Thesis

Bassam Kareem Kudhair
Structural and Biochemical studies of *Mycobacterium tuberculosis* WhiB1
PhD
September 2017

Structural and Biochemical studies of *Mycobacterium tuberculosis* WhiB1

Bassam Kareem Kudhair

A thesis submitted to the University of Sheffield in part fulfillment for the degree of Doctor of Philosophy

Department of Molecular Biology and Biotechnology, The University of Sheffield

September 2017

Summary

Mycobacterium tuberculosis is the causative agent of human tuberculosis (TB), which claimed 1.8 million lives in 2015 (WHO, 2016). A key component of TB pathogenesis is the ability of *M. tuberculosis* to enter a non-replicating persistent state following colonization of the human lung. Emergence from the persistent state upon immunosuppression, sometimes decades after the initial infection, results in an active infection (reactivation TB) that can be fatal if untreated. The fundamental role of Wbl-proteins in developmental processes in actinomycetes suggested that they could play a role in entry into and emergence from the non-replicative persistent state that is characteristic of *M. tuberculosis* infections. *Mycobacterium tuberculosis* possesses seven genes encoding Wbl proteins and several of these have been implicated in features of tuberculosis pathogenesis. The *M. tuberculosis whiB1* gene is essential and encodes a DNA-binding protein with an NO-sensitive [4Fe-4S] cluster. Nitric oxide is an important component of the host response to *M. tuberculosis* infection; high concentrations of NO generated by activated macrophages can kill the bacilli but low NO levels promote transition to the dormant non-replicating state. Here, the WhiB1 [4Fe-4S] cluster is shown to be essential for forming a complex with SigA and that reaction with NO disassembles the complex, indicating a role in reprogramming gene expression. Until now the structure of a Wbl protein had not been solved. An NMR structural model of WhiB1 is presented, revealing a core of three α-helices held together by the [4Fe-4S] cluster. The structure suggests that loss of the iron-sulfur cluster (by nitrosylation) frees a fourth *C*-terminal helix permitting positively-charged residues therein to engage in DNA-binding. The sensitivity of WhiB1 to nitric oxide, the ability to interact with SigA, and the structural characterization of the protein suggest that the protein has a role in regulating the pathogenicity of *M. tuberculosis*.

Acknowledgements

Top of the acknowledgements must be for my supervisor Professor Jeffrey Green, I would like to express my deepest gratitude and thankfulness for his exceptional levels of support, advice and guidance and for always managing to give me a time and listen to my questions. I have been so fortunate to have a supervisor who cared very much about my research over the last four years. Huge thanks also go to Professor Mike Williamson and Mrs Andrea Hounslow for their advice, support and engagement in my project, which due to their collaboration and their help the structure of WhiB1 protein is now available.

Many thanks for all people in F10, past and present, all of whom have been friendly and lovely, this lab was a great place to work in.

I would like also to thank our collaborators at the University of East Anglia, especially Dr Jason Crack for his help in ESI-MS. I have also to thank all people in X-ray protein crystallography group especially my supervisor Professor Peter Artymiuk, who died at the first year of my PhD course, for his assistance and for offering me a place with his team, and I am so sorry that he has not lived to see me reach this point; and also my second supervisor Dr Patrick Baker, my advisors Professor David Rice and Dr John Rafferty for their advice and discussions of my work, and Dr Svetlana Sedelnikova for giving me a huge information about protein purification.

Thank also to my family, my wife, Nooralhuda which was extremely supportive and made countless sacrifices to help me during all the period of the PhD course to get this point, and my parents for their encouragement, support, love and help.

Contents

List of figures

Chapter One

Chapter three

Chapter four

Chapter five

Chapter six

Appendix

List of tables

List of Abbreviations

BACTH	Bacterial Adenylate Cyclase Two-Hybrid System
cAMP	Cyclic adenosine monophosphate
CRP^{Mt}	*M. tuberculosis* cAMP receptor protein
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
TEVCS	Tobacco Etch Virus (TEV) protease cleavage site
THCS	Thrombin protease cleavage site
IAA	Iodoacetamide
IPTG	Isopropyl β-D-1-thiogalactopyranoside
Kb	kilobase
kDa	kilodalton
Mtb	*Mycobacterium tuberculosis*
NO	Nitric oxide
NOE	Nuclear overhauser effect
NOS2	Nitric oxide synthase
OD	Optical density
RF	Radio-frequency field
RNAP	RNA polymerase
RNI	Reactive nitrogen intermediates
ROI	Reactive oxygen intermediates
Sig	Sigma factor
TCEP	Tris(2-carboxyethyl)phosphine
Tris	Tris(hydroxymethyl)aminomethane
WbI	WhiB-like
$WhiB1^{Ec}$	WhiB1 overproduced in *E. coli* host
$WhiB1^{Ms}$	WhiB1 overproduced in *M. smegmatis* host
Xgal	5-Bromo-4-chloro-3-indolyl β-D-galactopyranoside

Chapter one

Introduction

1.0 Introduction

1.1 General characteristic of *M. tuberculosis*

Mycobacterium tuberculosis (*Mtb*) is the etiological agent of tuberculosis in humans. It is a species in the family of Mycobacteriaceae, it is rod-shaped with weak Gram-positive characteristics. It does not possess flagella or capsules, and cannot form spores. *Mycobacterium tuberculosis* grows slowly with a generation time about of 24 hours in both synthetic laboratory media and in infected animal cells. This pathogen is an aerobic bacterium that usually establishes an infection of the respiratory system. In 1882 Robert Koch isolated and cultured *Mtb* from crushed lung tubercles. The cell wall of *Mtb* is unique with up to 60% of unusual lipids such as glycolipids and mycolic acids (Ducati *et al.,* 2006 and references therein). Based on estimations of World Health Organization, one third of the earth's population is infected with *Mtb* with a probability of developing an active infection of about 10% during a lifetime. There are more than 9 million new cases each year and more than 1.5 million deaths per annum (Figure 1.1) (WHO, 2016).

