Dissertation

submitted to the
Combined faculty for the Natural Sciences and for Matematics
of the Rupuerto-Carola University of Heidelberg, Germany
for the degree of
Doctor of Natural Sciences

Presented by

Debora Lucarelli
born in Roma, Italia

Oral examination: ……..

Structural Studies on Ferric Uptake Regulator Proteins from *Mycobacterium tuberculosis*

Referees:

Prof. Dr. Irmgard Sinning, University of Heidelberg, Heidelberg

Dr. Elena Conti, EMBL Heidelberg, Heidelberg

This thesis is dedicated to my mum

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

1 Å	10^{-10} m
bp	Base pairs
B. subtilis	*Bacillus subtilis*
DNA	Deoxyribonucleic acid
dNTP	2'-deoxynucleoside 5'-triphosphate
EC	*Escherichia coli*
EDTA	Ethylene-diamine-tetra-acetic acid
EXAFS	Extended X-ray Absorption Fine Structure
Fur	Ferric uptake regulator
HIV	Human immunodeficiency virus
IPTG	Isopropyl-β-D-thiogalactoside
kDa	1000 Dalton, 1Dalton = Mass of one twelfth of a carbon atom
keV	1000 electron Volt, 1eV=1.602x10^{-19} Joule
LB	Luria Bertani
MAD	Multiple Wavelength Anomalous Diffraction
ml	Milliliter
μl	Microliter
mg	Milligram
μg	Microgram
mM	Millimolar
μM	Micromolar
Mtb	*Mycobacterium tuberculosis*
NiNTA	Nichel nitrilotriacetic
PA	*Pseudomanas Aeruginosa*
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
PDB	Protein Data Bank
PEG	Polyethylene glycole
SAD	Single Wavelength Anomalous Diffraction
SDS	Sodium dodecyl sulfate
Tris	Tris-(hydroxymethil)-aminoethane
UV	Ultraviolet
XANES	X-ray Absorption Near Edge Spectroscopy
Znu	Zn^{2+} uptake
Znt	Zn^{2+} transport
Zur	Zinc uptake regulator

PREFACE

This thesis includes the research work carried out at the European Molecular Biology Laboratory (EMBL), Hamburg outstation (Germany). The results collected in the last three and a half years are here reported. The project is focussed on structural studies on ferric uptake regulator proteins, named as FurA and FurB, from *Mycobacterium tuberculosis.* In order to achieve this purpose several methodologies, including EMSA (electrophoretic mobility shift assay), microPIXE (proton induced X-ray emission), ITC (isothermal titration calorimetry), EXAFS (extended X-ray absorption fine structure) and X-ray crystallography, have been used. In this thesis a description of these techniques and of relevant protocols is described. It was possible to distinguish different metal sites for each of the proteins and elucidate their biological functions. Furthermore the crystal structure of FurB has been solved and here reported.

The thesis is organised in seven chapters.
Chapter 1 introduces the reader to the topic. It starts with a brief description of metal ion homeostasis, an introduction to *Mycobacterium tuberculosis* and an overview of the importance of iron and zinc ions. The chapter continues describing these cations' regulation towards the more specific topic of the ferric and zinc uptake regulators.
Chapter 2 describes the materials and methods used in this project.
Chapter 3 and 4 present the experimental results for FurA and FurB respectively.
Chapter 5 and 6 collect discussions and conclusions for FurA and FurB respectively.
Chapter 7 summarizes my personal consideration of this project.

This thesis has been submitted for the degree of Doctor of Philosophy (Ph.D.) to the combined Faculties for the Natural Sciences and for the Mathematics of the Ruperto-Carola University of Heidelberg, ……2006

EMBL, Hamburg ……2006
Debora Lucarelli

ZUSAMMENFASSUNG

Eisen ist das häufigste Spurenelement im menschlichen Körper, Zink das zweithäufigste. Die Kontrolle der Ionenhomeostase ist von überlebenswichtiger Bedeutung für alle Organismen. Die Regulierung des Ionenflusses in oder aus der Zelle ist ein komplexer und und steng regulierter Mechanismus, der noch weitestgehend ungeklaert ist.

Der hochspezialisierte Krankheitserreger *Mycobacterium tuberculosis* muss sich mit der Sequestration von Eisen begnügen, um im menschlichen Körper überleben zu können. Der Eisenstoffwechsel wird durch die transktriptionelle Regulation von Genen kontrolliert, die mit der Aufnahme, dem Transport und der Speicherung von Eisen zusammenhängen. Der Mangel dieses Metalls löst eine weitreichende Reaktion hin zur erhöhten Aufnahme von Eisen aus, während ein Überschuss für die Zelle toxisch sein kann. Die Kontrolle der intrazellulären Eisenkonzentration ist mit anderen wichtigen Prozessen wie der Reaktion auf oxidativen Stress und der Regulation von Virulenzfaktoren. Neuere Studien haben gezeigt, dass bei TB/HIV-Patienten die hohe Mengen diätisches Eisen zu sich nehmen das Risiko einer akuten offenen Tuberkulose ansteigt. In *M. tuberculosis* wird der Eisenaufnahmeregulator A (engl. ferric uptake regulator A: FurA) durch Fe^{2+} zur spezifischen Bindung an seine DNA-Zielsequenz aktiviert, wodurch nachgeordnete Gene reprimiert werden.

Zink ist ein weiteres wichtiges Element für alle lebenden Organismen und dient als Kofaktor in allen sechs Enzymklassen sowie in verschiedenen Regulationsproteinen. Die intrazelluläre Konzentration dieses Metalls muss aufgrund seiner Toxizität sehr genau reguliert werden. Verglichen mit Eukaryonten ist nur sehr wenig über die Zinkhomeostase in Bakterien bekannt. Im Genom von *M. tuberculosis* wurden mehrere Gene fuer potentiell Zink bindende Proteine identifiziert, allerdings wurde bisher seltsamerweise kein Zinkregulator gefunden. Stattdessen wurden erstaunlicherweise zwei Fur-Gene identifiziert: Mtb *furA* und Mtb *furB*, allerdings wurde kein eindeutiger struktureller oder funktioneller Unterschied berichtet.

In dieser Arbeit wird eine sorgfältige und detaillierte strukturelle und biologische

Beschreibung der FurA- und FurB-Proteine dargestellt. Mit Hilfe einer Reihe von biochemischen und biophysikalischen Methoden, darunter Untersuchungen der elektrophoretischen Mobilitätsveränderungen (engl. electrophoretic mobility shift assay: EMSA), ortsspezifische Mutationen, isothermale Titrations Kalorimetrie (ITC), microPIXE, Röntgenabsorptionsspektroskopie (engl. extended X-ray absorption fine structure: EXAFS) und Röntgenkristallographie, haben wir die Metallbindungsstellen zusammen mit der Struktur und den Eigenschaften dieser Proteine charakterisiert. Die Kombination dieser Ergebnisse erlaubte es zwischen strukturell und funktionell unterschiedlichen Metallbindungsstellen zu unterscheiden, eine aeusserst genaue qualitative und quantitative Charakterisierung von ihnen zu erstellen. sowie eine Erklärung ihrer biologischen Aufgaben und zum ersten Mal eine dreidimensionale Darstellung eines Zinkaufnahmeregulators vorzustellen.

SUMMARY

Iron is the most abundant trace element in the human body and zinc is the second one. Control of ion homeostasis is of vital importance for mammals and bacteria. Regulation of the ion flux into or out of the cell is a complex and articulated mechanism that still needs to be elucidated.

The highly specialized pathogen *Mycobacterium tuberculosis* has to contend with iron sequestration in order to survive in the human body. Iron metabolism is regulated by controlling transcription of genes involved in iron uptake, transport and storage. Paucity of this metal triggers an extensive response to increase iron acquisition whereas an excess of it can be toxic for the cell. The control of intracellular iron concentration is also linked to other important processes including oxidative stress response and the regulation of virulence factors. Recent studies have shown that, in patients affected by TB/HIV exposed to high level of dietary iron, the risk of active pulmonary tuberculosis increases. In *M. tuberculosis* the ferric uptake regulator A (FurA) is activated by Fe^{2+} to bind specifically to its target DNA sequence thereby repressing the downstream genes.

Zinc is also an important element for all living organisms and serves as a cofactor in all six classes of enzymes and also in several regulatory proteins. The intracellular concentration of this metal must be carefully regulated due its toxicity. Compared with eukaryotes, little is known about zinc homeostasis in bacteria. In the tuberculosis genome several genes coding for zinc proteins have been classified but curiously no zinc regulator has been yet defined. Surprisingly, instead, two Fur genes were identified: Mtb *furA* and Mtb *furB,* but no clear structural or functional distinction has been reported.

In this thesis a careful and detailed structural and biological description of FurA and FurB proteins is presented. Using a variety of biochemical and biophysical methods - including electrophoretic mobility shift assay (EMSA), site directed mutations, isothermal calorimetry (ITC), microPIXE analysis, extended X-ray absorption fine spectroscopy (EXAFS) and X-ray crystallography - we investigated the metal binding sites together with the nature and the structure of these proteins. The combination of these results enable us to distinguish between structurally and functionally distinct metal binding sites,

provide a meticulous description and qualitative and quantitative characterization of them, propose biological roles and present for the first time a 3D picture of a zinc uptake regulator.

Chapter 1

INTRODUCTION

1.1 Metal ion homeostasis

The biological activity of one-third of all proteins requires metal ions to perform catalytic, structural or regulatory functions (Pennella and Giedroc, 2005). Involved in the cell physiology are zinc, copper, nickel, cobalt, iron, manganese, molybdenum, tungsten and a few more cations. In certain environments, such as the ocean, paucity of metal ions can limit primary production and affect the cellular metabolism. In other environments, particularly in the neighbourhood of industrial contamination, the concentration of metal ions exceeds the lethal limit for most organisms.

The cell membrane acts as a filter and avoids an uncontrolled flux of metal ions in to or out of the cytosol. Under depleting conditions, metal ions and selected chelating-complexes are internalized by active transport; conversely, when metal ions are in excess, systems for metal sequestration or efflux are induced.

The balance of metal ion import and export mechanisms in the cell is termed metal ion homeostasis. Tight regulation of metal homeostasis is fundamental for bacteria and their interaction with the hosts.

1.2 Introduction to the biology and pathogenesis of *Mycobacterium tuberculosis*

It is estimated that every year around eight million new cases of this disease can be counted, with an annual death toll of more than 2.5 million people (Kaufmann, 2000). Tuberculosis appears in the list of the top 10 major killers of our century. Even though short-course chemotherapy (DOTS) and the Bacille Guerin-Calmette (BCG) vaccine are now available for a large proportion of the populations of developed countries, the threat remains, since this pathogen has developed an extraordinary multi drug-resistance (MDR).

Mycobacterium tuberculosis appears to multiply both intracellularly, in mononuclear phagocytes - especially lung macrophages, and extracellularly, in lung cavities. Multiplication to very high levels in lung cavities is especially important for disease transmission. The pathogen is coughed up and disseminated to the new host (Gobin and Horwitz, 1996). This is a widespread human pathogen with a low speed of reproduction and a long generation-time. The organism shows an excellent ability to mutate and therefore adapt to challenges in its microenvironment. Recently it has been observed that many patients affected by HIV can easily contract tuberculosis due to a breaking down of the immune system (Dye et al., 2002). In underdeveloped countries, the lack of resources does not allow the spread of the disease to be controlled, and new mutant strains can be observed depending on the living conditions of the subject. Furthermore, none of the animal models for tuberculosis, except maybe non-human primates, reproduces the diversity of disease progression that is seen in humans (Boshoff and Barry, 2005). These aspects represent a limiting factor to the understanding of tuberculosis metabolism; nevertheless, the huge amount of studies carried out in the last decades lead scientists to postulate testable and reliable pictures of the bacterial metabolism.

The capacity to survive within a specialised phagosomal compartment is central to the pathogenicity of *Mycobacterium tuberculosis* (Agranoff and Krishna, 2004). Evidence for the significance of metal ion homeostasis comes from recent studies where the pathogen has been exposed to presence/absence of different metal ions. Patients exposed to deprivation or excess of iron have showed a diverse course of the disease and the attenuation of its severity (Gangaidzo et al., 2001). After iron, zinc is the second most

abundant trace element in eukaryotic and it is involved in a myriad of biological processes: catalytic function in enzymes, stabilization and induction of protein folding in storage proteins, transcription factors and replication proteins (Coleman, 1992).

1.3.1 Iron in Mycobacterium tuberculosis

In order to develop a defensive strategy against this bacterium, mammalian hosts have restricted the access of such organism to iron. It was observed that pulmonary tuberculosis patients are, in fact, often anemic, suggesting sequestration of this cation from the host (Rodriguez and Smith, 2003).

In *Mycobacterium tuberculosis,* iron is an obligate cofactor for at least 40 enzymes e.g. it is required for the cytochromes (involved in electron transport) and for hemoproteins (involved in oxygen metabolism) such as catalase-peroxidase KatG - a major virulence factor which activates host defences against nitric oxide, acid tolerance and oxidative stress. Most of the intracellular iron is bound to non-heme proteins, and it is for instance required for DNA synthesis in ribonucleotide reductase and cell deotoxification by superoxide dismutase. An examination of the tuberculosis genome revealed 155 iron-regulated genes (Rodriguez, 2006). In total the mycobacterial cell requires between 7 and 64 μg of Fe per gram of cell mass in order to support growth. Even though iron is the fourth most abundant element on earth, its avaibility in mammalian fluids as a free cation is limited. It is usually kept in solution bound to transferritin, lactoferritin and ferritin (Rodriguez and Smith, 2006), since free iron is potentially toxic, being responsible for the generation of oxygen radicals through the Fenton reaction (Bossmann et al., 2004):

$$Fe^{2+}+H_2O_2 \rightarrow Fe^{3+}+OH\bullet+OH^-$$

To overcome iron deficiency, *M. tuberculosis*, like all mycobacteria, has developed a highly sophisticated system of iron transporters to supplement and support the acquisition of this metal from the host. So-called siderophores are small iron-chelating molecules excreted to compete for the limited amount of environmental Fe^{+2} in the cell.

Siderophores are composed by four structurally distinct molecules: salicylic acid, citric acid (the simplest), mycobactin and exochelin (De Voss et al., 1999). In the host they

bind Fe^{3+}, and once the complex Fe^{3+}-siderophore is formed, it interacts with specific receptor proteins on the cell surface and is transported into the cytoplasm by ATP-binding cassette (ABC) transporters. Here Fe^{3+} is released and reduced to Fe^{2+} then quickly loaded by iron uptake, storage and transport proteins.

The synthesis of siderophores depends on the extracellular iron abundance. They are produced under iron limitation and the amount synthesised is directly correlated with the mRNA levels of genes encoding for proteins responsible for their synthesis. This is indicative of a regulatory mechanism that functions at the level of DNA transcription (Rodriguez, 2006).

Members of three families of iron-dependent DNA-binding proteins are encoded in the Tuberculosis genome (Cole et al., 1998): IdeR/DtxR, SirR and FurA/B. These proteins do not share any significant sequence identity and they are expected to bind to different target sequences. Thereby they regulate, at the transcriptional level, a large variety of genes involved in iron uptake and the oxidative stress response. The metal acts as a co-repressor by activating the DNA-binding capability of these proteins. IdeR, the best characterized iron dependent regulator, has multiple biological roles which includes controlling iron metabolism (Schmitt et al., 1995). This protein controls genes encoding putative transporters, transcriptional regulators, proteins involved in general metabolism, members of the PE/PPE family of conserved mycobacterial proteins and the virulence determinant MmpL4 (Rodriguez and Smith, 2003). Control of iron homeostasis and protection against oxidative stress is coupled in mycobacteria, with IdeR also playing a role as a positive modulator of the oxidative stress response (Rodriguez and Smith, 2003). The function of SirR (named as such due to its high homology with the staphylococcal iron regulator repressor) is still unknown, but some data suggests it to be a divalent metal cation-dependent transcriptional repressor of the promoter/operator region of sitABC (Hill et al., 1998).

In addition to these regulators, the *Mycobacterium tuberculosis* genome encodes two proteins FurA and FurB similar to the Ferric Uptake Regulator (Fur) proteins that are well characterized for the Gram-negative bacteria such as *Escherichia coli* and *Pseudomonas aeruginosa* (Cole et al., 1998).

1.3.2 Fur (ferric uptake regulator) family

Fur is the most studied and best characterized of all members of the family of the iron regulators. The *fur* gene is itself regulated by the concentration of iron in the cell (Wong et al., 1999), and it may act as either as a transcriptional activator or a repressor. Fur directly or indirectly regulates a substantial number of other genes encoding proteins with diverse functions: these proteins are involved in a variety of metabolic processes, such as the Krebs cycle, genes required to survive oxidative stress, genes necessary for scavenging iron and genes which contribute to virulence, such as Exotoxin A (Vasil and Ochsner, 1999). Furthermore, it was recently shown that in *Pseudomonas aeruginosa* and *Escherichia coli*, Fur has a positive but indirect effect on the abundance of a set of genes (fig. 1.1). This protein can repress the transcription of small regulatory RNAs that bind to and decrease the stability of several mRNAs encoding iron-containing proteins. As a result, these mRNAs are not stable in the absence of Fur (Masse and Gottesman, 2002; Wilderman et al., 2004).

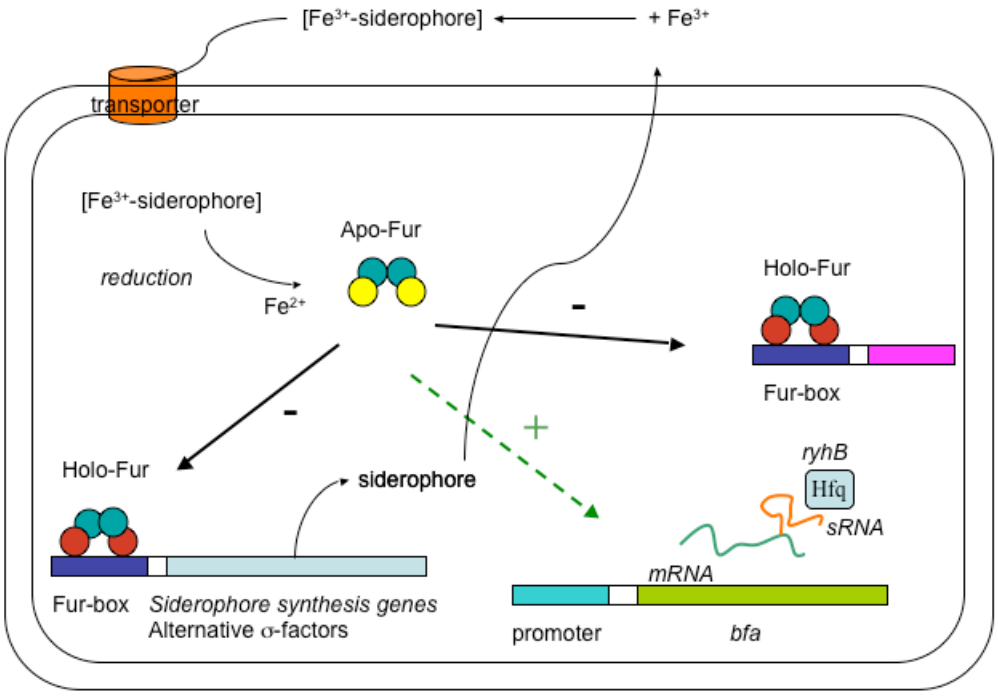

Figure 1.1: Schematic representation of Fur regulation in the cell

Many studies have recently been performed in order to shed light on this complex regulation system. Fur proteins from different organisms including *Escherichia coli, Pseudomonas aeruginosa, Helicobacter pylori* and *Anabaena* have been studied (Althaus et al., 1999; Hernandez et al., 2004; van Vliet et al., 2002; Vasil and Ochsner, 1999). A great deal of physiological, biochemical and structural information has been published, and different suggestions regarding the structure, function and mechanism of Fur have been proposed. Whereas *in vivo,* Fe^{2+} is the protein co-activator, it is possible to produce *in vitro* activity with a variety of divalent transition metals including Fe^{2+}, Zn^{2+}, Co^{2+}, Mn^{2+}, Cd^{2+} and Cu^{2+} (Mills and Marletta, 2005). All Fur proteins studied so far show a specific affinity for a DNA sequence, rich in A+T bases, called the Fur-Box: 5'-GATAATGATAATCATTATC-3' (Escolar et al., 1998). The Fur consensus box consists of a 19 bp (non-perfect) palindromic region suggesting the binding of at least one dimeric form of the protein. The emerging scenario suggests certain common features. In most bacteria, Fur contains tightly bound zinc, and often denaturation of the protein is necessary to remove this cation. A second metal co-factor, Fe^{2+}, is less strongly bound, but necessary for the DNA binding activity.

At present, the three-dimensional structure of only one member of this family has been determined: the *Pseudomonas aeruginosa* Fur (PA Fur) complexed with Zn^{+2} solved to 1.8 Å resolution (Pohl et al., 2003) (fig. 1.2).

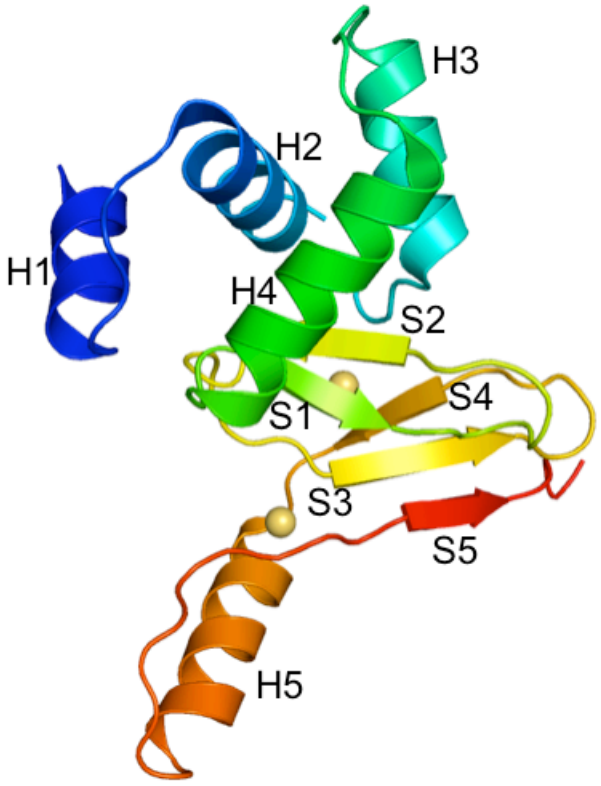

Figure 1.2: Ribbon representation of PA Fur

The protein is composed of 134 amino acids and has a molecular weight of 14435 Da. In solution Fur is a dimer and possesses two metal sites: one binds zinc ion and one binds iron ion. In the crystal structure, the protein exhibits two additional metal sites, both occupied by zinc ions. In the crystallization conditions the protein, purified with only one metal site (the structural zinc) occupied, was added to a mother liquor solution containing zinc-salt ($ZnSO_4$) and polyethylene. Thus the iron ion was replaced by zinc. This is confirmed by the EXAFS data collected in solution on a protein sample complexed with both metals and exhibiting the same coordination observed in the crystal structure. The structure is composed of two domains: a DNA-binding domain (residues 1-84) and a dimerization domain (residues 85-134). The DNA-binding domain is composed of four helices (H1, H2, H3 and H4) followed by two-stranded antiparallel β-sheets (S1 and S2). This domain exhibits a typical helix-turn-helix motif with a three-helix bundle (H2, H3 and H4) where the helix H4 is the putative DNA recognition region (Pohl et al., 2003).

The dimerization domain is constituted by three antiparallel β-strands (S3, S4 and S5) and one long α-helix (H5). The strand S3 and the helix H5 are involved in the protein dimer formation and they are kept in a rigid conformation by a metal site buried in the domain's central part. It is possible to observe hydrophobic and hydrophilic interactions between the helix H5 and the symmetry related helix H5' (fig. 1.3).

