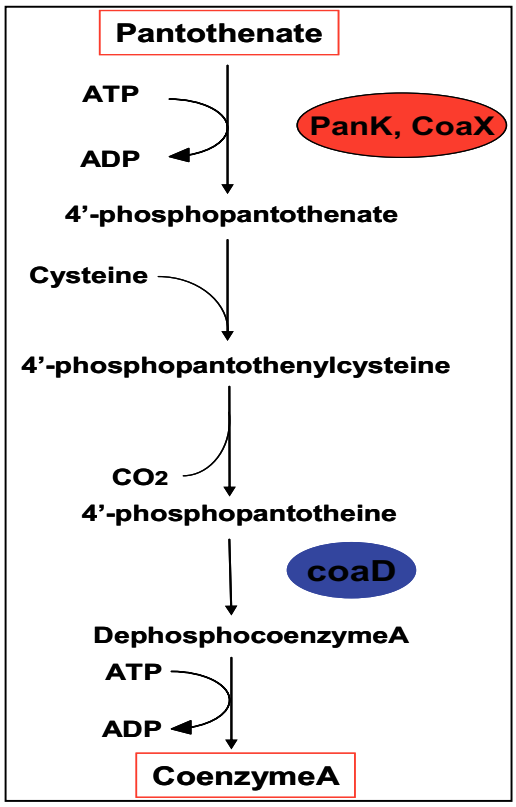

Fig. 1.10: PanK reaction

This is shown by the work done by Song *et al*, who discovered that a 76-fold overexpression of PanK only resulted in a 2.7 fold increase in the cellular concentration of CoA. This feedback inhibition of Pank by the different CoA molecular species controls the overall CoA availability in response to the cell's metabolic status.
The crystal structure of *E. coli* PanK in complex with either ATP or CoA has been determinate. Based on this structural data, Rock *et al*. set about designing three site-direct mutants of PanK that were predicted to be resistant to feedback inhibition by CoA, based on decreased binding efficiencies of this inhibitor. These mutants were shown to retain significant activity and be refractory to inhibition of CoA. In particular, the presence of the Arg106 which is replaced in one of mutants (this residue does not have a role in a catalysis reaction) is postulated to be an important and specific requirement for CoA binding since it forms a salt bridge with the phosphate attached to the 3-hydroxyl of the CoA ribose. The authors show that the mutants that are refractory to feedback inhibition accumulate intracellular phosphorylated pantothenate-derived metabolites, thus translating into a higher CoA content. This data confirms that the feedback inhibition is operating in vivo to limit the amount of CoA being produced. The cloning of the first eukaryotic PanK from fungus revealed a sequence completely different from *E. coli* PanK. However, this protein from *Aspergillus nidulans* is homologous to several proteins encoded by mammalian genes. This discovery (that the

bacterial PanKs were different from mammalian PanKs) led to the prediction that the pantothenate kinase step was a prime target for the identification and design of novel antibacterial drugs. The development of new antibiotics designed as inhibitors of CoA, utilizing enzymes (such as CoA antimetabolites) would be very beneficial.

1.4.2 Pantothenate kinases, different types

Currently two types of the enzyme have been characterized: the first (Type I) is found predominantly in prokaryotic organisms and is exemplified by the *Escherichia coli* PanK enzyme. The second (Type II) occurs mainly in eukaryotic systems, of which the murine enzyme has been the best characterized. These enzymes exhibit wide variations in their distribution, mechanisms of regulation and affinity for substrates. The Type I and II show very little sequence similarity and are predicted to be structurally distinct. Despite this, they share a common regulation mechanism based on feedback inhibition by CoA and its thioesters, although the degree of inhibition is system and inhibitor dependent. This feedback mechanism is primarily responsible for controlling the intracellular CoA concentration.

1.4.2.1 Type I Pantothenate kinase

The pantothenate kinase enzyme was first identified in *S. typhimurium* and *E. coli* and thereafter in numerous other bacteria by comparative genomics. Type I PanK from E. coli (EcPanK) is a homodimeric enzyme; it has been extensively structurally characterized through X-ray analysis of its complexes with CoA, with an analogue of ATP, and with ADP and pantothenate It is considered the prototypical bacterial PanK. A survey conducted by Cheek *et al.* has characterized all known kinases into specific groups and families based on their three dimensional structure. The PanK from *E. coli* is structurally distinct from its eukaryotic counterpart. Based on the crystal structure of the Type I PanK the authors place this type into the Rossmann-like fold group (group 2). Within this group they are classified into the P-loop kinase family. The P-loop family constitutes the largest family in the Rossmann-like fold group. Although the *E. coli* PanK is considered the model bacterial Pantothenate kinase, this type is not universally expressed in bacteria. For example, the eubacteria *Pseudomonas aeruginosa* and *Helicobacter pylori* do not have recognizable pantothenate kinases in their genomes, although all other components of the biosynthetic pathway are present.

1.4.2.2 Type II Pantothenate kinase

Mammalian Pantothenate kinase belongs to the group of Type II Pantothenate kinase. Of the mammalian PanKs, perhaps the best characterized is the murine PanK. The murine Patothenate kinase (*PanK1*) gene is located on Chromosome 19: two biochemical distinct isoforms named *PanK1α* and *PanK1β* are encoded by this PanK1

gene. The two isoforms differ with regards to their regulatory properties. The isoforms are differentially expressed in mouse tissue as a result of metabolic status of the tissue [Rock C.O., *et al*. 2002].
The *Asperigillus nidulans* PanK, also characterized as a Type II PanK, is sensitive to inhibition by acetyl-CoA but not effect, positively or negatively, by CoA. It is clear that the characteristics of the eukaryotic (Type II) PanK are in sharp contrast to the prokaryotic (Type I) PanK in terms of inhibition, with Type I PanKs being inhibited predominantly by CoA and to a lesser extent acetyl-CoA, and the Type II PanKs being inhibited more potently by acetyl-CoA. Unlike the Type I, which belong to the P-loop family of kinases, the Type II PanKs are predicted to belong to the Ribonuclease H-like kinase group. The ribonuclease H-like group belongs to the ASKHA (acetate and sugar kinase/hsc70/actin) superfamily. These structures are characterized by the duplicate domains of the Ribonucleasi H-like fold. Looking at the PanK from *Staphilococcus aureus* it is evident that this is an atypical Type II PanK. Even though *S. aureus* is a bacterium, its PanK sequence is distinctly different from that of E. coli. In fact, it is more closely related to the mammalian PanK than to those of its fellow bacteria. *S. aureus'* s PanK sequence shares 18% identity with the murine PanK isoform 1β and only 13% identity with *E. coli* PanK. The native molecular mass of *S. aureus* PanK is 59 kDa, which is consistent with the existence of a homodimer. All organisms characterized to date share a common mechanism to regulate CoA biosynthesis, feedback inhibition by CoA and/or its thioesters. The key difference between *S. aureus* PanK and other pantothenate kinases is that it is refractory to feedback inhibition by CoA and its thioesters. *S. aureus* produces CoA in proportion to the input of β-alanine with no indication of regulation at PanK or in downstream reactions. This observation may be understood in the light of *S. aureus* physiology. *S. aureus* does not contain glutathione: so CoA is the primary intracellular thiol and together with a unique CoA disulphide reductase functions as the reducing system. Because of this the concentration of CoA in *S. aureus* can reach millimolar levels, and the lack of CoA inhibition regulation in this organism means that the upper limit of CoA is set by the availability of the starting materials of CoA biosynthesis, namely pantothenate and cysteine.

1.4.2.3 Type III Pantothenate kinase

As discussed, at least two types of PanK exist in nature. These types show very little sequence similarity and are predicted to be structurally distinct. However, sequence analysis of bacterial genomes, using these two types as template, fails to locate a homologue of either analogue in some species. Curiously the genomes of certain pathogenic bacteria, do not contain a pantothenate kinase similar to either analogue, although these organisms possess all the other biosynthetic machinery required for CoA production. Since PanK activity is required for the biosynthesis of CoA in all living systems such as a gene must exist in these species. The group species lacking type of the currently identified PanKs can be separated into two subgroups: those belonging to the family of archeabacteria, and those that include well-known pathogenic eubacteria such as *Bordetella pertussis, Helicobacter pylori* and *Pseudomans aeruginosa*; whereas *Bacillus subtilis* and *Mycobacterium tuberculosis* have both type I and type III enzymes.

Since these organisms posses sequences that are homologues to all the other CoA biosynthetic genes, it is highly likely that at least one other analogue PanK exists that represents this activity in this bacteria. This type would have little or no sequence similarity to the well-characterized types. Through a studies based on genome sequence of *Bacillus subtilis* it was possible identify a new gene named *CoaX*, which encodes a protein that exhibits pantothenate kinase activity. Homology searches based on the *CoaX* gene sequence identified orthologues in several bacterial genomes, including those eubacteria that do not posses another PanK analogue. In spite of current knowledge of CoA biosynthesis in general, and PanK enzymes in particular, the identification of this activity remains elusive in a subset of pathogenic bacteria, including *H. pylori* [Brand L.A., *et al*.2005]. This fact was highlighted in two recent studies, both of which used a comparative genomics approach to reconstruct the universal biosynthetic pathway in representative organisms of all kingdoms. Among these, *H. pylori* and *P. aeruginosa* are examples of bacteria in which no putative PanK similar to either known analogue could be found, even though the four remaining CoA biosynthetic enzymes were clearly represented. Since PanK is an essential activity in these organisms, this suggests that at the least one additional, uncharacterized PanK exists. In comparison to the Type I and Type II PanKs, these CoaX enzymes exhibit distinctly different characteristics, suggesting that they are the first characterized examples of Type III PanK analogue. The CoaX enzymes exhibit a very low specificity for ATP. This is due to their surprisingly high k_M values, which is nearly 10 mM in the case of *HpCoaX*. These findings have important implications for the identification of new drug targets in pathogenic bacteria, as well as for the understanding of the convergent evolution of key metabolic enzymes.

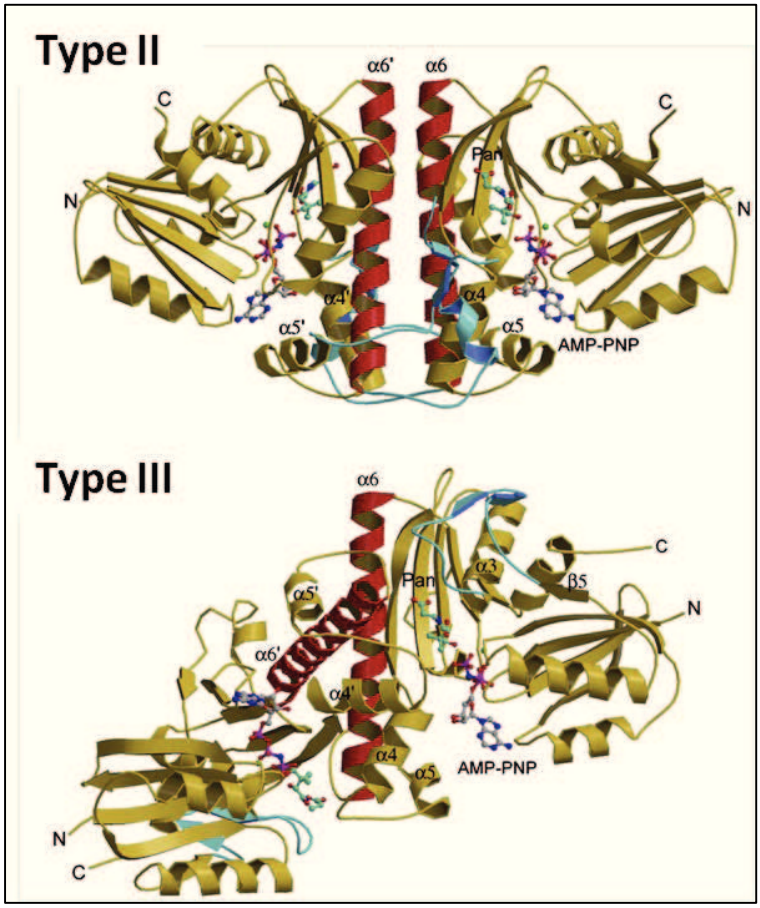

Figura 1.11: The monomer consists of two "duplicated" domains that are considered to be a result of gene duplication. The same monomer creates different dimer interfaces in type II and III PanKs. This provides a structural explanation for the lack of inhibition by CoA and inability to bind Pan analogues observed in type III PanKs.

1.4.2.4 Significance of the discovery of a third PanK analogue

The discovery of a third type of pantothenate kinase gives a significant contribution to our understanding of how identical enzyme activities arise through convergent evolution. Additionally, since genomic analysis studies suggest that the CoaX homologues account for the only PanK activity found in pathogenic bacteria like *H. pylori*, the immediate implications for the development of specific and directed antibacterial agents are evident.

1.4.2.5 Inhibition of Type III PanKs by CoA and acetyl-CoA

With the exception of the recently characterized *S. aureus* PanK enzyme, all other known PanKs are inhibited by CoA and/or its thio-esters. The extent of inhibition is system dependent: prokaryotic type I PanKs exhibit greater inhibition by free CoA than its thioesters, while the activity of eukaryotic Type II PanKs is effected more by the CoA thio-esters acetyl- and malonyl-CoA. In both cases the inhibition serves to regulate the intracellular CoA concentration by affecting the flux through the pathway. Differently neither of the CoaX enzymes is affected. *B. subtilis* PanK (Bs-PanK) and *H. pylori* PanK (Hp-PanK) exhibit unique kinetic characteristics, showing low specificity constants for ATP and no regulation by CoA and its thioesters. This is one of the main differences between them and the other PanK type classified to date. The only other PanK that shows a similar lack of inhibition is that of the recently classified *S. aureus*. They do not accept the pantotenic acid alkylopantenamides as substrates, nor are inhibited by them, which also distinguishes them from the atypical *S. aureus* enzyme. These properties all distinguish them from the other PanK type, and strongly suggest that they are the first characterized examples of the third analogue of PanK enzymes. It is still unclear what advantage, if any, these enzymes confer upon the mainly pathogenic organisms that harbour them.

1.4.2.6 Structural analysis of Kinase

Kinases are a group of enzymes that catalyzes the phosphoryl transfer reaction from a phosphor donor (usually ATP) to a receptor substrate. These were classified into structure/sequence families with evolutionary implications. Possible biochemical roles are suggested by the function of these protein's homologues. A total of 30 kinase families were classified. Of these kinase families, 19 contained at least one member with solved structure. These 19 families could be assembled into seven fold groups on the basis of similarity of structural fold. Families in the same fold group share structurally similar nucleotide-binding domains that are of the same architecture and topography for at least the core of domain. Comparative analysis of the protein structural folds allows for the inference of biochemical and biological functional properties among kinases [Cheek S., *et al*. 2002]. In the same way that the Type II PanK enzymes were assigned to the ribonuclease H-like family, so too were the Type III PanKs. This was based on the lack of sequence similarity between the Type I and Type III PanKs as well as their dissimilar predicted secondary structure. The absence of solved structure of either the Type II or CoaX-like (Type III) Pantothenate kinase does not allow to identify key residues through sequence alignments with protein of known structure. Several mutational studies carried out by Brand *et al*. indicate that the placement of the Type III PanKs into the ribonuclease H-like fold is justified. By multiple sequence alignment it was possible to observe the conservation motif of the ribonuclease H-like families between the CoaX and Type II proteins. This multiple sequence alignment helped to identify a number of conserved amino acids within the conserved motifs. Specifically there were identified key aspartate residues that are putative involved in the binding of Mg^{2+} and a stabilization of the phosphate group of ATP during the course of the PanK

reaction, and are based on the known interactions of other ribonuclease H-like proteins. A number of aspartate residues are highly conserved in many of sequence compared. Aspartate residues corresponding to *H. pylori* amino acid residues number 17, 87 and 102 are highly conserved and possibly involved in catalysis. By comparison within the ribonuclease H-like family, it is possible identify the presence of conserved aspartate residues within the active site of *Hp-PanK*. These residues are critical for the complete functionality of the protein. Mutation to glutamate allows residual activity to remain whereas the same mutation to asparagine causes the activity of the protein to decrease further. Increasing the substrate concentration does not affect the activity of the mutant proteins.

1.4.3 Inhibitors of Type III PanKs

The research is valuable insofar as the synthesis of new inhibitors which would be like a competitive substrates for the PanK enzymes. As of yet, no inhibitors have been found for Type III pantothenate kinases. As many of members belonging to this group are pathogenic bacteria and have no alternative pantothenate kinase enzymes in their genomes, finding an inhibitor that specifically targets this group and not other type of PanK would be invaluable to the formulation of new drugs against these pathogenic bacteria.
Microbes are becoming increasingly resistant to antibiotics. As bacteria evolve new means of defying the effects of antibiotics, so it needs the developing of new inhibitors. The increasing of bacterial resistance to commonly used antibiotics is of growing concern to the healthcare industry. The development and design of new inhibitors would go a long way in developing a means to overcome this dilemma. The ideal inhibitor must be able to cross the cell membrane and be inhibitory to the pathogenic *CoaX* proteins only. This would truly be a breakthrough in so far as the design of novel inhibitors for pathogenic bacteria, such as *Helicobacter pylori*. New compounds have just been designed against the PanK from *S. Aureus* [Choudhry A.E., et al. 2003]. Their role is the blocking of the CoA pathway, compromising the energetic pathogen metabolism. Comparison of substrate binding and catalytic sites of PanK-III with that of PanK-II and PanK I, revealed drastic differences in the binding modes for both ATP and pantothenate substrates, and suggests that these differences may be exploited in the development of new inhibitors specifically targeting PanK-III isoforms.

1.4.4 The role of *CoaX* proteins

The role of *CoaX* in organism that possess another PanK analogue is enigmatic. Does it serve as an additional, unregulated form of the enzyme, which is only active under certain conditions? Or is it a latent enzyme activity, which has only been adapted to PanK activity at a later evolutionary stage? Does the CoaX form of the PanK possess unique characteristics that set it apart from the other two types of PanK and in so doing:

make it ideally suited for adoption by mainly pathogenic organism for the synthesis of Coenzyme A?

1.4.5 Aim of the Project

This project was carried out by the principal expression, purification and characterization methodologies of proteins. In this context we have produced and characterized a recombinant HspA. In particular the study has been addressed on the Cys oxidised/reduced state, which are present in the B-domain. To understand how the target protein, HspA, can be involved in the redox environment modifications of the pathogen *H. pylori*, we have studied specifically the distribution pattern of disulphide bridges. So we have also designed and studied two particular mutants: a single mutant Cys94Ala and a double mutant Cys94Ala-Cys111Ala, to investigate on the correct disulphide bridge pattern. The disulphide bridge pattern has been assigned by integrating classical biochemical methodologies with mass spectrometry. These results can be related to the different redox environments that the protein experience inside and outside the bacterial cell. Another aim of this project has been the biochemical characterization of the Type III PanK from *H. pylori*; this new type III class posses very singular properties and is involved in a critical pathway for pathogen survival: that of coenzyme A (CoA) biosynthesis. This enzyme do not share sequence homology with any known PanK, and unlike the bacterial and eukaryotic PanK isoforms their activity is not regulated by either CoA or acetyl-CoA. They also do not accept the pantothenic acid antimetabolite N-pentylpantothenamide as a substrate or are inhibited by it.
In our study we have cloned the *HPCoaX* gene and expressed the recombinant protein. We have optimized the experimental conditions for purification and stability of the expressed protein. The purified protein was analyzed by gel filtration chromatography and dynamic light scattering to investigate its state of oligomerization.