1.2 Stages of disease

A- Initial stages

Inhaling the airborne droplets containing *Mtb* into the alveolar passages of a host is considered to be the initial step of infection. The dose and the status of transmitted tubercle bacilli causing the infection are still unclear (Fennelly *et al.,* 2012; Garton *et al.,* 2008). Resident macrophages are the first defences against the bacilli. However, other non-specific immune cells like dendritic cells and alveolar epithelial type II pneumocytes interact with *Mtb* as well. *Mycobacterium tuberculosis* is engulfed by phagocytic cells, and the engulfed bacilli are arrested within an endocytic vacuole called a phagosome (Smith, 2003). Normally, these vacuoles fuse with lysosomes in order to eliminate the pathogenic agents. However, the bacterium can block phagolysosome formation. Activation of autophagy and opsonization against the bacilli could lead to phagosome maturation which in turn depletes the

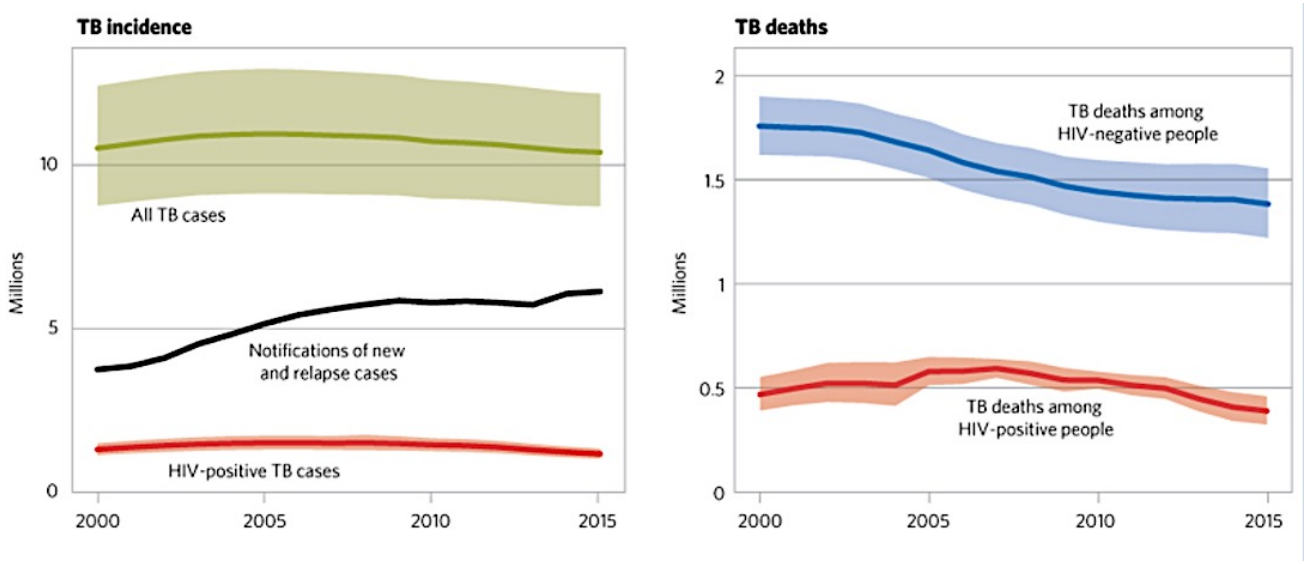


Figure 1.1: Estimated cases of TB incidence and deaths around the world each year from 2000 to 2015 according to World Health Organization estimations. The uncertainty intervals are shaded. This graph is from the WHO global tuberculosis report 2016 (WHO, 2016).

infection (Armstrong and Hart, 1975; Smith, 2003). *Mycobacterium tuberculosis* must survive in the harsh conditions within the macrophage. It has been found that activated macrophages in mice produce reactive nitrogen intermediates (RNIs), which are a significant antimicrobial weapon against TB. Other responses such as reactive oxygen intermediates (ROI), lysosomal enzymes and acid pH, also have potential roles in controlling the infection. The engulfed bacilli are either killed or grow within the cells, which leads to the latent stage of infection (Smith, 2003).

B- Latent stages

Macrophages infected with *Mtb* produce chemokines attracting other inactivated immune cells to form granulomatous focal lesions (van Crevel *et al.,* 2002). The typical mature human granulomas consist of a central region of infected macrophages surrounded by differentiated macrophages such as multinucleated giant cells, foamy macrophages and epithelioid cells. By development of cellular immune responses, infected macrophages are killed forming a caseous central mass. This central accumulation is surrounded later by lymphocytes CD4, CD8 T-cells and B-cells, blood-derived monocytes and fibroblasts forming a fibrotic capsule (Peters and Ernst, 2003; Russell *et al.,* 2010). Although, the caseous core of granuloma represents a hostile environment, due to the low level of oxygen, the presence of toxic fatty acids, and acidic pH, some *Mtb* survive in a dormant, asymptomatic, non-transmissible state for decades. The TB infection could be finished at this point or progress thereafter to cause an active pulmonary tuberculosis which depends on the quality of cellular immunity (Russell *et al.,* 2010; Tsai *et al.,* 2006). The caseous necrotic regions heal as small calcified lesions. Factors that suppress or depress the immunity like HIV infection, allow to these calcified lesions to be dissolved, allowing the persistent bacilli within its walls to grow quickly and spread through lungs causing an active infection (Figure 1.2) (O'Garra *et al.,* 2013).