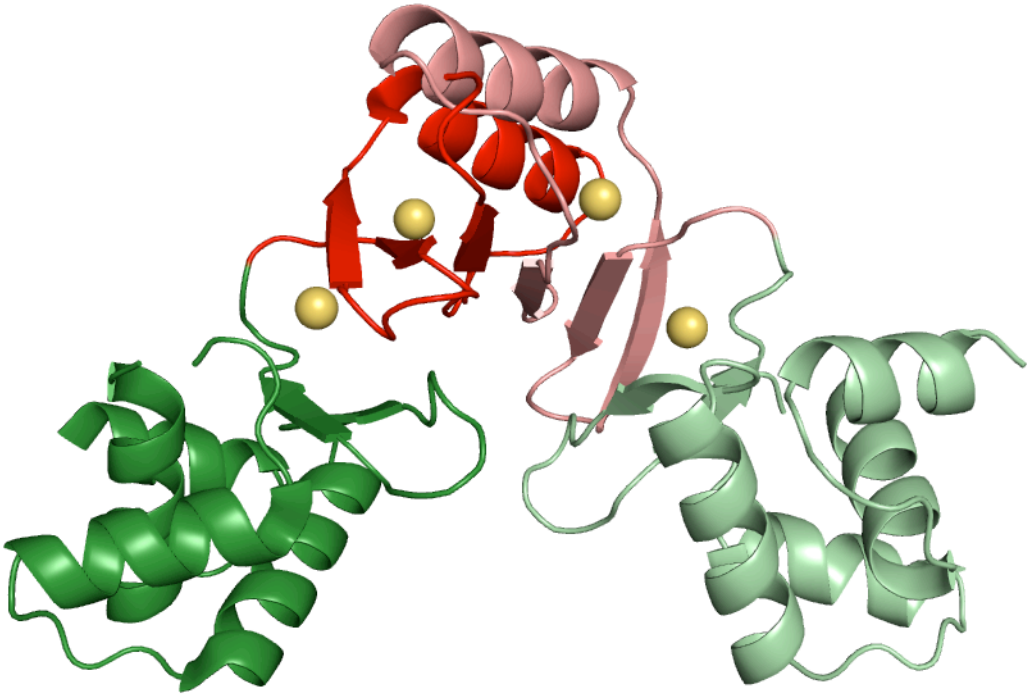

Figure 1.3: Ribbon representation of PA Fur dimer

The structural site shows a tetrahedral coordination and is surrounded by Asp88, Glu107, His86 and His124 (fig. 1.4)

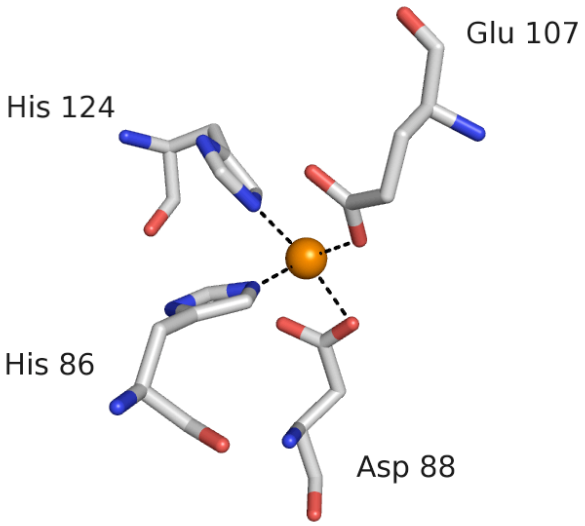

Figure 1.4: "Ball and sticks" representation of the structural site of PA Fur

In the regulatory site instead the cation is octahedrally coordinated by His32, Glu80, His89, Glu100, and 1 molecule of H_2O (fig. 1.5).

Both sites were also structurally analysed by absorption spectroscopy and data were in agreement with crystallographic coordinates.

EXAFS experiments on *Escherichia coli* Fur (EC Fur) (Jacquamet et al., 1998) reported the presence of more than one metal site also for this protein and suggested distinct functions for each of them. Recently the high-resolution structure of the DNA-binding domain of this protein was published by the same group (Pecqueur et al., 2006), confirming the previously published spectroscopic data. This domain is very similar to the related one in PA Fur and again the well-characterized motif, helix-turn-helix typical of DNA-binding protein, has been observed.

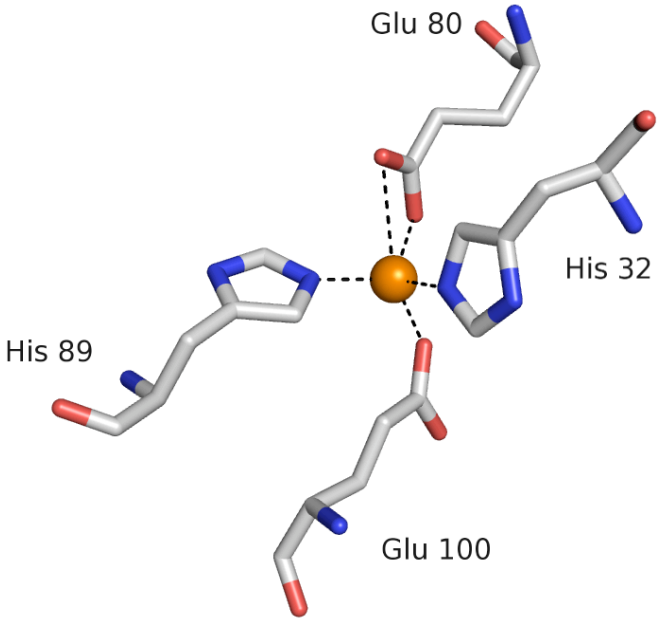

Figure 1.5: "Ball and sticks" representation of the regulatory site of PA Fur
The sixth ligand (H2O) is not reported in the PDB file (ID 1mzb).

For years long discussions and a few mechanisms have been proposed in order to explain the interaction between the protein and its DNA recognition sequence. As previously remarked, PA Fur binds the canonical Fur-box and is found in solution in a dimeric state, like all Fur proteins. These proteins interact with DNA by utilising two helices (of the helix-turn-helix motif) of which one (generally rich of electropositive residues) interacts with the negatively charged surface of the nucleotide, while the second is located in the major groove of the nucleotide. The dimeric protein form mainly acts as a "clamp" docking the palindromic DNA at least in two separate regions. The question of the number of Fur dimers involved in the Fur-box binding is still open. At present it is in fact only possible to confirm the DNA binding by using *in vitro* experiments, but determining the multiplicity of the dimer becomes more difficult.

1.3.3 Mtb FurA and Mtb FurB

As introduced above, the tuberculosis genome contains two genes coding for Fur-like proteins: *furA* and *furB* (Cole et al., 1998).

The *furA* gene is situated immediately upstream of the *katG* gene, which encodes for catalase-peroxidase enzyme. FurA negatively regulates (by an as yet unknown mechanism) the expression of *katG,* thereby modulating the response against oxidative stress (Rodriguez, 2006). In addition, FurA is thought to be involved in other mechanisms but no "official" data is yet available (Pym et al., 2001).

Very little is known about FurB. The only clear evidence available points to a strong affinity of this protein for zinc (and not for iron) and it was recently suggested (Canneva et al., 2005) that FurB could be involved in the control of Zn-sensitive genes.

After decoding the tuberculosis genome in 1998, many genes/proteins have been classified by homology or similarity to already well-defined and characterized proteins from other organisms. Less than 10 years have passed from that date and many studies have been conducted on tuberculosis targets, but more time is still required to cover the full genome and assign a major function to every single protein.

Despite unclear classifications of each of the individual targets in tuberculosis, a pattern of multiple metal dependent repressors is seen, similar to that found in *Bacillus spp.* Genes encoding proteins homologous to both classes of the major iron dependent repressors (i.e. DtxR and Fur) were found, as in most other prokaryotes. While DtxR and Fur are clearly involved in iron homeostasis in organisms where they function as the main regulator of iron homeostasis (i.e. corynebacteria or mycobacteria for DtxR and gram negatives for Fur), the functions of the additional Fur homologs like FurA and FurB are not very certain. In some cases they appear to be involved in resistance to redox stress or in the regulation of gene expression involving metals other than iron, for instance zinc or manganese. For example, one Fur homolog is Zur, which appears to be required for the control of zinc uptake in *Escherichia coli* and *Bacillus subtilis.* More recently, another Fur homolog, Nur (of *Streptomyces*), has been discovered to control the expression of a gene encoding a nickel dependent superoxide dismutase (Ahn et al., 2006). The sequences of the Fur homologs which are likely to be directly involved in iron homeostasis are highly conserved throughout their lengths, whereas the sequences of Fur proteins which have alternative functions (e.g. Zur and Nur) are most conserved in their putative metal binding regions and dimerization domains. In contrast, the amino terminal

DNA binding regions of these alternative Fur proteins are much less well conserved compared to the Fur proteins, which have a strong influence on iron homeostasis.

1.4 Zinc regulation

Zinc is an important element for all living organisms and serves as a cofactor in all six classes of enzymes (Coleman, 1998) and also in several classes of regulatory proteins (Outten and O'Halloran, 2001). The intracellular concentration of this metal must be carefully regulated due its toxicity. Compared with eukaryotes, little is known about zinc homeostasis in bacteria. In *Escherichia coli* a zinc efflux system (ZntA) and a zinc import system (ZnuABC) have been described (Hantke, 2001). ZntA (Zn^{2+} transport or tolerance) is an ion-motive P-type ATPase that exports Zn^{2+}, Cd^{2+} and Pb^{2+}. The Zn^{2+}-specific uptake system Znu (Zn^{2+} uptake) belongs to the ABC transporter family and is composed of the periplasmic binding protein ZnuA, the ATPase ZnuC, and the integral membrane protein ZnuB (fig. 1.6). In the *znu* gene cluster, the transcription of the *znuA* gene is divergent to that of the *znuCB* genes and the genes are separated by an unusually short intergenic region of 24 base pairs.

Today, zinc homeostasis gains an increasing interest. The zinc uptake system ZnuABC is regulated by Zur (Zn^{2+} uptake regulator) (Patzer and Hantke, 2000). Other zinc-specific metalloregulatory proteins from bacteria are known, but I would like to focus our attention on this particular system. Zur is considered to be a member of the Fur family (Moore and Helmann, 2005). Its structure is unknown but it is likely that, in homology to all members of the Fur family, this protein has a conserved structure with a N-terminal DNA-binding domain and a C-terminal dimerization domain. Recently, two metal binding sites per monomer have been described for the *E. coli* zinc uptake regulator (EC Zur). EXAFS data indicated a tetrahedral structural zinc site constituted of sulphur and oxygen atoms and a tetrahedral regulatory site coordinated to one sulphur and three histidines/oxygens (Outten et al., 2001). Most of the zinc-specific periplasmic binding proteins have a central histidine, aspartate and glutamate rich region, which seems to allow discrimination from the manganese transporters. Zinc has a preferred tetrahedral

coordination geometry that is quite different from the hexa-coordinated binding state preferred by manganese.

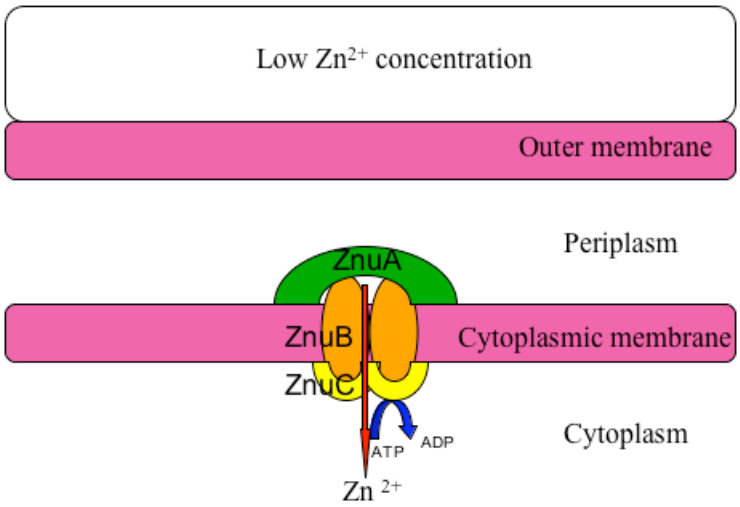

Figure 1.6: Schematic representation for the zinc uptake pathway

Depending on the Zn^{2+} concentration in the medium, different types of Zn^{2+} transporters are synthesized. At limiting Zn^{2+} concentrations, binding-protein-dependent ABC transporters are induced. Today there are still gaps in the understanding of this mechanism and little is known about the passage of divalent ions across the outer membrane and about which transport systems supply the cells under Zn^{2+}-replete conditions.

In many sequenced genomes of bacterial species, proteins designated as Fur homologues were recently reclassified as Zur (Hantke, 2001). Proteins belonging to the same family show various sequence similarities but evidently dissimilar biological functions. Curiously, two Fur genes were identified in the tuberculosis genome - *furA* and *furB* - but no Zur was yet reported. The metal selectivity of Fur-like repressors varies and is presumably determined by the precise spatial arrangement of the potential metal ligands around the regulatory site. Zur may have evolved the ability to sense zinc by modification of the iron-sensing site (Gaballa and Helmann, 1998).

1.5 Objectives

In recent years, many studies have identified several components of the homeostasis machinery. Metalloregulatory proteins are key to understanding the pathways of metal uptake, distribution, storage and efflux of metal ions.

Despite the amount of experimental data and knowledge available today, *Mycobacterium tuberculosis* homeostasis remains unclear and most of the pathways and the regulators involved are still unknown. In the past few years it became evident that this organism manifests a certain sensibility to the presence of metal ions. Vulnerability of this pathogen has been observed in the presence of iron and new therapy based on this knowledge can attenuate its progression (Gangaidzo et al., 2001). Today more and more studies are focussing on this topic, but still many gaps remain to be filled. Because of sequence and functional similarities between the major Fur regulators such as Fur from *P. aeruginosa* and those like FurA, FurB or Zur, it would be of interest to examine how metals interact with these subclasses of Fur. Furthermore, it could be of a certain importance to shed some light on the real functions of these two targets.

The work presented in this thesis is focused on structural and biological studies on Rv1909c (Mtb FurA) and Rv2359 (Mtb FurB) from *Mycobacterium tuberculosis*. The following chapters will demonstrate that through a variety of biochemical and biophysical methods - including cloning, expression and protein purification, EMSA (electrophoretic mobility shift assay) assays, isothermal calorimetry (ITC), microPIXE analysis, extended X-ray absorption fine structure (EXAFS) and protein crystallography – it was possible to investigate the nature and function of the different metals present in Mtb FurA and Mtb FurB. Furthermore it was also possible to unravel and clearly to state for the first time the biological functions of Mtb FurB.

Chapter 2

MATERIALS AND METHODS

In this chapter a brief description/introduction covering all methodologies utilized in this project is given.

In all cases, before starting any of the experiments, all containers were washed and sterilized and all buffers were pretreated with chelex-100 to avoid possible metal contamination. All chemicals were purchased from Sigma-Aldrich (Germany) except where explicitly indicated otherwise.

2.1 Cloning. expression and purification

2.1.1 FurA

A PCR amplification of the *furA* gene from the TB genome H37Rv was performed using the primers:

Fwd 5'GCATCGCCATGGCTTCTAGTGTGTCCTCTATACC3'

Rev 5'GCATCGAAGCTTTTACGGATGTGATCGCGAAGTG3'

(purchased from MWG, Ebersberg, Germany). The size of the insert was confirmed by agarose gel electrophoresis. Purification and gel extraction of the amplified product was performed using a gel extraction kit (Qiagen). The insert was subsequently digested at 37°C with restriction enzymes HindIII (New England Biolabs, Frankfurt, Germany) and XhoI (New England Biolabs, Frankfurt, Germany) in buffer NEB2 (New England

Biolabs, Frankfurt, Germany). Ligation into a pETM11 vector was performed using the Rapid DNA ligation kit (Fermentas) after cutting the plasmid with the same restriction enzymes.

DH5α cells were transformed with the product, spread on kanamycin-containing agar plate and left overnight at 37°C. One colony was picked from the plate and grown overnight in LB (Luria-Bertani) media in the presence of 50 mg kanamycin at 37°C. DNA was purified from the DH5α cells using a Qiagen kit and sent to MWG for sequencing. The purified plasmid was transformed into BL21* cells and over expressed in LB containing 50 mg kanamycin and 34 mg chloramphenicol. The cultures were grown at 37°C; when the OD reached a value of approximately 0.6, the cells were induced with 1mM IPTG (Isopropyl β-D-thiogalactoside) (ROTH) and left to grow overnight at 20°C.

Cells were harvested after centrifugation at 4000 xg for 20 min. Lysing buffer (50 mM Tris pH 8.0 HCl, 300 mM NaCl, 20 mM imidazole, 0.02%(v/v) of 1-thioglycerol) was added together with DNAase (BioLabs) and left for 20 min. After sonication the cells were spun down for 30 min at 20000 xg. All harvesting and lysing steps were performed at 4°C.

The supernatant was then passed through a 0.22 μm filter and loaded on a pre-equilibrated Ni-NTA column (resin purchased from Qiagen) for purification. The protein was eluted with a high imidazole concentration buffer (50 mM Tris pH 8.0 HCl, 300 mM NaCl, 500 mM imidazole, 0.02%(v/v) of 1-thioglycerol) and left with TEV protease, for overnight cleavage at 4°C, in dialysis in a buffer containing 50 mM Tris pH 8.0 HCl, 300 mM NaCl. It was then loaded once more onto the Ni-NTA column in order to remove the His_6-tagged protease and only the flow-through was collected. The eluted volume was reduced to 1ml and further purified on a Superdex 75 16/60 column (Amershan Bioscience) which had been previously equilibrated with 10 mM Tris pH 8.0 HCl, 150 mM NaCl, 10% glycerol and 5 mM DTT (ROTH).

The concentration was measured using either a Bradford assay (Biorad) or by measuring the absorption at 280 nm in 6.0 M guanidium chloride and 0.02 M phosphate buffer pH 6.5 and using the calculated absorption coefficient (obtained from the “Expasy protein

chart" at http://www.expasy.org/uniprot/P0A582)

2.1.2 Cloning, expression and purification of FurB

Methods and materials are similar to those described above in the FurA section; therefore only a short description is given.

The *furB* gene from the TB genome H37Rv was cloned into the pETM11 vector using the primers:

Fwd 5'GCATCGCCATGGCTAGTGCAGCCGGTGTCCG3'

Rev 5'GCATCGAAGCTTTTAGCTCCGGCAGTCTGAGC3'

and HindIII (New England Biolabs, Frankfurt, Germany) and XhoI (New England Biolabs, Frankfurt, Germany) as restriction sites. The protein was expressed in BL21* by overnight induction with 1mM IPTG (ROTH) at 20°C.

Purification steps included a Ni-NTA affinity chromatography (Qiagen) followed by cleavage of the His_6-tag with TEV protease and a size exclusion chromatography on a Superdex 75 16/60 column (Amersham Bioscience). In some cases an ion exchange step through a resource column (Amersham Boscience) was performed in addition.

Dynamic light scattering and mass spectroscopy were performed in order to confirm the purity of the sample. Concentration was measured using either a Bradford assay (Biorad) or by measuring the absorption at 280 nm in 6.0 M guanidium chloride and 0.02 M phosphate buffer pH 6.5 and using the calculated absorption coefficient.

2.2 Site directed mutagenesis

Single point-mutations of three cysteines (C76S, C126S and C129S) of FurB were prepared by using the QuikChange Site-Directed Mutagenesis Kit (Stratagene).

For each mutation a set of primers was designed as indicated in the kit's manual:

C76S_fwd 5' TCGGTCTACCGCAGA**TC**CTCGGAGCACCATCAC

C76S_rev 5’ GTGATGGTGCTCCGAGGATCTGCGGTAGACCGA
C126S_fwd 5’ GAGATCTTCGGCACC**TCC**TCAGACTGCCGGAGC
C126S_rev 5’ GCTCCGGCAGTCTGAGGAGGTGCCGAAGATCTC
C129S_fwd 5’ ATCTTCGGCACCTGCTCAGAC**TCC**CGGAGC
C129S_rev 5’ GCTCCGGGAGTCTGAGCAGGTGCCGAAGAT

The reagents for the PCR reaction were as follows:

5 µl of 10x reaction buffer
1 µl (25ng) of dsDNA template
1 µl (200 ng) of primer1
1 µl (200 ng) of primer2
1 µl of dNTP mix
1 µl of Pfu Turbo DNA polymerase
40 µl ddH_2O

After PCR, 1 µl of *Dpn I* was added directly to the amplification reaction and the tubes were incubated at 37°C for one hour to digest the parental DNA. 1 µl of the digested product was transferred into a tube containing supercompetent cells (Epicurian Coli XL1-Blue) and all left on ice for 30 minutes. The transformation was pulse heated to 42°C for 45 seconds and left on ice for two minutes. 0.5 ml of LB was added and the reaction was left at 37°C for one hour. Everything was then plated and left to grow overnight at 37°C. The following day, one colony was picked from the plate and grown in LB media with kanamycin and chloramphenicol. After 16 hours, cells were spun down and DNA was extracted using the Qiagen DNA-extraction kit (Stratagene). The DNA obtained was sequenced and mutations were confirmed.

All three FurB mutants were expressed following the purification protocol described above for FurB.

2.3 Bioinformatics analysis

At present, Zur (zinc uptake regulator) from *E. coli* is the most studied and best

characterized zinc uptake regulator. It regulates a *znu* gene cluster coding for the zinc import system ZnuABC (ZnuA is a periplasmic binding protein, ZnuC is an ATPase and ZnuB an integral membrane protein). The transcription of the *znuA* gene is divergent to that of *znuCB* genes. They are separated by an unusually short intergenic region of 24 base pairs. Protein searches against the *Mycobacterium tuberculosis* genome were executed with either BLAST or FASTA programs (Altschul et al., 1990; Pearson, 1990). A triplette of genes Rv2059 (2315172-2316707), Rv2060 (2316277-2316678), Rv2061c (2316679-2317083) respectively coding for a hypothetical adhesion protein, a possible conserved integral membrane protein and a hypothetical phosphatase protein was found. These proteins surprisingly match the related proteins in *E. coli* not only for high homology in the sequence alignments, but also for their functional annotations. Another similarity can be observed in the intergenic region of 70 base pairs present between Rv2059 and Rv2060-Rv2061c genes (these two are co-transcribed like in *E. coli*). Also, the molecular dimensions (related to the functions) of the proteins seem to be similar in tuberculosis. ZnuA shows, in fact, a bigger size compared to ZnuB and ZnuC which have instead smaller and approximately same molecular. The same pattern is observed for the newly-discovered genes: where Rv2059 is bigger than Rv2060 and Rv2061c which have similar masses.

2.4 DNA-binding assays

A 27 bp DNA sequence (referred to here as the Zur-box), almost palindromic and identical to the promoter region of Rv2059, was synthesized (MWG). Electrophoretic mobility shift assays were performed using the Zur-box and the canonical Fur-box. The oligonucleotides were mixed with the protein in the presence or absence of different metals in a low salt buffer (10 mM Tris pH 7.5, 100 mM NaCl, 10% Glycerol). The reaction was then loaded on a 4-20% native gel, stained with ethidium bromide and visualized under the UV lamp.

2.5 FurB-DNA complex purification

Once the DNA partner had been identified, the complex was formed by adding 1.2 equivalents of protein to the DNA in the presence of a zinc salt ($ZnSO_4$) solution. The reaction was left at room temperature for one hour and then passed through a gel filtration column, Superdex 75 16/60 (Amersham) for purification.

2.6 MicroPIXE

2.6.1 Brief introduction to microPIXE

Proton Induced X-ray Emission is a technique utilized to probe the amount of metal ions in a sample - particularly, as in this case, in proteins. The most significant effect produced by the absorption of the incident protron beam is ionization of atoms. Subsequent relaxation of an electron to a lower shell results in the emission of one or more photons or Auger electrons (fig. 2.1).

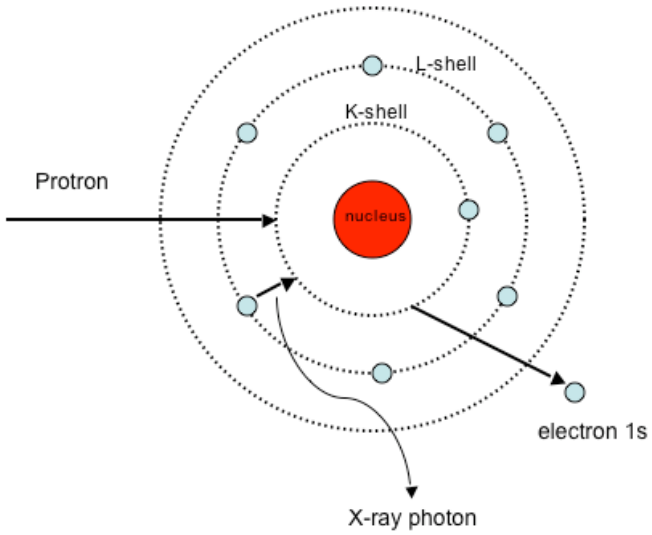

Figure 2.1: Schematic representation of proton-element interaction

An absorbing atom is photoionized by a proton beam. Electrons of a specific shell are excited to the continuum, followed by transition of electrons from higher shells (L,M) to the unoccupied lower shell (K). The energy of the X-ray quantum is element specific.

Each atom emits a characteristic energy depending on the atomic potential. This signal serves as the fingerprint of the element. PIXE is based on the detection of these

characteristic photons. Thus every element above a certain energy cut-off, in other words heavier than fluorine, can be detected simultaneously.

A complete description and introduction to this technique is reviewed in (Garman and Grime, 2005).

2.6.2 MicroPIXE sample preparation and data collection

All samples were left for overnight dialysis in 10 mM Tris pH 8.0, 150 mM NaBr, 10% glycerol and 5mM ascorbate. The buffer exchange is necessary to remove any traces of free S and Cl. Sulphur acts as an internal standard, and due to the proximity of the X-ray emission energies of S and Cl, strong chloride fluorescence can affect the accuracy with which the sulfur peak can be quantified.

The microPIXE measurements were carried out at the National Ion Beam Centre, University of Surrey, U.K. A 2.5 MeV proton beam of 1 μm in diameter was used to induce characteristic X-ray emission from the dried liquid droplet under vacuum. The X-rays were detected in a solid-state lithium-drifted silicon detector with high-energy resolution (Garman and Grime, 2005; Grime et al., 1991).

Spatial maps were obtained of all elements heavier than neon. Quantitative information was obtained by collecting 3 or 4 point spectra from the sample area. These spectra were analyzed with GUPIX (Johansson et al., 1995) within DAN32 (Grime, 1996) to extract the relative amount of each element of interest in the sample.

This part of the work has been done in collaboration with Dr. Elpseth Garman (Oxford university).

2.7 Isothermal Calorimetry

2.7.1 Brief introduction to the Isothermal Calorimeter

ITC (Isothermal Titration Calorimetry) allows the determination of binding constants by measuring the heat produced during a chemical reaction. The binding of a ligand to a protein can be an exothermic reaction. A calorimeter is an isolated system where the heat

produced or absorbed during a reaction is determined.