Chapter II

Materials and Methods

2.1 Strains, enzymes, reagents and instruments

Reagents used for preparation of buffers and growth media of *Escherichia coli* and the reagents for polyacrylamide gel electrophoresis (Acrylamide, APS, TEMED, SDS, Tris-Glycine) were supplied by Sigma Aldrich, Euroclone, Applichem and ICN Biomedicals. The molecular weight markers for proteins were from Sigma Aldrich. The restriction enzymes and the modification enzymes (calf intestine phosphatase and T4 DNA ligase) were supplied by New England Biolabs (NEB) and the molecular weight markers for nucleic acids were from NEB and Roche. The *Pfu Turbo* polymerase (2.5 U/μL) was supplied by Stratagene, while *Taq* DNA polymerase (5 U/μL) was from NEB; gene constructs cloned in pILL948 were kindly provided by Terradot Laurent (ESRF, Grenoble Cedex, France), while pROAEXTc/*HpCoaX* vector was constructed in our laboratory. *E. coli* TOPF'10 strain, used for cloning, *E. coli* BL21(DE3), *E. coli* BL21(DE3)pLysS, BL21(DE3)STAR competent cells used for screening and overexpression and also TEV protease used for recombinant protein cleavage were housemade. Complete Protease Inhibitor Cocktail were supplied by Sigma-Aldrich and used as a mixture of protease inhibitors, according to manufacturer's instruction. Ethanol, methanol, isopropilic alcohol and acetic acid were supplied by J.T. Baker. The protein samples identity were assessed by liquid chromatography mass spectrometry (LC-MS) performed on Phenomenex Jupiter C4 and C18 columns, connected to a Thermo Electron Corporation mass spectrometer, equipped with an ESI source. Light scattering studies were performed on a miniDAWN™ TREOS triple-angle light scattering detector, and a Shodex RI-101 refractive index detector, supplied by Wyatt Technology Corporation. All size exclusion and affinity chromatography columns were connected to AKTA Purifier and FPLC systems (GE-Healthcare).

2.2 Antibiotics

Ampicillin (supplied by Sigma Aldrich as ampicillin sodium salt) and kanamycin (supplied by Sigma Aldrich as kanamycin sulphate) were solubilized in deionized water at a concentration of 1000X, filter sterilized and stored at –20 °C until use. Ampicillin (Amp) was used at a concentration of 100 μg/mL and kanamycin (Kan) was used at a concentration of 50 μg/mL in both solid and liquid media. Chloramphenicol (CAM) and tetracycline were solubilized in ethanol 95% and used at a concentration of 33 μg/mL and 12.5 μg/mL, respectively.

2.3 *E. coli* cells transformation techniques

2.3.1 Preparation of *E. coli* TOPF'10 cells and transformation by electroporation

About 2.5 mL of an overnight culture of *E. coli* TOPF'10 cells were inoculated into 250 mL of LB medium. The cells were grown up to mid-log phase (OD_{600} ~ 0.6) at 37 °C, stored on ice for 30 min and then harvested by centrifugation (6000 rpm, 10 min, 4 °C). The pellet was washed in 250 mL of sterile water. After the second centrifugation the cells were washed in 125 mL of sterile water. The third washing was performed in 5 mL 10% glycerol; the cells were then harvested by centrifugation and the pellet was resuspended in 750 µL of 10% glycerol. Aliquots of 1010 cells/mL (40 µL) were mixed with 1 µL of the DNA ligase reaction, incubated for 1 min on ice and transferred into chilled plastic cuvettes with an electrode gap of 0.2 cm (M-Medical). High voltage electroporation (25 µF) was performed with a Bio-Rad Gene Pulser XcellTM at a field strength of 2.5 kV/cm and 200 Ohm. A shock pulse was applied to competent cells producing pulse length of ~5.0-5.5 ms. Immediately after electroporation cell mixtures were diluted to 1 mL with LB medium and incubated for ~1 h at 37 °C under shaking. The cells were then plated onto selective solid medium supplemented with appropriate antibiotic to isolate the recombinant clones. Single clones were inoculated in 5 mL LB medium with the same antibiotics and grown over night at 37 °C under shaking. Finally, were harvested by centrifugation (13000 rpm, 2 min, 4 °C) and processed to extract plasmidic DNA by using the *QIAprep Spin Miniprep Kit* or the *QIAGEN plasmid Maxi Kit* (both supplied by Qiagen).

2.3.2 Preparation of *E. coli* competent cells and transformation by heat shock

Single clones of *E. coli* strains, grown at 37 °C in LB agar, were inoculated into 2.5 mL of LB medium and incubated overnight at 37 °C on a shaker. The cells were inoculated into 250 mL of LB medium and the culture was grown up to mid-log phase (0.6 OD600nm) at 37 °C, stored on ice for 30 min and then harvested by centrifugation (6000 rpm, 10 min, 4 °C). The pellet was washed in 125 mL of cold 50 mM $CaCl_2$ and stored on ice for 30 min. Successively, the cells were harvested by centrifugation and the pellet resuspended in 16 mL of cold 50 mM $CaCl_2$. Aliquots of 200 µL of competent cells were mixed with 50 ng of plasmidic DNA target and stored on ice for 30 min. The cells mixtures were transferred at 42 °C for 2 min, on ice for 2 min (heat shock) and then diluted to 1 mL with LB medium. An incubation of 1 h at 37 °C under shaking was performed before plating the cells onto selective solid medium supplemented with the opportune antibiotics, depending on the strain.

2.4 Expression in *E. coli* of *HspA wild type*, *Cys94Ala and Cys94Ala-Cys111Ala*, and *HpCoax*

2.4.1 Cloning of different constructs

Genes were amplified by PCR from genomic DNA. The amplification was performed by using the specific couples of primers. All amplification reactions were performed in a final volume of 50 µL, using 50 ng of template DNA. The reaction mixture contained the specific primers (0.25 µM each), dNTPs (0.25 mM each), 10x amplification buffer, $MgSO_4$ (25mM), and the *Pfu turbo* polymerase (5U) with its buffer. PCR reaction was performed using an *Eppendorf Mastercycler personal* apparatus, following the procedure indicated below:

- Initial denaturation (step 1) 3 min at 95 °C
- Denaturation (step 2) 1 min at 95 °C
- Annealing (step 3) 1 min at the convenient temperature for each gene amplified
- Elongation (step 4) 1 min at 72 °C for 30 cycles, from step 2
- Elongation (step 5) 10 min at 72 °C

All amplification products were analyzed by 0.8% agarose (Euroclone) gel electrophoresis performed in TAE buffer (18.6 g/L EDTA, 242 g/L Tris base. Add Acetic acid until pH 7.8). PCR products were purified by using the *QIAquick PCR Purification Kit* (Qiagen), and digested with restriction enzymes. Each amplified fragment (1µg) was digested with 4 U of restriction enzymes for 3 h at 37 ºC in a buffer containing 50 mM NaCl, 10 mM Tris-HCl, 10 mM $MgCl_2$, 1 mM DTT pH 7.9 supplemented with BSA 100 µg/mL. Following the digestion, each fragment was cloned into the corresponding sites of the pETM11 expression vector (for HspA) and pPROEX-HTc expression vector (for HpCoaX) downstream to the His-tag sequence. To this purpose, the expression vector was previously digested with the same restriction enzymes (4 U/µg), and treated with calf intestine phosphatase (CIP, 10 U) (NEB) for 30 min at 37 ºC. CIP enzyme (10 U/µL) was then inactivated at 75 °C for 10 min. After digestion, PCR amplifications were purified by *QIAquick PCR Purification Kit*, while the digested pETM11 was purified by *QIAquick Gel Extraction Kit* (Qiagen). In order to determine the concentration of the digested plasmid and PCR products, 5µl of each were run on a gel. They were run against the ladders containing bands of known concentration. By comparing the intensity of the DNA sample with that of the ladder it is possible to determine the concentration of the DNA sample in ng/µl. This value is converted to fmoles/µl for the purpose of the ligation reaction. The concentration of the digested plasmid is not always equal to that of the digested PCR product. For ligation reactions was used a 1:3 molar ratio (vector/insert DNA). The reactions were performed using the Quick Ligation™ Kit. This kit ensures that the entire ligation protocol is completed in 5 minutes and is highly reliable: 20 U/µg DNA of the T4 DNA Ligase (400 U/µL), in a final volume of 10 µL, for 3 h at RT. The products of the ligation reaction were transformed into *E. coli* TOPF'10 strain electrocompetent cells. The transformation procedure is carried out by

electroporation. Thereafter 1mL of LB is added to the transformation mixture and it is incubated in a shaking incubator (200 rpm) for 1h at 37°C. The cell pellet is collected by centrifugation at 8000 rpm for 5 minutes. Using a pipette 800 µl of the supernatant is removed (being careful not to disturb the cell pellet) and the cells are then re-suspended in the remaining media. This is plated onto LB agar plate containing the relevant antibiotic. The cells are spread evenly on the plate by means of sterile glass pasteur. The plate is incubated at 37°C overnight. A screening is conducted to check for the insertion of the gene of interest into the plasmid. Firstly, a screening agarose gel is run. From the results of the screening gel, colonies are chosen and plasmids prepared from these colonies. Then, these plasmids are digested with the original restriction enzymes to check for an insert of the correct size. So, using a sterile toothpick, a single colony is scraped from the LB plate and re-suspended in a lysis solution (30 mM Tris pH 8.0, 5 mM Na_2EDTA, 50 mM NaCl, 20% sucrose, 0.05 mg/ml lysozime and 0.05 mg/ml RNAse). The process is repeated for each colony. The tubes are vortexed briefly followed by addition of 10µl of a second solution containing 1x TAE, 2% SDS, 5% sucrose and 2 mg/ml bromophenol blue. Then 10µl of the mixture from each tube is loaded on a 1% agarose gel. We can loaded an undigested pETM11 as a standard. The gel run at about 80V. Positive screens are those that are bigger than the plasmid standard loaded as a marker. A representative of the positive screens is used to inoculate 5 ml of media with the relevant antibiotic and grown in a shaking incubator at 37°C overnight. The above overnight culture is used to purify a plasmid and to make a Mini preparation of DNA. This is done with the help of the *Quiagen Miniprep Kit*. The purified plasmid is digested with the original restriction enzyme used to cut it before ligation reaction. The digestion products are run on a 1% agarose gel against the ladders. Positives clones are those that produce inserts the size of the gene of interest. The identity of the inserts in the resulting recombinant plasmids was confirmed by DNA automated sequencing (suppling by Sbm). On confirmation the clones are stocked at -20°C. After, positive plasmid is transformed into a appropriate *E. coli* strain cells for expression. (The best *E. coli* strain for each construct is chosen after an accurate screening of expression). For the cloning of *HspA wt* , *HspA single mutant Cys94Ala* and *HspA double mutant Cys94Ala-Cys111Ala* we have used pETM11 expression vector, while for *HpCoaX* we have used pPROEX-HTc expression vector.

2.4.2 Expression vector properties

Expression systems are designed to produce many copies of a desired protein within a host cell. In order to accomplish this, an expression vector is inserted into a host cell. This vector contains all of the genetic coding necessary to produce the protein, including a promoter appropriate to the host cell, a sequence which terminates transcription, and a sequence which codes for ribosome binding. One expression system was developed in 1986 by W. F. Studier and B. A. Moffatt, who created an RNA polymerase expression system which was highly selective for bacteriophage T7 RNA polymerase. The initial system involved two different methods of maintaining T7 RNA polymerase into the cell - in one method, a lambda bacteriophage was used to insert the gene which codes for T7 RNA polymerase, and in the other, the gene for T7 RNA polymerase was inserted into the host chromosome (Studier et al, 1986). This expression system has become known

as the pET Expression System, and is now widely used because of its ability to mass-produce proteins, the specificity involved in the T7 promoter which only binds T7 RNA polymerase, and also the design of the system which allows for the easy manipulation of how much of the desired protein is expressed and when that expression occurs. The pET System is the most powerful system yet developed for the cloning and expression of recombinant proteins in *E. coli*. The pET plasmids were chosen as they represent a powerful host-vector system for cloning and expression of recombinant proteins in *E. coli*. All of them contain a strong promoter from T7 bacteriophage, for chemical induction and overexpression of recombinant proteins, a multiple cloning site (MCS) and a TEV protease recognition site for cleaving the fusion protein (Fig. 2.1). Control of the pET expression system is accomplished through the *lac* promoter and operator. Before YFG can be transcribed, T7 polymerase must be present. The gene on the host cell chromosome usually has an inducible promoter which is activated by IPTG. This molecule, IPTG, displaces the repressor from the *lac* operator. Expression is induced by the addition of IPTG or lactose to the bacterial culture or using an autoinduction medium [Elvin C.S. *et al*. 1990], [Terpe, K. 2006].
In particular, pETM11 contains only a His_6-tag while to improve the solubility of the gene products (Fig. 2.2). The expression vector pPROAX-HTc is a bacterial plasmid which carries a Trc promoter, a His_6 protein tag at the N-terminous, and an Ampicillin resistance.(Fig. 2.3)

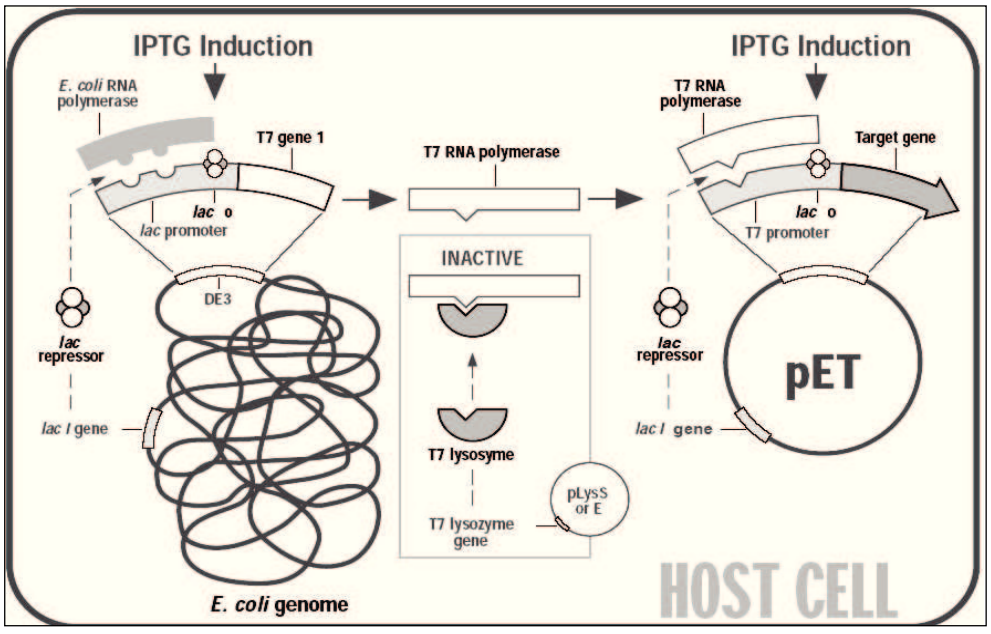

Fig. 2.1: General scheme of the gene expression regulation by the insertion into pET system

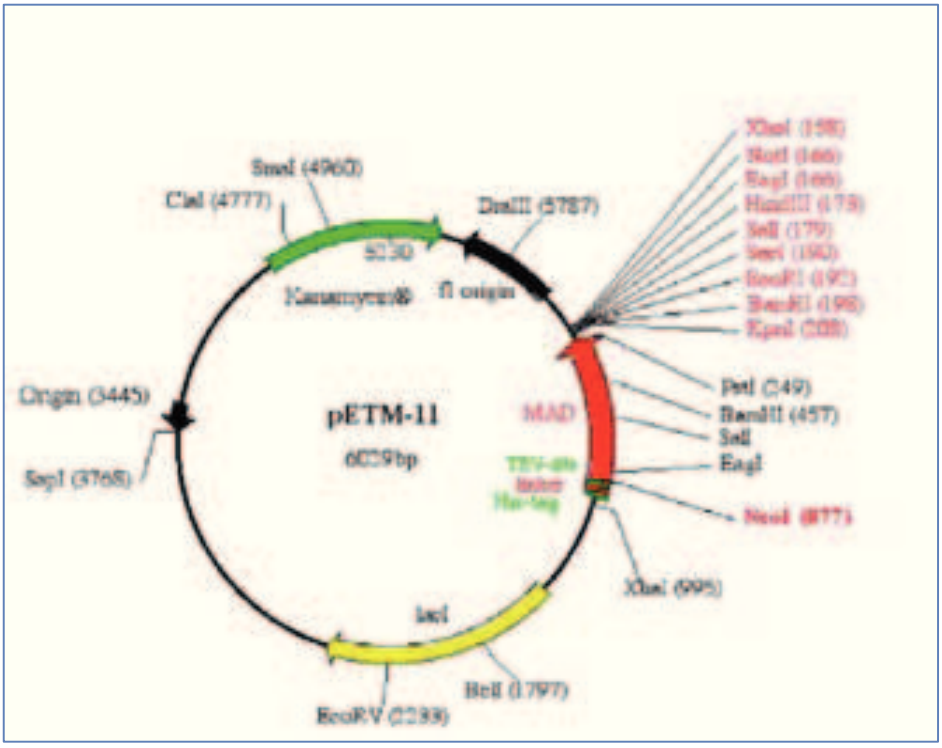

Fig. 2.2: Map of expression vector pETM11.

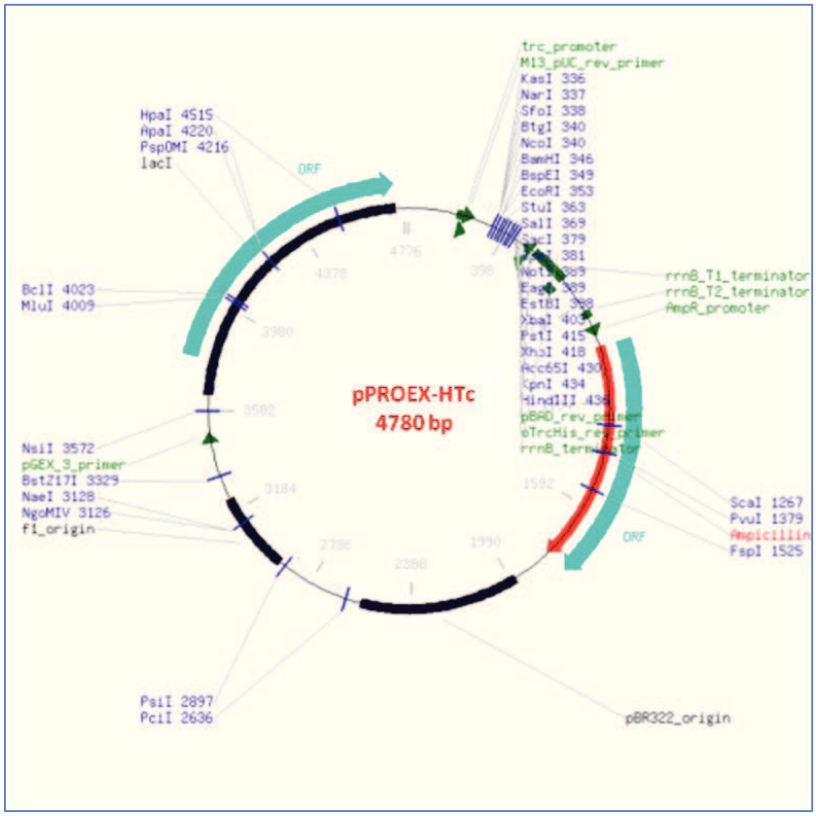

Fig. 2.3: Map of expression vector of pROEX-HTc vector

2.4.3 Scale up expression

Because of the vast fund of knowledge about its genetics, biochemistry, and molecular biology, *Escherichia coli* is the system of first choice for expression of many heterologous proteins. Genetic manipulations are straightforward, cultures of *E. coli* are easily and inexpensively grown, and many foreign proteins are well-tolerated and may be expressed at high levels [Sambrook and Russel, 2001]. *E. coli* expression systems can be used for the production of recombinant proteins either intracellularly or extracellularly. Recombinant expression plasmids require a strong transcriptional promoter to control high-level gene expression.