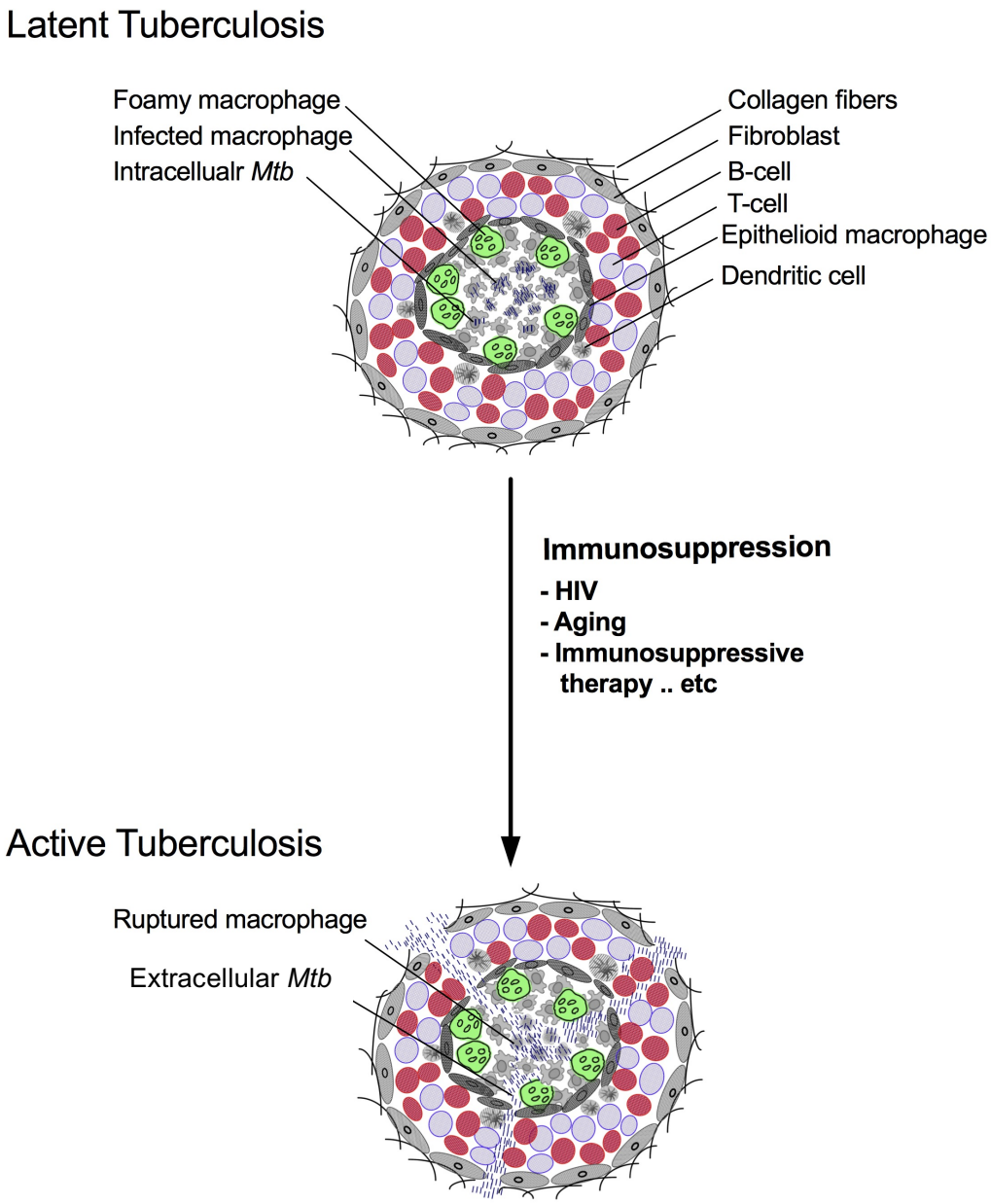


Figure 1.2: Tuberculosis granuloma structure in dormant (latent) and active states adapted from (Bhat et al., 2012). Under healthy immune system, mature human granulomas consist of a central region of macrophages infected by *Mtb* surrounded by differentiated macrophages and immune cells; immunosuppression leads to a disintegration of this structure and release of the active bacilli.

1.3 Redox stress during infection

Alveolar macrophages produce different redox agents in response to *Mtb* infection such as ROI and RNI. These agents act on different regulators that control the switching process between persistent and replicative bacilli (Bhat *et al.,* 2012). In this section the role of ROI, nitric oxide and hypoxia during infection is discussed.

NADPH oxidase (NOX) is an important enzyme of alveolar macrophages, generating superoxide radicals, which subsequently produce other ROI species by oxidizing NADPH (Jackett *et al.,* 1978). Reactive oxygen intermediates are considered important mycobactericidal agents because of their degradative role against lipids, proteins and nucleic acids (Kurthkoti and Varshney, 2012). Moreover, ROI play other roles in activating an inflammatory response and the induction of macrophage apoptosis. However, *Mtb* can neutralize the ROI stress by expressing defensive enzymes such as catalase (KatG), alkyl hydroperoxidase (AhpC), superoxide dismutase and peroxiredoxins (Bhat *et al.,* 2012; Kumar *et al.,* 2011).