This technique consists of measuring the heat absorbed or released during a titration where the ligand is added in small amounts allowing collection of a large number of data points to facilitate subsequent integration. These values are fitted to models and information about the stoichiometry, binding constant and enthalpy can be extrapolated. Correlation between the parameters is given by:

$$\Delta G = \Delta H - T\Delta S = -RT\ln K$$

where:

ΔG= Gibbs energy variation

ΔH= Enthalpy variation

T= Temperature

ΔS= Entropy variation

R= Gas constant (8.3145 J/mol K)

K= Association constant

More detailed information can be found in the "VP-ITC MicroCalorimeter – User's manual" MicroCal, LLC.

2.7.2 ITC sample preparation and data collection

ITC measurements were performed at 25 °C with a VP_ITC titration calorimeter (Microcal, Northampton, MA). After purification of FurA, the protein was left for overnight dialysis at 4°C in 10 mM Tris pH 7.5 HCl, 150 mM NaCl, 10% glycerol and 1mM TCEP (PIERCE). The concentration of the protein was carefully measured using the absorption coefficient, immediately before starting each data collection. The final protein concentration in the cell was 31.8 mM with a volume of 1.4 ml. The titration was performed by a total of 38 injections of 5 μl each using a solution of 1mM $ZnCl_2$. The salt was dissolved in the same buffer as described above. Blank titration of the zinc solution was performed in order to provide a control for possible effects due to the dilution during the protein titration.

The same protocol was followed for titration of FurB. In this case, the protein was

dialyzed at 4°C against a solution containing 10 mM Tris pH 7.5 HCl, 150 mM NaCl and 10% glycerol. The final protein concentration in the cell was 55.0 mM in a 1.4 ml volume. The titration was performed by a total of 50 injections of 5 μl each using a solution of 1mM $ZnCl_2$ and $CoCl_2$. Salts were dissolved in the same buffer described above. Control titration was performed as described previously.

Data reduction and analysis were carried out using the Origin software suite (Microcal, Northampton). The measured integrated heat values were fitted to the "single binding site" model for FurA and "double binding site" for FurB. The fitting procedure was iterated until the chi squared value (the parameter which estimates the goodness of the fit) reached a minimum.

2.8 EXAFS

2.8.1 Brief introduction to EXAFS

X-ray absorption spectroscopy (XAS) is a technique, which has developed over the last few decades due to the availability of increasingly sophisticated synchrotron radiation sources. It provides information about local atomic and electronic structure of the absorber, as well as high-accuracy parameters such as oxidation state, coordination number and first coordination shell distances. XAS can be applied to solid, liquid and amorphous samples; apart for its high accuracy for the first shell disctances this is considered one of the main advantages of the method.

X-ray Absorption Fine Structure (XAFS) is the modulation of the X-ray absorption coefficient (of metal ions) at energies near and above an X-ray absorption edge. When a metal ion absorbs a photon, one of the electrons from the core levels can be excited into the continuum. The atom is therefore left in an excited state with an empty electronic level. In order to rebalance the atomic energy, one of the electrons from the more external shells relaxes and occupies the inner shell. At the same time a fluorescent X-ray or an Auger electron is emitted (fig. 2.2)

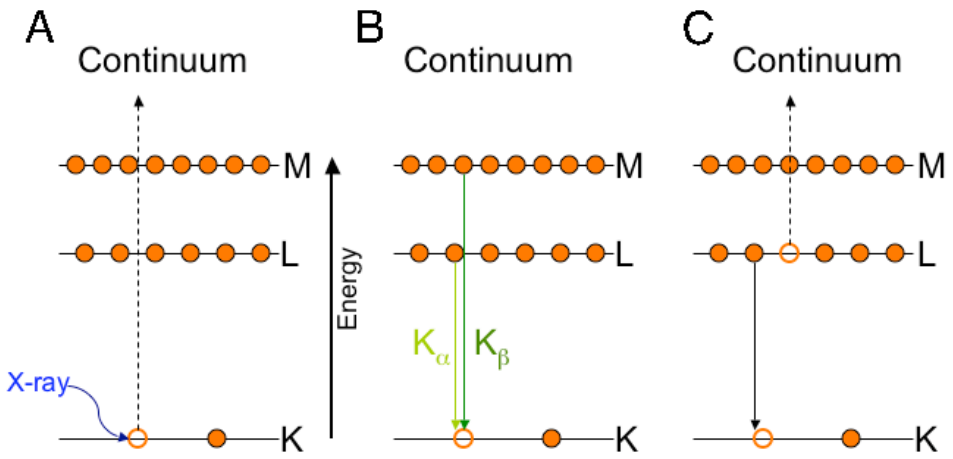

Figure 2.2: Schematic representation of the photo-electric effect

The X-ray is absorbed by an atom promoting a core-level electron (K) into the continuum. B X-ray fluorescence. X-ray fluorescence occurs at discrete energies that are characteristic of the absorbing atom, and can be used to identify the nature of the atom. C Auger effect.

Typically the XAS spectrum is divided in two main regions: the XANES (X-ray Absorption Near Edge Structure) covering the energy range around the edge and the EXAFS (Extended X-ray Absorption Fine Structure) covering the successive 30-2000eV (depending on the quality of the data) (fig. 2.3).

The EXAFS region provides information about type, number and distances of the ligands.

$$\lambda(k) = \sum_s \frac{N_s |f_s(\pi,k)|}{kR^2_{as}} \exp(-R_{as}/\lambda_f) \exp(-2\sigma^2_{as}k^2) \sin\left[2kR_{as} + \alpha_{as}(k)\right]$$

where:

k= wave vector

N_s= number of backscattering atoms in the shell

R_{as}= distance of the ligand-absorber

$|f_s(\pi,k)|$= backscattering amplitude

λ_f= free inelastic scattering

$\exp(-2\sigma^2_{as}k^2)$= Debye-Waller factor

$\alpha_{as}(k)$=phase shift

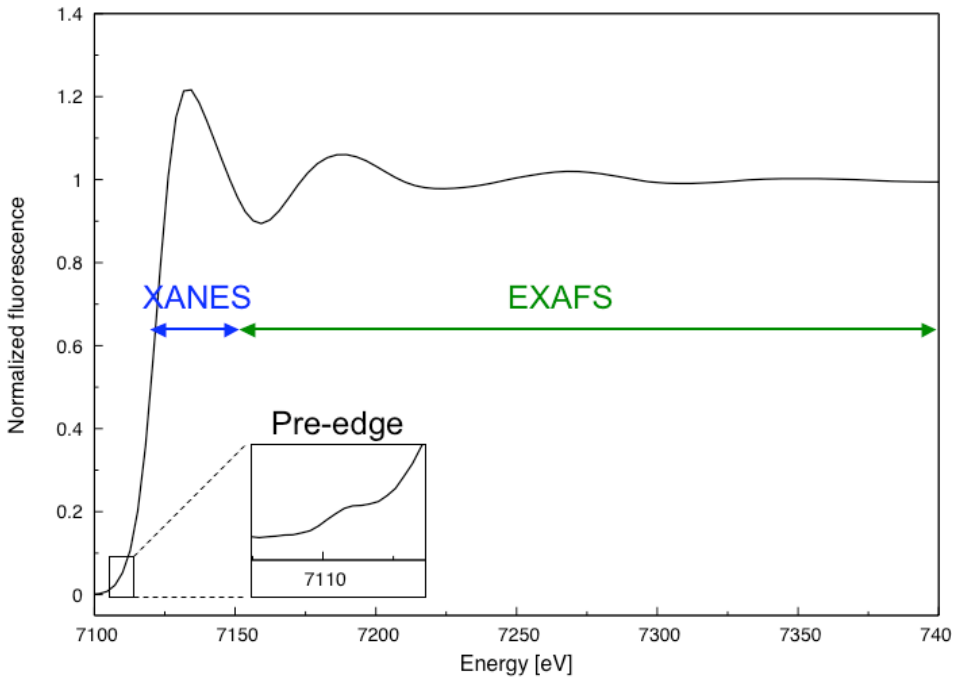

Figure 2.3: Iron X-ray absorption spectrum

The spectrum was collected at the iron K-edge. It is composed of different regions: "pre-edge" area due to a transition 1s→3d, XANES and EXAFS

Once the data have been collected, they are fitted to models using programs able to reproduce single or multiple scattering pathways (Scott, 2000). Further information related to this technique can be found in many reviews e.g. by Aksenov *et al.* (Aksenov et al., 2001).

2.8.2 X-ray absorption spectroscopy sample preparation and data collection

After concentrating the protein sample to 13mg/ml, FurA was dialyzed against a 1 mM Fe^{2+} solution ($Fe(SO_4)(NH_4)_26H_2O$, 10 mM Tris pH 8.0 HCl, 150 mM NaCl, 10% glycerol and 5 mM ascorbate). In addition, the same protein was washed for 24 hours with the corresponding iron free buffer to remove any excess of iron.

FurB protein was concentrated to 17mg/ml and dialyzed against a 0.5 mM Co^{2+} solution ($CoCl_2$), 10 mM Tris pH 8.0 HCl, 150 mM NaCl, 10% glycerol and 1 mM TCEP). The protein was washed for 24 hours with the corresponding cobalt free buffer to remove any

excess of the metal.

The samples were then transferred into plastic cuvettes (Hesar Glas) with Kapton® windows, frozen in liquid nitrogen and kept at 20K in a He closed-cycle cryostat (Oxford instruments) during the entire experiment.

X-ray absorption data at the iron, cobalt and zinc edges were recorded in fluorescence mode at the EMBL bending magnet beam line D2 (DESY, Hamburg, Germany) equipped with a Si(111) double crystal monochromator, a focusing mirror and a 13 element Ge solid-state fluorescence detector (Canberra). Harmonic rejection was achieved using a focusing mirror with cut-off energy of 21keV and a monochromator detuning to 70% (for the zinc) and 50% (for the iron and cobalt) the peak intensities. Dead time correction was applied to the fluorescence signal and saturation was not observed (detector dead time < 20%). The energy of each scan was calibrated by using the Bragg reflections of a static Si(220) crystal in back reflection geometry (Pettifer and Hermes, 1985). Data reduction analysis of FurA and FurB were performed with the software program Kemp (M. Korbas, 2006) using E_0=9660eV for the zinc, E_0=7120eV for the iron edge and E_0=7716eV for the cobalt. The EXAFS spectra were analyzed with EXCURV98 (Binsted et al., 1991).

Two data sets for each of the proteins, FurA and FurB, were collected on two independently prepared samples. Results were reproducible and only the best are reported. The quality of the fit can be assessed quantitatively using the Fit index (Φ_{EXAFS}), the R factor (R_{EXAFS}) and the reduced χ^2 (ε_ν^2). The fit index (the parameter that is minimized during the least-squares refinement) is of the sum of the square of residuals

$$\Phi_{EXAFS} = \sum_i^N w_i^2 \left(\chi_i^{\exp}(k) - \chi_i^{th}(k) \right)^2$$

where $\chi^{exp}(k)$ and $\chi^{th}(k)$ are the experimental and theoretical EXAFS respectively, and

$$w_i = \frac{k_i^n}{\sum_i^N k_i^j \left| k_j^{\exp}(k_j) \right|}$$

for which n=3 in the data analysis presented here. This weighting is typical for biological EXAFS data analysis.

R_{EXAFS} gives an indication of the quality of the fit in *k* space

$$R_{EXAFS} = \sum_{i}^{N} w_i \left(\left| \chi_i^{\exp}(k) - \chi_i^{th}(k) \right| \right) \times 100\%$$

Neither Φ_{EXAFS} nor R_{EXAFS} take into account the number of parameters that have been refined (N_{pars}) or the number of independent data points (N_{ind}). However, ε_v^2 gives an indication of the overall goodness of the fit, considering the degree of over-determination of the system. Reduced χ^2 is not an absolute value but relative and thus can be used to decide which is the best statistical fit of

$$\varepsilon_v^2 = \left[\frac{1}{(N_{ind} - N_{pars})} \right] \left(\frac{N_{ind}}{N} \right) \sum_{i}^{N} w_i \left(\chi_i^{\exp}(k) - \chi_i^{th}(k) \right)^2$$

where the number of independent parameters is given by

$$N_{ind} = \frac{2\Delta k \Delta R}{\pi} + 1$$

Δk is the data range being fitted in k space (between k_{min} and k_{max}), and ΔR is the range of data being fitted in R space (between R_{min} and R_{max}) by the model where the data are substantial.

For each model, the interatomic distances and the Debye-Waller factors were refined by least-square techniques starting from a variety of initial positions and converging to the results summarized in tables 3.3, 3.4, 4.2 and 4.3.

2.9 Homology modelling

A FurA homology model was generated using the Swiss Model suite (http://swissmodel.expasy.org/). The crystal structures of PA Fur (PDB ID 1mzb) and FurB were used as a templates, and the amino acid sequence of the Mtb FurA was submitted for modeling (Guex and Peitsch, 1997; Peitsch, 1995; Schwede et al., 2003).

2.10 Crystallization and crystal structure determination

FurB crystals were obtained by hanging-drop vapor diffusion by mixing the protein at a concentration of 13 mg/ml with a reservoir solution containing 14% glycerol and 0.3 M

$(NH_4)_2HPO_4$. Crystals are tetragonal with a=b=51.6, c=133.4 Å and space group $P4_12_12$. They were mounted directly from the mother liquor and flash-cooled. X-ray diffraction data were collected at 100 K at the beam line X06SA at the Swiss Light Source using a MarCCD 225-mm detector (Villigen, Switzerland). Data were indexed and scaled using XDS (Kabsch, 1993). Three zinc sites per monomer were located with the Patterson superimposition minimum function implemented in SHELXS (Sheldrick, 1998) and refined using CNS (Brunger et al., 1998).

The refinement against the 2.7 Å native dataset was completed by iteractive cycles of interactive model building using O (Jones et al., 1991) and crystallographic refinement using a maximum likelihood target function with CNS. The final crystallographic model includes residues 2-130, three zinc ions and 7 solvent molecules. The metal sites were refined without any target values for the distances, and the Van der Waals' electrostatic terms were switched off. The side chains of residues 1 and 131 are invisible in the electron density map and were thus refined as alanines. All residues are in the allowed region.

The atomic coordinates and the structure factors have been deposited at the Protein Data Bank, Research Collaboratory for Structural Bioinformatic, Rutgers University, New Brunswich, NJ, USA (http://www.rcsb.org/) with access code XXX.

Chapter 3

Mtb FurA RESULTS

3.1 Purification

The final yield of pure protein was approximately 10 mg per liter of culture. From the last purification step (size exclusion column), the protein was eluted mainly as a dimer and could easily be separated from the aggregated forms (fig. 3.1). Only pure FurA dimer was used for the experiments detailed below. The purity and nature of the protein were confirmed by ESI mass spectroscopy (16361Da) and SDS-PAGE (fig. 3.1).

3.2 Metal analysis

Three different FurA samples were analyzed by microPIXE (table 3.1). In two of them (FurA and FurA+Fe), the amount of Zn^{2+} was nearly constant at half a zinc atom per protein monomer and this site thus appears to be only around 50% occupied. In the third sample (FurA+EDTA), it was possible to completely remove the metal using EDTA as a chelating agent.

After dialysis against an iron salt solution, complete occupation of a second metal site resulted; this was not reversed by consecutive dialysis against a Fe-free solution. The amount of Fe^{2+} in FurA is very close to 1 atom per monomer.

From these results it is possible to distinguish two different metal binding sites in FurA: one zinc site, where the metal comes directly from the protein expression; and one iron site, where the metal had to be added to the protein after the purification.

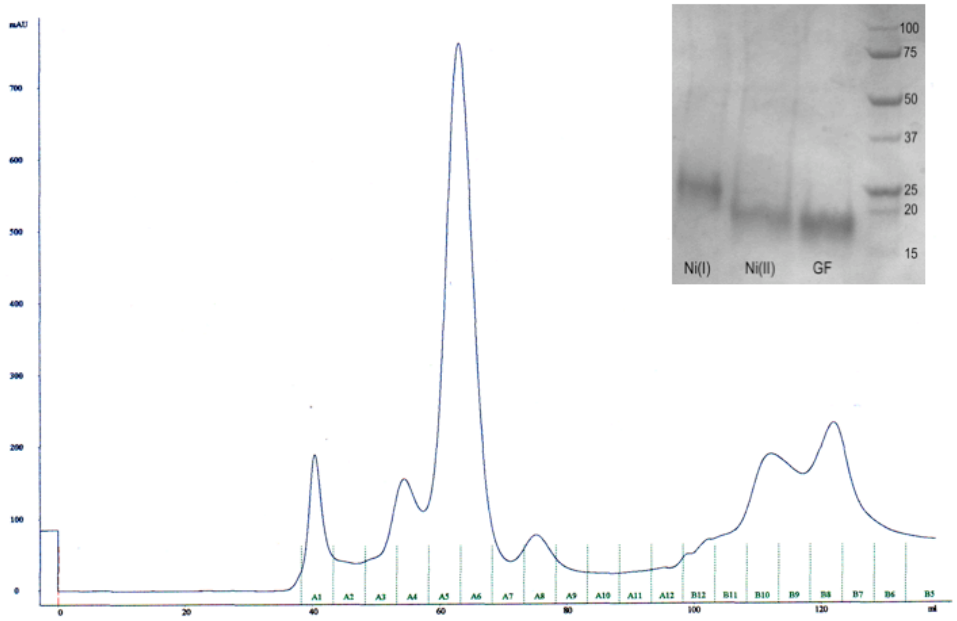

Figure 3.1: FurA purification
Gel filtration (Sup 75 16/60) profile of FurA. In the grey box a SDS PAGE with purification steps; lane 1: first Ni-NTA column [indicated as Ni((I)]; lane2: second Ni-NTA column (after TEV cleavage) [indicated as Ni(II)]; lane 3: after gel filtration [indicated as GF]

Table 3.1: FurA metal quantification by microPIXE

	Zn^{2+}	Fe^{2+}	Ni^{2+}	Cu^{2+}
FurA	0.5(2)	n.d.	0.3(1)	0.4(1)
FurA+Fe	0.5(1)	1.2(1)	0.2(0.5)	0.3(1)
FurA+EDTA	n.d.	0.1(0.5)	n.d.	0.5(2)

In each cell the number of metal ions per protein molecule is reported. FurA is the protein as isolated, FurA+Fe is the protein dialyzed against iron salt solution and FurA+EDTA is the protein EDTA treated. n.d. = not determined. The quoted errors are largely due to low counting statistics. The presence of Ni and Cu results from the purification protocol

The presence of Ni and Cu in the metal analysis results can be explained by the purification protocol, as a Ni-containing column was used and the anti-oxidants contain Cu as an impurity. The microPIXE technique has a sensitivity of between 1 and 10 parts per million of dry sample weight, so these contaminants are readily detected, It is interesting to note that within the experimental error, the occurrence frequencies of Zn, Ni and Cu sums to one. Thus, it appears likely that the zinc ion in this metal binding site may be partially exchanged with another divalent transition metal during purification.

In vivo this site is most likely occupied by zinc ion, hence this is denoted as Zn-FurA. The activated form of the protein containing an additional iron cation is designated Fe,Zn-FurA.

3.3 Isothermal titration

ITC experiments were performed to determine the number of additional metal binding sites and their affinities. In the protein sample used for the experiment, the protein as isolated (ZnFurA) was used, where the zinc-binding site was already occupied. Zinc chloride was added to the sample, thus causing the second metal site (the regulatory site) to become occupied with zinc ions. Figure 3.2 shows the heat of binding which resulted from each injection. The best fit to the experimental data was achieved using a "single metal binding site" model. From this analysis it is possible to obtain the dissociation constant of the reaction as well as thermodynamic parameters such as enthalpy and entropy (table 3.2). The association constant, $1.27 \cdot 10^5$ M, underlines the metal-binding affinity of the FurA for a second more loosely bound cation which is therefore exchangeable, and the value obtained is in the range of the binding constants for the regulatory sites (Mills and Marletta, 2005).

ITC titration shows a single binding event for the ZnFurA sample and confirms the metal analysis results (fig. 3.2).

Table 3.2: Thermodynamic parameters of FurA-zinc titration

K_a[M]	$1.27(7)\cdot10^5$
K_d[μM]	7.9(4)
ΔH[cal]	$-1.99(5)\cdot10^4$
ΔS[cal]	43.6

Thermodynamic parameters based on the ITC measurements depicted in figure 3. K_a is the association constant and K_d is the dissociation constant extracted from a fit to integrated-binding heats.

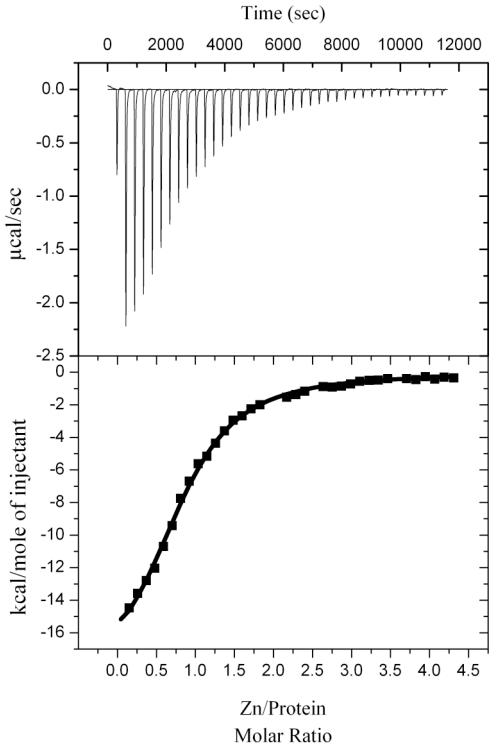

Figure 3.2: FurA-zinc ITC

FurA ITC titration against zinc: in the top panel the experimental data are plotted. In the bottom panel experimental data (squares) are fitted to a "one metal binding site model" (solid line)

Titration with Fe^{2+} solution failed to give reasonable results due to the oxidation of the ferrous iron to the ferric state. The use of a reducing agent led to a large heat of dilution and did not completely prevent oxidation of Fe^{2+} to Fe^{3+} during the titration

3.4 EXAFS analysis

3.4.1 Zinc site

The EXAFS spectrum and its Fourier transform (FT) of Fe,Zn-FurA, collected at the zinc edge, are shown in figure 3.3. The fine structure is dominated by one frequency with intense oscillations at high *k*. This corresponds to a strong first shell peak at 2.3 Å in the Fourier transform (FT) typical for metal-sulfur coordination. A second contribution visualized by the FT at about $2Å^{-1}$ indicates the presence of a light ligand such as oxygen or nitrogen. Models based on these considerations yielding coordination numbers of 4, 5, and 6 were compared with the data. The major results are summarized in table 3. Only two of the fits can be retained for further consideration: 3S+1(N/O) (model B_{Zn} table 3.3) and 2S+2(N/O) (model C_{Zn} table 3.3).

3.4.2 Iron site

The edge region (XANES), in particular the intensity of the 1s-3d transition (pre-edge peak), is sensitive to the electronic and geometric structure of the iron site and serves as a measure for the coordination and/or ligand sphere homogeneity. The 1s-3d transition becomes more intense as the metal environment deviates from centrosymmetry or an ideal octahedral coordination. Here the edge position is typical for an oxidation state of +2 (fig. 3.4). An octahedral iron coordination is supported by the area of the pre-edge peak (~11eV) (Roe et al., 1984).

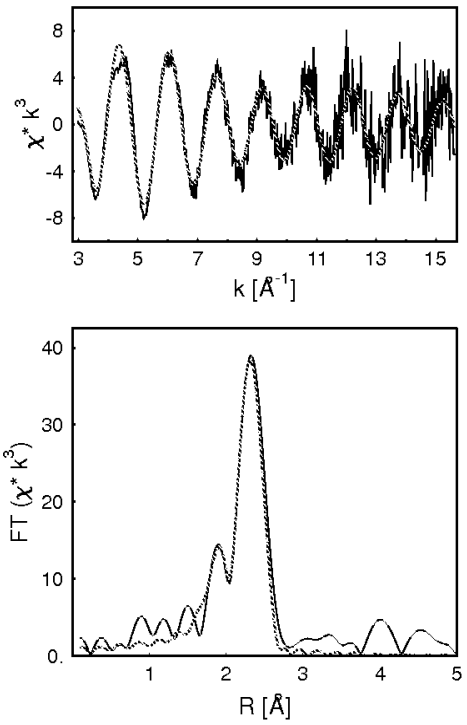

Figure 3.3: FurA zinc EXAFS
The figure shows the Zn K-edge EXAFS (top panel) and the corresponding Fourier transform (bottom panel) for Fe,Zn-FurA. The solid line represents the experimental data; the dashed line the best fit (C_{Zn}). Details of the model are stated in Table 3. $\chi(k)$ is the EXAFS amplitude; R[Å], the metal–ligand distance corrected for first-shell phase shifts; FT, the Fourier transform amplitude.

Table 3.3. Models for zinc coordination in FurA

Model	N	R[Å]	$2\sigma^2$[Å^2]	EF	R_{EXAFS}
A_{Zn}	4S	2.32(1)	0.012(1)	-9.6(9)	49.36
B_{Zn}	3S	2.32(1)	0.009(1)	-9.3(7)	41.80
	1O/N	1.99(1)	0.009(1)		
C_{Zn}	2S	2.32(1)	0.005(1)	-9.5(5)	38.64
	2O/N	2.02(1)	0.010(1)		

N is the co-ordination number, R the mean interatomic distance, $2\sigma^2$ the Debye-Waller parameter and EF the shift of the energy origin (Fermi energy) with respect to E_0=9660eV. Numbers in parentheses represent the uncertainties of the last digit.