2.4.3.1 Expression of HspA and HpCoaX from *H. pylori*

The gene encoding HspA (*hp0011)* was inserted into the pETM11a expression vector (Novagen) to give pILL948 as described (Kansau *et al*., 1996). Single (C94A) and double (C94A/C111A) HspA mutants were constructed using the *QuickChange* site directed mutagenesis protocol (Stratagene). The vector was introduced into the *E. coli* BL21(DE3)pLysS expression strain (Invitrogen) using heat-shock: 1ul of vector (50

ng/ul) was added to 100ul of cells. After a first screening of small-scale expression cultures, performed using different strains, temperatures, IPTG concentration and induction length, recombinant constructs were transferred into *E. coli* strain that assured the best expression level in soluble phase. The transformed cells were transferred on LB-Agar plates (15 g/L Agar, 10g/L NaCl, 10g/L Tryptone, 5g/L Yeast extract) supplemented with antibiotics (100ug/mL Ampicillin, 50ug/mL Cloramphenicol) and incubated overnight at 37°C. This pre-culture was then used to inoculate 10mL of fresh LB broth (Luria-Bertani growth medium: 10g/L NaCl, 10g/L Tryptone, 5g/L Yeast extract) supplemented with antibiotics (100ug/mL Ampicillin, 50ug/mL Cloramphenicol), using selected single colonies from LB-Agar plates. The cultures were then incubated overnight at 37°C. For protein preparation, 1L of pre-warmed LB medium supplemented with 100 mg/l ampicillin and 50 mg/l chloramphenicol was inoculated with 10mL of pre-inoculum described above and the resulting culture was grown at 37°C. When the OD_{600} reached 0.5-0.6, expression of recombinant protein was induced by the addition of isopropyl-*d*-thiogalactopyranoside (IPTG) to 1 mM final concentration and cells were incubated at 22° or 37 °C as appropriate. After 3 or 16 h (if we work at the overnight induction) the solubility of each protein was assessed by harvesting the cells (centrifugation at 6000 rpm, 15 min, 4 °C). Cell pellets were resuspended in a lysis buffer appropriate and optimized for each target protein studied. For HspA and its mutants the lysis buffer containing 0.02M Na_2HPO_4, 0.5M NaCl, 1mM $MgCl_2$, pH 7.4 and protease inhibitors. While HpCoaX cell pellets were resuspended in a very accurate buffer contains 50mM TRIS pH 8.0, 500 mM NaCl, 5% Glycerol, 0.2 % TRITON X_100, 2mM DTT, PMSF, Lysozime, DNase, Protease Inhibitors , 10mM Imidazole. The suspension was sonicated for about 10 min (time depends on the lysate volume), by using a Misonix Sonicator 3000 apparatus with a micro tip probe and an impulse output of 1.5/2 (=9/12 Watt). Bacterial lysates were then centrifuged (16000 rpm, 30 min, 4 °C) and the supernatant (soluble fraction) was collected and analyzed by SDS-PAGE to assess the presence of recombinant products of interest.

2.4.3.2 Construction of expression vector of Type III Pank from *H. pylori* (HpCoaX)

The *H.* pylori HpCoaX gene (HP0862) was amplified by PCR from *H. pylori* genomic Dna using *Taq* DNA polymerase and the following primers: 5'-ATAAGAAGTAGGCATATGCCAGCTAGGC-3' (forward primer), introducing an *Nde*I site (underlined) at the start of the gene, and 5'-ATGCCCAAAAAACTCGAGTTGTGCATC-3' (reverse primer), introducing an *Xho*I site (underlined) at the end of the gene. The resulting PCR products were digested with *Nco*I and *Xho*I and ligated to *Nco*I/*Xho*I-digested pROAEXTc expression vector using the NEB Quick Ligation Kit. The sequence of the resulting plasmid, was verified by automated DNA sequencing. HpCoaX protein was prepared by cloning the target gene into overexpression vector coding for N-terminal His_6-tagged fusion protein. The vector was transformed into BL21(DE3)pLysS; culture was grown at 37°C to an A_{600} of about 0.6, and subsequently induced by the addition of IPTG. After growing overnight at 37°C, the cells were harvested, suspended in sonication buffer, disrupted by sonication and centrifuged at

15.000x g for 30 min to clarify the cell-free exstract. The protein was overexpressed and purified using immobilized metal affinity chromatography.

2.4.4 Purification of 6xHis tagged proteins

Lysates containing tagged recombinant proteins were purified by affinity chromatography onto HisTrap columns connected to an AKTA-FPLC system. Before loading the lysate, the resin was extensively washed with water and then equilibrated in buffer A (20 mM Na_2HPO_4, pH 7.4, 500 mM NaCl, for HspA; and 50mM TRIS pH=8, 500mM NaCl, 10mM Imidazole, for HpCoax). The lysate was loaded in presence of 10 mM imidazole to avoid non-specific binding of *E. coli* contaminants. Flow-through containing all unbound proteins was collected and the resin was washed with buffer A; finally His-tagged proteins were eluted with high concentrations of Imidazole (250-300 mM) applying a gradient 0-100% of eluition buffer B (buffer A and 500mM Imidazole). All collected pools (flow-through, washes and elutions fractions) were analyzed on SDS–PAGE gels and stained with Coomassie Brilliant Blue R-250. Pools of interest were dialyzed to remove Imidazole by using Spectra/Por membranes with the appropriate MWCO. The purity and homogeneity was tested by SDS-PAGE and by mass spectrometry. Antigenicity of the recombinant proteins was analyzed by western blot.

2.4.5 TEV digestion of tagged proteins

TEV protease is a very useful reagent for cleaving fusion proteins. It is also relatively easy to overproduce and purify large quantities of the enzyme. After purification by affinity chromatography, all partner-tagged proteins were dialyzed overnight against TEV buffer (50 mM Tris/HCl, 500 mM NaCl, 5mM EDTA, 1mM DTT, pH 8.0) at 4 °C. TEV protease was added to protein substrates, using a molar ratio (protease : substrate) of 1 : 50. Cleavage products were analyzed by polyacrylamide gel electrophoresis; then, mixture was loaded onto an HisTrap column and purification was performed in the same conditions described above: protein product of interest was eluted in flow-through, while the fusion partner and un-cut complexes were bound to resin and removed with imidazole.

2.4.6 Size Exclusion Chromatography

A gel filtration column separates proteins based on their size and shape. A gel filtration column is packed with a gel which comprises porous beads, *e.g.* highly cross-linked agarose. When a sample is passed down the column, separation depends on the different abilities of the sample components to enter the pores within the gel beads. Larger molecules, which cannot enter even the largest pores, pass through the column fastest. Smaller molecules, which can enter the pores freely, are delayed to different degree during the passage through the gel, depending on their size and shape. Proteins are therefore eluted in order of decreasing size. The advantages of gel filtration are that

the method is easy to use, it provides free choice of eluent, works almost always, and the result is easy to predict. Due to limitations such as low resolution and small sample volume, this chromatography technique is preferably used as an intermediate, or final purification step. Also, it should be noted that influence of flow rate, and column efficiency, plays important roles. Size exclusion chromatography was performed to purify all recombinant proteins previously expressed and produced in *E. coli* strains, by removing aggregates and contaminants. Runs were carried-out at a flow of 0.5 ml/min onto a Superdex _200 [10/30] (Amersham Biosciences) and Superdex-75 [10/30] columns connected to an AKTA Purifier system, in 20mM TRIS pH 7.4, 200mM NaCl. Molecular weight standards from GE-Healthcare were used to calibrate column. A calibration curve is prepared by measuring the elution volumes of several standards, calculating their corresponding K_{av} values and plotting their K_{av} values versus the logarithm of their molecular weight. The molecular weight of a protein is determined from the calibration curve once its K_{av} value is calculated from its measured elution volume.

$$K_{av} = \frac{V_e - V_0}{V_t - V_0}$$

where V_e is the elution volume for the protein, V_0 is the column void volume (elution volume for Blue Dextran 2000) and V_t is the total bed volume. Pools of interest were concentrated on Amicon-Ultra membranes 5k (Millipore) and on Centriplus centrifugal filter (Millipore). Aliquots of purified proteins were stored at -80°C.

2.4.7 Light scattering analysis (SEC-MALS)

The MW of protein can also be determined by analytical size exclusion chromatography combined with multi-angle light scattering (SEC-MALS). To perform the SEC-MALS analysis we have used a SEC-LS system consisting of a semi-preparative size exclusion chromatography column (Superdex-200 10/30, GEHealthcare) coupled to a light scattering detector (miniDAWN TREOS) and a differential refractive index detector (Shodex RI-101).
The Astra (5.3.4 version, Wyatt Technology Corporation) software allowed us to the collect, record, and process the scattering data. Data is processed assuming a specific refractive index increment (dn/dc) of 0.185 ml/g. To determine the detector delay volumes and the normalization coefficients for the MALS detector, a BSA sample (Sigma) is used as a reference.

2.5 Protein Analyses

2.5.1 Determination of the protein concentration

The concentration of the proteins in solution was determined according to the Bradford's method (Bradford, 1976). The Coomassie Brilliant (Bio-Rad) reagent was added to the samples and the absorbance at 595nm was monitored. A solution of bovine serum albumin (BSA) was used as standard. Protein concentration was also measured by UV

spectroscopy, reading the absorbance at 280nm. The protein concentration value is obtained applying the Lambert-beer law and knowing the specific ellipticity coefficient (ε). The measure is carry out using a Jasco V-550 UV-VIS spectrophotometer, in a 1 cm quartz cell.

2.5.2 Electrophoretic analysis of proteins (SDS-PAGE)

SDS-polyacrilamide gels determined the purity of the proteins (Fig. 2.4). The gel consists of a running gel (40% v/v 30% acryl-bisacrylamide mix, 0.75M Tris pH 8.8, 0.1%SDS v/v, 0.1% APS v/v and 0.04% TEMED v/v) and a stacking gel (17% v/v 30% acryl-bisacrylamide mix, 0.063M Tris pH 6.8, 1%SDS v/v, 1% APS v/v and 0.33% TEMED v/v). The electrophoresis on 12% or 15% polyacrylamide gel in denaturing conditions was performed according to Laemmli's protocol (Laemmli, 1970). The samples were denatured in the heating block at 95°C for 5 min and pre-treated with loading buffer: 1% SDS (Applichem), 5% β-mercaptoethanol (Sigma), 0.001% bromophenol blue (ICN Biomedicals) and 10% glycerol (Applichem). The samples were then loaded on a polyacrylamide gel and the electrophoresis was performed in 0.025 M Tris/HCl, 0.2 M glycine pH 8.3 and 0.1% SDS, at 30 mA until the front of the samples reached the bottom of the gel . The proteins were then revealed by Coomassie Brilliant-Blue (Applichem) staining; the gel was submerged in the staining solution (0.1% Coomassie Brilliant-Blue R250, 25% isopropilic alcohol and 10% acetic acid) for 30 min with gentle agitation. The gel was washed in a solution containing 30% ethanol and 10% acetic acid to remove the excess of Coomassie and then washed again in deionized water for 2 times for 10 min each. Alternatively, the proteins were transferred by electroblotting from gel to a polyvinylidene fluoride (PVDF) membrane as described in the paragraph "Western blot analysis".

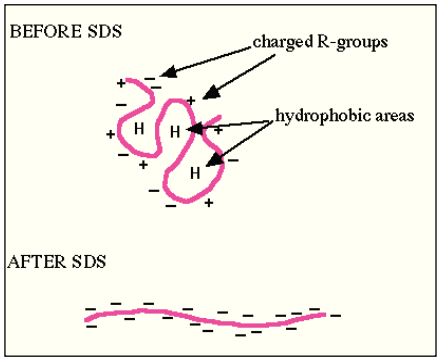

Fig. 2.4: SDS effect on the protein conformation

2.5.3 Western Blot analysis

After gel electrophoresis, proteins were transferred to a polyvinylidene fluoride (PVDF) membrane (BioRad) in Washing Buffer (25 mM Tris/HCl pH 8.0, 190 mM Glycine and 10% methilic alcohol). The PVDF membrane was submerged in methilic alcohol for 2 min before using, to eliminate the hydrophobic properties, washed in H2O for 2 min and equilibrated in Washing Buffer for 5 min. After the electro-blotting using a powerfully system (BioRad), membrane was stained with Ponceau Red to verify proteins transferring and incubated in the blocking solution (0.5% NFDM in TTBS buffer, containing 0.05% Tween 20) at room temperature for 1 h. It was washed with TTBS buffer, and after incubated in a mouse anti-His peroxidase-conjugated anti-mouse antibody (Sigma Aldrich) 1:1000 dilution, at room temperature for 1 h. Membrane was finally washed with TTBS buffer. The detection of immunopositive species by enzyme-linked chemi-luminescence (enhanced chemi-luminescence: ECL) was performed according to the manufacturer's instructions, using a specific detection kit (Bio-Rad).

2.5.4 Mass spectrometry

LC-MS, using a Phenomenex Jupiter C18 column (4.6 x 250 mm, 5 µm, 300Å), was performed to confirm the mass of HspA wt, its mutants and of HpCoax. The method developed at 1 mL/min with a linear gradient of 0.05% TFA in H_2O (buffer A) in 0.05% TFA in CH_3CN, from 5% to 70% for 30 min.

2.5.4.1 Vinyl pyridine assay (4-VP)

The 4-Vinyl-pyridine assay was performed using specific reagent and monitoring the reaction by mass spectrometry. At the first the sample was treated with a denaturing buffer; after it was incubated with the 4-VP; the reaction of alkylation was carried out in a thermo-block, at the temperature of about 65°C. The samples were analyzed by LC/MS.

2.5.5 Cys distribution in GroES proteins

The search was carried out against all sequences in the non redundant database RefSeq (www.ncbi.nlm.nih.gov). The database was accessed in February, 2007. The keyword *GroES* in the Gene Name field restricted to master sequences (master of set of segmented sequences) resulted in 313 hits; 4 sequences consisting of a residue number less than 50 or greater than 220 were discarded; the 3 HspA sequences from HELPY, HELPJ, HELAC were also discarded. The working subset of GroES sequences contained 306 sequences consisting of 30145 residues in total. The search for Cys residue revealed only 14 cysteines in 13 sequences. For comparison the frequency of all 20 naturally occurring aminoacids was computed; the most abundant amino acid is valin with 3603 residues (12.0%) whereas the less represented – apart from Cys – is tryptophan, with 66 residues (0.2%), as expected. A query in the PDB database for the 3D-structure of any of the 13 Cys-containing sequences resulted in 0 hits, so the 3D-structure is unknown.

2.5.5.1 Chemical modification of cysteine residues

Reduction of Cys residues in HspA was carried out in 0.25 M Tris-HCl, 1.25 mM EDTA, pH 8.0 containing 6 M guanidinium chloride by incubation with a 10:1 molar excess of dithiothreitol (DTT) over the total -SH groups at 37°C for two hours under nitrogen atmosphere (N_2 flow). Alkylation of native HspA was performed by incubation of the protein in tha same buffer with a 10-fold molar excess of iodoacetamide (IAM) over the total -SH groups for 10 min at room temperature in the dark. For reduction and carboxyamidomethylation of the Cys residues, HspA was incubated with a 10:1 molar excess of DTT over the total -SH groups at 37°C for two hours under nitrogen atmosphere followed by alkylation with a 10-fold molar excess of iodoacetamide over the total -SH groups for 10 min at room temperature in the dark. In all cases, the protein samples were desalted by reverse phase chromatography (RP-HPLC) on a C-4 column using a HP 1100 chromatograph (Agilent Technologies). The elution was performed with a speed gradient of 0.1% TFA (buffer A) and 0.07% TFA in 95% acetonitrile (buffer B). HPLC-desalted HspA samples were directly analysed by ES/MS using a Quattro Micro triple quadrupole mass spectrometer (Waters). Data were recorded and after analyzed by the Mass Lynx. Software.

2.5.5.2 Mass spectrometry measurements

HPLC-desalted HspA samples were directly analysed by ES/MS using a Quattro Micro triple quadrupole mass spectrometer (Waters). Samples were injected into the ion source by a Harward syringe pump at a flow rate of 10 μl/min. Data were acquired and processed by the Mass Lynx software provided by the manufacturer. The instrument was calibrated by a separate injection of horse heart myoglobin (average molecular mass 16,951.5 Da); all masses are reported as average mass.

2.5.5.3 Enzymatic Hydrolysis

Native HspA was digested with trypsin in 50 mM ammonium acetate, pH 6.0 at 37°C overnight, using an enzyme/substrate ratio of 1:50 (by mass). The mixtures of tryptic peptides were either directly analyzed by MALDI mass spectrometry using a Voyager DE PRO instrument (Applied Biosystem).

2.5.5.4 Analysis by MALDI mass spectrometry (MALDI/MS)

Mass mapping experiments were performed by MALDI mass spectrometry analysis using a Voyager DE PRO instrument (Applied Biosystem) operating both in linear and reflectron mode. Typically, 1 μl of analyte solution was mixed with 1 μl of a-cyano-4-hydroxycinammic acid 10 mg/ml in acetonitrile/50 mM ammonium citrate, 70:30 v/v, containing 250 fmol of bovine insulin. The mixture was applied onto the metallic sample plate and air dried. Mass calibration was performed using the quasi-molecular ions of insulin at m/z 5734.5 and a matrix peak at m/z 379.1 in linear mode and by multipoint calibration in reflectron mode using a peptide calibration mixture provided by the manufacturer. All mass values in linear mode are reported as average masses, whereas those measured in reflectron mode are reported as monoisotopic values.

2.5.6 Disulphide bridges pattern

The protein dilsuphides bridge pattern assessment can be carried out by mass spectrometry techniques. The peptidic mix obtained by enzymatic hydrolysis has been analyzed to minimize the possibility of dilsuphide bridge reduction and arrangement. Protein hydrolysis can isolate each cysteine into a different peptide to facilitate the data interpretation. Mass spectra signals are been attributed at specific peptides basing on their molecular weight. It is possible to individuate specific signals from two peptides linked by a disulphide bridge, and consequently the cysteine residues involved in the S-S bond. The assessment is confirmed treating the mix with a reducing agent and analyzing it again by MALDI/MS. The signals from peptides linked by disulphides bridges disappear to generate higher mass values corresponding to reduced peptides. The assessment of distribution disulphides bridges pattern in HspA wt and its mutants is been executed essentially by enzymatic digestion and then by MALDI/MS analysis on the peptidic mix, according to the well established mass-mapping strategy [Amoresano A., and Pucci P., 2003], [Amoresano A., and Pucci P., 2001], [Morris H.R., and Pucci P., 1985]. This provides to define a peptidic map of a protein: specific signals are attributed basing on the peptides molecular mass and on the specific proteolytic enzymes used.

2.5.7 Modelling of HspA

The structure of the A domain of HspA was generated by homology modelling using the automatic modelling Swiss Modeller server using *Escherichia coli* GroES as a template (pdb code 1wnr.pdb) Ref). Three models were generated and validated using verify 3D (refs). In addition, the loop consisting of residue 50-53 was rebuilt and minimized to fit the biochemical data obtained in the study (disulphide bridge between Cys51 and Cys53). To illustrate the results obtained in our study (disulphide bridge pattern) we added the B domain to the A domain by manual building in COOT [Emsley P., Cowtan K., 2004]. Minimization was carried out using GROMOS96 implementation of Swiss-PdbViewer [van Gunsteren *et al*., 1996].

2.5.8 Circular dicroism analysis (CD)

All CD spectra were recorded with a Jasco J-810 spectropolarimeter equipped with a Peltier temperature control system (Model PTC-423-S) in a 0.1 mm quartz cell. The spectropolarimeter was calibrated with an aqueous solution of 1S-(+)-10-camphorsulfonic acid at 290 nm. Molar ellipticity per mean residue, [θ] in degrees cm^2 $dmol^{-1}$, was calculated from the equation: $[\theta] = [\theta]_{obs} mrw(10/C)^{-1}$, where $[\theta]_{obs}$ is the ellipticity measured in degrees, mrw is the mean residue molecular mass (111.5 Da), *C* is the protein concentration in g L^{-1}, and *l* is the optical path length of the cell in centimeters. Spectra were acquired according the following parameters: far UV range of 190-260nm, band width of 1nm, response of 8 sec, data pitch of 0.2nm and scanning speed of 10 nm/min. Spectra were carried out at 20 °C using a 0.01 cm optical path-length cell and a protein concentration of 10-15 mg mL^{-1}. Cuvettes with path lengths of 0.2 and 0.5 cm and protein concentrations of 0.2 and 0.8 mg/mL were used in the far-

UV and near-UV regions, respectively. CD spectra were recorded with a time constant of 4 s, a 2 nm bandwidth, and a scan rate of 5 nm/min, and were signal-averaged over at least three scans, and baseline corrected by subtracting a buffer spectrum. Thermal unfolding curves were recorded in the temperature mode, over the range of 5-110 °C, with a scan rate of 1.0 °C/min. Samples were rapidly cooled after the first heating run and scanned for a second time to estimate the reversibility of the unfolding transition. CD data were expressed as mean residue ellipticity (). Spectra processing was obtained by using the Spectra Manager software. Proteins concentrations was determined by monitoring Tyr absorbance at 278nm and using the molar extinction coefficient derived by the calcule from ProtParaM tool (By expasy database). CD intensity is expressed as mean residue ellipticity (deg ×cm2×dmol−1 of amino acid) calculated referring to the total amino acid concentration.