The level of O_2 in macrophages and granulomas represents another stress factor. It is clear that *Mtb* faces a hypoxic environment in the granuloma, because this structure is an avascular assemblage. Oxygen availability is an important factor for *Mtb* growth. Rapid depletion of O_2 causes death, but the steadily decreasing in O_2 supply allows *Mtb* to respond and enter a dormant state (Wayne and Hayes, 1996). *Mycobacterium tuberculosis,* in response to the new conditions changes its cell wall thickness, becomes elongated, stops the synthesis of proteins, RNA and DNA, loses its acid fast characteristics and exhibits resistance to antibiotics (Wayne and Hayes, 1996; Wayne and Sohaskey, 2001). Oxygen absence as a terminal electron acceptor in the electron transport chain causes the accumulation of reduced electron carriers such as NADH (Bhat *et al.,* 2012).

High levels of NO cause *Mtb* death while lower concentrations impair the ability of *Mtb* to replicate and induce dormancy (Bhat *et al.,* 2012; Firmani and Riley, 2002). Nitric oxide production is catalyzed by NO synthase (NOS2) using L-

arginine as a substrate. This happens when suitable inducers activate macrophages and other monocytes. However, it has been shown that inhibition of NOS2 in murine macrophages during acute phase of *Mtb* infection leads to a dramatic increase in replication ability, indicating the bacteriostatic and bactericidal role of NO in host defence (MacMicking *et al.,* 1997; Raja, 2004).

1.4 Role of immunocytes during *M. tuberculosis* infection

Cytokines play an important role in keeping TB infections under control. CD4 T cells are the main producers of IFN-γ, and CD8 T cells also have the ability to produce this cytokine. Thus CD4 T cell depletion in mice causes CD8 T cells to produce more IFN-γ to a level that is enough to activate macrophages (Flynn and Chan, 2001). Interferon gamma (IFN-γ) is an important cytokine that can induce infected macrophages through activation of nitric oxide synthase (NOS2) leading to increase concentrations of RNI. However, IFN-γ seems to have other functions, since NOS2 deficient mice are less susceptible than mice deficient in IFN-γ (Dalton et al., 1993; Flynn and Chan, 2001).

On the other hand, depletion of CD4 T cells impairs the ability to control the infection in mice, even when the same level of IFN-γ is produced by CD8 T cells, suggesting that CD4 T cells may control the *Mtb* infection by other mechanisms. In addition, increasing the risk of *Mtb* reactivation and the extrapulmonary dissemination when coinfection with HIV happens, indicates the role of CD4 T cells in controlling TB. The other role of T cell types appears during the formation of granuloma in mice and humans (Flynn and Chan, 2001; O'Garra *et al.,* 2013).

Tumor necrosis factor alpha (TNFα) is another cytokine involved in controlling TB through activation of macrophages and RNI production, formation of granulomas, and effects on expression of other cytokines. The important role of TNFα and IFN-γ is in preventing *Mtb* reactivation (Kindler *et al.,* 1989). The continuous secretion of these cytokines during the period of *Mtb* dormancy suggests roles in keeping the bacilli in an inactive state (Flynn et al., 1998).

1.5 The *M. tuberculosis* responses to stress

Mycobacterium tuberculosis has the ability to invade the human immune system and remain within host cells in a dormant state for decades. Co-infection with other pathogens could result in an immuno-compromised state, allowing *Mtb* to switch from the persistent state to an active infection. This process necessitates regulation of the *Mtb* genetic network (Wang *et al.,* 2011).

Two-component systems consist of a sensor protein and a cytoplasmic response factor. These systems respond to environmental changes by eliciting a phosphotransfer reaction causing an adaptive transcriptional response (He *et al.,* 2006).

Repeated alternations in the level of oxygen within macrophages suggested that *Mtb* during the periods of infection has the ability to respond quickly switching its metabolism to and from non-respiring pathways (Leistikow *et al.,* 2010). The Dos two-component system regulates about 48 genes and is composed of two histidine kinase sensors (DosS and DosT) and one response regulator (DosR). This system responds to hypoxia, NO, and CO and plays a crucial role in metabolic processes, by sustaining the level of energy and keeping redox balance during conditions that hinder aerobic respiration (Kumar *et al.,* 2007; Shiloh *et al.,* 2008). Furthermore, the DosR regulon plays a role in transition of *Mtb* between respiring active and non-respiring dormant states (Leistikow *et al.,* 2010).

PhoP-PhoR is another two-component system, consisting of a sensor histidine kinase (PhoR) and a transcriptional regulator (PhoP), playing a role of *Mtb* in both dormant and active states. Disruption of *Mtb* PhoP regulator attenuates the ability to replicate intracellularly in mice models, alters the nature of *Mtb* cell envelope due to the systems role in regulating the biosynthesis of polyketide-derived lipids, and impairs *Mtb* virulence (Gonzalo Asensio *et al.,* 2006). In addition, it has been suggested that PhoP regulates *rv3614c-rv3616c* (*espACD*) gene cluster, a genetic region required to secrete ESAT-6 and CFP-10 proteins, which are substrates for the ESX-1 secretion system that is essential for virulence of *Mtb*. This has been indicated through the down

regulation of *espA* gene cluster in a PhoP mutant (Fortune *et al.,* 2005; Frigui *et al.,* 2008; Walters *et al.,* 2006). Therefore, the regulation of the ESX-1 secretion system could respond to environmental conditions sensed by the PhoP/R system.

1.6 Structure and general features of Wbl proteins

WhiB-like (Wbl) proteins are a distinctive group of proteins that are only found in actinomycetes and play a crucial role in developmental processes (Soliveri *et al.,* 2000). The first Wbl protein (WhiB) was discovered in *Streptomyces coelicolor,* and was shown to play a role in sporulation (Davis and Chater, 1992). Wbl proteins are small (75 to 130 amino acid). Their *C*-terminal regions have weakly predicted helix-turn-helix motifs with a loop composed of a conserved motif (GV/IWGG) between the two helices. These *C*-terminal regions are positively-charged, increasing the possibility of interacting with DNA. The *N*-terminal domains contain four conserved cysteine residues that act as anchors for iron-sulfur clusters that could change the conformation of the proteins in response to redox stresses (Larsson et al., 2012; Soliveri et al., 2000).