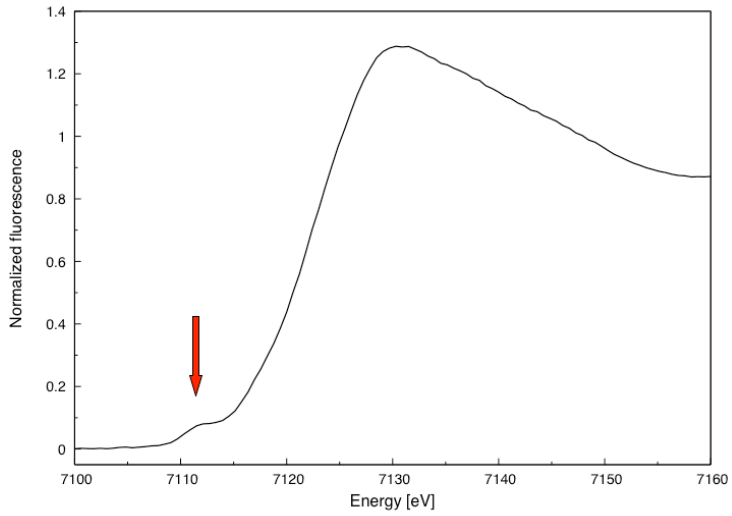

Figure 3.4: FurA iron XANES

Normalized Fe XANES spectrum of Fe,Zn-FurA. The pre-edge peak indicates the 1s→3d transition and its area (11eV) is at the upper limit for octahedral iron coordination. This is consistent with the wide white line at 7120eV indicating the presence of a heterogeneous metal coordination.

The extracted fine structure and its Fourier transform (FT) of Fe,Zn-FurA, collected at the iron edge, are shown in figure 3.5.

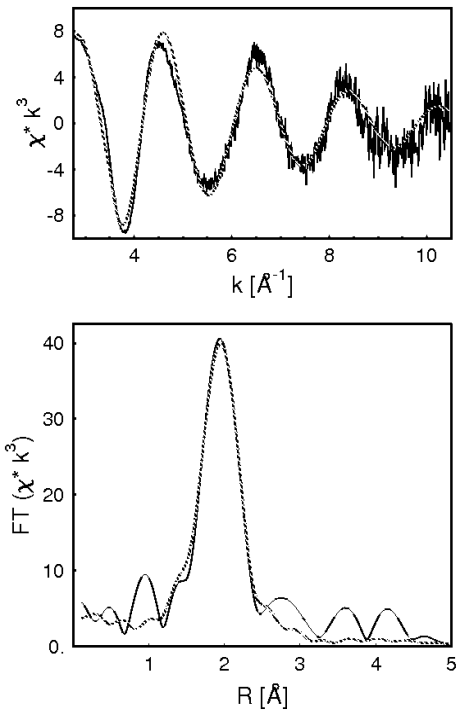

Figure 3.5: FurA iron EXAFS
Fe K-edge EXAFS (top panel) and the corresponding Fourier transform (bottom panel) for Fe,Zn-FurA. The solid line represents the experimental data, the dashed line the best fit. Details of the model are stated in Table 4. χ(k), the EXAFS amplitude; R[Å], the metal–ligand distance corrected for first-shell phase shifts; FT, Fourier transform amplitude.

In order to achieve satisfactory fits, sulfur backscattering had to be excluded. In fact, whenever such a contribution was introduced in the refinement procedure, it resulted in an unreasonably short Fe-S distance and a low quality fit. As already indicated by the XANES analysis, the hexa-coordinated model is the only plausible one because both the penta- and the tetra-coordination lead to unrealistically high R-factors.

Three models were evaluated in detail and the results are shown in table 3.4. Only the models (see table 3.4) resulting in good quality indicators are discussed further: 3O+3N and 4O+2N. These models are identical within the XAS error range.

Table 3.4: Models for iron coordination in FurA

Model	N	R[Å]	$2\sigma^2$[Å^2]	EF	R_{EXAFS}
A_{Fe}	5O/N	1.99(1)	0.020(1)	-13.0(5)	31.6
	1S	2.27(1)	0.020(1)		
B_{Fe}	4O	2.04(1)	0.014(1)	-13.0(3)	27.30
	2N	1.96(1)	0.010(1)		
C_{Fe}	3O	2.06(1)	0.007(1)	-12.7(3)	27.04
	3N	1.96(1)	0.007(1)		

N is the co-ordination number, R the mean interatomic distance, $2\sigma^2$ the Debye-Waller parameter and EF the shift of the energy origin (Fermi energy) with respect to E_0=7120eV. Numbers in parentheses represent the uncertainties of the last digit

Strong back-scattering from neighboring atoms between 2.5 and 4 Å in both spectra was not observed. This has to be highlighted as it suggests the absence of a metal back-scattering, ruling out a binuclear centre.

3.5 DNA affinity

Electro-mobility shifts assays were performed using FurA protein and different lengths of Fur-box nucleotides:

20_fwd 5'-GATAATGATAATCATTATCT-3'

20_rev 5'- AGATAATGATTATCATTATC-3'

22_fwd 5'-GAGATAATGATAATCATTATCT-3'

22_rev 5'- AGATAATGATTATCATTATCTC-3'

25_fwd 5'-CTAGAGATAATGATAATCATTATCT-3'

25_rev 5'- AGATAATGATTATCATTATCTCTAG-3'

From the gels (fig. 3.7) it is possible to observe DNA-binding with fairly weak affinity. FurA is able to bind the canonical nucleotide recognised by all ferric uptake regulators.

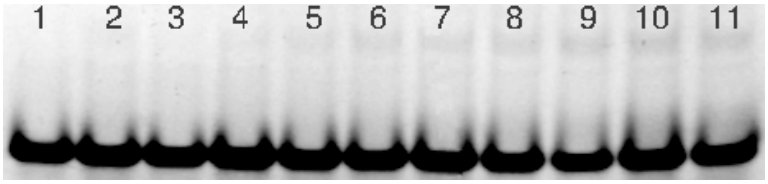

Figure 3.7: FurA DNA-binding assay
Affinity of FurA for the 20 bp Fur-box. DNA concentration was ~ 0.7 nM. The protein was mixed with an equimolar amount of zinc ion and the concentration was as follows: lane 1, 0 nM (only DNA); lane 2, 1 nM (no zinc ion); lane 3, 10 nM (no zinc ion); lane 4, 5 nM; lane 5, 10 nM; lane 6, 20 nM; lane 7, 30 nM; lane 8, 40 nM; lane 9, 50 nM; lane 10, 60 nM; lane 11, 70 nM.

3.6 Homology model

A homology model was obtained after submitting the amino acid sequence of FurA and the PDB file of Fur from *Pseudomonas aeruginosa* to the SwissModel database. Conservation of the DNA binding domain and particularly of the regulatory site can be observed (fig 3.7). The residues involved in the metal coordination (His 32, Glu 80, His 87, His 89, Glu 100) are also highly conserved in the sequence alignment through the species (see fig. 5.2; numbering of the amino acids referring to the sequence alignment). Unfortunately it is not possible to observe the same correspondence for the structural site (see discussions in chapter 5).

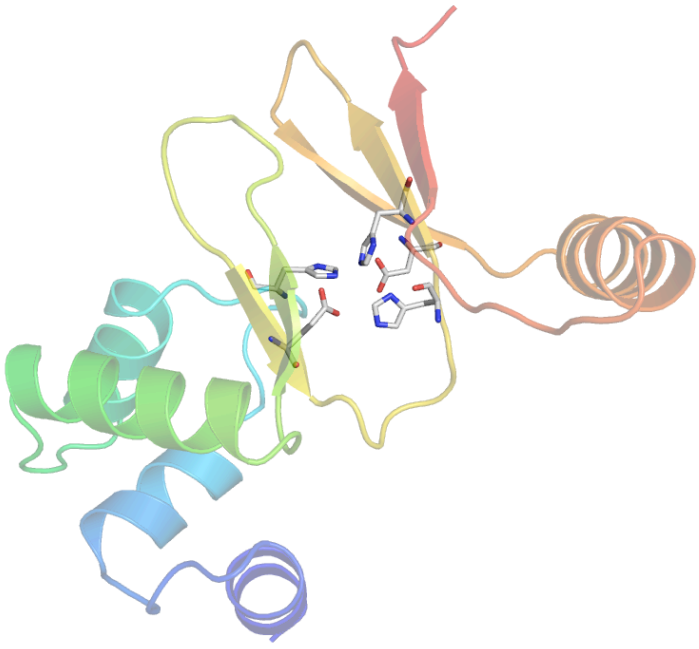

Figure 3.7: FurA homology model

3.7 Proteolysis

Limited digestion proteolysis was performed as a guide for potential crystallization conditions. The following enzymes were used: chemotrypsin, elastase, trypsin and thermolysin. The protein was left for 24 hours in digestion and samples were taken after 5, 10, 30, 60, 120, 300 minutes and after 24 hours (some examples are reported in figure 3.8 and 3.9).

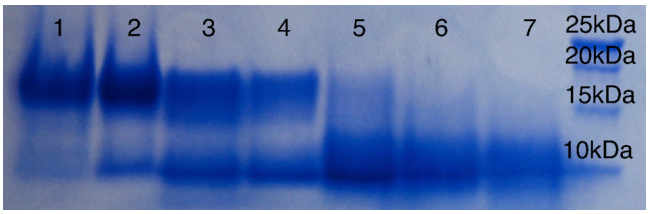

Figure 3.8: FurA digested with chemotrypsin
The reaction has been checked after: 5min, lane 1; 10 min, lane 2; 30 min, lane 3; 60 min, lane 4; 120 min, lane 5; 300 min, lane 6; 24 h, lane 7.

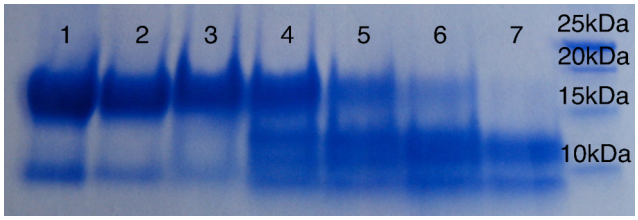

Figure 3.9: FurA digested with elastase
The reaction has been checked after: 5min, lane 1; 10 min, lane 2; 30 min, lane 3; 60 min, lane 4; 120 min, lane 5; 300 min, lane 6; 24 h, lane 7.

Most of the results described in this chapter have been collected in a manuscript recently submitted:

"The regulatory and the structural metal binding sites in FurA *from Mycobacterium tuberculosis*"

Debora Lucarelli, Ehmke Pohl, Matthew Groves, Michael L. Vasil, Elspeth Garman, Wolfram Meyer-Klaucke.

Chapter 4

Mtb FurB RESULTS

4.1 Purification

Mtb FurB was purified as a dimer with a molecular mass of 14596Da per monomer and final yield of 20 mg per litre of culture. The purity of the protein was confirmed by ESI mass spectroscopy and SDS PAGE (fig. 4.1).

4.2 Metal analysis

MicroPIXE analysis was used to determine the type and number of metal ions in the protein. Two samples were prepared: the protein as isolated from the purification and protein dialyzed for 24 hours with a strong chelating agent (1M EDTA). Very interestingly both samples showed one bound zinc ion per monomer (Table 4.1).

It was not possible to remove the metal ion even after treatment with 1M EDTA. The copper found is presumably an impurity from the ascorbate used as mild reducing agent. Surprisingly, it proved impossible to complex the protein with Fe^{2+}, the natural cofactor of all ferric uptake regulators. Only Co^{2+} and Zn^{2+} could be added to the protein and a particularly strong affinity was observed for Zn^{2+}.

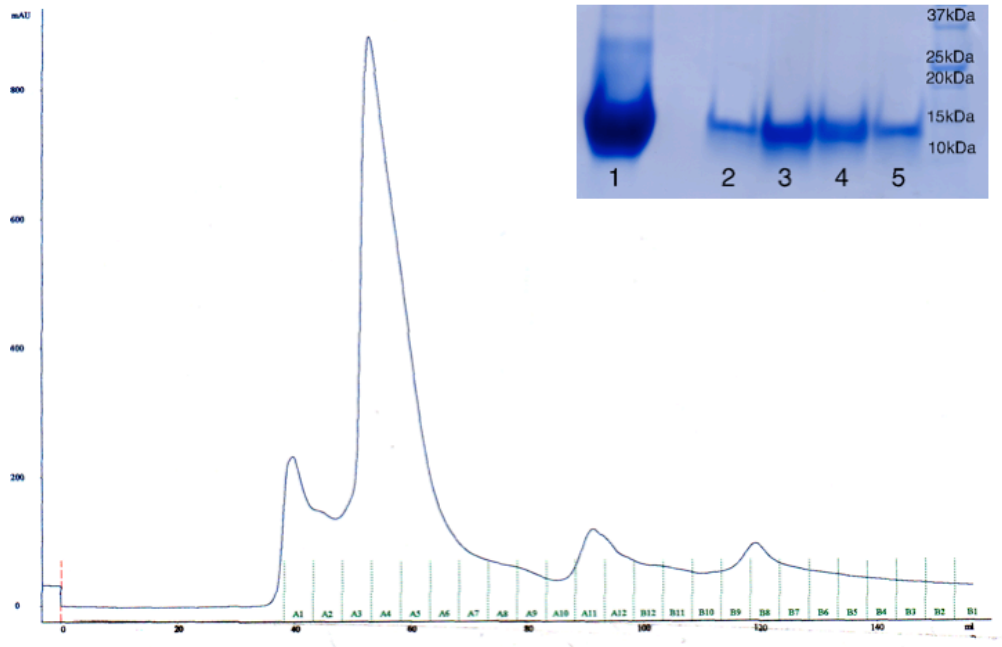

Figure 4.1: FurB purification

Gel filtration (Sup 75 16/60) profile of FurB. In the blue box SDS PAGE purification steps; lane 1: FurB before injection into the column; lane2: fraction A3; lane 3: fraction A4; lane 4: fraction A5; lane 5: fraction A6.

Table 4.1: FurB metal quantification by microPIXE

	Zn^{2+}	Fe^{2+}	Cu^{2+}
FurB	0.8(1)*	0.1(1)	0.3(1)
FurB (EDTA)	1.1(1)	n.d.	0.5(2)

*The relatively large errors quoted are due to low counting statistics. The presence of copper is due to contaminations during the purification protocol. N.d.: not determined as the concentration was below the limit of detection

4.3 Isothermal titration

ITC experiments were performed to determine the number of additional metal binding sites and their affinities. Protein as isolated from the last purification step was used for this experiment. Here a presumably very tightly bound zinc ion was detected by microPIXE, and this site should be added to the ITC results in order to obtain the final

number of metal sites occupied. Zinc chloride was added to the sample, thus causing occupation of two more sites. In order to reproduce this result, cobalt chloride was also utilized and the same number of binding sites was observed. In figures 4.2 and 4.3, the heat released after each injection is shown. The best fit to the experimental data was achieved using a "double metal binding site" model in both experiments.

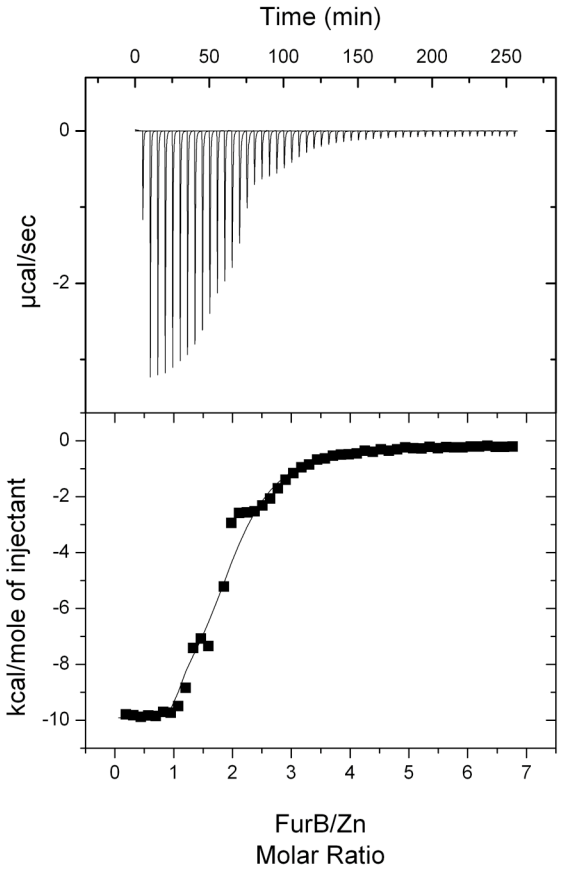

Figure 4.2: FurB-zinc ITC

Titration of FurB with a zinc solution. In the top panel the experimental data are plotted, in the bottom experimental data (squares) are fitted to a "double metal binding site model" (solid line).

From both experimental curves it is possible to observe two metal binding sites. This number has to be added to the already occupied locus, therefore it is possible to conclude that FurB can bind at least three metal ions.

The collected and presented data need to be optimised in order to achieve clear information on the binding constant and thermodynamic parameters, but they are certainly sufficient for estimating the stoichiometry of the reaction.

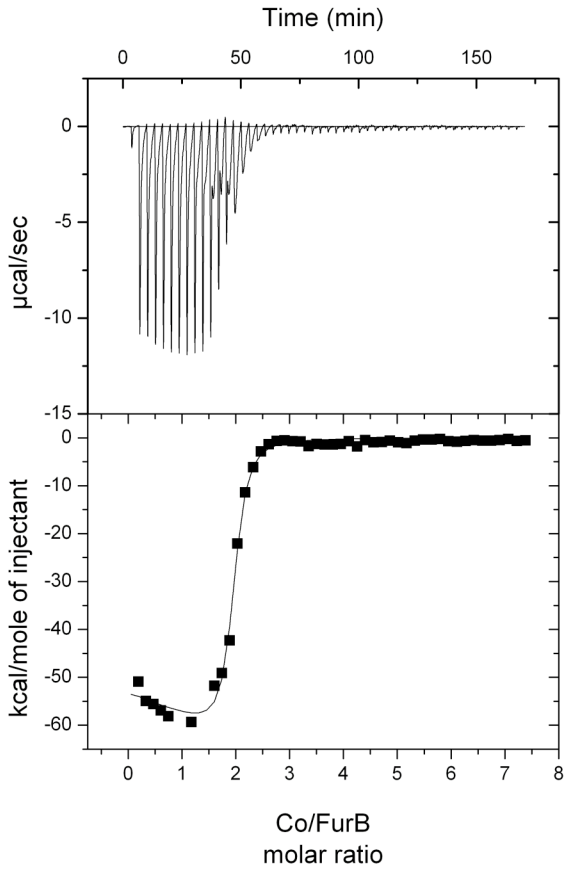

Figure 4.3: FurB-cobalt ITC
Titration of FurB with a cobalt solution. In the top panel the experimental data are plotted, in the bottom experimental data (squares) are fitted to a "double metal binding site model" (solid line). Blank solution is not subtracted and glitches (due to external perturbations) are present during injections 15-20.

4.4 DNA affinity

Electrophoretic mobility shift assays (EMSA) were performed in order to clarify the biological role of FurB. As shown in figure 4.4, the protein was unable to bind the canonical Fur-Box (5'-CTAGAGATAATGATAATCATTATCT-3'). Considering the strong affinity of the protein for zinc, we identified possible DNA targets by locating gene clusters in the Mtb genome that show similarity to the Zur-regulated *znuABC* operon in *E. coli*. In the Mtb genome, a triplet of such genes was found. These genes, Rv2059 (2315172-2316707), Rv2060 (2316277-2316678), and Rv2061c (2316679-2317083), coding respectively for a hypothetical adhesion protein, a possible conserved integral membrane protein and a hypothetical phosphatase protein, show high sequence similarity with the *E. coli znuABC* operon. Another analogy can be seen in the intergenic region of 70 base pairs between the Rv2059 gene and the Rv2060-Rv2061c genes that are co-transcribed. Even the molecular dimensions of the proteins seem to be conserved in Mtb. ZnuA is larger than ZnuB and ZnuC which have approximately the same molecular weight. This same pattern is observed for Rv2059, which is larger than the similarly sized Rv2060 and Rv2061c. In the promoter region of this operon an almost perfectly palindromic 27 base pair sequence was identified (5'-CAATAATGAAAACTGTTATCGATAAGG-3') and consecutively used for the gel shifts (fig. 4.4). The shift results showed that, remarkably, Mtb FurB only binds to this operator sequence and not to the canonical Fur-box (fig. 4.4).

Once the DNA partner was found, the metal affinity/specificity of FurB was tested. Interestingly, only Zn^{2+} can activate the regulator (fig. 4.5). Mtb FurB appears to be more selective than the Fur proteins from other organisms, which can be activated *in vitro* by a variety of divalent transition metal ions (Mills and Marletta, 2005)

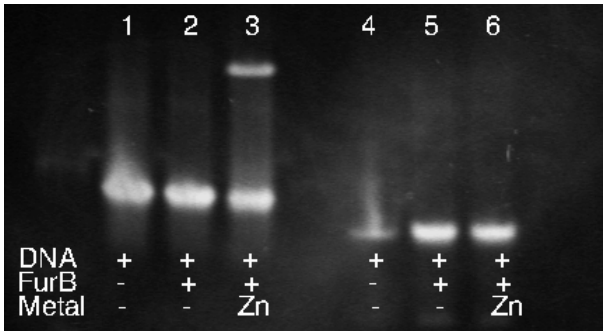

Figure 4.4: FurB EMSA with DNA boxes
EMSA of FurB binding to the promoter region of Rv2059 (lane 1-3) and to the canonical Fur-Box (lane 4-6). Lane 1: DNA only; lane 2: DNA + FurB; lane 3: DNA + FurB + Zn^{2+}; lane 4: Fur-Box only; lane 5: Fur-Box + FurB; lane 6: Fur-Box +FurB + Zn^{2+}.
Promoter region of Rv2059: 5'-CAATAATGAAAACTGTTATCGATAAGG-3'.
Fur-box: 5'-CTAGAGATAATGATAATCATTATCT-3'.

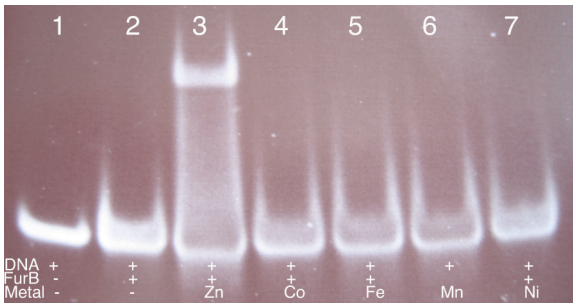

Figure 4.5: FurB EMSA with several metal ions
EMSA of FurB binding to the Rv2059 promoter region with different divalent transition metals. Lane 1: DNA only; lane 2: DNA + FurB; lane 3: DNA + FurB + Zn^{2+}; lane 4: DNA + FurB + Co^{2+}; lane 5: DNA + FurB + Fe^{2+}; lane 6: DNA + FurB + Mn^{2+}; lane 7: DNA + FurB + Ni^{2+}.

4.5 FurB-DNA complex

FurB-DNA complexes were purified with either the 27 base pairs sequence

5'-CAATAATGAAAACTGTTATCGATAAGG-3'

or the 25 base pairs sequence:

5'-AATAATGAAAACTGTTATCGATAAG-3'

in the presence of zinc (fig. 4.6).

The pure complex was set up in crystallization trays.

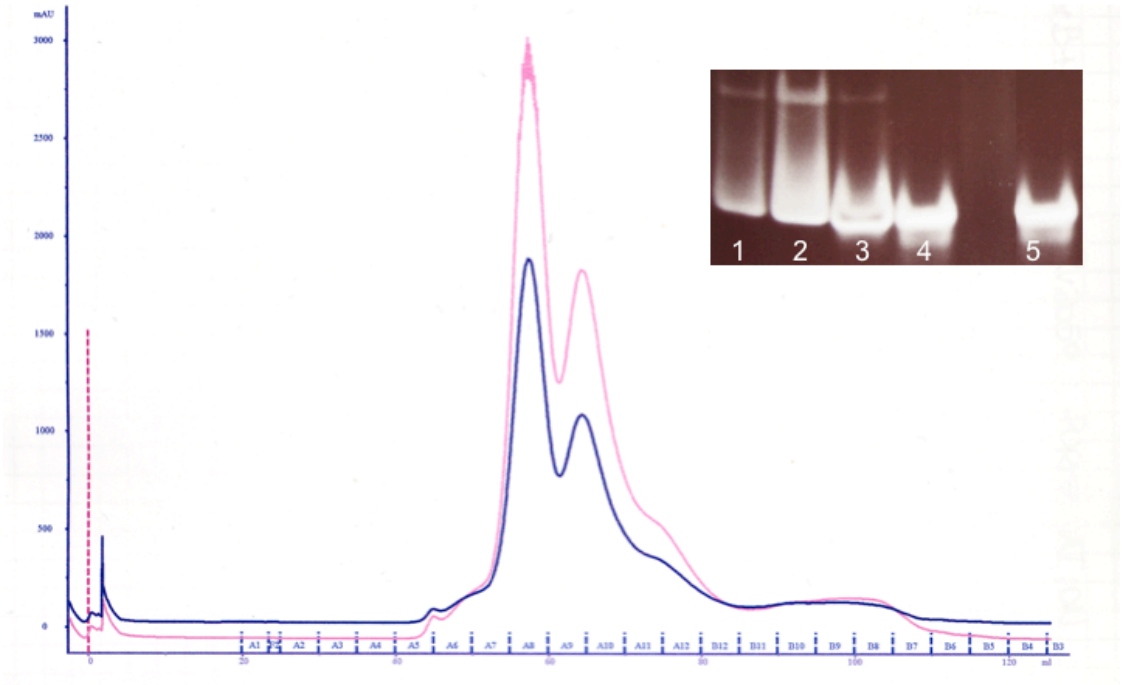

Figure 4.6: FurB-DNA complex purification
Profile of the gel filtration purification of FurB-DNA (27bp) complex. The blue line shows the absorbance at 280 nm and in pink, the absorbance at 260 nm. The first peak corresponds to the complex the second to free DNA and free protein. Fractions have been loaded on the gel shown aside. Lane 1: A8; lane 2: A9; lane 3: A10; lane 4: A11; lane 5: DNA only.