2.5.9 Crystallization tests

Protein samples of HpCoaX were concentrated to 12 mg/mL, in 20 mM Tris/HCl pH 8.0 150 mM NaCl, 3 mM DTT using Centriplus centrifugal filter (Millipore). Protein concentration was estimated by the method of Bradford, using BSA as standard (Bradford, 1976). Crystallization experiments were carried out at 20 °C, and at 4°C, using the hanging and/or sitting-drop vapour-diffusion technique. Initial crystallization conditions were found using commercially available crystallization-screening kits (Crystal Screen I and II, INDEX I and II, Hampton Research, Riverside, CA). Drops containing equal volumes (1 µL) of protein and reservoir solution were equilibrated against 1 mL of reservoir solution. Several crystal forms were obtained only at 4°C, using a reservoir solution containing 0.1 M Bis-Tris pH 6.5, 0.2 M. Only diffraction data at low resolution were obtained from these crystals (about 5 Å). Several parameters such as crystallization technique, buffer composition, protein concentration and ionic strength were varied in order to improve the quality of the crystals. We have also carried out some co-crystallization experiments, or soaking, using nonhydrolysable analogue of ATP (AMPPNP), and substrate (PAN), in the presence of $MgCl_2$. Unlikely we did not observe an improving of crystal's quality.

Chapter III

Results and discussion

3.1 Expression and purification of HspA

In order to obtain and characterize the HspA we have produced this protein by recombinant strategy. The gene encoding HpHspA (*hp0011)* was inserted into the pETM11a expression vector. To analyze cysteine residues located on B-domain we have also expressed two mutant variants of HspA: Cys94Ala (single mutant) and Cys94Ala, Cys111Ala (double mutant). The strain BL21(DE3) pLysS was chosen for the optimal expression that was carried out at 22 °C for 16 h. as described before (see 2.4.3.1 section). The expression was assessed by SDS-PAGE, running the samples before and after the induction (Fig. 3.1). HspA possess a natural His_8-tag at the C- terminus, which can be used for the purification affinity steps. Following induction with IPTG, all recombinant proteins were purified from *E. coli* whole-cell extracts by affinity chromatography, loading them onto nitrilo-tri-acetic acid-chelating nickel (Ni-NTA) columns, frequently used to purify recombinant histidine-rich proteins. This first step of purification is showed in figures 3.2 and 3.3. HspA-wt and its mutants bound to the Ni-NTA resin and, following elution, exhibited the expected molecular weights (13 kDa) as visualized by SDS-PAGE analysis (Fig. 3.4).

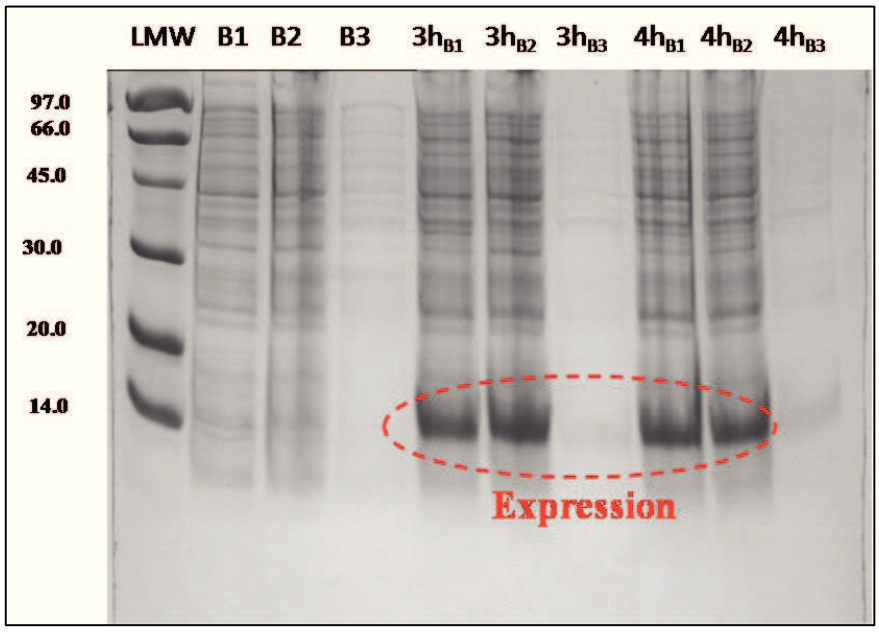

Fig. 3.1: SDS-PAGE before the IPTG induction, after 3h and 4h

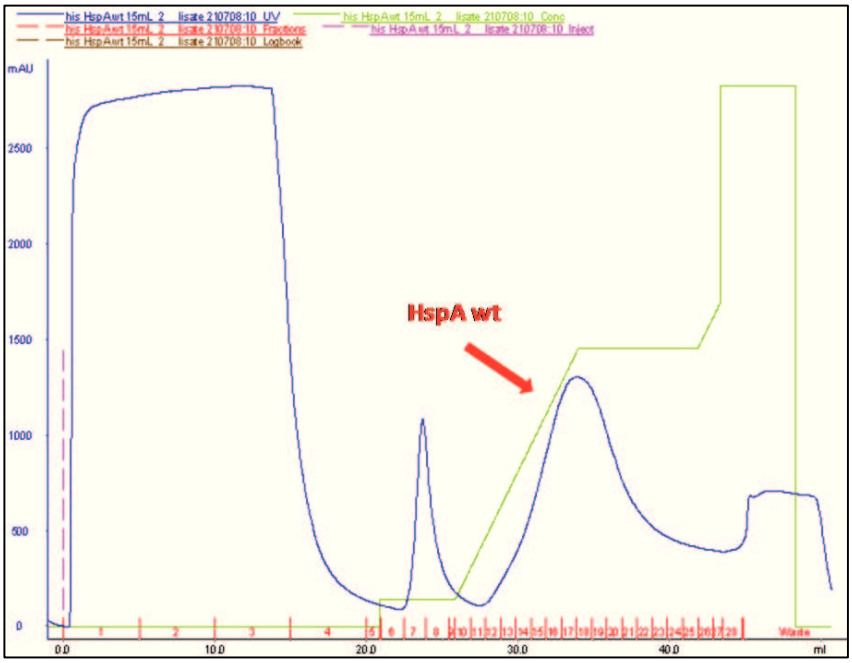

Fig. 3.2: His-trap profile of HspA wt

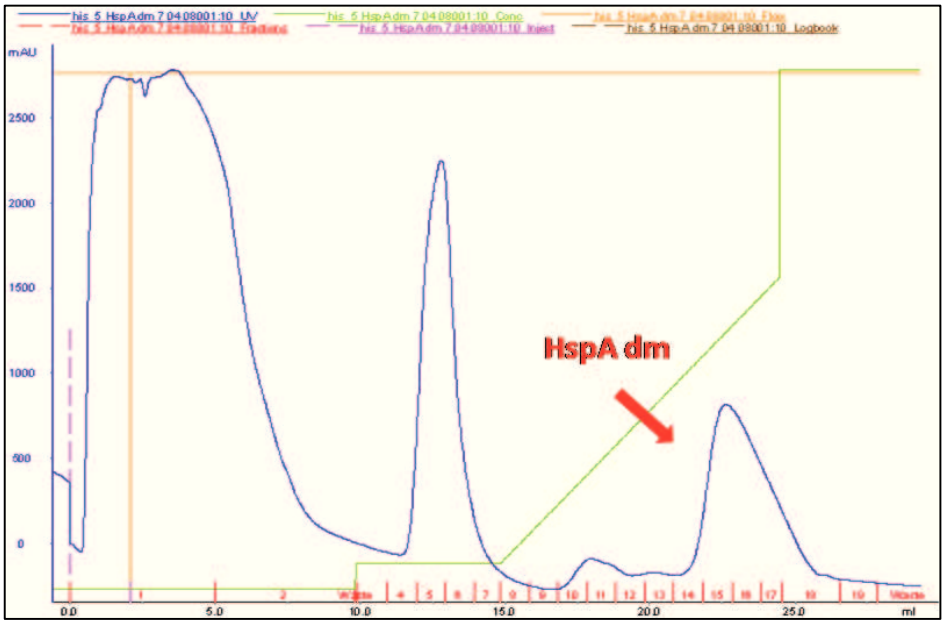

Fig. 3.3: His-trap profile of HspA double mutant

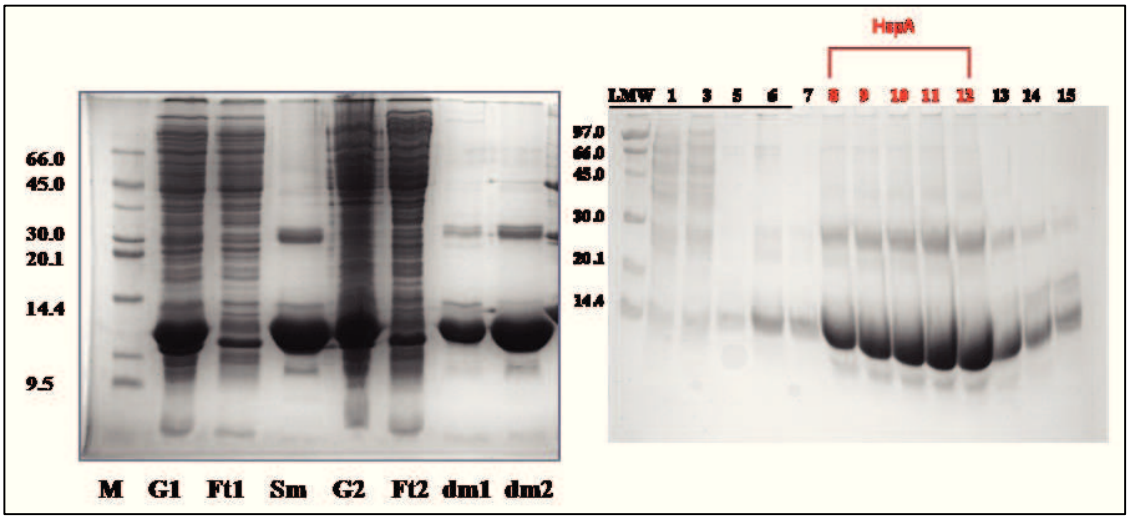

Fig. 3.4: SDS-page analysis after His-trap affinity. Gel on the left shows different fractions eluted: Cell extract (G1,G2), Flow trough (Ft1, Ft2),single mutant (sm) and double mutant (dm1, dm2). Gel on the right shows HspA wt fractions.

After, we have carried out exclusion chromatography assay (gel filtration); the relative profiles are reported in figures 3.5 and 3.6. Gel filtration analysis was also conducted in the presence of DTT. HspA was treated with the reducing agent, and then it was loaded on Superdex_200 column. The multimeric rate decreses significantly and the protein can be eluted as a heptameric form. The formation of aggregates is time dependent and is strongly dependent on the cysteines oxidation (Fig. 3.7).

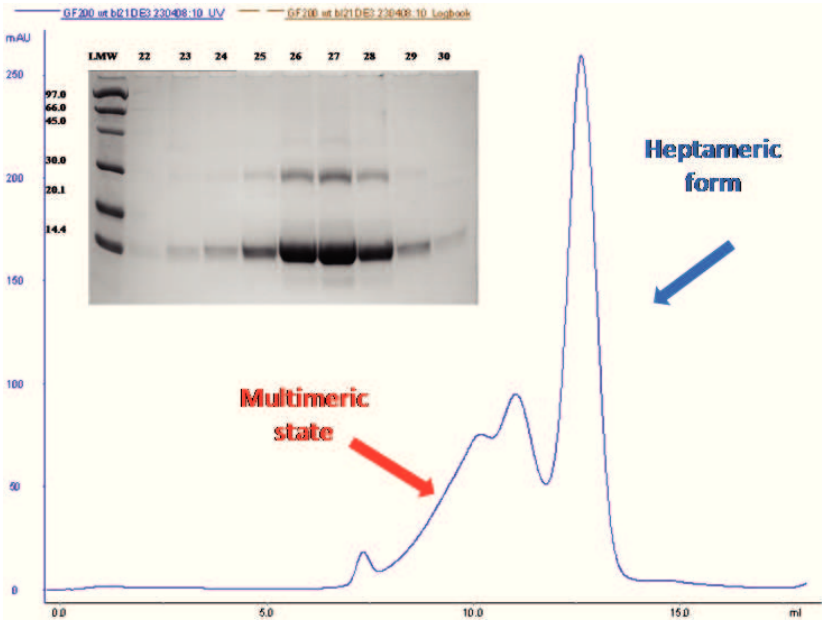

Fig. 3.5: Gel Filtration profile SDS-Page analysis of HspA wt.

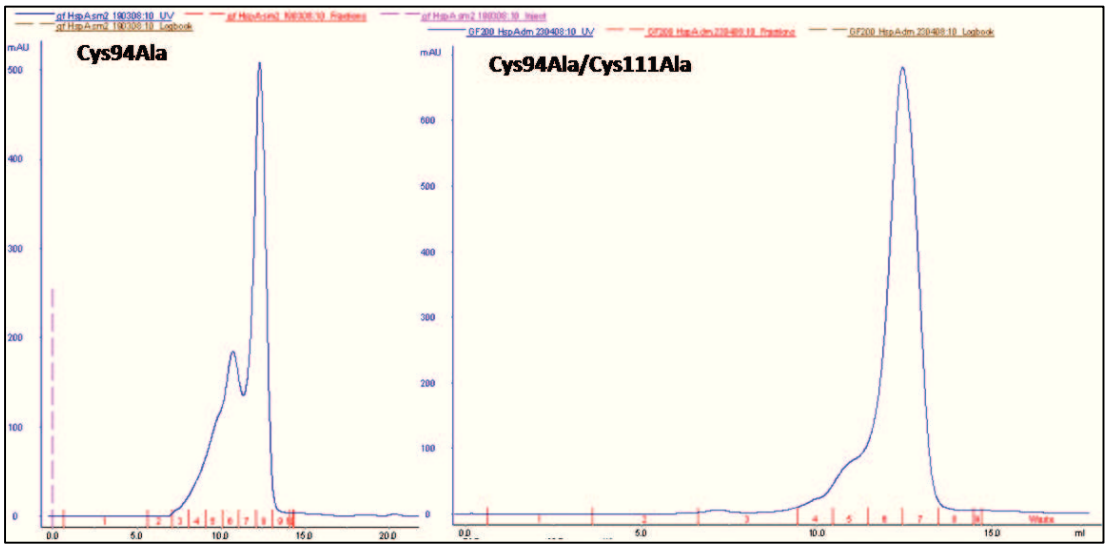

Fig. 3.6: Gel filtration profiles of two mutants: the wt and Cys94Ala show the same profile while the Cys94Ala/Cys111Ala presents a single peak

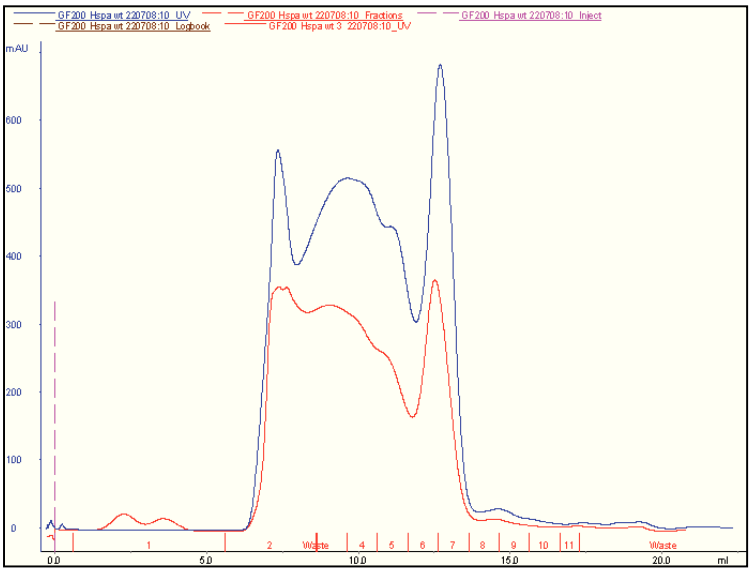

Fig. 3.7: Gel Filtration profile of HspA wt stored at 4°C for two days. In red is showed the first injection on Superdex_200, in blue the second injection, after about 1h.

3.1.2 Gel filtration analysis in presence of Ni^{2+} and Bi^{3+}

The HspA wt was first incubated with TCEP (24 equivalents) for about 1h. Then the samples were treated for about 30 minutes with a buffer containing $NiSO_4$ 1mM and $Bi(NO_3)_3$ 5mM respectively and loaded onto a Superdex_75 column. The profile in the presence of these metals shows only one peak in both cases (Fig. 3.8). The resulting protein is concentrated and then used for crystallization experiments. Unfortunately the drops did not show the presence of any crystals.

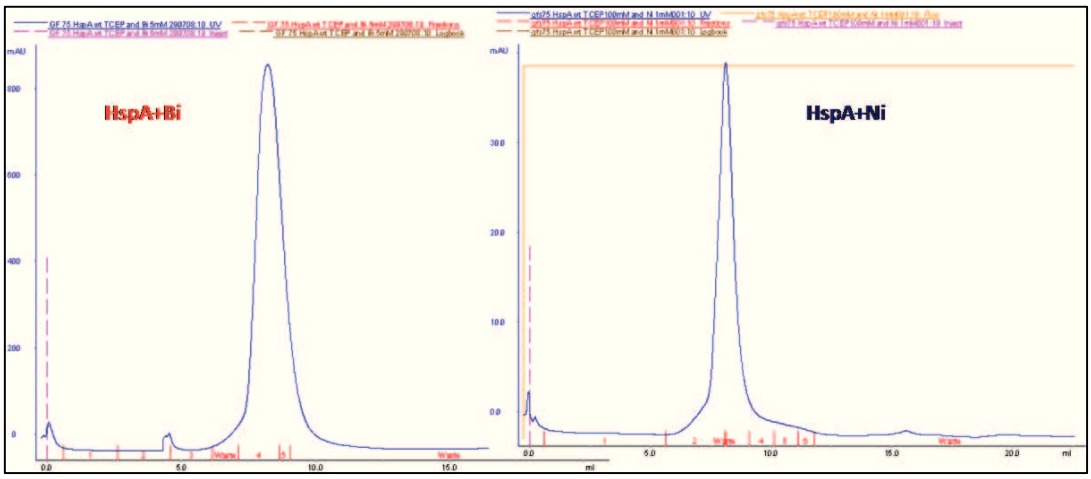

Fig. 3.8: Gel filtration profiles of HspA with nickel and bismuth

3.1.3 Sequence analysis on HspA Cys residues

A statistical survey of cysteine distribution in GroES proteins was carried out using the available sequences in the non redundant database RefSeq (release of February, 2007). Following the selection criteria defined in the Method section, and excluding HpHspA and its homologues, we identified 306 GroES sequences. displaying a Cys frequency of 0.05% (13/30165). Notably, only 13 out of them contain Cys residue (a list is available in the supplementary material). Its frequency across GroES proteins is ~ 0.1% revealing that cysteine is the less represented amino acid within the selected sequence subset (see table 3.1). As a result, it can be deduced that Cys is not an essential residue for the function of GroES chaperonins. With this data at hand, the high Cys content (6/118) requires Hp (and HELAC) HspA play additional and still unknown roles, peculiar of GroES like proteins from bacteria that attach the gastric mucosa.