In *Mtb*, the seven Wbl (WhiB1-7) proteins share a common arrangement of the four conserved cysteine residues (Figure 1.3). The general alignment is C-X19-36-C-X-X-C-X5-7-C except in WhiB5, which has three amino acids between the two central cysteines instead of two (Alam *et al.,* 2009; Soliveri *et al.,* 2000). The four conserved cysteine residues of *Mtb* Wbl proteins can bind [2Fe-2S] clusters when the Wbl proteins are aerobically expressed and purified, and [4Fe-4S] clusters when anaerobically reconstituted. The iron-sulfur clusters of *Mtb* Wbl proteins disassemble gradually under oxidizing conditions and are relatively stable under reducing conditions, suggesting a possible role of these proteins as sensors of redox stress (Alam *et al.,* 2009). Furthermore, all *Mtb* Wbl proteins, except WhiB2, have the ability to reduce the insulin disulfide, suggesting a role as protein reductase. It is thought that the enzymatic activity is controlled by the two central cysteine residues and the disassembly of the iron-sulfur cluster to switch the protein from holo- to apo-form is important to possess this activity (Alam et al., 2009; Garg et al., 2007a).

1.7 Role of Wbl proteins in *M. tuberculosis*

Mycobacterium tuberculosis undergoes many modifications in gene expression in order to adapt to the harsh environment of the macrophage. These changes keep the bacteria in a non-replicating persistent state. However, how *Mtb* switches between replicating and non-replicating forms is poorly characterized. It is possible that Wbl proteins play a vital role in *Mtb* dormancy and reactivation processes by responding to stresses and modulating the transcription of genes.

WhiB1 protein has the ability to bind at its own promoter, suggesting its role as a self regulator. Apo-WhiB1 in both reduced and oxidized forms is able to bind DNA, while the holo forms cannot. Interestingly, the [4Fe-4S] cluster interacts rapidly with eight NO molecules to form a dinitrosyl iron complex (DNIC). The nitrosylation process can switch the protein from a non-binding form to a DNA-binding protein. Although, there is no direct evidence of a WhiB1 response to NO within macrophages, the *in vitro* reaction suggests that WhiB1 could regulate gene transcription in response to NO within activated macrophages (Smith *et al.,* 2010). WhiB1 expression is also regulated by CRP^{Mt} (Rv3676). It has been shown that *whiB1* transcription is regulated by direct binding of cAMP- activated CRP^{Mt} upstream of the *Mtb whiB1* promoter (Agarwal *et al.,* 2006). Further study showed that the CRP^{Mt} homdimer binds two molecules of cAMP, and there are two adjacent binding sites within the *whiB1* promoter. Interestingly, *whiB1* transcription is repressed when both binding sites are occupied by CRP^{Mt} at high levels of cAMP, while the only upstream binding site is occupied by CRP^{Mt} at low concentrations of cAMP resulting in activating of *whiB1* transcription (Stapleton *et al.,* 2010). It seems that the high level of NO and cAMP in macrophage could repress *whiB1* expression suggesting the possibility of entering into the persistent state (Smith *et al.,* 2010).

WhiB2 is functionally homologous to *Mycobacterium smegmatis* WhmD, and the identity between amino acid sequences is 70%. Moreover, the *C*-terminal regions of both proteins are similar to the WhiB protein of *Streptomyces coelicolor* (Figure 1.3B) (Cole *et al.,* 1998; Raghunand and Bishai, 2006a). WhmD is important for mycobacterial septation, and when *whmD* is mutated in

a conditionally complemented system, *M. smegmatis* loses the ability to form septa, leading to irreversible filamentous branched growth. In contrast, the overexpression of *whmD* causes hyperseptation (Gomez and Bishai, 2000). Both *Mtb* WhiB2 and *S. coelicolor* WhiB proteins can complement the defect in an *M. smegmatis whmD* conditional mutant demonstrating the homology between these three proteins (Raghunand and Bishai, 2006b). Unlike WhiB1, WhiB2 in *M. smegmatis* uses three of the four conserved cysteines with a conserved aspartate to anchor [4Fe-4S] cluster. Site-directed mutagenesis shows that C90, C93, C99 and D71 are indispensable for the activity of the protein while C67 is not essential. It was suggested that D71 could hold the cluster along with the other three cysteines in the absence of C67 (Raghunand and Bishai, 2006b). WhiB2 like other Wbl proteins is thought to act as a transcriptional regulator. This suggestion comes from the ability of the apo-WhiB2, but not the holo-form, to bind its promoter, indicating its role as an autoregulator. The expression of *whiB2* is downregulated in the present of mycobacteriophage TM4 protein (WhiBTM4), which is able to bind upstream of *whiB2* to hinder septation within mycobacterial host. Interestingly, WhiBTM4 can be reconstituted to anchor [2Fe-2S] cluster; however holo-forms of WhiB2 and WhiBTM4 must be switched to apo-forms in order to be able to bind on the *whiB2* promoter region (Rybniker *et al.,* 2010). Also, it has been shown that WhiB2 has the characteristics of a chaperone, which can prevent the aggregation of many substrates like citrate synthase, rhodanese and luciferase without the need for ATP, and the redox state of the cysteines does not have any effect on this activity (Konar *et al.,* 2012).