4.6 EXAFS analysis

The initial EXAFS experiments in which the protein sample was incubated with Fe^{2+}-containing solution did not show any significant Fe-fluorescence. This result further supported the microPIXE and gel shift experiments described above, indicating that iron can not bind to FurB. This sample was used to record the Zn-fluorescence in order to determine the chemical environment of the strongly bound zinc cation. To investigate the surrounding of the exchangeable metal site a protein sample was incubated with Co^{2+}. Presumably, the divalent cobalt ion can occupy the site normally reserved for the activating Zn^{2+}. The amount of Co^{2+} in the dialysis buffer, however, had to be kept very low in order to prevent protein precipitation and therefore its final amount was slightly less than equimolar ratio.

4.6.1 Zinc site

The EXAFS spectrum and its Fourier transform (FT), collected at the zinc edge, are shown in figures 4.7 and 4.8. The EXAFS fine structure is dominated by one frequency with intense oscillations at high *k.* This corresponds to a strong first shell peak at 2.3Å in the Fourier transform (FT) typical of metal-sulphur coordination. The absence of a large detectable second shell in the FT between 2.5 and 4Å suggests the absence of metal back-scattering, ruling out a binuclear centre. Two possible models are taken under consideration after data analysis (table 4.2). (a) Due to the small peak at about 2Å, the data could be interpreted as a zinc ion coordinated with three sulphurs and one light atom (O/N); see figure 4.7. (b) The zinc ion could also be surrounded by four sulphur atoms, with the small peak at 2Å regarded as an artefact generated by the Fourier transformation (fig. 4.8).

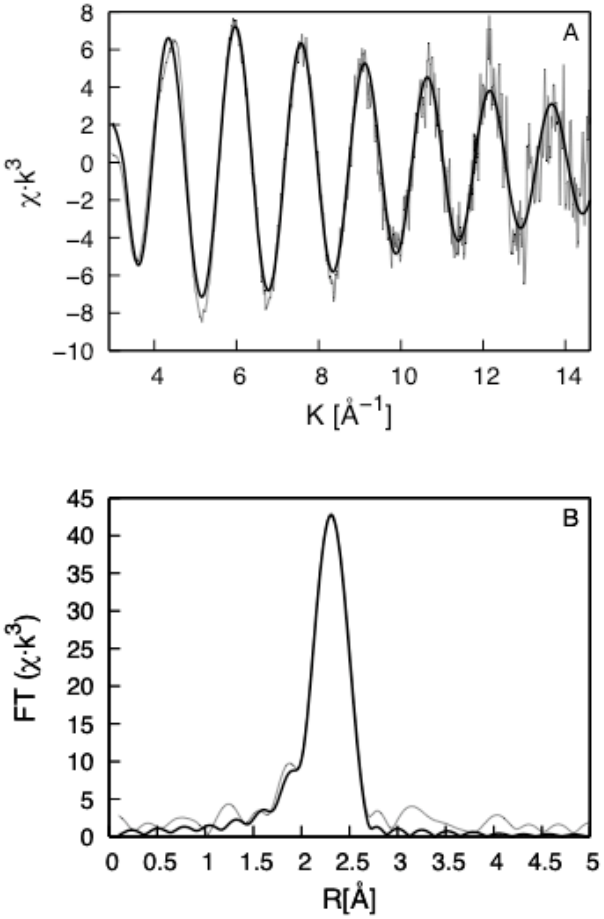

Figure 4.7: FurB zinc EXAFS (A)
Zn K-edge EXAFS (panel A) and the corresponding Fourier transform (panel B) for Zn/CoFurB. The solid line represents the experimental data, the dashed line the best fit. χ(k), the EXAFS amplitude; R[Å], the metal–ligand distance corrected for first-shell phase shifts; FT, Fourier transform amplitude.

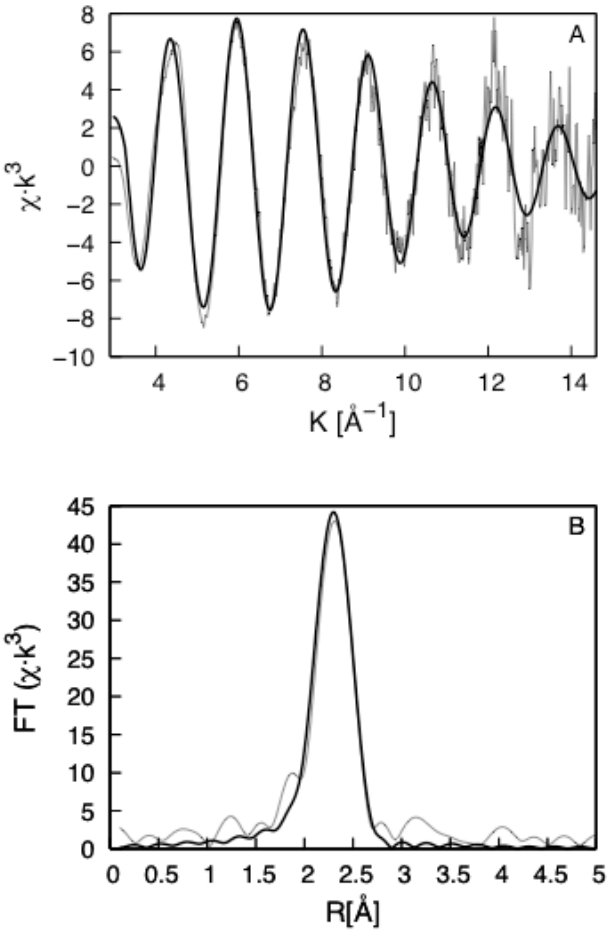

Figure 4.8: FurB zinc EXAFS (B)
Zn K-edge EXAFS (panel A) and the corresponding Fourier transform (panel B) for Zn/CoFurB. The solid line represents the experimental data, the dashed line the best fit. χ(k), the EXAFS amplitude; R[Å], the metal–ligand distance corrected for first-shell phase shifts; FT, Fourier transform amplitude.

Table 4.2 Models for zinc coordination in FurB

Model	N	R[Å]	$2\sigma^2$[Å^2]	EF	R_{EXAFS}
A_{Zn}	4S	2.33(1)	0.009(1)	-11.6(4)	29.73
B_{Zn}	3S	2.33(1)	0.006(1)	-11.0(7)	25.45
	1O/N	2.00(1)	0.010(2)		

N is the co-ordination number, R the mean interatomic distance, $2\sigma^2$ the Debye-Waller parameter and EF the shift of the energy origin (Fermi energy) with respect to E_0=9660 eV and R_{EXAFS} gives an extimation of the goodness of the fit. Numbers in parentheses represent the uncertainties of the last digit.

Both models resulted in similar fit parameters; even thought model B_{Zn} shows slightly better parameters, the results of the crystal structure analysis suggest and confirm the tetragonal environment with four sulphur atoms from cysteines (ZnS_4).

4.6.2 Cobalt site

The area of the 1s →3d transition in the normalized Co XANES spectrum is shown in figure 4.9 and is compatible with a 4-coordinated Co^{2+}. The extracted fine structure and its Fourier transform (FT), collected at the cobalt edge, are shown in figure 4.10. The first shell FT peak width and asymmetry suggests the presence of two shells of nearest neighbours. The split peak observed in the EXAFS spectrum indicates the presence of histidines in the metal environment. In all the models evaluated, sulphur backscattering had to be included in order to achieve satisfactory fits, but introduction of more than one sulphur contribution in the coordination sphere produced unreasonable results and Debye-Waller factors. The best fit obtained shows one sulphur, one oxygen and two histidine nitrogens (table 4.3).

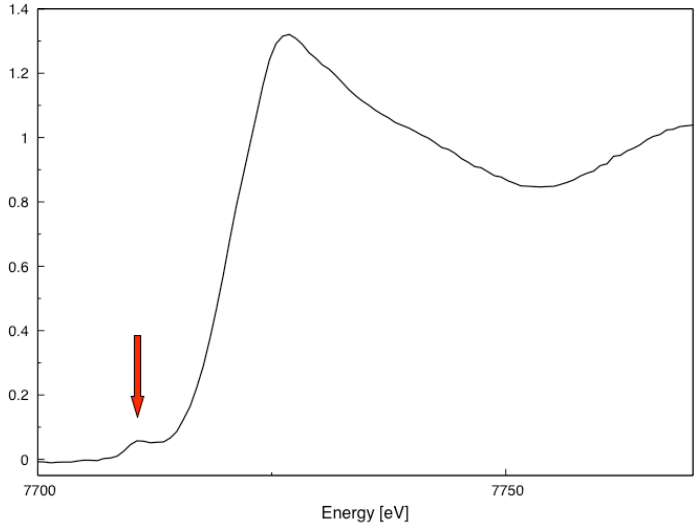

Figure 4.9: FurB cobalt XANES
Normalized Co XANES spectrum of Co/ZnFurB. The pre-edge peak indicates the 1s→3d transition and its area (15eV) is compatible with a tetrahedral coordination.

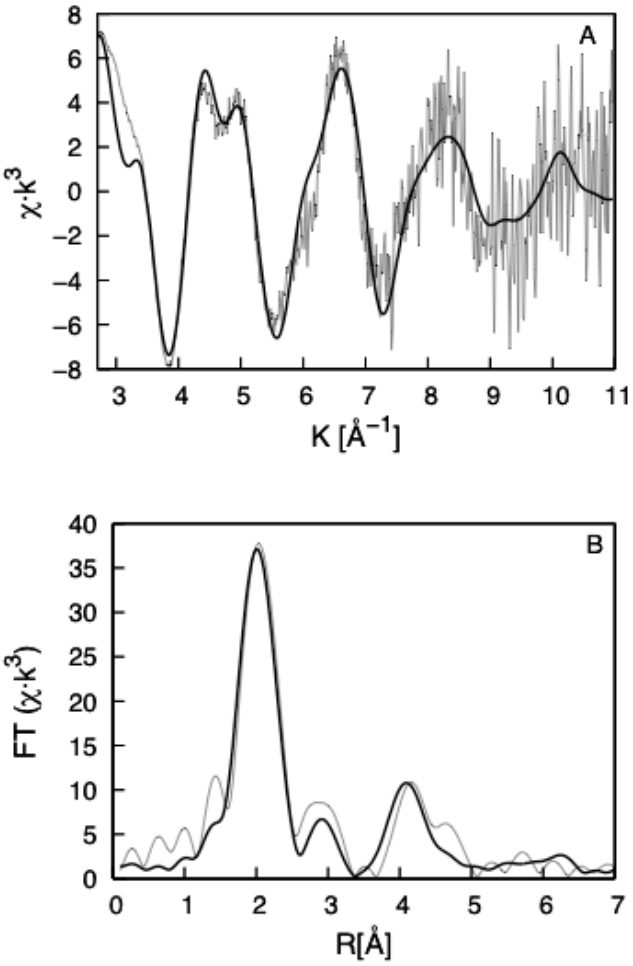

Figure 4.10: FurB cobalt EXAFS
Co K-edge EXAFS (panel C) and the corresponding Fourier transform (panel D) for Zn/CoFurB. The solid line represents the experimental data, the dashed line the best fit. χ(k), the EXAFS amplitude; R[Å], the metal–ligand distance corrected for first-shell phase shifts; FT, Fourier transform amplitude.

Table 4.3: Models for cobalt coordination in FurB

Model	N	R[Å]	$2\sigma^2$[Å^2]	EF	R_{EXAFS}
A_{Co}	1O	1.98(2)	[0.012(3)]	-4.25(7)	38.9
	2His	1.99(2)	0.008(2)		
	1S	2.26(1)	0.011(2)		
B_{Co}	3His	1.98(1)	0.008(1)	-4.5(5)	41.73
	1S	2.26(1)	0.011(2)		

N is the co-ordination number, R the mean interatomic distance, $2\sigma^2$ the Debye-Waller parameter and EF the shift of the energy origin (Fermi energy) with respect to E_0=7716eV and R_{EXAFS} gives an estimation of the goodness of the fit. Numbers in parentheses represent the uncertainties of the last digit.

4.7 Proteolysis

Limited digestion proteolysis was performed in order to gain additional information for possible crystallization conditions. The following enzymes have been used: chemotrypsin, elastase, trypsin and thermolysin. The protein was left for 24 hours in digestion and samples were taken after 5, 10, 30, 60, 120, 300 minutes and after 24 hours (some examples are reported in figure 4.11 and 4.12).

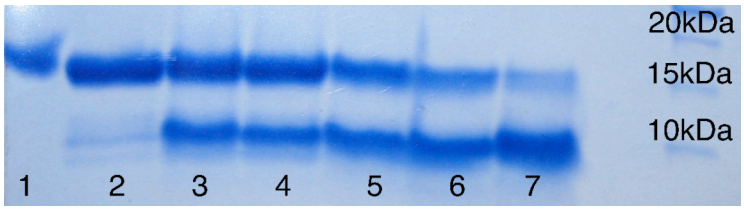

Figure 4.11: FurB digested with elastase
The reaction was checked after: 5min, lane 1; 10 min, lane 2; 30 min, lane 3; 60 min, lane 4; 120 min, lane 5; 300 min, lane 6; 24 h, lane 7.

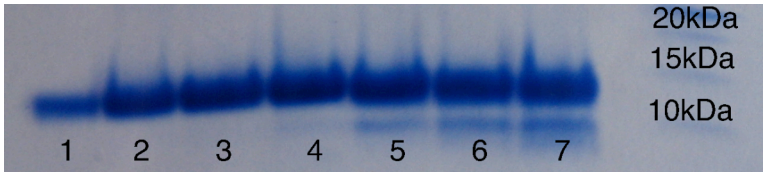

Figure 4.12: FurB digested with thermolysin
The reaction was checked after: 5min, lane 1; 10 min, lane 2; 30 min, lane 3; 60 min, lane 4; 120 min, lane 5; 300 min, lane 6; 24 h, lane 7.

None of the experiments shown here were taken forward for crystallization, as the cuts resulted in too long protein fragments and therefore lose of the integrity of protein domains.

4.8 Mutants

Crystallization of this protein was a long and irksome job. One of the main problems was the presence of a heavy skin. It appeared seemingly at random after minutes, hours or days independently of the conditions. In order to prevent skin formation, some of the cysteine residues were mutated. Due to sequence analogies with FurA (fig. 4.13), we assumed Cys86 and Cys89 to be involved in the metal binding. EXAFS data confirmed that some of the cysteines (at least four) are involved in metal binding. The full sequence of FurB presents a total of five cysteines; therefore we decided to mutate each of them separately starting with Cys76, Cys126 and Cys129.

Each of these three mutants was expressed and purified (fig. 4.14) and set in crystallization trays. After less than one day the drops containing the mutants Cys126Ser and Cys129Ser showed heavy precipitation.

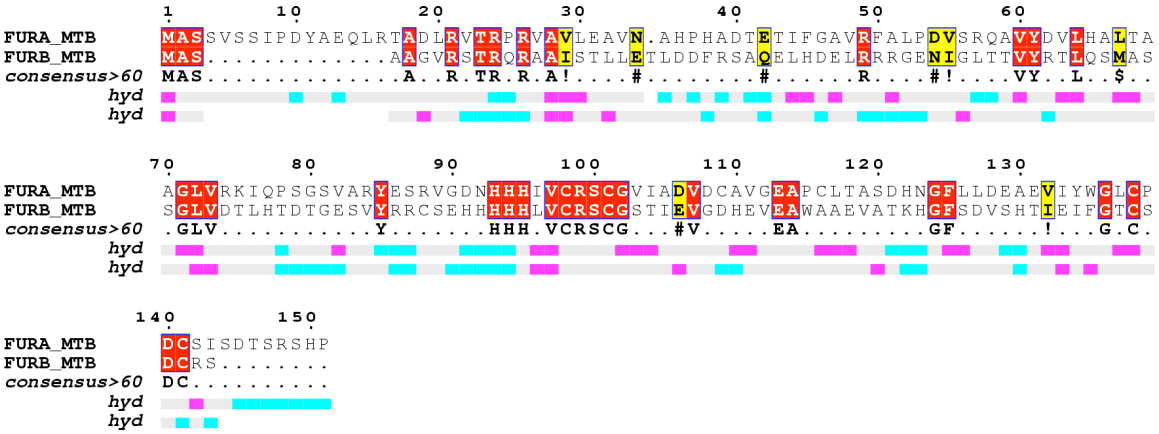

Figure 4.13: Sequence alignment of Mtb FurA and Mtb FurB
It is possible to observe that 2 cysteine-motifs (Cys98XXCys101 and Cys139XXCys141) are conserved in both targets. The yellow boxes represent the conserved regions. In red are highlighted the identical residues. `hyd' indicates the calculated hydropathy (pink: hydrophobic, grey: intermediate and cyan: hydrophilic). The alignment has been calculated with ClustalW (http://www.ebi.ac.uk/clustalw/index.html) (Pearson, 1990; Pearson and Lipman, 1988) and the graphically displayed with ESPript (http://espript.ibcp.fr/ESPript/cgi-bin/ESPript.cgi) (Gouet et al., 1999).

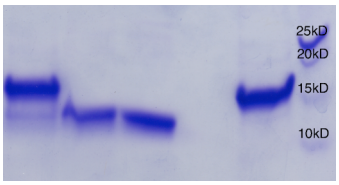

Figure 4.14: SDS PAGE of FurB mutant Cys129Ser.
Purification profile of FurB Cys129Ser: lane 1 (and 4): first Ni-NTA column; lane2: second Ni-NTA column (after TEV cleavage); lane 3: after gel filtration. (All mutants resulted in similar gels)

4.9.1 Crystallization

As previously mentioned, crystallization of this protein required a lot of time and resources. It was not possible to obtain crystals by using any of the commercial screens. Considering that this protein binds DNA, a phosphate screen was created. The first microcrystals appeared after three months and diffracted to 3.5Å (fig. 4.15).

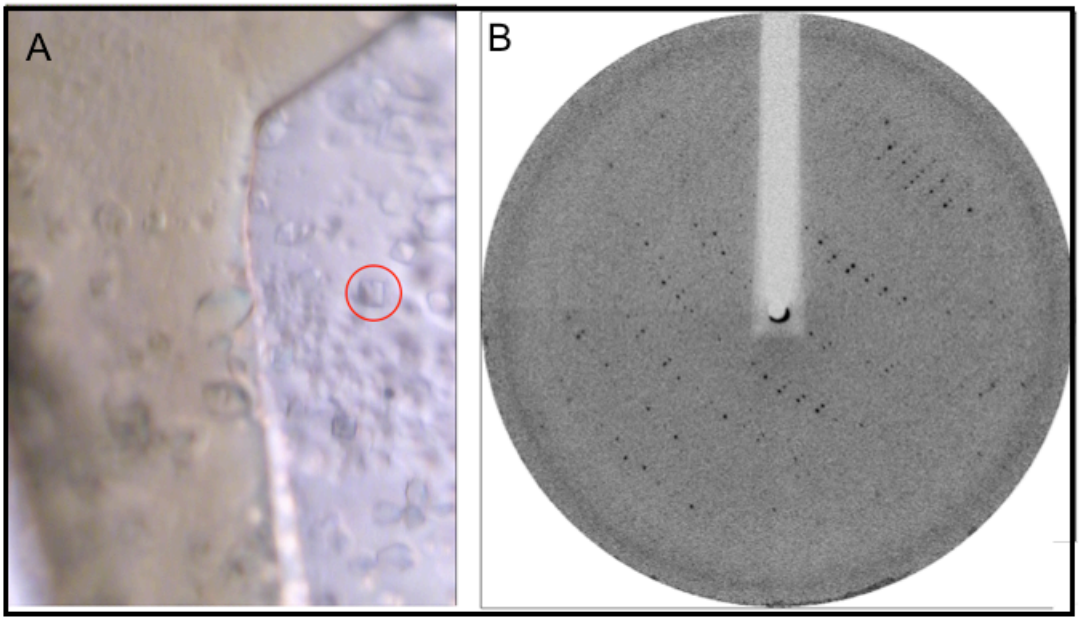

Figure 4.15: First FurB microcrystals
In Panel A the crystals (10x10 μm) are mounted on a micromount. In panel B is reported the diffraction pattern produced by these crystals.

From that date different optimization methods were tried: pH screening, PEG (Polyethylene glycol in different sizes) screening, glycerol screening, phosphate screening, hanging drops, sitting drops, micro and macro seeding, additive screen. More than 200 optimization conditions were tested and more than 300 trays set up (fig. 4.16).

The final crystals were obtained using hanging-drop vapour diffusion by mixing the protein at a concentration of 13 mg/ml to a reservoir solution containing 14% glycerol and 0.3 M $(NH_4)_2HPO_4$. (fig. 4.17).

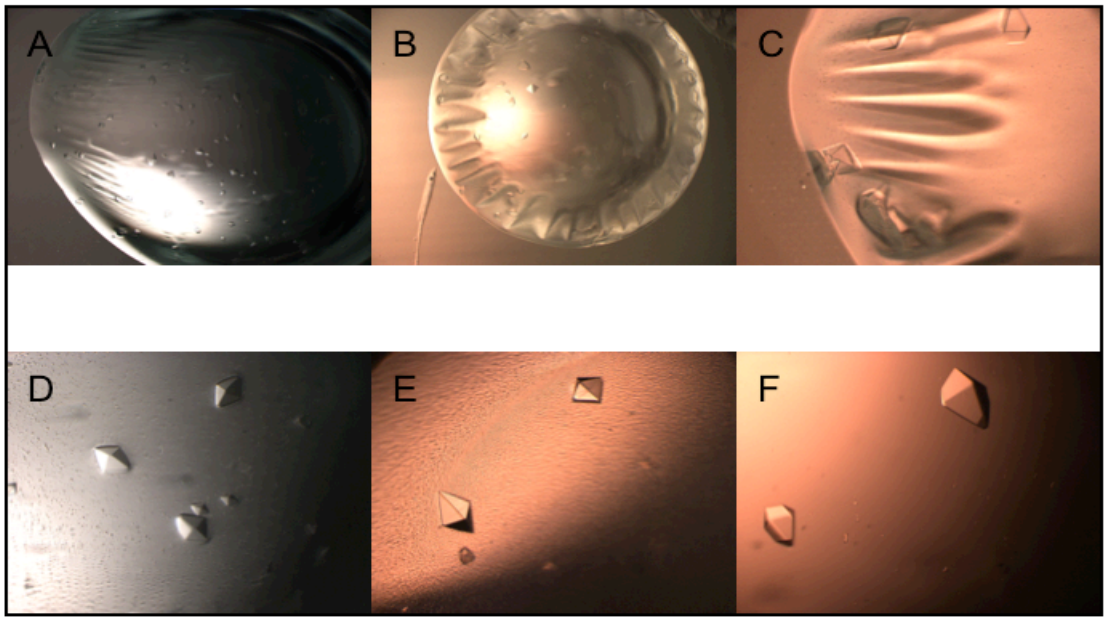

Figure 4.16: Crystals' history
A) First crystals appearing after two months. B) Initial optimization conditions, heavy skin creates a lot of difficulties in growing and fishing the crystals. C) Crystal obtained by steak seeding. D) Crystal obtained with the additive screen. E) Crystal from PEG400 screen. F) Screen obtained form a glycerol screen.

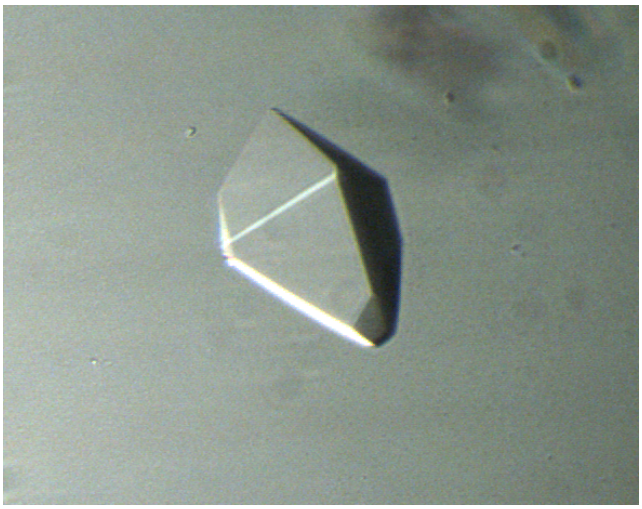

Figure 4.17: FurB crystal diffracting to 2.7 Å
Crystal dimensions: 80x80x200 μm

4.9.2 Crystal structure

The crystal structure of FurB was determined by multi-wavelength anomalous diffraction (MAD) to a minimum Bragg spacing of 2.7 Å and refined to an R-factor of 0.23 (Rfree = 0.25) with excellent geometry (see Table 4.4).