AMINOACID	Frequency (%)	N° of residues
A	7.1	2178
C	0.1	38
D	6.7	2074
E	9.6	2950
F	1.5	463
G	10.3	3172
H	0.9	276
I	8.0	2482
K	9.3	2862
L	8.1	2502
M	2.2	675
N	2.5	763
P	3.7	1130
Q	2.0	626
R	4.5	1378
S	5.0	1536
T	4.9	1500
V	11.8	3654
W	0.2	67
Y	1.8	553

Tab.1

Table 3.1: The table reports the frequency (%) for each amino acid residue analyzed. In red is reported the less represented amino acid, as cysteine, while in blue are reported the most represented amino acids

3.1.4 Assessment of the oxidation state of HspA Cys residues

The oxidation state of the cysteines occurring in HspA and in its mutants was assessed by determination of the accurate molecular mass of the proteins by electrospray mass spectrometry (ES/MS), following a procedure previously developed [Amoresano A. and Pucci P., 2003]. Aliquots of the HspA samples were directly analysed by ES/MS producing the spectra shown in figure 3.9. Wt-HspA exhibited a molecular mass of 12984.0±0.8 Da, about 6 Da lower than expected for the fully reduced protein (theoretical mass value 12990.9), thus suggesting that the 6 Cys residues were involved in 3 disulphide bridges. Accordingly, the molecular mass of HspA shifted to 12989.9±0.4 Da following reduction with DTT, exhibiting the expected increase of about 6 Da, whereas no difference in the mass value was observed following direct alkylation of the protein with an excess of IAM, confirming the absence of free Cys residues (data not shown). Finally, when the wt protein was first reduced with DTT and then carboxyamidomethylated with IAM, the corresponding ES/MS spectrum exhibited a molecular mass of 13331.5±0.8, as shown in figure 3.9(B), with an increase of 340.6 Da,

corresponding to the alkylation of 6 Cys residues. Similar experiments were also carried out on the HspA mutants. Figure 3.9(C) shows the ESMS spectrum of the Cys94Ala/Cys111Ala double mutant exhibiting a molecular mass of 12921.0±0.4 Da, about 4 Da lower than expected for the fully reduced protein, suggesting the occurrence of two S-S bridges. This result was confirmed by the reduction and reduction/alkylation experiments analogous to those described above for wt-HspA indicating that very likely the two mutated cystein residues, Cys94 and Cys111, were joined by a disulphide bond in the wt-HspA molecule. Finally, the ESMS analysis of the Cys94Ala single mutant yielded a mass value of 25903.8±1.0 Da, as shown in figure 3.9(D), corresponding to a dimeric species stabilised by an intermolecular disulphide bridge very likely involving the free Cys111 residue of each HspA molecule.

3.1.5 Disulphide bridge pattern

Assignment of the disulphide bridges pattern in both wt-HspA and the double mutant was accomplished according to the well established mass mapping strategy described in 2.5.6 section. Wt-HspA was digested with trypsin at pH 6.0 to avoid scrambling of S-S bridges and the resulting peptide mixture was directly analysed by MALDI-MS. Figure 3.9(A) shows the high mass region of the resulting MALDI spectrum recorded in linear mode where a series of mass signals was attributed to S-S bridged peptides. The peaks at m/z 4234.8 and 3969.6 were interpreted as arising from the peptide 78-103, originated by aspecific cleavage at Tyr77, joined to the fragments 107-118 and 107-116, respectively by two disulphide bridge involving the four Cys residues at positions 94, 95, 111 and 112. Both mass signals were accompanied by satellite peaks at about 335 Da higher (m/z 4571.8 and 4304.5) corresponding to the same peptide 78-103 linked to the fragments 104-118 and 104-116 respectively via the same two S-S bridges. Finally the mass signals at m/z 3891.1 and 3626.3 were attributed to the peptide 81-103 bridged to the fragments 107-118 and 107-116 by two disulphides involving the above mentioned four cysteines (see Table 3.2). All these signals disappeared following incubation with DTT. The mass spectral investigation of wt-HspA provided the overall scheme of cysteine pairings with Cys94, 95, 111 and 112 originating a network of two disulphide bridges and Cys51 and Cys53 forming the third S-S bridge. This interpretation was supported by the weak mass signal at m/z 2404.3 observed in the MALDI reflectron spectrum and was further confirmed by the double mutant analysis (see below). This peak occurred 2 Da lower than the peptide 42-64, suggesting the presence of an intermolecular disulphide bridge joining Cys51 and Cys53.

The correct pairings of the Cys residues in HspA was obtained by the MALDI mass spectral analysis of the tryptic digest of the C94A/C111A double mutant. As shown in figure 3.10, mutation of the two Cys residues seemed not to affect the C-terminal pattern of disulphide bridges. The mass signals at m/z 5344.7, 5053.7 (peptides 65-103 and 67-103 joined to 107-116 respectively) and 3907.7, 3564.6 and 3322.4 (peptides 78-103, 81-103 and 83-103 linked to the fragment 107-116) essentially corresponded to the previously observed S-S bridged peptides. This strongly suggests that Cys94 and Cys111 are held together by a S-S bridge in wt-HspA. The C-terminal fragments are then still joined by the second disulphide bond involving Cys95 and Cys112. The third S-S bridge involving Cys51 and Cys53 was confirmed by the mass signal at m/z 964.5 (peptide 47-55 containing an intramolecular disulphide bridge) leading to the unambiguous assignment of the S-S bridge pattern in HspA.

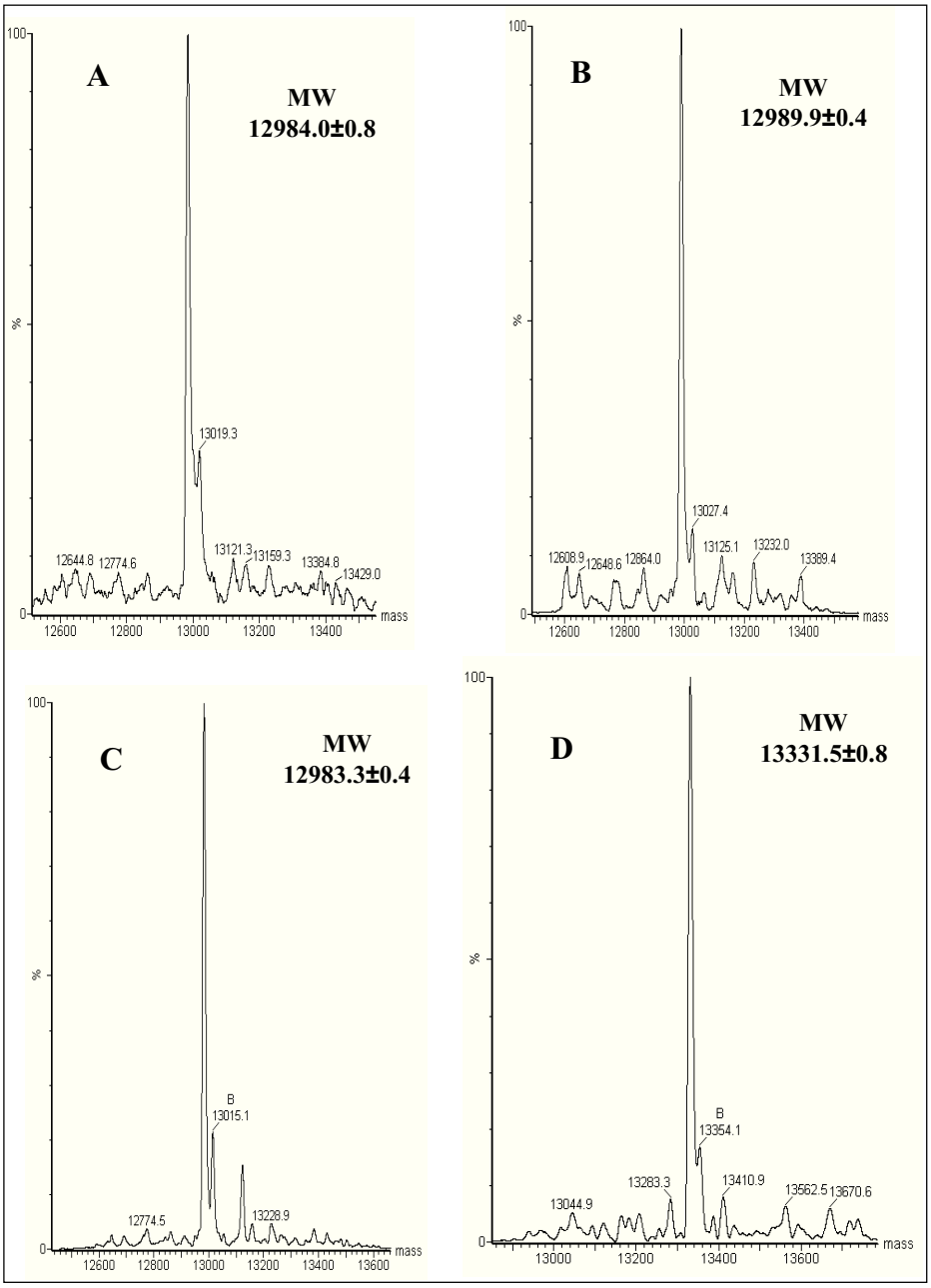

Fig. 3.9: ESMS spectra. a)WT HspA; b) HspA carboxymethylated at position 51, 53, 94, 95, 111, 112, after reduction with DDT; c) C94A/C111A double mutant; d) dimeric C94A mutant.

Peptides Involved In S-S bridges	Observed Mass (Da)	Theorical Mass (Da)
78-103+ 107-116	3969.6	3969.7
78-103+ 107-118	4234.8	4234.88
78-103+ 104-116	4304.5	4304.7
78-103+ 104-118	4571.8	4569.88
81-103+ 104-116	3891.1	3962.73
81-103+ 107-116	3626.3	3626.53

Table 3.2: Table of peptides obtained from trypsine digestion and involved in the S-S bonds formation. It is reported the observed Mass, as experimental result, and the theoretical Mass calculated with the software MS-digest

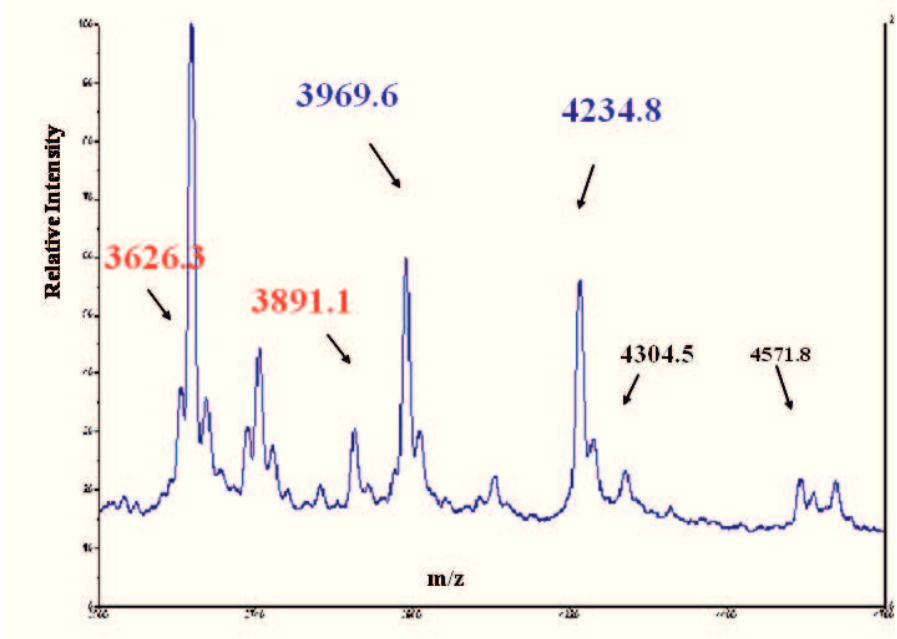

Fig. 3.10: Maldi spectra obtained from peptide mixtures following the mass mapping strategy

3.1.6 Role of S-S bonds

As previously described, we found that, under non reducing conditions, HspA contains 3 intramolecular disulfide bridges. The bonds involve either the GroES-like domain A (Cys51-Cys53) or the nickel-binding domain B (Cys94–Cys111 and Cys95–Cys112). The latter 2 bridges dramatically affect the topology of the C-terminal domain generating a very unusual "lasso-like" structure: a small rigid ring of 16 atoms built up from the two disulfide bridges and a large flexible ring formed by 18 residues of the B domain. This large closed loop, embodies 5 His residues leaving out of the ring the neighbouring 3 C-terminal His (Fig. 3.11). It is tempting to speculate that this lasso-like structure would be ideal to position the hystidine side chain in conformation suitable to bind nickel ions. Previous data showed that, under non-reducing conditions, MBP-fused HspA binds 2 Ni(II)/molecule with an apparent Kd of 1.8 μM, whereas free HspA binds only 0.6 Ni(II)/molecule with a lower Kd (0.6 μM). We surmise that the difference in stechiometry and equilibrium constants may be related to the Cys state, as the MBP may protect the small HspA chain from Cys reduction/oxidation.

HspA occurs at approximately 50% either in the soluble protein fraction in the structure bound protein fraction, suggesting that the subcellular localization might depend on the environment conditions. Along this line, our findings suggest that HspA may switch from an oxidised to a reduced state depending on redox control of the environment. This holds particularly for the cell surface, where the protein is embedded in an oxidizing environment and then contains 3 S-S bridges; whereas the cytoplasmatic protein may exhibit a diverse Cys redox-state (Fig. 3.12). In conclusion, our results suggest that two of the disulphide bridges, located in the B domain force the C-term domain to adopt a unique closed loop structure built-up from vicinal disulfide bridges which is relevant to the Ni dependent functioning of HspA in *H. pylori*. This structural determinant would be optimal for binding up to 2 Ni ions as suggested by the different redox environments that the protein experiences inside and outside the bacterial cell [Loguercio S., *et al.* 2008].

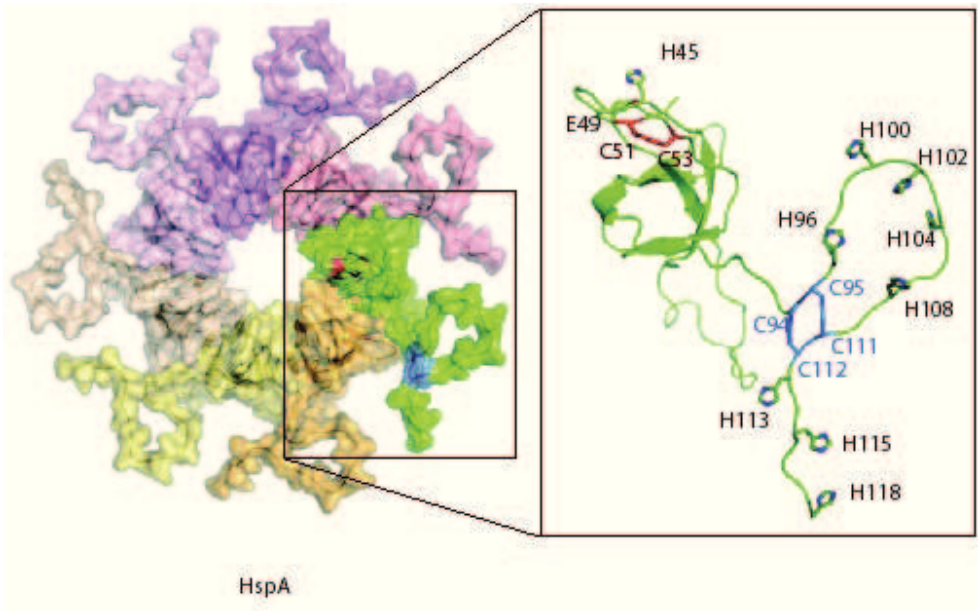

Fig. 3.11: Space filled representation of the A domain of HspA obtained by homology modelling. B domain has been built by COOT and represented as a green coil; S-S bonds are showed in blue.

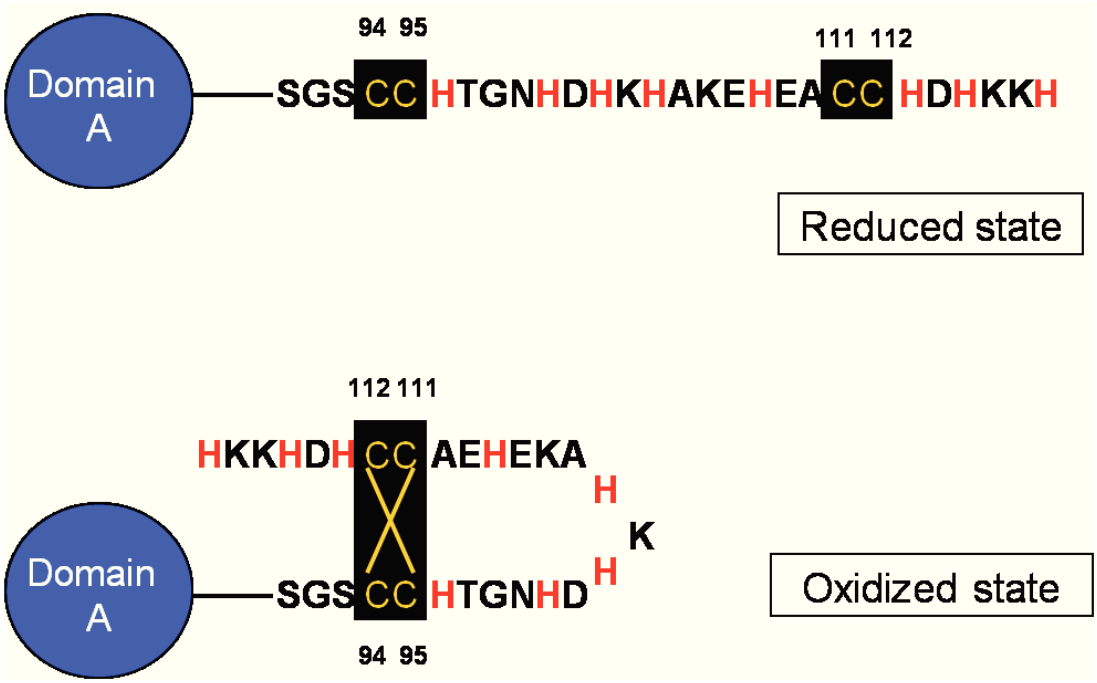

Fig. 3.12: Schematic representation of B-domain in two different states: reduced and oxidized. In the oxidized state, two of the disulphide bridges, located in the B domain, force the C-term domain to adopt a unique closed loop structure.

3.2 Expression and purification of Type III PanK, HpCoaX

The CoaX protein from *H. pylori* was prepared by cloning into a pPROEX-HTc expression vector, according to the methods reported in 2.4.3.2 section. The vector was transformed into *E. coli* BL21DE3pLysS strain. The protein of interest was overexpressed and then the cell extract containing recombinant Hp0862 was loaded onto a His-trap affinity column pre-equilibrated with buffer A for the first purification step (Fig. 3.13 A). Proteins fractions were eluted applying an Imidazole gradient, with the elution buffer B. The eluted fractions were resolved by SDS-PAGE (Fig. 3.13 B). Protein fractions were pooled, digested with TEV protease and finally applied to a gel filtration column for the following purification steps. The purity and the homogeneity of HpCoaX was firstly verified by SDS-PAGE analysis and after by mass spectrometry.

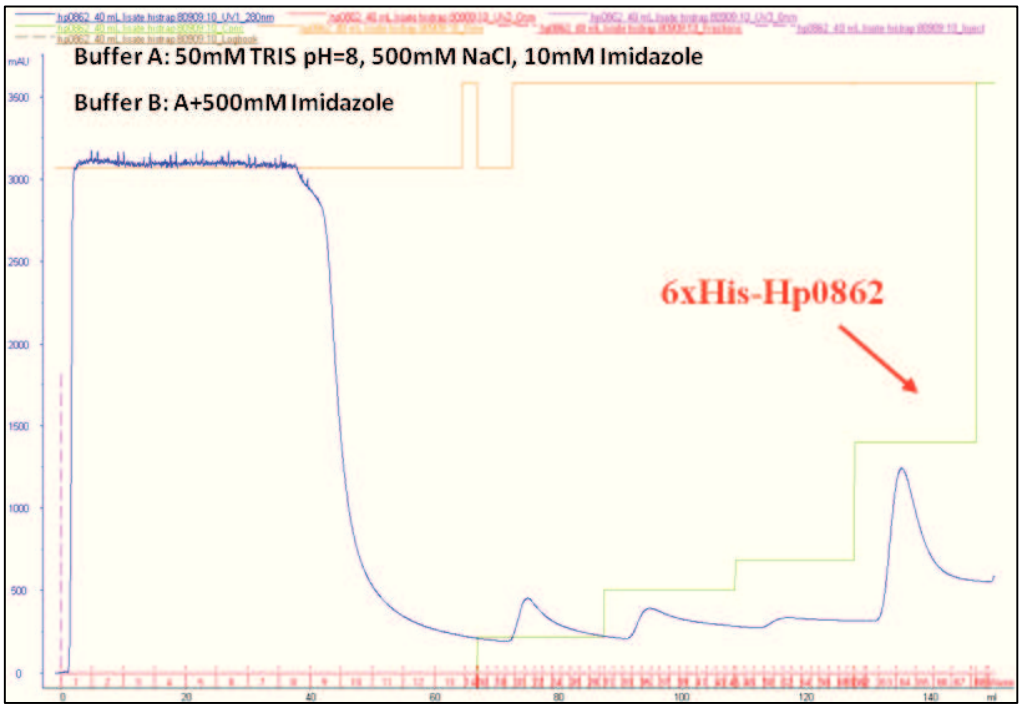

Fig. 3.13(A): Chromatogram from purification of *H. pylori* CoaX using Ni affinity chromatography. Hp0862 is detected in fractions 63 to 68, These fractions represent the 6xHis-Hp0862 peak marked by the arrow in the chromatogram. The Y-axis indicates the absorbance given in mAU (milli absorbance unit). The X-axis indicates the elution volume in ml.