WhiB3 like other Wbl proteins is thought to act as a DNA-binding regulatory protein. Interestingly, WhiB3 regulates *Mtb* lipid metabolism during infection in responding to redox stresses, and the iron-sulfur cluster is the key factor in this response (Singh *et al.,* 2009). The cluster can react specifically with dormancy signals e.g. O_2 and NO, which could lead to changes in the protein's conformation (Singh *et al.,* 2007). WhiB3 holo form with [4Fe-4S] cluster binds DNA, and the redox state of this cluster does not change its DNA-binding ability. Moreover, reduction of cysteine residues in apo-WhiB3 abolishes DNA-binding, while oxidation stimulates a more stable binding than that observed for

the holo-form. WhiB3 can bind at *pks2* and *pks3* promoter regions, that are responsible for regulating polyketide synthesis (Singh *et al.,* 2009). This has been supported by other findings that suggest the *Mtb*Δ*whiB3* suffers alterations in lipid composition of the cell wall, which effect colony rugosity (Singh *et al.,* 2007; Steyn *et al.,* 2002). It is also been found that WhiB3 regulates the anabolism of other virulence lipids such as polyacyltrehaloses (PAT), diacyltrehaloses (DAT) and sulfolipids (SL-1), and deletion of WhiB3 decreases the expression of these lipids resulting in attenuation in *Mtb* and *M. bovis* in mouse and guinea pig models (Steyn *et al.,* 2002). WhiB3 seems to be important to maintain redox homeostasis. High levels of NADH and NADPH are produced during the β-oxidation of fatty acids when they are utilized as a primary energy source in *Mtb*. In macrophages, it has been noticed reductive stress increases upon infection in *Mtb*Δ*whiB3* compared with wild type. It was also found that WhiB3 dissipates the toxicity of propionate produced from the oxidation of fatty acids (Singh *et al.,* 2007, 2009). To sum up, WhiB3 could play a regulatory role in lipid metabolism in response to environmental redox signals.

WhiB4 is not much different in its biophysical and biochemical characteristics to the other *Mtb* Wbl proteins. However, WhiB4 responds to O_2 and loses its [4Fe-4S] cluster more quickly than other Wbl proteins (Chawla *et al.,* 2012). WhiB4 was also found to react with NO forming a DNIC, similar to WhiB1 and WihB2. The WhiB4 holo-form [4Fe-4S] in both oxidized and reduced states does not show DNA binding ability. Interestingly, cysteine residues in the protein are required for DNA-binding (Chawla *et al.,* 2012). Like WhiB3, reduction of apo-WhiB4 with DTT abolishes DNA-binding ability, while oxidation with diamide induced this ability, suggesting that the oxidation of the four cysteines is essential to possess this activity. It is still unknown whether WhiB4 can bind to specific regions in *Mtb* genome. However, WhiB4 can bind non-specifically on minor grooves of GC-rich regions. The most important finding was *M. tuberculosis* WhiB4 regulates the expression of stress responsive and antioxidant genes (Chawla *et al.,* 2012). Microarray analysis showed that expression of 25 genes are up-regulated in *Mtb*Δ*whiB4* under normal conditions, and qRT-PCR results indicated that several of them are

associated with antioxidant systems. Surprisingly, WhiB4 and WhiB6 were up-regulated in a *Mtb*Δ*whiB4* mutant, suggesting that WhiB4 could be a self-repressing protein, and that WhiB6 could compensate for the role of WhiB4 as a redox sensor (Chawla *et al.,* 2012).

Little is known about WhiB5 protein function. However, one study showed that the WhiB5 is a regulatory protein that is important in virulence and reactivation of *Mtb* during infection (Casonato *et al.,* 2012). WhiB5 mutant showed attenuation during infection with poorer dissemination within infected mouse lungs but with a high inflammatory response. Losing the ability of growth resumption in the mutant after reactivation from chronic infection was also recorded (Casonato *et al.,* 2012). This suggested that the role of WhiB5 is an immunomodulator and a regulatory protein controlling the expression of *Mtb* genes involved in the reactivation process. Moreover, a number of genes such as *sigM*, encoding an alternative sigma factor, and other genes needed for ESX-2 and ESX-4 secretion systems are influenced by the WhiB5 protein, indicating its role in virulence (Casonato *et al.,* 2012).