The molecule is located on a crystallographic twofold axis that generates the functional dimer. The basic fold of the monomer, shown in figure 4.18, consists of two domains: the N-terminal DNA-binding domain (residues 1-77) and a C-terminal dimerization domain (residues 77-131). The DNA-binding domain is composed of three helices followed by a two-stranded antiparallel β-sheet. This domain exhibits the typical winged-helix motif with a three-helix bundle (H1, H2 and H3), where helix H3 (residue 34 to 54) is the putative DNA-recognition region (Brennen GR and BW, 1989).

The dimerization domain is constituted of three antiparallel β-strands (S3, S4 and S5) and one long α-helix (H4) further stabilized by one of the metal binding sites. The β-strands S3-S5 form a six-stranded anti-parallel sheet with its symmetry-mates S5'-S3'. In addition, there are a number of hydrophilic and hydrophobic contacts from helix H4 to the symmetry-related H4'. This arrangement results in a tight dimer interface with a buried surface of approximately 2600 $Å^2$ (fig. 4.19).

Table 4.4 Crystallographic data and refinement statistics

Data collection				
	Inflection	Peak	Remote	SAD
Wavelength(Å)	1.28441	1.28358	1.26924	1.28358
Resolution(Å)	39-3.1	39-3.1	39-3.1	33.9-2.7
No. of reflection	52518	52381	52163	90395
No. of unique refl.	6138	6130	6115	9344
Redundancy	8.55	8.53	8.53	9.67
Completeness	98.3(100)	98.4(100)	98.4(100)	98.9(99.1)
I/I(σ) (last shell)	16.6 (4.3)	18.9 (6.3)	18.4 (5.4)	16.2 (3.2)
R_{obs}	0.12 (0.54)	0.10 (0.38)	0.10 (0. 43)	0.09 (0.63)
Refinement				
Resolution [Å]	33.7-2.7			
Space group	$P4_12_12$			
No. reflections	5076			
No. reflection R_{free}	268			
No. of atoms	964			
No. of water	7			
R	0.237			
R_{free}	0.249			
Rmsd bond length	0.015			
Rmsd angles	1.35			

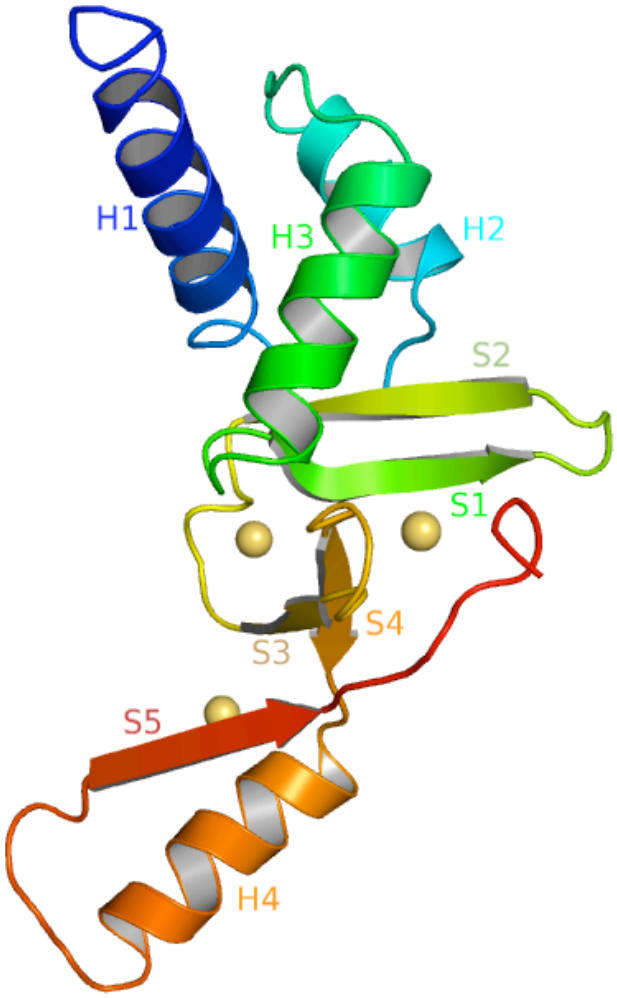

Figure 4.18: Ribbon representation of FurB
Ribbon diagram of the FurB monomer with secondary structural elements annotated. The metal sites are shown as yellow spheres. The DNA-binding domain is shown with colours changing from the N-terminus in blue to green, the dimerization domain from yellow to the C-terminus in red / orange.

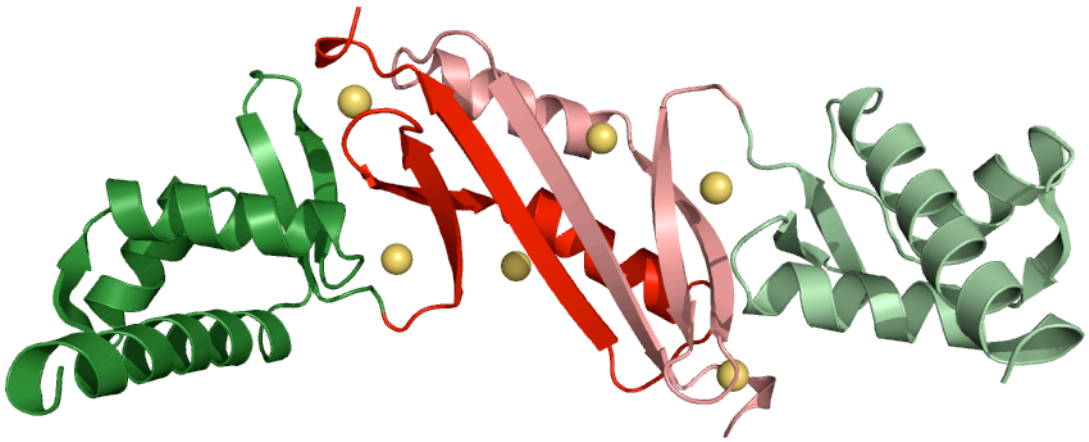

Figure 4.19 Ribbon representation of FurB dimer
Ribbon diagram FurB dimer with the DNA-binding domains in blue and the dimerization domain in orange. The view shown is approximately perpendicular to the crystallographic 2-fold axis.

4.9.3 Metal binding sites in the crystal structure

The crystal structure of FurB possesses three functionally important zinc sites (Zn1-Zn3). Two of the metal ions are in the dimerization domain and one is located in the hinge region between the dimerization and the DNA-binding domains. Each of the metal sites exhibits different structural and chemical environments depending on its functional role. Site 1 is surrounded by Asp 62, Cys 76, His 81 and His 83 (fig. 4.20). This site occupies a strategic position in the protein structure, since two of the coordinating amino acids are part of the DNA domain (Asp 62 and Cys 76) and two belong to the dimerization domain (His 81 and His 83).

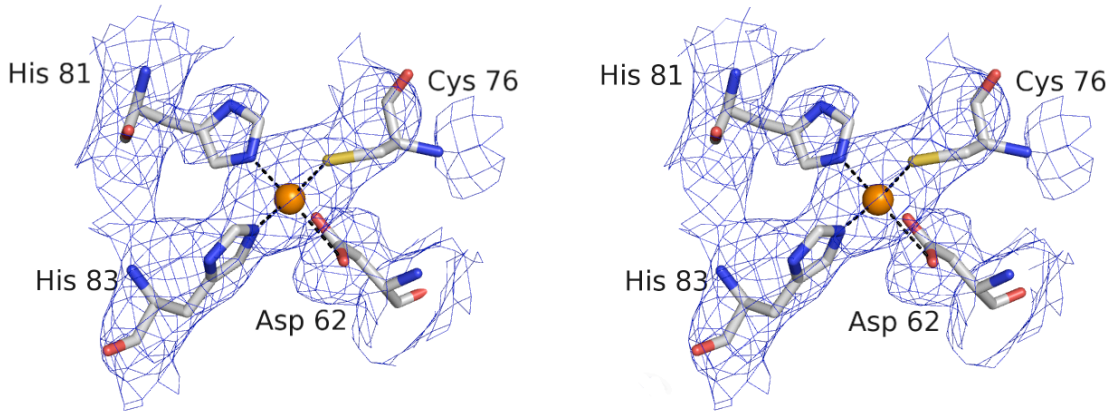

Figure 4.20: FurB zinc site 1
Stereo view of the experimental MAD density depicted at 1σ with the final model in a ball-and-stick representation centred on Zn1.

Site 2 is tetrahedrally coordinated to a cluster of sulphur ligands: Cys86 and Cys89 from the dimerization domain and Cys126 and Cys129 from the C-terminus (fig. 4.21). This geometry was already observed in many other zinc proteins, and the geometrical parameters agree with those previously reported in the literature (Harding, 2004).

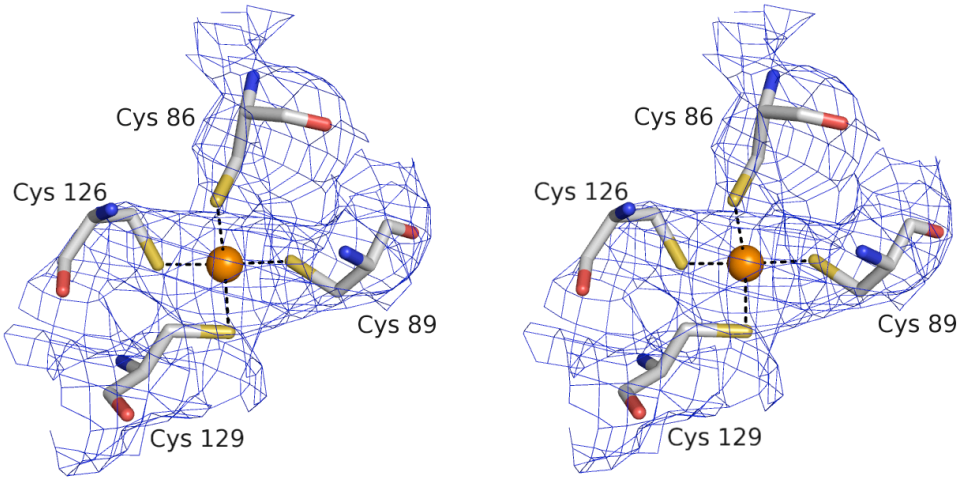

Figure 4.21: FurB zinc site 2
Stereo view of the experimental MAD density depicted at 1σ with the final model in a ball-and-stick representation centred on Zn2.

Site 3 is located in core of the dimerization domain. The Zn^{2+} is tetrahedrally coordinated by the three histidines (His 80, His 82, His 118) and one glutamate (Glu 101) (Fig. 4.22). This metal site links the β-sheet with the C-terminal α-helix of the dimerization domain.

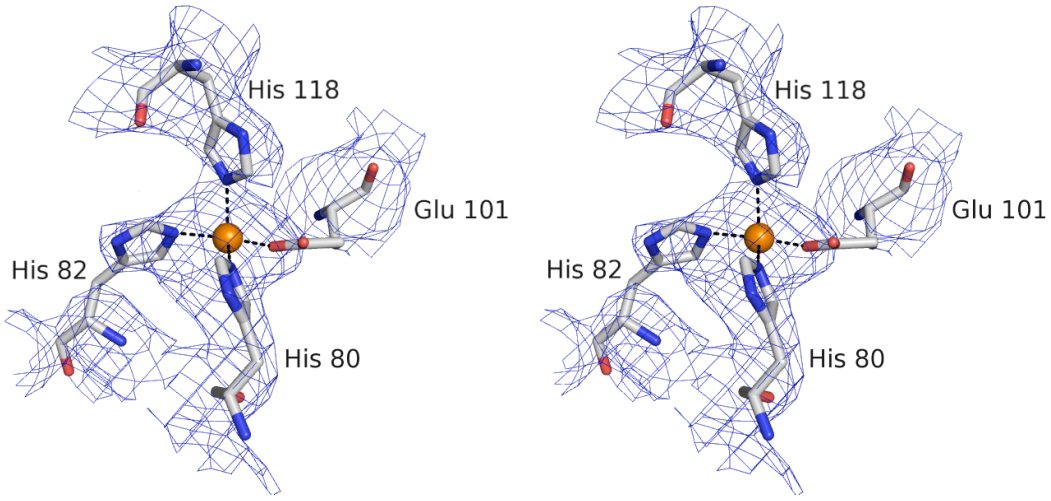

Figure 4.22: FurB zinc site 3
Stereo view of the experimental MAD density depicted at 1σ with the final model in a ball-and-stick representation centred on Zn3.

Most of the results described in this chapter have been collected in a manuscript recently submitted:

"Structure and function of FurB from *M. tuberculosis* – the first Zinc uptake regulator"

Debora Lucarelli, Elspeth Garman, Santina Russo, Wolfram Meyer-Klaucke, Ehmke Pohl.

Chapter 5

Mtb FurA DISCUSSIONS

5.1 FurA possesses two metal sites

The protein was isolated as a dimer containing a partially exchangeable zinc ion site in each monomer. MicroPIXE was performed in order to estimate the amount of metal ions bound to FurA. Results show the presence of 0.5 zinc atom and 0.3 nickel atom per protein monomer in both samples: FurA as isolated and incubated with Fe^{2+}-solution (table 1, chapter 3). No metal presence is observed after EDTA treatment except for the copper impurity introduced by buffer exchange. Considering that one zinc site per monomer has been reported for a number of Fur proteins, it is likely that FurA also possesses a single zinc site per monomer which is partially exchanged by nickel during the purification. Furthermore, a second metal site can be filled with Zn^{2+} as shown by ITC titration, or with Fe^{2+} as shown by microPIXE and EXAFS.

5.2 The structural site

Zinc can have four-, five- and six-coordination number, but in proteins, zinc structural sites always exhibit four ligands and no bound water (Auld, 2004). The ligands are generally located near (within, before or after) β-sheets and often supplied in pairs by short sequences of three to six (mainly located in loop regions) amino acids. The majority of these loop regions is composed of two β-sheets. The sulphur (of a cysteine) is the ligand which most frequently coordinates zinc, but it can also be found in combination with a nitrogen (from a histidine) and/or and oxygen (from aspartic/glutamic acid) (Auld,

2004).

Mtb FurA presents six cysteines in pairs located in relatively short linear sequence of 15 to 33 amino acids and organized in CysXXCys motifs. A recent study on EC Fur (Pecqueur et al., 2006), based on a homology model, proposes the zinc tetrahedrally coordinated by four sulphur atoms. Two homology models have been created for Mtb FurA (fig. 5.1).

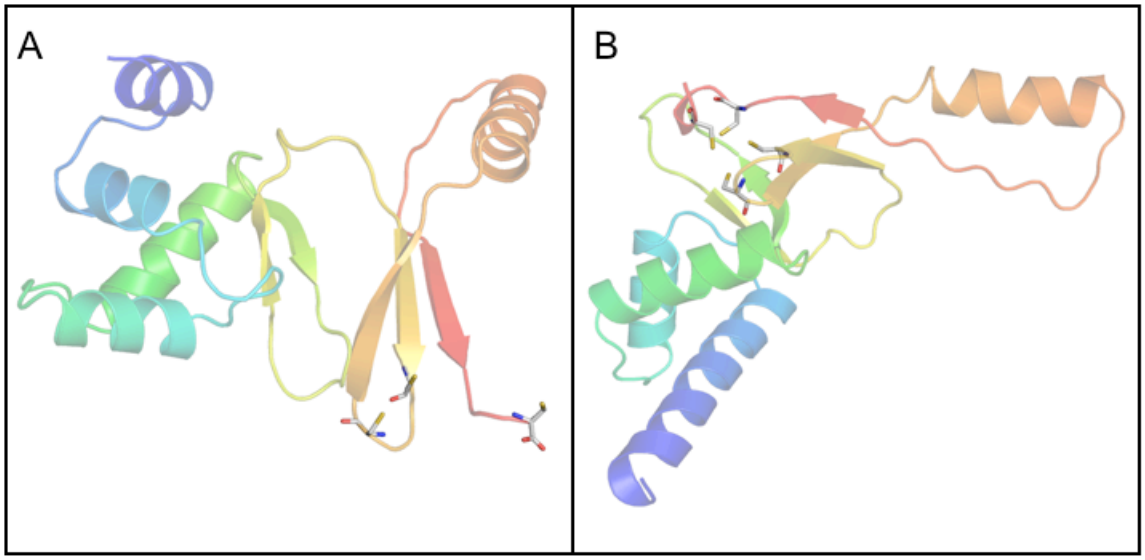

Figure 5.1: Homology models
A) Homology model of FurA based on PAFur. B) Homology model of FurA based on FurB

In figure 5.1 A is reported a ribbon representation of a homology model based on PA Fur crystal structure. In sticks are underlined three of the cysteine residues (Cys 92, Cys 95, Cys Y - third cysteine highlighted in green in figure 5.2: numbering of the amino acids referring to the sequence alignment reported in the figure) conserved in the sequence alignment. It was not possible to add the fourth one (Cys 130) as the Mtb FurA amino acid sequence is longer than PA Fur sequence. Most likely this cysteine is located two residues apart on the C-terminus β-sheet. Considering a possible flexibility of this sheet, it could be possible to imagine a movement of this secondary element so that the four cysteines could assume a typical conformation of a zinc structural site (Auld, 2004).

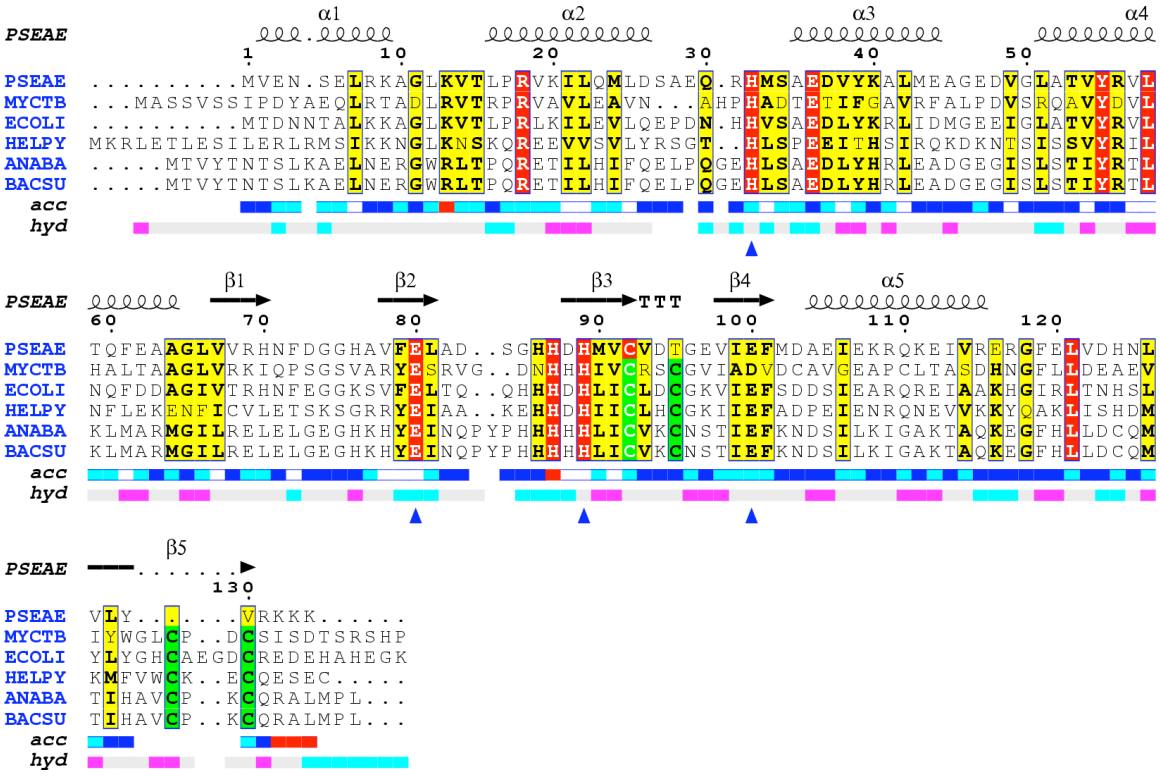

Figure 5.2: Fur proteins alignment
The numbering refers to the PA Fur sequence. The yellow boxes represent the conserved regions. In red are highlighted the identical residues, and in green those cysteine residues potentially involved in metal binding. The arrows indicate the residues forming the regulatory site in *Pseudomonas aeruginosa*. `acc' represents the accessibility to PA Fur (blue: accessible, cyan: intermediate and white: not accessible). `hyd' indicates the calculated hydropathy for Mtb FurA (pink: hydrophobic, grey: intermediate and cyan: hydrophilic). The alignment has been calculated with ClustalW (http://www.ebi.ac.uk/clustalw/index.html) (Pearson, 1990; Pearson and Lipman, 1988) and the graphically displayed with ESPript (http://espript.ibcp.fr/ESPript/cgi-bin/ESPript.cgi) (Gouet et al., 1999).

In figure 5.1 B a ribbon representation of a homology model based on Mtb FurB is shown. Also in this case the four cysteines (Cys 98, Cys 101, Cys 138, Cys 141: numbering of the amino acids referring to the sequence alignment reported in figure 5.3), highly conserved, in the sequence alignment (fig. 5.3) are organized in a typical zinc structural site.

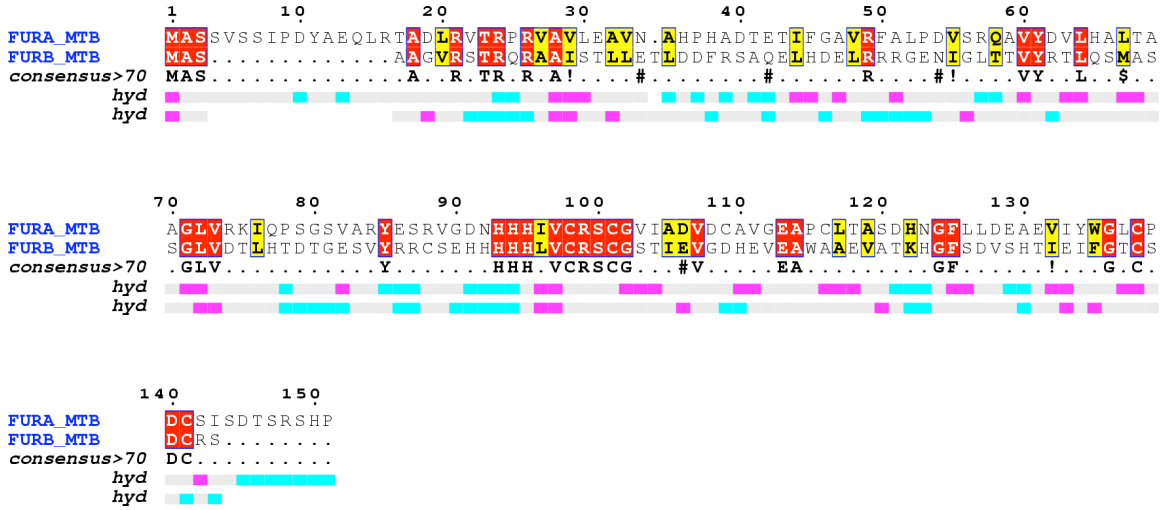

Figure 5.3: Sequence alignment of Mtb FurA and Mtb FurB
The numbering refers to the Mtb FurA sequence. The yellow boxes represent the conserved regions and in red are highlighted the identical residues. `hyd' indicates the calculated hydropathy for Mtb FurA (pink: hydrophobic, grey: intermediate and cyan: hydrophilic). The alignment has been calculated with ClustalW (http://www.ebi.ac.uk/clustalw/index.html) (Pearson, 1990; Pearson and Lipman, 1988) and the graphically displayed with ESPript (http://espript.ibcp.fr/ESPript/cgi-bin/ESPript.cgi) (Gouet et al., 1999).

From both homology models, a zinc ion coordinated to four cysteines seem to be suggested but experimental data show a different coordination.

Spectroscopic analysis of Fe,Zn-FurA reveals a tetrahedral zinc coordination by sulphur and light atoms (O/N). Of all models considered during the EXAFS refinement, the one with two S at 2.32Å and two O/N at 2.01Å is the most persuasive. Distances and Debye-Waller factors are consistent with the values reported for other zinc proteins containing sulphur atoms in the metal coordination sphere (Harding, 2004). It is noteworthy that the same coordination, with similar distances, has also been proposed by EXAFS studies for the zinc binding site of EC Fur (Adrait et al., 1999; Althaus et al., 1999).

An alignment of some of the Fur proteins previously described is depicted in figure 5.2. It is clear that most of the cysteine residues are conserved in all targets except for *P. aeruginosa*. In particular, Fur from this organism shows a peculiarity: only one cysteine is present in the whole sequence. In contrast, Fur proteins from more than twenty other organisms contain two highly conserved motifs: Cys92-(2/5)X-Cys95 and CysY-XX-Cys130 in the C-termini (numbering according to the alignment in figure 5.2). The

cysteine residues in these motifs were, for some time, proposed to play a role in the metal binding properties of the protein. Mutational studies on EC Fur (Coy et al., 1994) have demonstrated the essential nature of Cys92 and Cys95 for the activity of the protein. Furthermore, EC Fur mutants, in which the C-terminal cysteines were mutated (Cys132Gly and the double mutant Cys132Gly/Cys137Gly) demonstrated that these residues are not involved in the zinc coordination (D'Autreaux et al., 2004). Mtb FurA contains two additional cysteines: Cys103 and Cys110. However, based on the homology model of Mtb FurA with PA Fur, these cysteines appear to be located on the fifth helix and to be solvent exposed. We therefore propose that Cys92 and Cys95 are likely to be involved in the formation of the structural zinc site. The two remaining zinc-binding amino acids could be aspartate, glutamate or histidine. Although these residues are also highly conserved in Fur proteins, it is not possible to unambiguously identify them based on the spectroscopic data alone.