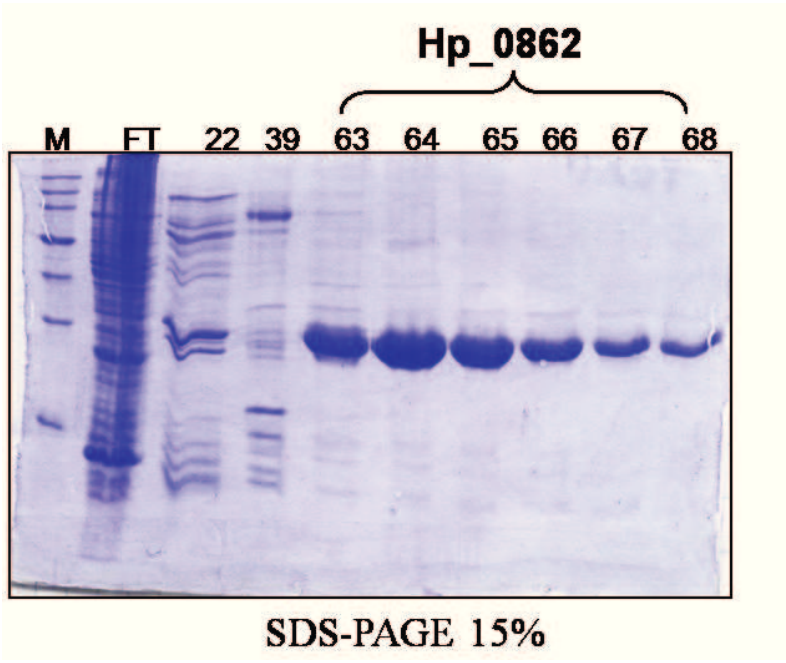

Fig. 3.13(B): SDS-PAGE analysis. the fractions. The brace indicates the fractions eluted from His-Trap column, containing the recombinant Hp_0862, corresponding to the size marker of 25kDa.

In order to obtain the pure HpCoax without its artificial tail, we have utilized the best conditions for the TEV cleavage, after a long protocol's screening. Digestion with TEV protease was performed at 4°C overnight, in the presence of DTT, directly setting the sample into the appropriate membrane for dialysis (Fig. 3.14). After 16h we can observe from the gel, that is present a lower molecular weight band which confirms that the Tag-cleavage was performed successfully. After this step, we have re-loaded Hp_0862 cleaved on His trap column and concentrated the sample for the following analysis.

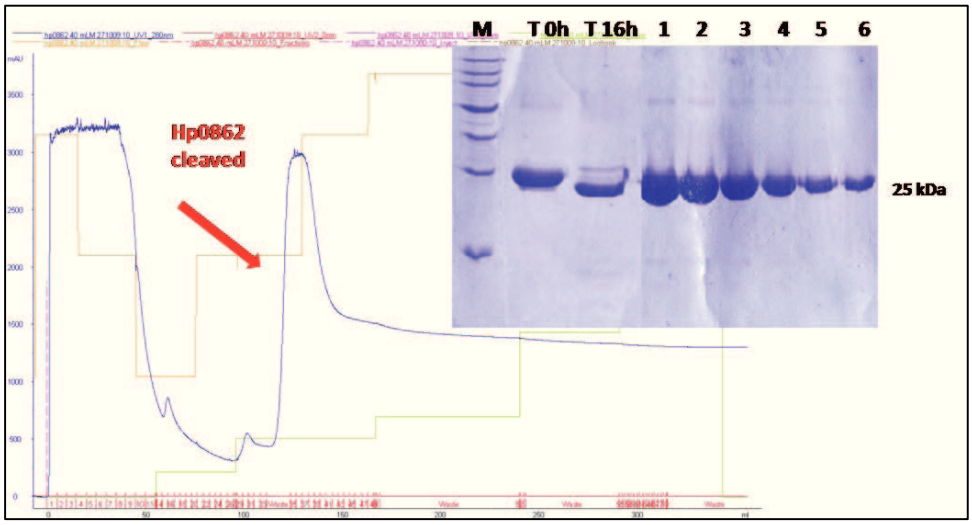

Fig. 3.14: Histrap profile and SDS-PAGE analysis to verify the correct digestion of Hp_0862. After 16h the band on the gel, is lower than the band at the time 0 (before the cleavage)

3.2.1 Sequence analysis of type III PanK, HpCoaX

A comparison among various Pank III sequences has revealed that Hp_0862 contains a high amount of Cys residues as showed in figure 3.15 (3,6%). The Cys presence within a sequence can interfere with the protein stability.

1 MPARQSFTDLKNLVLCDIGNTRIHFAQNYQLFSSAKEDLKRLGIQKEIFYISVNEENEK
ALLNCYPNAKNIAGFFHLETDYVGLGIDRQMACLAVNNGVVVDAGSAITIDLIKEGKHL
GGCILPGLAQYIHAYKKSAKILEQPFKALDSLEVLPKSTRDAVNYGMVLSVIACIQHLA
KNQKIYLCGGDAKYLSAFLPHSVCKERLVFDGMEIALKKAGILECK 223

Fig. 3.15: Amino acid sequence of Hp_0862. 8 cysteines are highlighted in red.

Indeed during the purification steps, we have observed a different behavior of the HpCoaX samples, depending on temperature of the storage and also on the presence of a reducing agent. We can observe the possible presence of non native S-S bonds which interfere with the protein stability (Fig. 3.16).

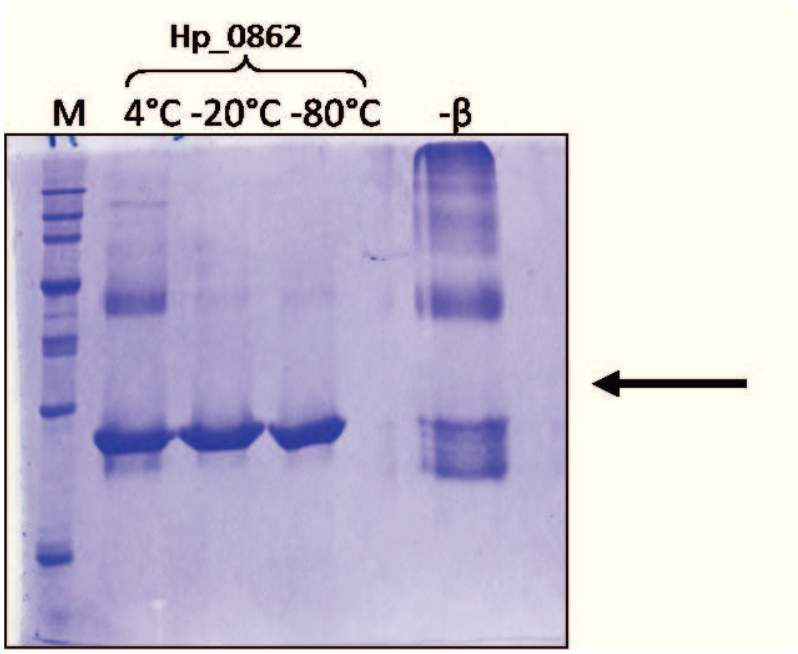

Fig. 3.16: Protein sample is stored at various temperature: 4°C. -20°C and -80°C. In the absence of the reducing agent, SDS-PAGE shows multiple bands pattern as indicate by the black arrow on the right.

To obtain a stable protein sample, it was necessary the treatment with a reducing agents (use of 2-3 mM DTT) during each purification steps. To limit the aerobic oxidation, both of DTT and

of Cys residues present in the sample, the protein preparation and the final step of storage were carried out under N_2 atmosphere. This specific wariness is important to control the aggregates formation and also to obtain a stable sample, in particular to investigate on the protein oligomerization state and to carry out crystallization trials.

3.2.1.2 Free SH- groups analysis

To confirm that the height cysteines present in the amino acid sequence of HpCoaX were not involved into any intra-molecular disulphide bridge, we have carried out a 4-Vinyl-pyridine assay. Hp_0862 was denatured and then alkylated with the 4-VP. This step is important for the incorporation of 4-VP molecule into the protein. To visualize the molecular mass increase and so to determine the effectively linkage between the Cys residues and the 4-VP, the samples were submitted to a LC-MS analysis. Looking at the Figure 3.17 we can observe the real increase of molecular mass: 24.971 Da+ [105x8].

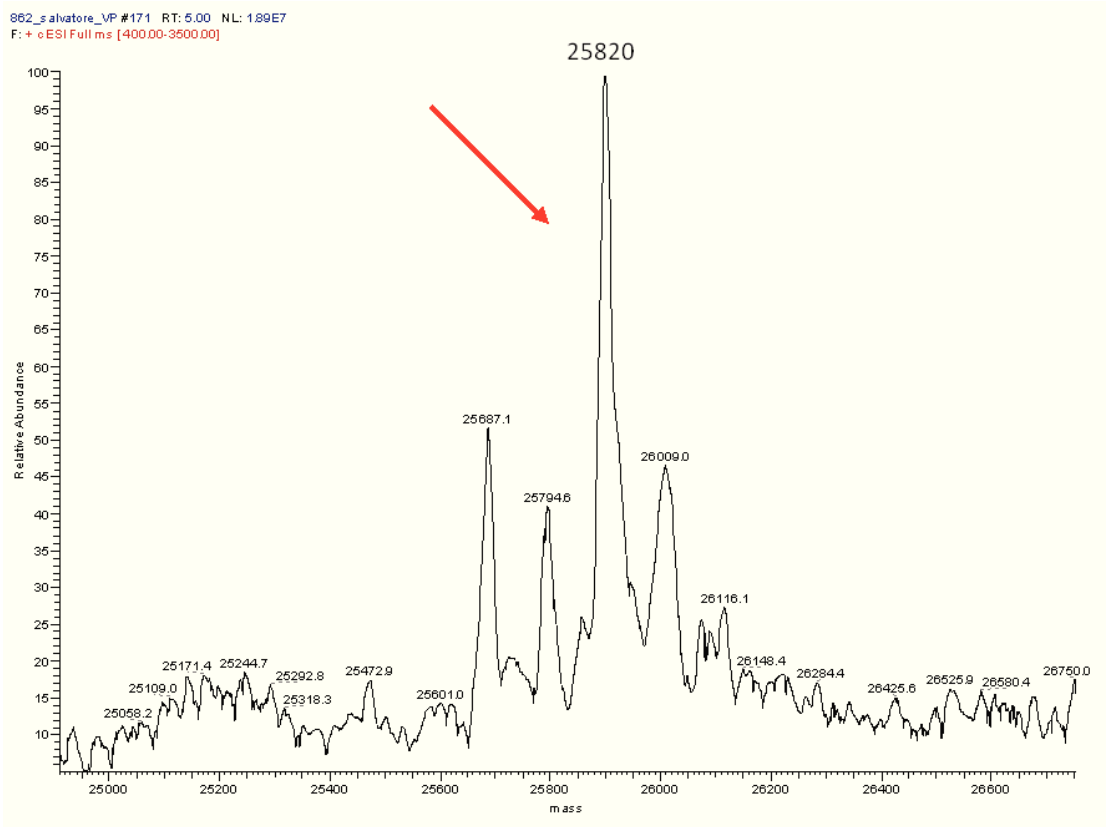

Fig. 3.17: LC-MS spectra of Hp0862

3.2.3 Biochemical characterization of Type III PanK, HpCoaX

Many biochemical and biophysical methods can be used to characterize the oligomerization state of proteins. Accurate information about the quaternary structure of HpCoaX protein was obtained combining molecular exclusion chromatography with light scattering analysis. The purified HP_0862 was analyzed by gel filtration chromatography and dynamic light scattering to investigate its state of oligomerization. The Gel filtration was performed using a Superdex_200 column, pre-equilibrated in the buffer containing 50 mM Tris pH 8.0, 200 mM NaCl and 2 mM DTT, and calibrated with the reference standards at V_e known. The PanKIII from *H. pylori* reports a V_e of 15.2 mL/min, which corresponds to a molecular weight of about 50 kDa, comparable to that expected for a dimeric organization (Fig. 3.18). This result is in according to the homology prediction reported for the other PanK enzymes which function as a dimer for their catalysis.

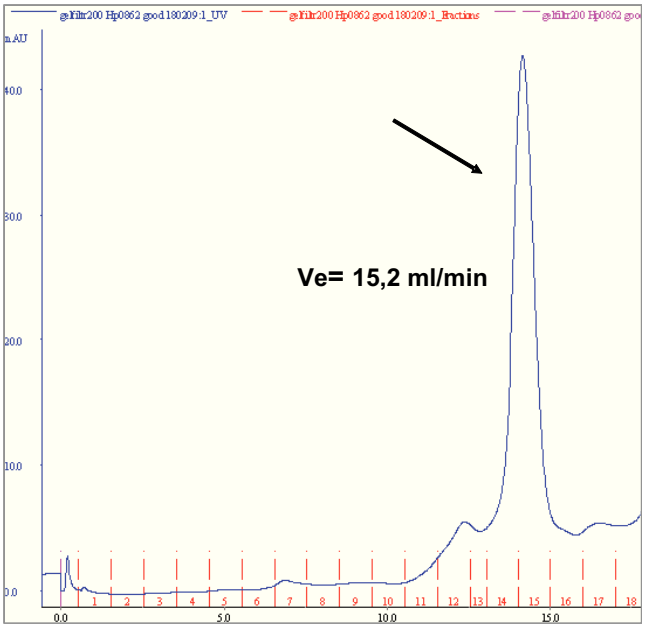

Fig. 3.18: Gel filtration profile of Hp0862 obtained by size exclusion chromatography**.** Results show that HP0862 exists as a homodimer, as predicted by its homology with the other PanK III bacterial proteins.

To confirm this data the purified protein was also submitted to an accurate analysis of DLS (Dynamic Light Scattering), using the same column Superdex_200 and the same running buffer. The figure 3.19 shows a SEC-MALS profile. The polydispersity, expressed as Mw/Mn, is about of 1.001(with an error of 4%). The Molar mass moments (g/mol) is Mn 4.790e+4 (error 3%) and Mp is 4.820e+4(error 2%).
The protein identity was assessed by liquid chromatography and mass spectrometry: data give a 24.971Da molecular mass of the recombinant HpCoaX, which is in good agreement with the theoretical molecular mass calculated according to the amino acid sequence.

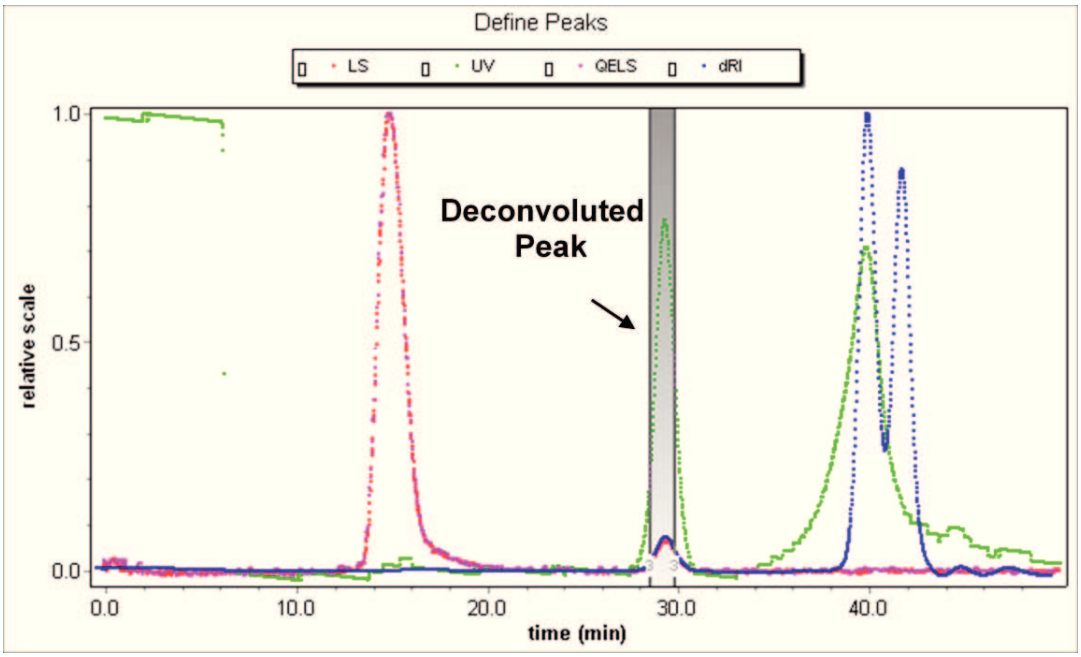

Fig. 3.19(A): SEC-time course of Hp0862 sample. The black arrow shows the deconvoluted peak corresponding to protein area signals.

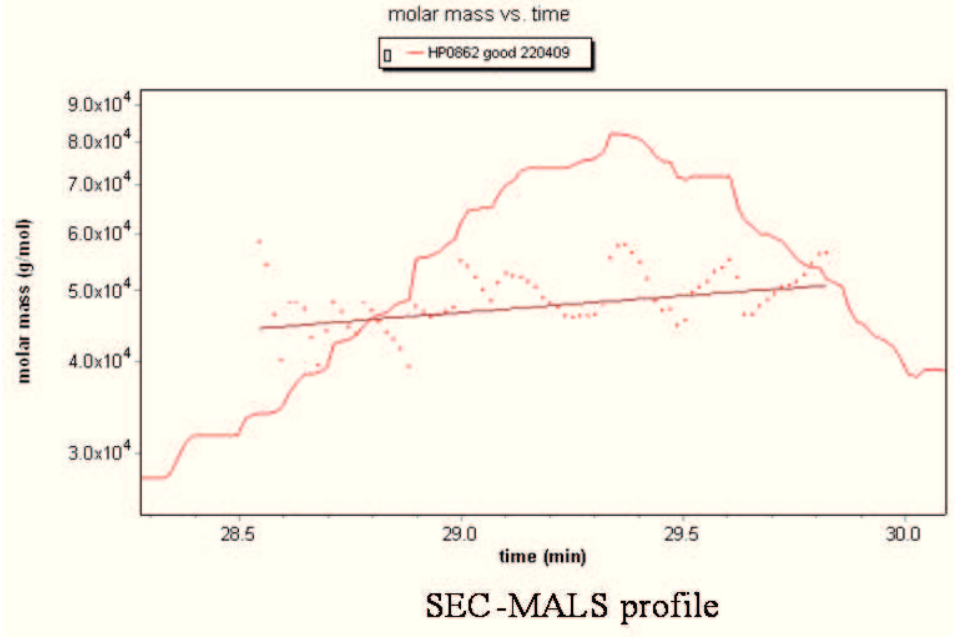

Fig. 3.19(B): SEC-MALS profile obtained from derivation calculus. Molar mass plot versus time.

3.2.4 Crystallization experiment

The crystallization conditions were performed as described in 2.5.9 section. Several crystal forms were obtained only at 4°C, not at 20°C, confirming the intrinsic protein difficulty to remain in a stable form. Only diffraction data at low resolution were obtained from these crystals (about 5 Å).
In order to improve the crystal quality we have also carried out the crystallization experiment in presence of Pan and AMPNP. Unfortunately, also these crystals diffracted a 5.0 Å of resolution, not enough to solve the structure

3.3 *H. pylori* PanK III, a new target for antimicrobical agents

There are several functional advantages that may be conferred by oligomerization and that may have been selected for through evolution. The study of protein oligomerization can provide insights into the early protein environment and the evolution of modern proteins. So, the advises from the HpCoaX oligomerization state represent a good starting point for future investigations among this PanK enzyme. Belonging to the Type III class, HpCoax posses indeed specific features in the mode of binding the substrate and also exhibits particular characteristics in its catalytic action. The Type III is drastically distinct from the Type II class in the substrate binding and dimer formation modes; although the SaCoaA (Type II) and PaCoaA (Type III) monomers have very similar structures, they each contain a loop region that is absent from the other enzyme, which results in their assembly into dimers with distinctly different architectures. Although both loops contribute to the monomer-monomer interface, they do so in different ways. Due to the different structures of the insertion elements which comprise a large portion of the dimer interface in both PanK-II and PanK-III, the relative arrangement of two monomers in the dimer are very different for the two enzymes, resulting in distinct and mutually exclusive Pan-binding pockets. [Hong B.S., *et al*. 2006], [Yang K., *et al*. 2009].
The alarming rise of antibiotics resistance among key bacterial pathogens is stimulating an urgent need to discover novel antibacterial agents acting on new drug targets. Fortunately, the accomplishment of *H. pylori* genome-sequencing project [Tomb G.F., *et al* . 1997] has heralded a new era for antibacterial chemotherapy against the pathogenic bacterium. At the same time, comparison of bacterial target genes with human genes will also be necessary because, to avoid adverse effects, a good antimicrobial drug target should have no homolog in mammalian cells. Clearly, alternative strategies are needed to develop new types of specific *H. pylori* inhibitors that target PanK enzymes in general, and the type III PanKs that occur in a wide range of pathogenic bacteria in particular. I hope that further research into HpCoaX protein will be able to answer some of these burning questions.