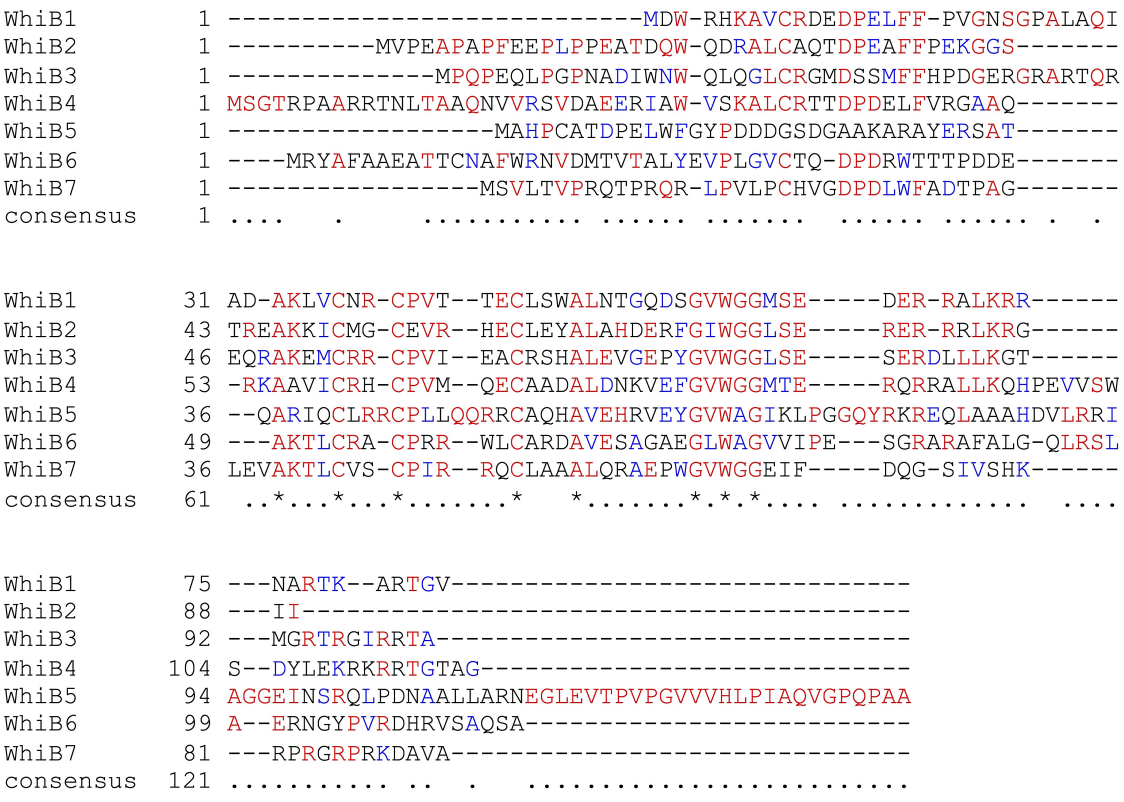
WhiB6 has been suggested to play a role in regulating the virulence factor (ESAT-6) via the PhoP regulator (Solans *et al.,* 2014). ESAT-6 and its partner CFP10 in *Mtb* are secreted by ESX-1. Inactivation of one of these two factors leads to extreme attenuation of virulence (Fortune *et al.,* 2005). WhiB6 expression is differentially regulated in clinical and laboratory strains via the PhoP regulator. Moreover, the clinical isolates of *Mtb* (GC1237 and MT103) produce and secrete ESAT-6 in larger amounts than the H37Rv strain. Interestingly, an insertion mutation has been found in WhiB6 promoter region in the H37Rv and H37Ra strains at the PhoP binding site. This insertion could form an inverted repeat that can produce a loop structure (Solans *et al.,* 2014). The PhoP regulator up-regulates the expression of *whiB6* in clinical strains, but it down-regulates the expression in H37Rv. In addition, reintroducing a wild type copy of *whiB6* in H37Rv causes production and secretion of ESAT-6 to the same level as in the clinical strains, suggesting the role of WhiB6 as a regulator of ESX-1 system via PhoP in clinical strains (Solans *et al.,* 2014).

ESAT-6 secretion can be controlled by PhoP in two dependent pathways. First, as mentioned previously, PhoP regulates ESAT-6 through *espACD* gene cluster in the extended ESX-1 region. Secondly, through WhiB6 leading to regulation of the *esxA* and *esxB* genes in the RD1 locus (Solans *et al.,* 2014). Genes found in both regions are needed for ESX-1 function (Fortune *et al.,* 2005). However, it is not clear whether WhiB6 can directly regulate genes of RD1 or/and extended ESX-1 regions or whether it acts through other regulators (Figure 1.4). Expression of *Mtb whiB6* is up-regulated within macrophages and when the *in vitro* cultures are treated with NO (Larsson et al., 2012). This suggests the possibility of PhoPR working with WhiB6 in response to nitrosative/oxidative stresses within macrophages leading to control of ESAT-6 secretion.

WhiB7 is an important regulator required for antibiotic resistance in *Mtb*. The expression of *whiB7* is upregulated in response to various structural classes of antibiotics including macrolides, tetracyclines, lincosamides and some aminoglycosides. In addition, a *whiB7* mutant shows hypersensitivity to these antibiotics (Burian *et al.,* 2012; Morris *et al.,* 2005). Interestingly, 86 antibiotics induce *whiB7* expression, but not all of them were effective against the mutant, suggesting the possible role of WhiB7 in other functions (Burian *et al.,* 2012). WhiB7 is also upregulated in response to physiological stresses such as iron-starvation, heat shock, and entry into stationary phase (Geiman *et al.,* 2006). Three different domains in WhiB7 defined its probable function in response to redox stress leading to intrinsic drug resistance. The four conserved cysteine residues in the *N*-terminal region form the binding domain for an iron-sulfur cluster that stabilizes the tertiary structure of the protein, which in turn induces binding with SigA through the middle domain. The AT hook motif in the *C*-terminal region is predicted to bind at AT-rich sequences, aid the RNAP recruitment and regulate gene expression (Burian et al., 2013). The *whiB7* promoter and promoters of its regulon (*eis*, *erm* and *tap*) contain a conserved AT-rich sequence motif suggesting that this is the recognition sequence for the AT-hook motif (Burian et al., 2012, 2013; Morris et al., 2005). Moreover, *whiB7* transcription is induced under reducing conditions, while expression is decreased under oxidizing conditions. This suggests that *whiB7* expression

under antibiotic induction could be dependent on a shift in metabolism which causes a change in redox potential (Burian *et al.,* 2012). Although, WhiB3 dissipates the reductive stresses generated by accumulation of NADH/NADPH, which result from catabolism of fatty acids within macrophages, the induction of *whiB7* expression immediately after infection of macrophages by *Mtb* could represent a response to the initial reductive burst (Homolka *et al.,* 2010).