5.3 The regulatory site

In addition to the zinc site, Mtb FurA also contains a second metal binding site. The nature of this site was first analyzed by ITC. Due to the high tendency of Fe^{2+} to oxidize, it was not possible to determine the affinity of Mtb FurA for ferrous iron. Instead, Zn^{2+} was used, since it can generally mimic the octahedral Fe^{+2} coordination (d^6). This electronic configuration confers important characteristic to this metal ion:

a) Since all orbitals are fully occupied, this element is not susceptible to any crystal field stabilization energy. This electronic energy term can favor certain arrangements of ligands over the others (Berg and Shi, 1996).

b) In terms of hard-soft acid-base theory, Zn^{2+} is regarded as a borderline acid and can interact with ligands like cysteines, nitrogens and oxygens (Berg and Shi, 1996).

c) Divalent zinc is not redox active, therefore forms like Zn^{+} or Zn^{3+} are not achievable under physiological conditions (Berg and Shi, 1996).

d) Zn^{2+} is relatively labile in kinetics and can easily participate in ligand exchange

reactions (Berg and Shi, 1996).

The binding constant obtained from the ITC titration is $1.27 \cdot 10^5$M (dissociation constant of ~8 μM). This value is comparable to those reported previously for the regulatory site of EC Fur (dissociation constants ranging from 10 and 80 μM).

The iron site was further characterized by X-ray absorption spectroscopy. As discussed above, two possible models can be considered to describe its coordination: 4O+2N and 3O+3N. It is important to note that only light atoms surround the cation and that the presence of sulphur atoms in the metal coordination sphere can be completely excluded. The chemical environment proposed here is frequently found in non-heme iron-proteins. Previous research on the Fur iron binding sites in *E. coli* (Jacquamet et al., 1998) and *Anabaena* (Hernandez et al., 2005) has pointed out a high conservation of particular residues. An EXAFS and crystallography study on PA Fur (Pohl et al., 2003) has reported a coordination of 4O+2N where three of the oxygen donors belong to two glutamate residues and one belongs to a molecule of solvent water. In addition, the high resolution structure of the DNA domain of EC Fur was recently published (Pecqueur et al., 2006), and the metal site present in this domain exhibited a coordination very similar to that observed in PA Fur. According to the sequence alignment, the residues involved in the coordination of Fe^{2+} of PA Fur and EC Fur are highly conserved in the Fur family (as indicated by the blue arrows in figure 5.2). As the chemical environment of Fe^{2+} in Mtb FurA, PA Fur and EC Fur is very similar, we are inclined to suggest that the amino acids involved in the metal coordination of Mtb FurA are identical to those in the other organisms - namely His32, Glu80, His89 and Glu100. This is sustained by the observation that mutation of His32 and His89 in EC Fur completely abolished protein activity (Adrait et al., 1999).

All this evidence is indeed consistent and in agreement with the models B_{Fe} and C_{Fe} reported in table 3.4, where nitrogens are coordinated at 1.96 Å and oxygens at 2.04/2.06Å.

5.4 DNA affinity

Figure 3.7 shows a faint binding of FurA to the canonical oligonucleotide Fur-box. Even though the interaction is not very strong, it is possible to observe a certain affinity of the protein for the DNA. This result represents preliminary information for the investigation of the physiological role of Mtb FurA. The protein could, in fact, be classified as a ferric uptake regulator, but further experiments have to be undertaken before claiming a biological function.

5.5 Summary and conclusion

The results presented in this work amount to a thorough characterization of the two metal-binding sites of Mtb FurA, including a detailed description of their chemical environments. The zinc site, to which the metal is already bound after purification, is presumably involved in the stabilization of the protein architecture and serves a purely structural role. The function of the zinc ion is most likely the maintenance of the tertiary structure. Previous studies have already reported that the cysteine-motif (CysXXCys) can participate in forming structural sites to hold helical sections, thus stabilizing the tertiary structure (Harding, 2004). The iron site, in contrast, is readily exchangeable and fully occupied after incubation with Fe^{2+}-solution. This metal is essential for the activation of the protein and the its DNA-binding function (Mills and Marletta, 2005). The amino acids coordinated to this site (histidine and glutamic acids) are the most frequently statistically counted in the coordination sphere of the Fe^{2+} (Harding, 2004). Same electronic octahedral configuration of the iron can be achieved by the next closest elements, such as Co, Cu, Mn, and Zn. When iron binds to Fur proteins, it is likely that a hinge motion of the C-terminal domain takes place in order to achieve the conformational changes that allow the DNA recognition motif to assume its active configuration. All this evidence finally leads us to postulate that the zinc site serves a structural function and the iron site a regulatory function. Notably, through this work, we have been able to shed some light on questions still open about the role and the involvement of some specific residues such as cysteines, considered crucial for the activity of the protein. We have also provided a meticulous description and a qualitative and quantitative characterization of the metal

sites in FurA from *Mycobacterium tuberculosis*. It is notable that the metal binding features of the Fur proteins, which have a major effect on iron homeostasis (e.g. PA Fur) are very similar to those with more limited functions (i.e. Mtb FurA), such as regulating specific genes involved in oxidative stress. Consequently, it appears that the specific DNA binding region of these regulators plays a major role in the determination of the specific genes that are ultimately regulated by the particular type of Fur.

Chapter 6

Mtb FurB DISCUSSIONS

6.1 Structure overview

FurB from *Mycobacterium tuberculosis* is a DNA binding protein with a high affinity for Zn^{2+}. The protein is expressed and purified with a very tightly bound zinc atom. Through the binding of additional zinc FurB reaches its active status. It is so able to bind the promoter region of the *znuABC* gene cluster encoding for proteins involved in zinc uptake.

The first crystal structure of one of the members of the Zur (zinc uptake regulator) family has been presented in this work: Mtb FurB solved to 2.7Å resolution. The overall fold of Mtb FurB with its N-terminal DNA-binding and C-terminal dimerization domain is similar to the PA Fur one - for long time the only available crystal structure of any member of the Fur family (Pohl et al., 2003). Domain conservation can be observed and only the presence of an additional helix of 10 residues in the N-terminus of PA Fur could be regarded as one of the main differences between (fig. 6.1).

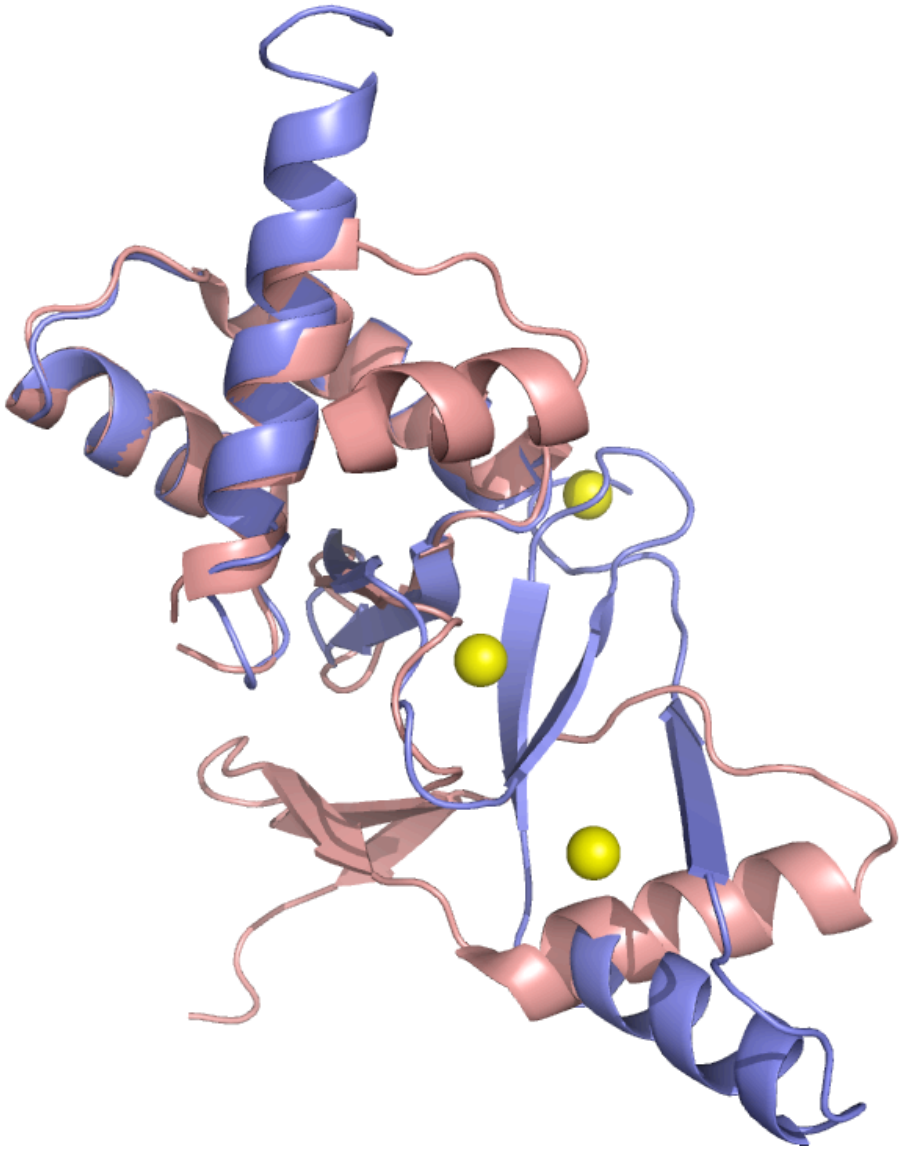

Figure 6.1: DNA-binding domains superposition
Ribbon representation in blue of Mtb FurB and in salmon of PA Fur.

The individual domains superimpose well with an rmsd of 1.6 Å for 67 equivalent Cα-atoms in the DNA-binding domain and 2.0 Å for 50 Cα-atoms dimerization domain, respectively (calculated using the DALI server (Holm and Sander, 1993)). The largest and more important difference between these two proteins is the orientation and therefore the angle between the two domains (fig. 6.2). Mtb FurB adopts a much wider conformation where the DNA-binding domains have moved in a hinge motion of approximately 77˚ with respect to the dimerization domain. The hinge region is located between residues 77-79.

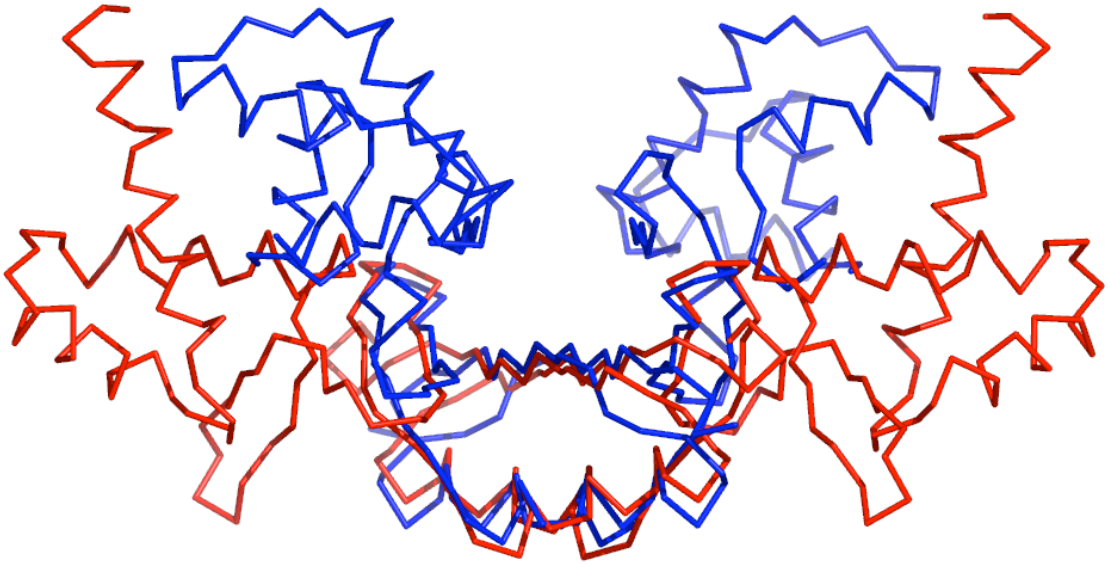

Figure 6.2: Dimerization domains superposition
Least-squares superposition of dimerization domains of Mtb FurB dimer (red) and PA Fur dimer (blue). Only the Cα-atoms of the dimerization domain were used to calculate the superposition matrix which was consecutively applied to the full dimer.

6.2 Metal sites

From the crystal structure three distinct Zn^{2+}-binding sites were identified in FurB.

Zinc 1 is surrounded by two histidines (His 81 and His 83), one cysteine (Cys 76) and one aspartic acid (Asp 62) (see fig. 4.20). The position of this site is presumably important for the activation of the protein as the metal ion is located at the hinge region between the two domains and may thus be responsible for the orientation of these two units with respect to each other. In fact two of the zinc ion ligands (Aps 62 and Cys 76) are located in the DNA-binding domain and two (His 81 and His 83) in the dimerization domain. The role of this metal site is presumably important for the function of the protein. The chemical environment corresponds exactly to that described for the EXAFS cobalt spectrum of Mtb FurB, where the ion presumably occupies the putative regulatory

site. This cation is often used in biology to replace and thus mimic zinc ion. This expedient allowed us to distinguish between the different metal sites in solution. It is important to note that it was not possible to introduce Fe^{2+}, which prefers an octahedral coordination sphere with only N and O atoms (Harding, 2004).

Zinc 2 is tetrahedrally coordinated by four cysteines arranged in a motif frequently reported in literature: CXXC (Auld, 2001) (see fig. 4.21). The metal ties the dimerization domain (via Cys86 and Cys 89) and the C-terminus (Cys126 and Cys 129) together. A similar metal environment was also described by EXAFS data collected at the zinc edge in Mtb FurB solution. This cation can only be removed from its locus by denaturing the protein therefore it is very likely that this high-affinity site serves a structural role. Tetragonal ZnS_4-clusters are by far the preferred motifs for structural zinc site (Auld, 2004).

In the crystal structure Mtb FurB exhibits a third Zn^{2+} site, surrounded by three histidines (His 80, His 82 and His 118) and one glutamate (Glu 101) (see fig. 4.22). This site could not be analyzed by X-ray absorption spectroscopy, due to the fact that it was not possible to introduce a third metal without observing precipitation. Although the exact biological function of this site remains open the following possibilities are at present proposed:

a) Metal site 3 could be a second structural site. There are a number of examples (Auld, 2001) in literature where a structural zinc site is composed of one aspartate/glutamate and three histidines. This motif is highly conserved in the metalloproteinases family where the metal ion is coordinated to four protein ligands. Remarkably, like in this case, no-bound water and a relatively short sequence of the amino acids involved in the cation coordination is observed (Auld, 2001). Location of zinc 3 in the crystal structure of Mtb FurB, buried in the dimerization domain and coordinating to residues placed in the helix S3 and the strands S3 and S5, seems to play a stabilization role for the structure. Its presence, in fact, brings together all the secondary structure elements of the dimerization domain and confers a certain rigidity and stability to this unit. However, in contrast to the first structural site this site is not necessary for structural integrity as it can easily be removed by EDTA treatment.

b) This site could be an artefact of crystallization. There are a number of cases where additional metal sites were found in crystal structures - sometimes they are responsible for crystal contacts as in the case of the crystal structure of PA Fur (Pohl et al., 2003). However this site well superimposes with the putative regulatory Fe-binding site of the PA-Fur crystal structure. After least-squares superposition of the dimerization domain (Cα-atoms only) the position of the two metals differ by only 1Å. In addition, three of the four residues binding the metal are conserved but whereas in PA Fur this site constitutes an octahedral ligation sphere, well suited for Fe^{2+}, in our structure the site forms a tetrahedral environment preferred by Zn^{2+}.

It is clear that proteins belonging to the Fur family exhibits a conserved overall structure with a DNA-binding domain and a dimerization but the metal ion selectivity of Fur-like repressor varies and is presumably determined by the spatial arrangement of the ligands involved in the metal coordination sphere (Moore and Helmann, 2005).

Obviously, further experiments are needed to elucidate the exact role of this Zn^{2+} binding site.

6.3 Remarkable points

6.3.1 EXAFS similarity

The EXAFS results reported here are interesting if compared to the X-ray absorption spectroscopy analysis of EC Zur metal sites (Outten et al., 2001). As with the protein studied here, the authors were able to add one cobalt ion to Zur already complexed with a very tightly bound zinc ion. Both cations, Zn^{2+} and Co^{2+}, exhibited significantly similar environments to those identified by EXAFS for the Mtb FurB (Outten et al., 2001). It is important to note that, although the best-fitting model proposed for EC Zur was a 3S+1N/O model, the second best model, whose goodness of fit was not significantly worse, was a 4S model, as proposed here. Moreover, looking at the sequence alignment of Mtb FurB with the Zur proteins (including that from *E. coli*), it is possible to observe that the four cysteines coordinating the Mtb FurB zinc structural site in the crystal structure are highly conserved among the species (fig. 6.3).

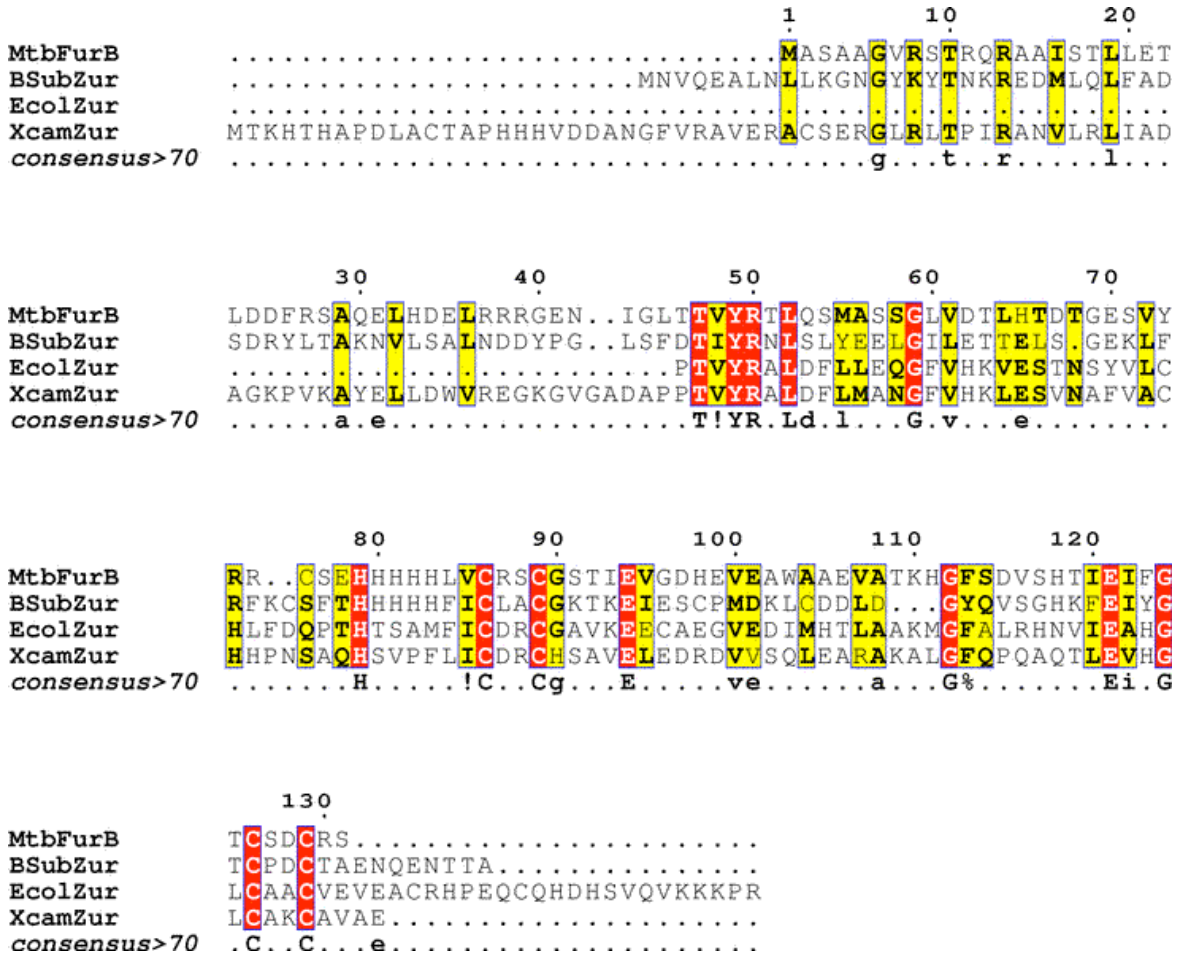

Figure 6.3: Zur proteins alignment
The numbering refers to Mtb FurB sequence. The yellow boxes represent the conserved regions and in red are highlighted the identical residues. Consensus > 70 has been chosen to underline the conservation of the residues in at least three/four cases. The alignment has been calculated with ClustalW (http://www.ebi.ac.uk/clustalw/index.html) (Pearson, 1990; Pearson and Lipman, 1988) and the graphically displayed with ESPript (http://espript.ibcp.fr/ESPript/cgi-bin/ESPript.cgi) (Gouet et al., 1999). Mtb: *Mycobacterium tuberculosis*, Bsub: *Bacillus subtilis*, Ecol: *Escherichia coli*, Xcam: *Xanthomonas campestris*

Considering that the difference in the parameters between the two best EXAFS models (in both cases EC Zur and Mtb FurB) is not enough to completely rule out one or the other model, along with the crystallographic results and the alignment information, I am inclined to propose a more generally valid structural site coordinated by four cysteines.

The chemical environment proposed for the regulatory site occupied by Co^{2+} in EC Zur is instead identical to metal site 1 of Mtb FurB crystal structure. These similarities further support the notion that Mtb FurB is in fact the zinc uptake regulator.

6.3.2 DNA-binding

The promoter region of Rv2059 is similar to the promoter of ZnuA and to the Zur-boxes from other organisms, all sequences are pseudo-palindromic and rich in AT base pairs. In the tuberculosis genome this is not common as more than 66% of its base pairs is constituted by CG (Cole et al., 1998). The presence of such a particular sequence of nucleotides, regularly present in the promoter region of a specific class of genes, can not be regarded just as a mere coincidence and most likely it is an important feature for the recognition and then regulation of genes involved in zinc homeostasis.

Although the crystal structure of Mtb FurB presented here shows a detailed picture of the overall architecture as well as the metal binding sites, the model of DNA recognition is not immediately obvious. The dimer forms a widely open structure where the two putative DNA-recognition helices (H3 and H3', respectively) are oriented almost parallel to each other on opposite sides. Hence the binding of each of the DNA-recognition elements in the major groove of one half site of the (pseudo-)palindromic DNA-target sequence as observed in the case of a number of prokaryotic DNA-binding proteins (Huffman and Brennan, 2002) including DtxR (Pohl et al., 1999; White et al., 1998) cannot be accomplished without major conformational change upon the hinge region of the protein. The electrostatic potential representation depicted in figure 6.4 shows that helices H1 and H3 with a number of surface arginines possess a mainly positively charged surface. This may be an indication that H1 and its symmetry-mate H1' are also involved in DNA-backbone binding.

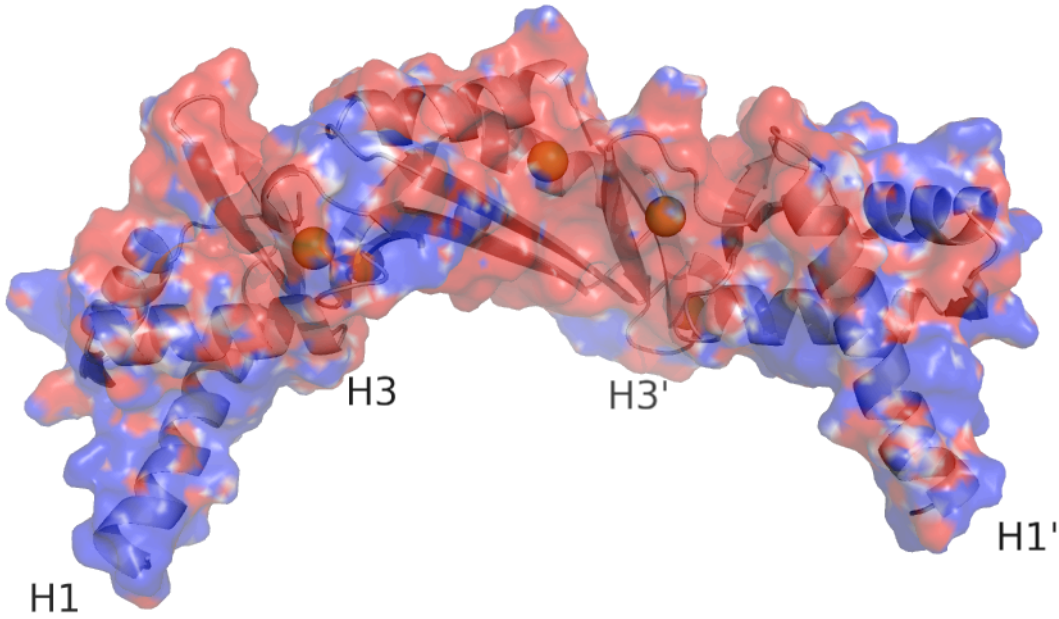

Figure 6.4: Electrostatic representation of Mtb FurB dimer
Surface representation with the calculated electrostatic potential (DeLano Scientific LLC, 1998-2003) varying from positive (blue) to negative (red).