Chapter IV

Bibliography

Appelmelk B.J., Simoons-Smit In.A., Negrini R., Moran A.P., Aspinall G.O., Forte J.G., De Vries T., Hu Quan, Verboom T., Maaskant J.J., Ghiara P., Kuipers E.J., Bloemena E., Tadema T.M., Townsend R.R., Tyagarajan K., Crothers J.M., Monteiro M.A., Savio A. And De Graff J. "Potential Role of Molecular Mimicry between *Helicobacter pylori* Lipopolysaccharide and Host Lewis Blood Group Antigens in Autoimmunity" Infection and Immunity, 1996; **64:** 2031–2040.

Amoresano, A., Pucci, P., Duro, G., Colombo, P., Costa, M.A., Izzo, V., Lamba, D. and Geraci, D. "Assignment of disulphide bridges in Par j 20101, a major allergen of Parietaria judaica pollen." Biol. Chem. 2003; **384:** 1165–1172.

Amoresano, A., Orru, S., Siciliano, R.A., De Luca, E., Napoleoni, R., Sirna, A. and Pucci, P. "Assignment of the complete disulphide bridge pattern in the human recombinant follitropin beta-chain." Biol. Chem. 2001; **382:** 961–968.

Baddiley J., Thain E.M., Novelli G.D., and Lipmann F. "Structure of coenzyme A." Nature.1953; **171:** 76.

Bizzozero G. Ueber die schlauchformigen drusen des magendarmkanals und die bezienhungen ihres epithels zu dem oberflachenepithel der schleimhaut. Arch. Mikr Anat. 1893; **42:** 82.

Blaser M.J.,Perez G.I., Kleanthous H., Cover T.L., Peek R.M., Chyou P.H., Stermmemman G.M. and Nomura A. "Infection with *Helicobacter pylori* Strains Possessing cagA Is Associated with an Increased Risk of Developing Adenocarcinoma of the Stomach." Cancer Research. 1995; **55**: 2111-2115.

Brand L.A., and Strauss E. "Characterization of a New Pantothenate Kinase Isoform from Helicobacter pylori. "J.Biol.Chem. 2005; **21:** 20185-20188.

Bravo L.E, Doom LJ, Realpe JL, Correa P. "Virulence associated genotypes of *Helicobacter pylori*: do they explain the African enigma?" Am J Gastroenterol. 2002; **97**:2839–42.

Cheek S., Zhang H., and Grishin N.V. "Sequence and structure classification of Kinases." J. Mol. Biol. 2002; **320**: 855-881.

Choudhry A.E., Mandichak T.L., Broskey J.P., Egolf R.W., Kinsland C., Begley T.P., Seefeld M.A., Ku T.W., Brown J.R.,Zalacain M., and Ratnam K. "Inhibitors of pantothenate kinase: Novel antibiotics for staphylococcal infections." Antimicrob. Agents Chemother. 2003; **47**: 2051-2055.

Coghlan, J.G., Gilligan, D., Humphries, H., Kc Kenna, D., Dooley, C., Sweeney, E., Keane, C., O'Morain, C. "Campylobacter pylori and recurrence of duodenal ulcers - a 12 month follow-up study." Lancet 1987; **Ii**: 1109-1111.

Covacci A, Telford JL, Del Giudice G, Parsonnet J, Rappuoli R. "*Helicobacter pylori* virulence and genetic geography." Science 1999; **284**:1328–33.

Cun S., Li H., Ge R., Lin M.C.M. and Sun H. "A Histidine-rich and Cysteine-rich Metal-binding Domain at the C Terminus of Heat Shock Protein A from Helicobacter pylori." J. Biol. Chem, 2008; **283**: 15142-15151.

Eaton K.A., Brooks C.L., Morgan D.R., and Krakowka S. "' Essential Role of Urease in Pathogenesis of Gastritis Induced by Helicobacter pylori in Gnotobiotic Piglets". Infection and Immunity, 1991, **59:** 2470-2475.

Elvin, C.M., Thompson, P.R., Argall, M.E., Hendry, P., Stamford, N.P., Lilley, P.E., Dixon, N.E. "Modified bacteriophage lambda promoter for overproduction of proteins in *E. coli*.". Gene. 1990; **87:** 123-126.

Emsley, P. and Cowtan, K. "Coot: model-building tools for molecular graphics." Acta Crystallogr. D Biol. Crystallogr. 2004; **60:** 2126–2132.

Gebert, B., W. Fischer, et al. "*Helicobacter pylori* vacuolating cytotoxin inhibits T lymphocyte activation." Science 2003; **301:** 1099-1102.

Goodwin, C.S., "*Helicobacter pylori* gastritis, peptic ulcer and gastric cancer: clinical and molecular aspects." Clin. Infect. Dis. 1997; **25**: 1017-1019.

Graham, D.Y., Lew, G.M., Evans, D.G., Evans, Jr D.J., Klein, P.D. "Effect of triple therapy (antibiotics plus bismuth) on duodenal ulcer healing. A randomized controlled trial." Ann. Intern. Med. 1991. **115:** 266-269.

Hong B.S.,Yun M.K., Zhang Y.M., Chohnan S., Rock C.O., White S.V., Jackowski S. Park H.W., and Leonardi R. "Prokaryotic Type II and Type III Pantothenate Kinases: The Same Monomer Fold Creates Dimers with Distinct Catalytic Properties." Structure. 2006; **14:** 1251-1261.

Horwich A.L, Farr G.W. and Fenton W.A. "GroELGroES-Mediated Protein Folding." Chem. Rev. 2006; **106:** 19171930.

Horwich, A.L., Fenton, W.A., Chapman, E. and Farr, G.W. "Two families of chaperonin: physiology and mechanism.". Annu. Rev. Cell Dev. Biol. 2007; **23:** 115–145.

Jackowski S., and Rock C.O. “Regulation of coenzyme A biosynthesis.” J. Bacteriol. 1981; **148:** 926-932.

Kansau I., Guillain F., Thiberge J-M., and Labigne A. “ Nickel binding and immunological properties of the C-terminal domain of *Helicobacter pylori* GroES homologue (HspA).” Mol Microbiol. 1996; **22**: 1013-23.

Lara L.F, Cisneros G, Gurney M, et al. “One-day quadruple therapy compared with 7-day triple therapy for *Helicobacter pylori* infection.” Arch Intern Med 2003; **163:** 2079–84.

Leonardi R., Zhang Y.-M., Rock C.O., and Jackowski S.” Coenzyme A: Back in action.” Progr.Lipid Res. 2005; **44:** 125-153.

Loguercio S., Dian C., Flagiello A., Scannella A., Pucci P., Terradot L., Zagari A. “In HspA from *Helicobacter pylori* vicinal disulfide bridges are a key determinant of domain B structure**.”** FEBS Letters 2008; **582**: 3537-3541.

Mahdavi, J., B. Sonden, *et al*. "*Helicobacter pylori* SabA adhesin in persistent infection and chronic inflammation." Science. 2002; **297:** 573-578.

Marshall, B., 2002. Helicobacter pioneers: Firsthand Accounts from the Scientists who discovered Helicobacters 1892-1982. Wiley-Blackwell.

Marshall BJ, Warren JR. “Unidentified curved bacilli in the stomach of patients with gastritis and peptic ulceration.” Lancet. 1984; **1**:1311–5.

Mobley, H.L.T., Mendz, G.L., Hazell, S.L., 2001. *Helicobacter pylori*: Physiology and Genetics. AMS Press
(Available as e-book at http:/www.ncbi.nlm.nih.gov/books/bv.fcgi?rid=hp).

Morris, H.R. and Pucci, P. “A new method for rapid assignment of S–S bridges in proteins.” Biochem. Biophys. Res. Commun. 1985; **126**: 1122–1128.

Odenbreit, S., H. Kavermann, *et al*. "CagA tyrosine phosphorylation and interleukin-8 induction by Helicobacter pylori are independent from alpAB, HopZ and bab group outer membrane proteins." Int J Med Microbiol. 2002; **292**: 257-266.

Phadnis S.H., Parlow M.H., Levy M, Ilver D., Caulkins C.M., Connors J.B. and BRUCE Dunn B.E. “Surface Localization of Helicobacter pylori Urease and a Heat Shock Protein Homolog Requires Bacterial Autolysis.” Infection and Immunity, 1996; **64:** 905–912.

Rauws, E.A., Tytgat, G.N., “Cure of dudenal ulcer associated with eradication of Helicobacter pylori.” Lancet 1990; **335:** 1233-1235.

Rock C.O., Karim M.A., Zhang Y., and Jackowski S." The murine pantothenate kinase (Pank1) gene encodes two differentially regulated pantothenate kinase enzymes." Gene. 2002; **291**: 35-43.

Rothenbacher D, Brenner H. "Burden of *Helicobacter pylor*i and *H. pylori*-related diseases in developed countries: recent developments and further implications." Microbes Infect. 2003; **5:** 693-703.

Sambrook, J., Russel, D., 2001. Molecular Cloning, A Laboratory Manual, 3rd edition.

Sauerbaum, S., Josenhans, C., *Helicobacter pylori* evolution and phenotypic diversification in a changing host. Nature rev. microbiol. 2007; **5**: 441-452.

Sauerbaum, S., Thiberge, J.M., Kansau, I., Ferrero, R.L. and Labigne, A. "*Helicobacter pylori* hspA-hspB heat-shock gene cluster: nucleotide sequence, expression, putative function and immunogenicity." Mol. Microbiol. 1994; **14**: 959–974.

Song W.J., and Jackowski S. "Cloning, sequencing and expression of the pantothenate kinase (coaA) gene of *Escherichia coli*." J.Biol.Chem. 1992; **174:** 6411-6417.

Sun, X., Ge, R., Chiu, J.F., Sun, H. and He, Q.Y. " Identification of proteins related to nickel homeostasis in *Helicobater pylori* by immobilized metal affinity chromatography and two-dimensional gel electrophoresis." Met. Based Drugs. 2008; 289490.

Svennerholm, A.-M. and Lundgren, A. "Progress in vaccine development against *Helicobacter pylori*." FEMS Immunology & Medical Microbiology. 2007; **50:** 146–156.

Szabo` I., Brutsche S., Tombola F.,Moschioni M., Satin B., Telford J.L., Rappuoli R., Montecucco C., E., and Zoratti M. **"**Formation of anion-selective channels in the cell plasma membrane by the toxin VacA of *Helicobacter pylori* is required for its biological activity". The EMBO Journal, 1999; **18**: 5517–5527.

Terpe, K. " Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems." Appl Microbiol. Biotechnol. 2006; **72:** 211-222.

Tomb J.F., *et al.* "The complete genome sequence of the gastric pathogen *Helicobacter Pylori*." Nature. 1997; **388:** 539–547.

van Gunsteren, W. F., S. R. Billeter, *et al*. (1996). Biomolecular Simulations: The GROMOS96 Manual and User Guide. Zürich, VdF Hochschulverlag ETHZ.

Vanet A. and Labigne A. “Evidence for Specific Secretion Rather than Autolysis in the Release of Some *Helicobacter pylori* Proteins.” Infection and Immunity. 1998; **66:**1023–1027.

Yang K., Strauss E., Huerta C., and Zhang H. “Structural Basis for Substrate Binding and the Catalytic Mechanism of Type III Pantothenate Kinase. “ Biochemistry. 2008; **47**: 1369-1380.

List of publications inherent to the Doctorate research activity of Alessandra Scannella

Publications:

- Loguercio S., Dian C., Flagiello A., Scannella A., Pucci P., Terradot L., Zagari A. **In HspA from *Helicobacter pylori* vicinal disulfide bridges are a key determinant of domain B structure.** FEBS Letters Epub 2008 sep. 19; **582**: p. 3537-3541.

Communications:

- Scannella Alessandra, Dathan Nina, Monti Simona M., Terradot Laurent and Adriana Zagari "Biochemical characterization of HP0862, a new Pantothenate Kinase isoform from *Helicobacter pylori.*"P57 on Abstract's book. *12th Naples Workshop on Bioactive Peptides, Naplesi 4-7 June 2010*

- Scannella Alessandra, Dathan Nina, Loguercio Salvatore, Monti Simona, Pucci Piero, Flagiello Angela, Terradot Laurent and Adriana Zagari. "HspA, the GroES homologue from *Helicobacter pylori*, is able to bind nickel ions." Page 37, P2 on abstract's book. *School of National Bioinorganic chemistry. Naples 14-16 September 2008.*

- Scannella Alessandra, Dathan Nina, Loguercio Salvatore, Monti Simona, Pucci Piero, Flagiello Angela, Terradot Laurent and Adriana Zagari. "Singolar properties of HspA, the GroES homologue from *Helicobacter pylori.***"** Page 84, P31 on abstract's book.*11th Naples Workshop on Bioactive Peptides, Naples 24-27 May 2008.*

Research activity in Scientific Institutions abroad

- From October to December 2008, the research activity of Alessandra Scannella has been carried out at the European Synchrotron Radiation Facility (ESRF), Grenoble-Cedex, France.

FEBS Letters 582 (2008) 3537–3541

In HspA from *Helicobacter pylori* vicinal disulfide bridges are a key determinant of domain B structure

Salvatore Loguercio[a,b], Cyril Dian[b], Angela Flagiello[c], Alessandra Scannella[a], Piero Pucci[c,d], Laurent Terradot[b,*], Adriana Zagari[a,c,*]

[a] *Department of Biological Sciences and CNISM, University of Naples "Federico II", Via Mezzocannone 16, I-80134 Naples, Italy*
[b] *European Synchrotron Radiation Facility, BP 220 F-38043, Grenoble Cedex 9, France*
[c] *CEINGE – Biotecnologie Avanzate Scarl, Naples, Italy*
[d] *Department of Organic Chemistry and Biochemistry, University of Naples "Federico II", Naples, Italy*

Received 18 August 2008; accepted 10 September 2008

Available online 19 September 2008

Edited by Stuart Ferguson

Abstract ***Helicobacter pylori*** **produces a heat shock protein A (HspA) that is unique to this bacteria. While the first 91 residues (domain A) of the protein are similar to GroES, the last 26 (domain B) are unique to HspA. Domain B contains eight histidines and four cysteines and was suggested to bind nickel. We have produced HspA and two mutants: Cys94Ala and Cys94Ala/Cys111Ala and identified the disulfide bridge pattern of the protein. We found that the cysteines are engaged in three disulfide bonds: Cys51/Cys53, Cys94/Cys111 and Cys95/Cys112 that result in a unique closed loop structure for the domain B.**

Keywords: Small heat shock protein; Chaperonin GroES; Nickel binding protein; Mass spectrometry; *Helicobacter pylori*

1. Introduction

Helicobacter pylori is a gram negative bacteria that colonise the human stomach. It is present in more than half of the world's population and causes major diseases such as gastritis, peptic ulcers and stomach cancer [1]. To colonise this highly acidic niche, this neutrophile bacterium has developed peculiar strategies. Among them the production of the urease is the most important one [2,3]. Urease is a nickel-dependent enzyme that hydrolyses urea to produce ammonia, which allows the bacteria to survive at low pH [4,5]. The activity of urease strongly depends on the availability of nickel ions [6]. Other proteins such as Hpn, Hpn-like, NixA, Frpb4 and a heat shock protein, namely HspA, have been demonstrated to bind the metal playing the role of importing, providing or sequestering nickel ions [7–9].

Heat shock proteins (Hsp) synthesis is stimulated in response to abnormal environmental conditions like such as heat shock, oxidative stress, and also exposure to heavy metals. The proteins are found in all domains of life where they actually act as molecular chaperons [10]. A subgroup, referred as to small Hsp, shares common features such as small molecular mass (10–30 kDa).

The small heat shock protein A (HspA) from *H. pylori* is an essential protein with a double localization. Unexpectedly this cytoplasmic protein, has been localised also on the bacterium surface. Whether this protein is released through a specific secretion mechanism or a programmed bacterium autolysis is an interesting matter of debate [11,12].

Owing to its external localization, HspA presents strong antigenic properties [13–15]. A protective immunity from the infection was induced in mice after immunization with HspA. The protein has been then considered appealing as a vaccine component [13].

HspA consists of 118 amino-acids divided in two domains; the A domain (1–91) and the B domain (92–118). The domain A of the protein shares sequence similarity to the GroES sequence [13]. The domain B of HspA is unique to *H. pylori* and *Helicobacter acinonuchis* and contains eight histidines and four cysteines.

This unusual sequence plays a role in Ni uptake and release [9,13,15,16] to assist nickel homeostasis in vivo. The B domain does exhibit high affinity for nickel ions with dissociation constants of 1.8 μM [15] and 1.1 μM [16]. In vitro, the binding to nickel is pH dependent, being the nickel released at low pH ($pH_{1/2} = 3.8$) and was demonstrated by several complementary methods such as dialysis equilibrium, UV–Vis spectroscopy, ICP-MS mass [15,16].

Cys is the less represented amino-acid in the GroES protein family, because it does not serve for the co-chaperonin function. Vice versa, no substitution affects the His/Cys residues of HspA sequence even though several substitutions occur among clinical isolates, suggesting that these amino-acids are specifically involved in nickel homeostasis [9,15]. Therefore, the role played by the B domain should be mediated by a high number of cysteines with capacity of different redox state.

To get insight into Cys role, we have produced the wild-type HspA as well as two mutants: Cys94Ala and Cys94Ala/Cys111Ala. By combining biochemical methods with mass spectrometry, we have determined the Cys oxidation state, assigned the disulfide bridge pattern and provided evidence for a unique lasso-like structure of the B domain in HspA.

*Corresponding authors. Address: Department of Biological Sciences and CNISM, University of Naples "Federico II", Via Mezzocannone 16, I-80134 Naples, Italy. Fax: +39 081 2536 603 (A. Zagari).
E-mail addresses: terradot@esrf.fr (L. Terradot), zagari@unina.it (A. Zagari).

Abbreviations: HspA, heat shock protein A; IAM, iodoacetamide; DTT, dithiothreitol

0014-5793/$34.00
doi:10.1016/j.febslet.2008.09.025

2. Materials and methods

2.1. Expression and purification of HspA from H. pylori

The wild-type HspA was expressed and purified using the vector pILL948 described in [15]. Single (C94A) and double (C94A/C111A) HspA mutants were constructed using the QuickChange site directed mutagenesis protocol (Stratagene). *E. coli* BL21(DE3)pLysS, carrying the expression vector, was grown in LB medium at 37 °C, and the expression of the recombinant product was induced by 1 mM isopropyl-D-thiogalactopyranoside (IPTG). The cells were harvested by centrifugation, resuspended in lysis buffer A (20 mM Na_2HPO_4 pH 7.4, 500 mM NaCl and 1 mM $MgCl_2$) and disrupted by sonication. The soluble fraction was isolated from the cell debris by centrifugation and applied to a HiTrap HP column (Amersham Biosciences) charged with 100 mM $NiSO_4$ and equilibrated with buffer A. The protein was eluted from the column using a gradient of 0–100% buffer B (buffer A and 500 mM imidazole), and further purified by size exclusion chromatography (Superdex 200 (10/30); Amersham Biosciences) in a buffer (20 mM Tris pH 7.4, 200 mM NaCl).The purity and homogeneity was first tested by SDS–PAGE and then by mass spectrometry.

2.2. Mass spectrometry measurements

Reduction with DDT and alkylation with IAM were carried out according to [17]. HPLC-desalted samples of wt-HspA and mutants were directly analyzed by ES/MS using a Quattro Micro triple quadrupole mass spectrometer (Waters). Wt-HspA and both mutants samples were digested with trypsin in 50 mM ammonium acetate, pH 6.0 at 37 °C overnight, and the mixtures of tryptic peptides were directly analyzed by MALDI mass spectrometry using a Voyager DE PRO instrument (Applied Biosystem) operating both in linear and reflectron mode.

3. Results and discussion

3.1. Assessment of the oxidation state of HspA Cys residues

The oxidation state of the cysteines occurring in HspA and in its mutants was assessed by determination of the accurate molecular mass of the proteins by electrospray mass spectrometry (ESMS), following a procedure previously developed [17]. Aliquots of the wt-HspA and both mutants samples were directly analyzed by ESMS producing the spectra shown in Fig. 1. Wt-HspA exhibited a molecular mass of 12984.0 ± 0.8 Da, about 6 Da lower than expected for the fully reduced protein (theoretical mass value 12990.9), thus suggesting that the six Cys residues were involved in three disulfide bridges. Accordingly, the molecular mass of HspA shifted to 12989.9 ± 0.4 Da following reduction with DTT, exhibiting the expected increase of about 6 Da, whereas no difference in the mass value was observed following direct alkylation of

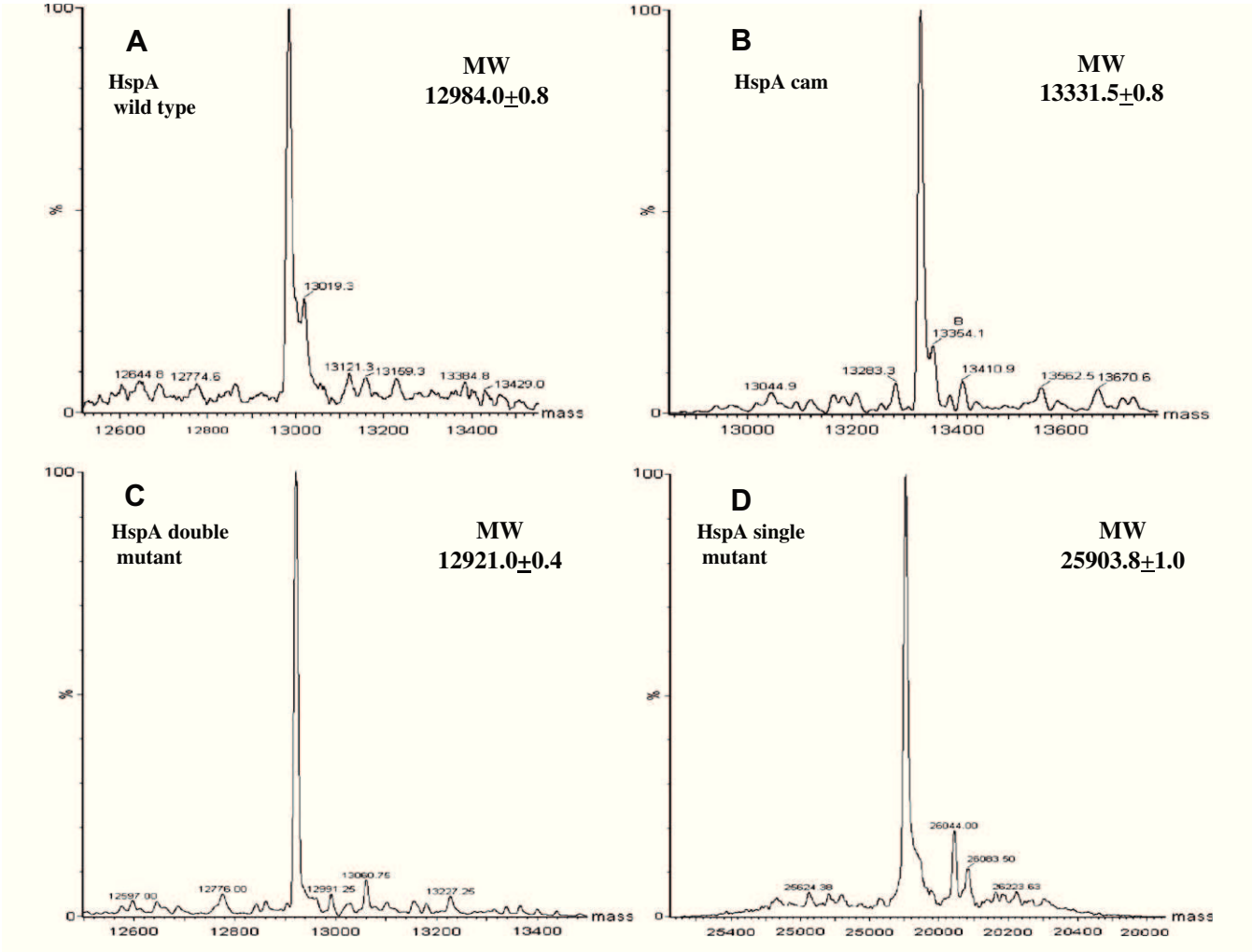

Fig. 1. ESMS spectra. (A) Wt-HspA; (B) HspA carboxymethylated at position 51, 53, 94, 95, 111, 112, after reduction with DDT; (C) C94A/C111A double mutant; (D) dimeric C94A mutant.

the protein with an excess of IAM, confirming the absence of free Cys residues (data not shown). It is worth noting that, after reduction, disulfides quickly reform in air. Finally, when the wt protein was first reduced with DTT and then carboxyamidomethylated with IAM, the corresponding ESMS spectrum exhibited a molecular mass of 13331.5 ± 0.8, as shown in Fig. 1B, with an increase of 340.6 Da, corresponding to the alkylation of six Cys residues.

Similar experiments were also carried out on the HspA mutants. Fig. 1C shows the ESMS spectrum of the Cys94Ala/Cys111Ala double mutant exhibiting a molecular mass of 12921.0 ± 0.4 Da, about 4 Da lower than expected for the fully reduced protein, suggesting the occurrence of two S–S bridges. This result was confirmed by the reduction and reduction/alkylation experiments analogous to those described above for wt-HspA, indicating that the two mutated cystein residues, Cys94 and Cys111, were joined by a disulfide bond in the wt-HspA molecule. Finally, the ESMS analysis of the Cys94Ala single mutant yielded a mass value of 25903.8 ± 1.0 Da, as shown in Fig. 1D, corresponding to a dimeric species stabilised by an intermolecular disulfide bridge likely involving the free Cys111 residue of each HspA molecule.

3.2. Disulfide bridge pattern

Assignment of the disulfide bridges pattern in both wt-HspA and the double mutant was accomplished according to the well established mass mapping strategy [17–20]. Wt-HspA was digested with trypsin at pH 6.0 to avoid scrambling of S–S bridges and the resulting peptide mixture was directly analyzed by MALDI-MS. Fig. 2A shows the high mass region of the resulting MALDI spectrum recorded in linear mode where a series of mass signals was attributed to S–S bridged peptides. The peaks at *m/z* 4234.8 and 3969.6 were interpreted as arising from the peptide 78–103, originated by aspecific cleavage at Tyr77, joined to the fragments 107–118 and 107–116, respectively by two disulfide bridges involving the four Cys residues at positions 94, 95, 111 and 112. Both mass signals were accompanied by satellite peaks at about 335 Da higher (*m/z* 4571.8 and 4304.5) corresponding to the same peptide 78–103 linked to the fragments 104–118 and 104–116, respectively via the same two S–S bridges. Finally the mass signals at *m/z* 3891.1 and 3626.3 were attributed to the peptide 81–103 bridged to the fragments 107–118 and 107–116 by two disulfides involving the above mentioned four cysteines. All these signals disappeared following incubation with DTT.

The mass spectral investigation of wt-HspA provided the overall scheme of cysteine pairings with Cys94, 95, 111 and 112 originating a network of two disulfide bridges and Cys51 and Cys53 forming the third S–S bridge. This interpretation was supported by the weak mass signal at *m/z* 2404.3 observed in the MALDI reflectron spectrum and was further confirmed by the double mutant analysis (see below). This peak occurred 2 Da lower than the peptide 42–64, suggesting the presence of an intermolecular disulfide bridge joining Cys51 and Cys53.

The correct pairings of the Cys residues in HspA was obtained by the MALDI mass spectral analysis of the tryptic digest of the C94A/C111A double mutant. As shown in Fig. 2B, mutation of the two Cys residues seemed not to affect the C-terminal pattern of disulfide bridges. The mass signals at *m/z* 5344.7, 5053.7 (peptides 65–103 and 67–103 joined to 107–116, respectively) and 3907.7, 3564.6 and 3322.4 (peptides 78–103, 81–103 and 83–103 linked to the fragment 107–116) essentially corresponded to the previously observed S–S bridged peptides. This strongly suggests that Cys94 and

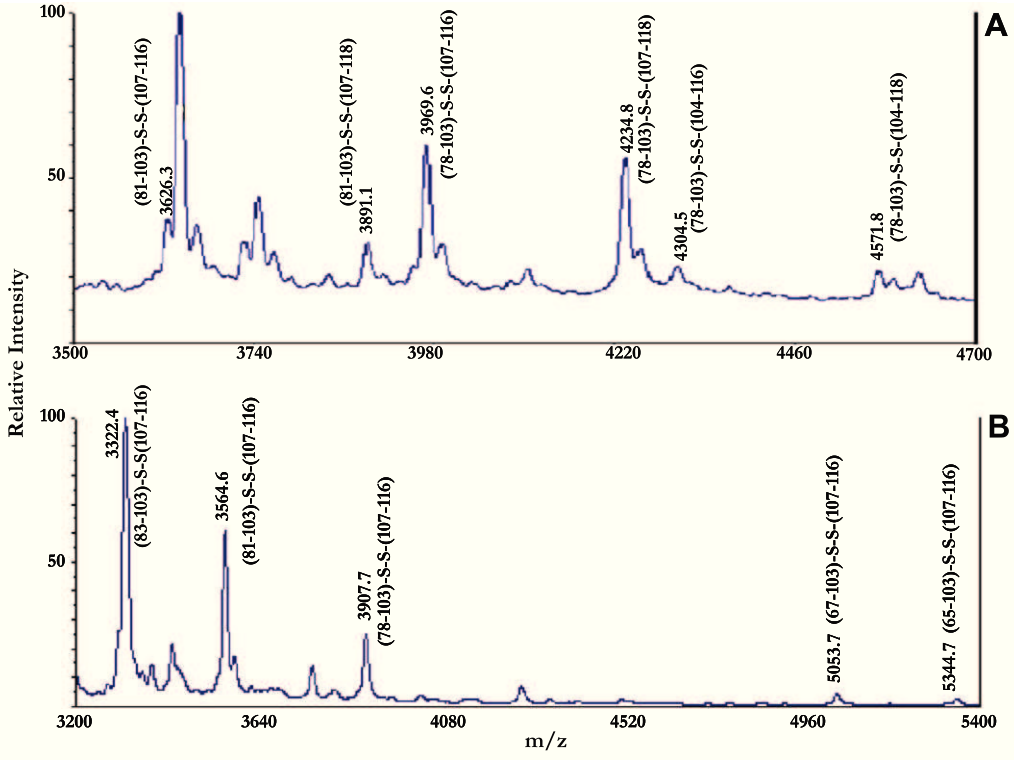

Fig. 2. MALDI-MS spectra after tryptic digestion of: (A) Wt-HspA; (B) C94A/C111A double mutant.

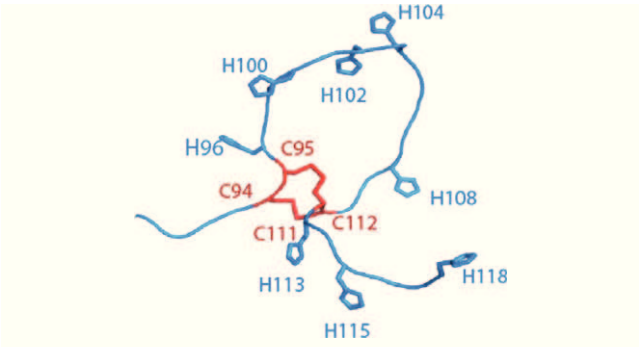

Fig. 3. Model of domain B (92-GSCCHTGNHDHKHAKE-HEACCHDHKKH-118) in HspA. The disulfide bridge pattern generating the lasso-like structure is indicated in red (Cys94-Cys111 and Cys95-Cys112). The side chains of cysteines and histidines are displayed in sticks. Coordinates were obtained by manual building in COOT [22].

Cys111 are held together by an S–S bridge in wt-HspA. The C-terminal fragments are then still joined by the second disulfide bond involving Cys95 and Cys112. The third S–S bridge involving Cys51 and Cys53 was confirmed by the mass signal at *m/z* 964.5 (peptide 47–55 containing an intramolecular disulfide bridge) leading to the unambiguous assignment of the S–S bridge pattern in HspA.

4. Conclusion

In this study, we found that HspA contains three intramolecular disulfide bridges. The bonds involve either the GroES-like domain A (Cys51–Cys53) or the nickel-binding domain B (Cys94–Cys111 and Cys95–Cys112). In the latter domain, the small 4-Cys ring we identified constitutes a strong and very rare structural determinant, being observed only in the conotoxin, a short peptide (pdb code 1wct) [21]. Indeed, in HspA the two consecutive S–S bridges form a rigid knot that dramatically affects the topology of the C-terminal domain generating a "lasso-like" structure, here described for the first time (Fig. 3). A large and flexible ring, formed by 18 residues, embodies five His residues leaving out of the lasso the neighboring three C-term His. This lasso-like structure would be ideal to position the ligand side chains in conformation suitable to bind up to two nickel ions [15,16]. The Ni binding role may be played by HspA either through intracellular or extracellular Ni sequestration, as the protein was also found in association with the outer membrane and as a subcellular material [11,12]. Along this line, it is tempting to speculate that HspA might switch from an oxidised to a reduced state depending on the environment. This holds for the cell surface, where the protein is embedded in an oxidizing environment and then possesses the S–S bridges; whereas, in cytoplasm, the disulfide bridges may exhibit a diverse stability, depending on the conditions (pH and Ni depletion or overload).

In conclusion, being the cysteines strongly conserved in different strains they may have the role to assist the Ni binding/release. Then the unique lasso-like structure, built-up from vicinal disulfide bridges is relevant to the Ni dependent functioning of HspA in *H. pylori*.

Acknowledgments: We are indebted to Prof. A. Labigne and Dr. H. De Reuse for communicating some of their results on HspA prior to publication and for providing us with the vector pILL948. We thank Ulrike Kapp for technical assistance with the purification of HspA. This work was supported by FIRB RBRN07BMCT and the ESRF.

References

[1] Covacci, A., Telford, J.L., del Giudice, G., Parsonnet, J. and Rappuoli, R. (1999) *Helicobacter pylori* virulence and genetic geography. Science 284, 1328–1333.
[2] Eaton, K.A., Brooks, C.L., Morgan, D.R. and Krakowka, S. (1991) Essential role of urease in pathogenesis of gastritis induced by *Helicobacter pylori* in gnotobiotic piglets. Infect. Immunol. 59, 2470–2475.
[3] Stingl, K. and De Reuse, H. (2005) Stayng alive overdosed: how does *Helicobater pylori* control urease activity? Int. J. Med. Microbiol. 295, 307–315.
[4] Bauerfeind, P., Garner, R., Dunn, B.E. and Mobley, H.L. (1997) Synthesis and activity of *Helicobacter pylori* urease and catalase at low pH. Gut 40, 25–30.
[5] Stingl, K., Altendorf, K. and Bakker, E.P. (2002) Acid survival of *Helicobacter pylori*: how does urease activity trigger cytoplasmic pH homeostasis? Trends Microbiol. 10, 70–74.
[6] van Vliet, A.H., Kuipers, E.J., Stoof, J., Poppelaars, S.W. and Kusters, J.G. (2004) Acid-responsive gene induction of ammonia-producing enzymes in *Helicobacter pylori* is mediated via a metal-responsive repressor cascade. Infect. Immunol. 72, 766–773.
[7] Bury-Moné, S., Thiberge, J.M., Contreras, M., Maitournam, A., Labigne, A. and De Reuse, H. (2004) Responsiveness to acidity via metal ion regulators mediates virulence in the gastric pathogen *Helicobacter pylori*. Mol. Microbiol. 53, 623–638.
[8] Schauer, K., Gouget, B., Carriere, M., Labigne, A. and de Reuse, H. (2007) Novel nickel transport mechanism across the bacterial outer membrane energized by the TonB/ExbB/ExbD machinery. Mol. Microbiol. 63, 1054–1068.
[9] Sun, X., Ge, R., Chiu, J.F., Sun, H. and He, Q.Y. (2008) Identification of proteins related to nickel homeostasis in *Helicobater pylori* by immobilized metal affinity chromatography and two-dimensional gel electrophoresis. Met. Based Drugs 2008, 289490.
[10] Horwich, A.L., Fenton, W.A., Chapman, E. and Farr, G.W. (2007) Two families of chaperonin: physiology and mechanism. Annu. Rev. Cell Dev. Biol. 23, 115–145.
[11] Vanet, A. and Labigne, A. (1998) Evidence for specific secretion rather than autolysis in the release of some *Helicobacter pylori* proteins. Infect. Immunol. 66, 1023–1027.
[12] Phadnis, S.H., Parlow, M.H., Levy, M., Ilver, D., Caulkins, C.M., Connors, J.B. and Dunn, B.E. (1996) Surface localization of *Helicobacter pylori* urease and a heat shock protein homolog requires bacterial autolysis. Infect. Immunol. 64, 905–912.
[13] Suerbaum, S., Thiberge, J.M., Kansau, I., Ferrero, R.L. and Labigne, A. (1994) *Helicobacter pylori* hspA-hspB heat-shock gene cluster: nucleotide sequence, expression, putative function and immunogenicity. Mol. Microbiol. 14, 959–974.
[14] Ferrero, R.L., Thiberge, J.M., Kansau, I., Wuscher, N., Huerre, M. and Labigne, A. (1995) The GroES homolog of *Helicobacter pylori* confers protective immunity against mucosal infection in mice. Proc. Natl. Acad. Sci. USA 92, 6499–6503.
[15] Kansau, I., Guillain, F., Thiberge, J.M. and Labigne, A. (1996) Nickel binding and immunological properties of the C-terminal domain of the *Helicobacter pylori* GroES homologue (HspA). Mol. Microbiol. 22, 1013–1023.
[16] Cun, S., Li, H., Ge, R., Lin, M.C. and Sun, H. (2008) A histidine-rich and cysteine-rich metal-binding domain at the C terminus of heat shock protein A from *Helicobacter pylori*: implication for nickel homeostatsis and bismuth susceptibility. J. Biol. Chem. 283, 15142–15151.
[17] Amoresano, A., Pucci, P., Duro, G., Colombo, P., Costa, M.A., Izzo, V., Lamba, D. and Geraci, D. (2003) Assignment of disulphide bridges in Par j 20101, a major allergen of *Parietaria judaica* pollen. Biol. Chem. 384, 1165–1172.
[18] Amoresano, A., Orru, S., Siciliano, R.A., De Luca, E., Napoleoni, R., Sirna, A. and Pucci, P. (2001) Assignment of the complete

disulphide bridge pattern in the human recombinant follitropin beta-chain. Biol. Chem. 382, 961–968.
[19] Galliano, M., Minchiotti, L., Campagnoli, M., Sala, A., Visai, L., Amoresano, A., Pucci, P., Casbarra, A., Cauci, M., Perduca, M. and Monaco, HL. (2003) Structural and biochemical characterization of a new type of lectin isolated from carp eggs. Biochem. J. 376, 433–440.
[20] Morris, H.R. and Pucci, P. (1985) A new method for rapid assignment of S–S bridges in proteins. Biochem. Biophys. Res. Commun. 126, 1122–1128.
[21] Rigby, A.C., Lucas-Meunier, E., Kalume, D.E., Czerwiec, E., Hambe, B., Dahlqvist, I., Fossier, P., Baux, G., Roepstorff, P., Baleja, J.D., Furie, B.C., Furie, B. and Stenflo, J. (1999) A conotoxin from *Conus textile* with unusual posttranslational modifications reduces presynaptic Ca^{2+} influx. Proc. Natl. Acad. Sci. USA 96, 5758–5763.
[22] Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. 60, 2126–2132.