Clearly, a crystal structure of a Fur/Zur DNA complex is needed to unravel the exact mode of DNA-binding.

6.3.3 Alignment considerations

Sequence alignment of MtB FurB with Zur from Gram-positive bacteria (i.e *Bacillus subtilis*) shows 25% of identity (fig. 6.5).

Unfortunately it is still very difficult to distinguish between Fur and Zur proteins only from sequence alignments. A clear classification, supported by experimental results, can be stipulated only based on a few organisms. In fact today Fur and Zur proteins with their own respective functions and target specificity can be assigned only in *Escherichia coli*, *Bacillus subtilis*, *Xantomonas campestris* and with this work also in *Mycobacterium tuberculosis*. All other targets, classified as Zur, are assigned only by sequence homology and are not supported by experimental proves. However when a comparison between Zurs from different species is required it is important to remember that Zur proteins form

at least three unrelated groups: enterobacterial Zur, Zur proteins of low-GC Gram-positive bacteria and high CG Gram-positive bacteria (Hantke, 2005).

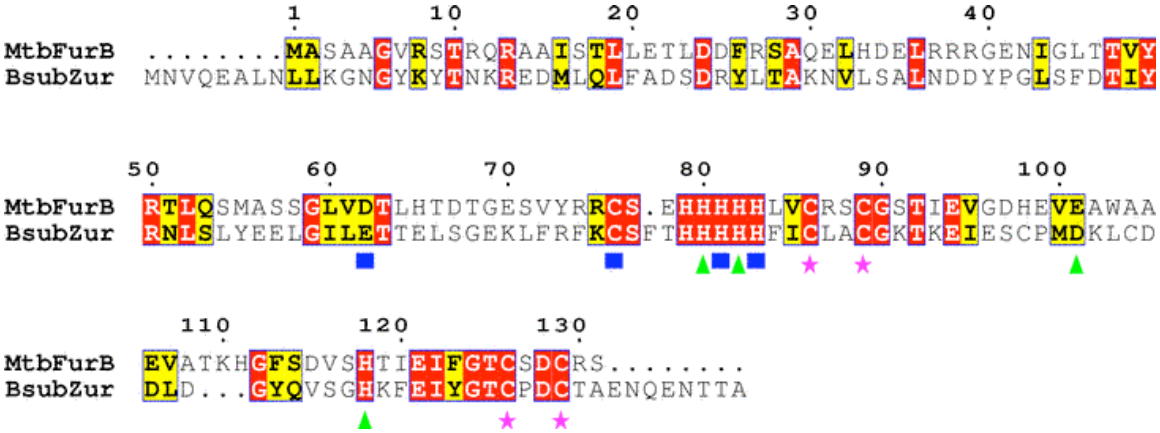

Figure 6.5: Sequence alignment of Mtb FurB and Zur from *Bacillus subtilis*
Bacillus subtilis is a Gram-positive bacteria like *Mycobacteirum tuberculosis*. The numbering refers to the Mtb FurB sequence. The blue square indicated site 1 in Mtb FurB, the pink stars the site 2 and the green arrows the site 3. These sites seem to be conserved in *Bacillus subtilis* too. The yellow boxes represent the conserved regions, in red are highlighted the identical residues. The alignment has been calculated with ClustalW (http://www.ebi.ac.uk/clustalw/index.html) (Pearson, 1990; Pearson and Lipman, 1988) and the graphically displayed with ESPript (http://espript.ibcp.fr/ESPript/cgi-bin/ESPript.cgi) (Gouet et al., 1999).

6.4 Conclusion

Mtb FurB shows many structural and functional similarities with Zur. The metal sites are likely to be conserved through the Zur species. The protein is very sensitive to the presence of zinc in the media and an overload of this metal causes precipitation. Same phenomena has happened to other Zur proteins (Outten et al., 2001). FurB binds the promoter regions of a triplet of genes involved in zinc transport, these genes are regulated in other already characterized organisms by Zur (Gaballa and Helmann, 1998; Patzer and Hantke, 2000). Furthermore this protein does not show any iron binding affinity and the ferric uptake regulator of *Mycobacterium tuberculosis* has been already identified as FurA (Pym et al., 2001). In summary, a combination of X-ray absorption spectroscopy and crystallographic data have allowed the investigation and distinction between the

structural and the regulatory site of zinc sites in Mtb FurB. All experimental results presented here suggest a clear biological function for FurB as a Zinc uptake regulator.

Chapter 7

CONCLUSIONS

When this project was started, very little was known about the proteins FurA and FurB from *Mycobacterium tuberculosis.* At that time, many uncertainties existed about the nature of ferric and zinc uptake regulators even though the field was already very prolific. While many of the same questions remain today, knowledge of the importance of these proteins has grown with exponential velocity/rate in recent years and they have become the focus of new interest for many laboratories. In the last few years, scientists have discovered the vital importance of metal homeostasis for humans and bacteria. Today, unravelling the mechanisms that control these highly sophisticated regulation systems has become a priority in the battle against pathogens.

This project started with knowledge of two Fur genes in the tuberculosis genome, assigned, like for many other bacteria, only by sequence homology to the Fur protein previously known from *E. coli*. A number of techniques have been utilized and a set of protocols established in order to characterise both systems and to shed some light on the real nature and function of these proteins. The work was all carried out starting from the very beginning and proceeded along three and a half years.

A detailed and careful description of all metal sites in both targets is provided in this work. For years, many interpretations and discussions have been reported in literature about the possible evolution, in the same family, of the metal sites, and often a comparison between Fur and Zur has been required. Finally it is possible to propose models and to permit a clarification based only on experimental evidence. Specifically,

this work has allowed estimation of the number of metal sites, their binding constants, coordination environment and nature of the metal ions. By solving the crystal structure of FurB it was possible to present for the first time a 3D picture of a zinc uptake regulator. No member of this family has ever been crystallized before. On the basis of these results it is possible to assign functions to both targets, recognising FurA as the ferric uptake regulator and FurB as the zinc uptake regulator. This can be regarded as an important step especially if it is considered that, despite the presence of several zinc-dependent genes in the tuberculosis genome, until today no zinc regulator had been classified in that organism. The respective DNA partners were also identified and studies on the nature of these complexes are ongoing.

At present several steps forward have been made from the original basic knowledge of *furA* and *furB* gene numbers. The project has assumed a well-defined shape and more importantly, the ground is ready for a more careful and advanced study of these targets. Both systems can now easily be expressed and stabilized, although a lot of work still remains to be done. Crystallization of the FurB-DNA complex, FurA and FurA-DNA complex has already started as well as a biological characterization of the DNA binding constants. The system is there, just waiting to be disclosed.

All results reported here could be achieved only by involving a combination of different techniques. This proves that there is no certain nor unique method universally valid for studying and understanding nature but that this is Science, where each system represents a new “universe” which, by being “observed” from different sides, reveals its true nature to the investigator.

REFERENCES

Adrait, A., Jacquamet, L., Le Pape, L., Gonzalez de Peredo, A., Aberdam, D., Hazemann, J.L., Latour, J.M. and Michaud-Soret, I. (1999) Spectroscopic and saturation magnetization properties of the manganese- and cobalt-substituted Fur (ferric uptake regulation) protein from Escherichia coli. *Biochemistry*, **38**, 6248-6260.

Agranoff, D. and Krishna, S. (2004) Metal ion transport and regulation in Mycobacterium tuberculosis. *Frontiers in Bioscience*, **9**, 2996-3006.

Ahn, B.E., Cha, J., Lee, E.J., Han, A.R., Thompson, C.J. and Roe, J.H. (2006) Nur, a nickel-responsive regulator of the Fur family, regulates superoxide dismutases and nickel transport in Streptomyces coelicolor. *Mol Microbiol*, **59**, 1848-1858.

Aksenov, V.L., Kuzmin, A.Y., Purans, J. and Tyutyunnikov, S.I. (2001) EXAFS Spectroscopy at Synchrotron-Radiation Beams. *Physics of particles and nuclei*, **32**, 1-33.

Althaus, E.W., Outten, C.E., Olson, K.E., Cao, H. and O'Halloran, T.V. (1999) The ferric uptake regulation (Fur) repressor is a zinc metalloprotein. *Biochemistry*, **38**, 6559-6569.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) Basic local alignment search tool. *J Mol Biol*, **215**, 403-410.

Auld, D.S. (2001) Zinc coordination sphere in biochemical zinc sites. *Biometals*, **14**, 271-313.

Auld, D.S. (2004) Structural zinc sites. *Handbook of Metalloproteinases*, **3**, 403-415.

Berg, J.M. and Shi, Y. (1996) The Galvanization of Biology: A Growing Appreciation for the Roles of Zinc. *Science*, **271**, 1081-1085.

Binsted, N., Gurman, S.J., Campbell, J.W. and Stephenson, P. (1991) *SERC Daresbury Laboratory Program EXCURVE. SERC Daresbury Laboratorty, Warrington. U.K.*

Boshoff, H.I.M. and Barry, C.E., 3rd. (2005) Tuberculosis-metabolismand respiration in the absence of growth. *Nature Reviews*, **3**, 70-80.

Bossmann, S.H., Oliveros, E., Kantor, M., Niebler, S., Bonfill, A., Shahin, N., Worner, M. and Braun, A.M. (2004) New insights into the mechanisms of the thermal Fenton reactions occurring using different iron(II)-complexes. *Water Sci Technol*, **49**, 75-80.

Brennen GR and BW, M. (1989) The Helix-Turn-Helix DNA Binding Motif. *The Journal of Biological Chemistry*, **264**, 1093-1906.

Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., Read, R.J., Rice, L.M., Simonson, T. and Warren, G.L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr*, **54**, 905-921.

Canneva, F., Branzoni, M., Riccardi, G., Provvedi, R. and Milano, A. (2005) Rv2358 and FurB: two transcriptional regulators from Mycobacterium tuberculosis which respond to zinc. *J Bacteriol*, **187**, 5837-5840.

Cole, S.T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S.V., Eiglmeier, K., Gas, S., Barry, C.E., 3rd, Tekaia, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J., Moule, S., Murphy, L., Oliver, K., Osborne, J., Quail, M.A., Rajandream, M.A., Rogers, J., Rutter, S., Seeger, K., Skelton, J., Squares, R., Squares, S., Sulston, J.E., Taylor, K., Whitehead, S. and Barrell, B.G. (1998) Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. *Nature*, **393**, 537-544.

Coleman, J.E. (1992) Zinc proteins: enzymes, storage proteins, transcription factors, and replication proteins. *Annu Rev Biochem*, **61**, 897-946.

Coleman, J.E. (1998) Zinc enzymes. *Curr Opin Chem Biol*, **2**, 222-234.

Coy, M., Doyle, C., Besser, J. and Neilands, J.B. (1994) Site-directed mutagenesis of the ferric uptake regulation gene of Escherichia coli. *Biometals*, **7**, 292-298.

D'Autreaux, B., Horner, O., Oddou, J.L., Jeandey, C., Gambarelli, S., Berthomieu, C., Latour, J.M. and Michaud-Soret, I. (2004) Spectroscopic description of the two nitrosyl-iron complexes responsible for fur inhibition by nitric oxide. *J Am Chem Soc*, **126**, 6005-6016.

De Voss, J.J., Rutter, K., Schroeder, B.G. and Barry, C.E., 3rd. (1999) Iron acquisition and metabolism by mycobacteria. *J Bacteriol*, **181**, 4443-4451.

DeLano Scientific LLC, S.C.C., USA. (1998-2003) Pymol.

Dye, C., Williams, B.G., Espinal, M.A. and Raviglione, M.C. (2002) Erasing the world's slow stain: strategies to beat multidrug-resistant tuberculosis. *Science*, **295**, 2042-2046.

Escolar, L., Perez-Martin, J. and de Lorenzo, V. (1998) Binding of the fur (ferric uptake regulator) repressor of Escherichia coli to arrays of the GATAAT sequence. *J Mol Biol*, **283**, 537-547.

Gaballa, A. and Helmann, J.D. (1998) Identification of a zinc-specific metalloregulatory protein, Zur, controlling zinc transport operons in Bacillus subtilis. *J Bacteriol*, **180**, 5815-5821.

Gangaidzo, I.T., Moyo, V.M., Mvundura, E., Aggrey, G., Murphree, N.L., Khumalo, H., Saungweme, T., Kasvosve, I., Gomo, Z.A., Rouault, T., Boelaert, J.R. and Gordeuk, V.R. (2001) Association of pulmonary tuberculosis with increased dietary iron. *J Infect Dis*, **184**, 936-939.

Garman, E.F. and Grime, G.W. (2005) Elemental analysis of proteins by microPIXE. *Prog Biophys Mol Biol*, **89**, 173-205.

Gobin, G. and Horwitz, M.A. (1996) Exochelins of Mycobacterium tuberculosis Remove Iron from Human Iron-binding proteins and Donate Iron to Mycobactin in the M.tuberculosis Call Wall. *J. Exp. Med.*, **183**, 1527-1532.

Gouet, P., Courcelle, E., Stuart, D.I. and Metoz, F. (1999) ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics*, **15**, 305-308.

Grime, G.W. (1996) The "q factor" method: Quantitative microPIXE analysis using RBS normalisation. *Nucl. Instrum. Methods Phys. Res., Sect. B*, **109**, 170-174.

Grime, G.W., Dawson, M., Marsh, M., McArthur, I.C. and Watt, F. (1991) The Oxford submicron nuclear microscopy facility. *Nucl. Instrum. Methods Phys. Res., Sect. B*, **54**, 52-63.

Guex, N. and Peitsch, M.C. (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis*, **18**, 2714-2723.

Hantke, K. (2001) Bacterial zinc transporters and regulators. *Biometals*, **14**, 239-249.

Hantke, K. (2005) Bacterial zinc uptake and regulators. *Curr Opin Microbiol*, **8**, 196-202.

Harding, M.M. (2004) The architecture of metal coordination groups in proteins. *Acta Crystallogr D Biol Crystallogr*, **60**, 849-859.

Hernandez, J.A., Meier, J., Barrera, F.N., Ruiz de Los Panos, O., Hurtado-Gomez, E., Bes, M.T., Fillat, M.F., Peleato, M.L., Cavasotto, C.N. and Neira, J.L. (2005) The conformational stability and thermodynamics of Fur A (ferric uptake regulator) from Anabaena sp. PCC 7119. *Biophys J.*

Hernandez, J.A., Peleato, M.L., Fillat, M.F. and Bes, M.T. (2004) Heme binds to and inhibits the DNA-binding activity of the global regulator FurA from Anabaena sp. PCC 7120. *FEBS Lett*, **577**, 35-41.

Hill, P.J., Cockayne, A., Landers, P., Morrissey, J.A., Sims, C.M. and Williams, P. (1998) SirR, a novel iron-dependent repressor in Staphylococcus epidermidis. *Infect Immun*, **66**, 4123-4129.

Holm, L. and Sander, C. (1993) Protein structure comparison by alignment of distance matrices. *J Mol Biol*, **233**, 123-138.

Huffman, J.L. and Brennan, R.G. (2002) Prokaryotic transcription regulators: more than just the helix-turn-helix motif. *Curr Opin Struct Biol*, **12**, 98-106.

Jacquamet, L., Aberdam, D., Adrait, A., Hazemann, J.L., Latour, J.M. and Michaud-Soret, I. (1998) X-ray absorption spectroscopy of a new zinc site in the fur protein from Escherichia coli. *Biochemistry*, **37**, 2564-2571.

Johansson, S.A.E., Campbell, J.L. and Malmqvist, K.G. (1995) *Particle induced X-ray Emission Spectroscopy*. *Wiley and Sons, New York.*

Jones, T.A., Zou, J.Y., Cowan, S.W. and Kjeldgaard. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A*, **47 (Pt 2)**, 110-119.

Kabsch, W. (1993) Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Cryst.*, **26**, 795-800.

Kaufmann, S.H. (2000) Is the development of a new tuberculosis vaccine possible? *Nat Med*, **6**, 955-960.

M. Korbas, D.F.M.W.M.-K. (2006) A program suite for automated biological X-ray absorption spectroscopy data reduction. *Submitted*.

Masse, E. and Gottesman, S. (2002) A small RNA regulates the expression of genes involved in iron metabolism in Escherichia coli. *Proc Natl Acad Sci U S A*, **99**, 4620-4625.

Mills, S.A. and Marletta, M.A. (2005) Metal Binding Characteristics and Role of Iron Oxidation in the Ferric Uptake Regulator from Escherichia coli. *Biochemistry*, **44**, 13553-13559.

Moore, C.M. and Helmann, J.D. (2005) Metal ion homeostasis in Bacillus subtilis. *Curr Opin Microbiol*, **8**, 188-195.

Outten, C.E. and O'Halloran, T.V. (2001) Femtomolar sensitivity of metalloregulatory proteins controlling zinc homeostasis. *Science*, **292**, 2488-2492.

Outten, C.E., Tobin, D.A., Penner-Hahn, J.E. and O'Halloran, T.V. (2001) Characterization of the metal receptor sites in Escherichia coli Zur, an ultrasensitive zinc(II) metalloregulatory protein. *Biochemistry*, **40**, 10417-10423.

Patzer, S.I. and Hantke, K. (2000) The zinc-responsive regulator Zur and its control of the znu gene cluster encoding the ZnuABC zinc uptake system in Escherichia coli. *J Biol Chem*, **275**, 24321-24332.

Pearson, W.R. (1990) Rapid and sensitive sequence comparison with FASTP and FASTA. *Methods Enzymol*, **183**, 63-98.

Pearson, W.R. and Lipman, D.J. (1988) Improved tools for biological sequence comparison. *Proc Natl Acad Sci U S A*, **85**, 2444-2448.

Pecqueur, L., D'Autreaux, B., Dupuy, J., Nicolet, Y., Jacquamet, L., Brutscher, B., Michaud-Soret, I. and Bersch, B. (2006) Structural changes of E. coli Fur during metal dependent dimerization and activation explored by NMR and X-ray crystallography. *J Biol Chem*.

Peitsch, M.C. (1995) Protein modeling by E-mail. *Bio/Technology*, **13**, 658-660.

Pennella, M.A. and Giedroc, D.P. (2005) Structural Determinants of Metal Selectivity in Prokaryotic Metal-responsive Transcriptional Regulators. *Biometals*, **18**, 413-428.

Pettifer, R.F. and Hermes, C. (1985) Absolute energy calibration of X-ray radiation form synchrotron sources. *J. Appl. Cryst.*, **18**, 404-412.

Pohl, E., Haller, J.C., Mijovilovich, A., Meyer-Klaucke, W., Garman, E. and Vasil, M.L. (2003) Architecture of a protein central to iron homeostasis: crystal structure and spectroscopic analysis of the ferric uptake regulator. *Mol Microbiol*, **47**, 903-915.

Pohl, E., Holmes, R.K. and Hol, W.G. (1999) Crystal structure of the iron-dependent regulator (IdeR) from Mycobacterium tuberculosis shows both metal binding sites fully occupied. *J Mol Biol*, **285**, 1145-1156.

Pym, A.S., Domenech, P., Honore, N., Song, J., Deretic, V. and Cole, S.T. (2001) Regulation of catalase-peroxidase (KatG) expression, isoniazid sensitivity and virulence by furA of Mycobacterium tuberculosis. *Mol Microbiol*, **40**, 879-889.

Rodriguez, G.M. (2006) Control of iron metabolism in Mycobacterium tuberculosis. *Trends Microbiol*, **14**, 320-327.

Rodriguez, G.M. and Smith, I. (2003) Mechanisms of iron regulation in mycobacteria: role in physiology and virulence. *Mol Microbiol*, **47**, 1485-1494

Rodriguez, G.M. and Smith, I. (2006) Identification of an ABC transporter required for iron acquisition and virulence in Mycobacterium tuberculosis. *J Bacteriol*, **188**, 424-430.

Roe, A.L., Schneider, D.J., Mayer, R.J., Pyrz, J.W., Widom, J. and Que, L.J. (1984) X-Ray Absorption Spectroscopy of Iron-Tyrosinate Proteins. *J. Am. Chem. Soc*, **106**, 1676-1681.

Schmitt, M.P., Predich, M., Doukhan, L., Smith, I. and Holmes, R.K. (1995) Characterization of an iron-dependent regulatory protein (IdeR) of Mycobacterium tuberculosis as a functional homolog of the diphtheria toxin repressor (DtxR) from Corynebacterium diphtheriae. *Infect Immun*, **63**, 4284-4289.

Schwede, T., Kopp, J., Guex, N. and Peitsch, M.C. (2003) SWISS-MODEL: An automated protein homology-modeling server. *Nucleic Acids Res*, **31**, 3381-3385.

Scott, R.A. (2000) X-ray absorption spectorscopy. *In: Que , Jr. (ed) Physical methods in bioinorganic chemistry: Spectroscopy and Magnetism. University Books, Sausalito*, 465-503.

Sheldrick, G. (1998) SHELX: Application to macromolecules. *Direct Methods for Solving Macromolecular Structures*, 401-410.

van Vliet, A.H., Stoof, J., Vlasblom, R., Wainwright, S.A., Hughes, N.J., Kelly, D.J., Bereswill, S., Bijlsma, J.J., Hoogenboezem, T., Vandenbroucke-Grauls, C.M., Kist, M., Kuipers, E.J. and Kusters, J.G. (2002) The role of the Ferric Uptake Regulator (Fur) in regulation of Helicobacter pylori iron uptake. *Helicobacter*, **7**, 237-244.

Vasil, M.L. and Ochsner, U.A. (1999) The response of Pseudomonas aeruginosa to iron: genetics, biochemistry and virulence. *Mol Microbiol*, **34**, 399-413.

White, A., Ding, X., vanderSpek, J.C., Murphy, J.R. and Ringe, D. (1998) Structure of the metal-ion-activated diphtheria toxin repressor/tox operator complex. *Nature*, **394**, 502-506.

Wilderman, P.J., Sowa, N.A., FitzGerald, D.J., FitzGerald, P.C., Gottesman, S., Ochsner, U.A. and Vasil, M.L. (2004) Identification of tandem duplicate regulatory small RNAs in *Pseudomonas aerugionosa* involved in iron homeostasis. *PNAS*, **101**, 9792-9797.

Wong, D.K., Lee, B.Y., Horwitz, M.A. and Gibson, B.W. (1999) Identification of fur, aconitase, and other proteins expressed by Mycobacterium tuberculosis under conditions of low and high concentrations of iron by combined two-dimensional gel electrophoresis and mass spectrometry. *Infect Immun*, **67**, 327-336.

ACKNOWLEDGEMENTS

After more than three years it is my pleasure and my duty to thank all those whose help, support, advice, encouragement, friendship and patience has made this thesis possible.

Firstly I would like to thank my supervisor, Ehmke Pohl, for believing in me from the first day we met, for supervising me constantly with advice and suggestions, and for providing always encouragement and support. All my gratitude goes to him, as I would have never reached this point without his patience and understanding. I would like to thank also my supervisor Wolfram Meyer-Klaucke for allowing me to join his group and enabling me to work in a good and very stimulating environment, for participating and keeping his interest in this project. I would like to include my TAC members, Elena Conti Irmgard Sinning and Manfred S. Weiss for all their useful suggestions and interest in my project, for always coming to my annual reviews and for following me with particular attention. I also would like to thank my collaborator, Elspeth Garman, who actively participated to this project and had some good words for me.

Special thanks go to those who took care of all the administrative work during these years: Margret Fischer, Silvia Bertini, Rika Pfaff, Ivanka Arujo, Nicola Porter from EMBL-Hamburg Tiziana Novarini and Milanka Stojkovic from EMBL-Heidelberg. They were constantly ready to sort out every single document, ticket or translation. They took care of all the bureaucratic work at any time and without delay, no matter how important it was.

My time at EMBL has been very busy and enjoyable due to all the friends I met during these years. The time has come to thank all of them: Simone Weyand, my first friend at EMBL and the person with whom I started this path from the very beginning; Robert Janowski, with whom I had a very good time and many personal chats, Matthew Groves and Rositsa Jordanova, who were there every time I needed them; Brenda Kostelecky,

with whom I shared most of my social life and interests during these years - hopefully our ways will still cross in the future; Esben Lorentzen, who beside the distance was always able to be a good friend - the good times we had together and the great scientific discussions have been a constant stimulation to do science and to enjoy his company; Daniel Fulla Marsa, who has helped me any time I needed.

I also would like to thank Stefanie, Ehmke, Katharina and Annabelle for opening the door of their house and their life to me. They always made me feel welcome. Their friendship has been and is very important to me; with them all I spent beautiful moments of real life.

I can not forget of Susanne and René Hartmann, who “sponsored” me during my PhD and who tried to spend as much time as possible with me while I lived with them.

To my parents, Alfonso and Saskia, goes all my love. They constantly encouraged me to realise my dreams and they always believed I would have made it to the end. To my sister Sabina all my gratitude and love for all the time she spent travelling to come to Hamburg and keep up with her sister’s life; we had beautiful moments together. She also stayed with me in the most difficult time of my life and she never let me down. Thanks mum, dad and Sabina.

And last but not least, a special thanks goes to Gareth Stockwell. A PhD time is a very “special time” with plenty of emotions and both good and bad moments. Gareth understood that and stayed with me, he tried to be there for me especially in the darkest periods, he never gave up on me and helped me to get to the end. Thanks Gareth!