[A]



[B]

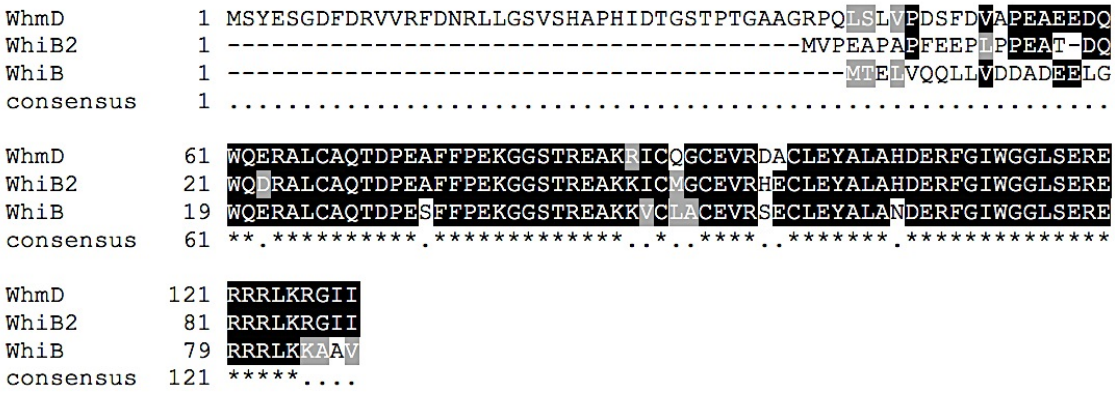


Figure 1.3: Multiple sequence alignment of Wbl proteins. [A] *Mycobacterium tuberculosis* Wbl proteins, note, WhiB5 lacks the first Cys residue. **[B]** *Mycobacterium smegmatis* WhmD, *M tuberculosis* WhiB2, and *Streptomyces coelicolor* WhiB proteins. The sequences of the proteins were obtained from NCBI website, aligned by multalin tool, shaded and coloured by Boxshade tool.

1.8 The role of *M. tuberculosis* WbI proteins in regulation of type VII virulence secretion system

Mycobacterium tuberculosis has five types of secretion systems to produce and secrete proteins that are important for its virulence and viability. These systems are named as ESX-1 to ESX-5. The best-known system is the type VII secretion system (ESX-1), which is implicated in secreting two major antigens, EsxA and EsxB (Houben *et al.,* 2014). These two proteins are also known as ESAT-6 and CF10 respectively, together they form a heterodimer and the structure of this dimer has been solved (Renshaw *et al.,* 2005). ESX-1 gene clusters are present in the region of difference 1 (RD1), a chromosomal locus of about 9.5 kb that is absent in the *M. bovis* BCG vaccine strain (Guinn *et al.,* 2004). It has been found that the deletion of this region attenuates the growth of virulent mycobacterial species within macrophages. Moreover, the absence of this region in *M. bovis* BCG is responsible for attenuation of this strain; since, introducing the extended RD1 into *M. bovis* BCG increases virulence (Lewis *et al.,* 2003; Pym *et al.,* 2002). ESAT-6 and CFP-10 are small proteins with masses of 6 kDa and 10 kDa, respectively. These two factors play diverse roles in pathogenicity and interaction with the host immune system, inhibition of T cell, host cell lysis, preventing the phagosomal maturation process and modulating macrophage cell death. Deletion of these proteins attenuates *Mtb*, while deletion of other genes in the cluster impairs the function of ESX-1 (Aguiló *et al.,* 2013; Samten *et al.,* 2011).

Another region is also needed for the function of ESX1. This extended ESX-1 locus is located outside the RD1, and composed of an operon containing five genes: *espA* (*Rv3616c*), *espC* (*Rv3615c*), *espD* (*Rv3614c*), *Rv3613c* and *Rv3612c* (Gao *et al.,* 2004). The most important Esp protein is EspA because its secretion is dependent on ESAT-6, this means that deletion of *espA* prevents *Mtb* from secreting ESAT-6. Furthermore, EspC is required for secretion of EspA and ESAT-6; while, EspD is important to modulate the levels of EspA and EspC (Fortune *et al.,* 2005). Upstream of the *espACD* operon is a large intergenic region of about 1.5 kb, and several regulatory factors have been found to regulate the expression of the *esp* operon directly and indirectly (Figure 1.4). The most important factors are: the positive regulator, EspR; the

negative regulators, CRP, MprAB, PhoP and WhiB1 (Cao et al., 2015; Pang et al., 2013; Rickman et al., 2005). In addition, according to Solans *et al.,* (2014), qRT-PCR data show that WhiB6 is a very important regulator of the ESX-1 secreted virulence factor (ESAT-6) in clinical strains of *Mtb*. This type of regulation is mainly under the control of the PhoP regulator. Although PhoP can regulate the expression of ESX-1 through indirect regulation of the *espACD* operon*,* the study showed that PhoP positively regulates *whiB6*, which in turn increases the activity of ESX-1. *Mycobacterium tuberculosis* lacking the PhoP or WhiB6 regulators exhibit decreased production and secretion of ESAT-6. Moreover, as mentioned above MprAB stimulates the expression of *espA* through regulation of EspR, it has been found that MprAB also binds upstream of *whiB6* (Pang *et al.,* 2013), which might lead to regulation of the *espACD* operon. According to this information, WhiB6 is thought to play a very important role in regulation of ESX-1 secretion system. However, there is a lack of information about the nature and mechanism of this regulation.

In *Mtb,* expression of the *espA* operon was upregulated when the *crp* gene was deleted (Rickman et al., 2005). Unpublished data (by Dr. Laura Smith) showed that CRP binds at the *espA* promoter region. Moreover, RT-PCR showed that the expression of *espA* was induced when *whiB1* expression was reduced (Roger Buxton, NIMR, unpublished data). Therefore, it has been hypothesized that WhiB1 and CRP regulate the *espA* operon.

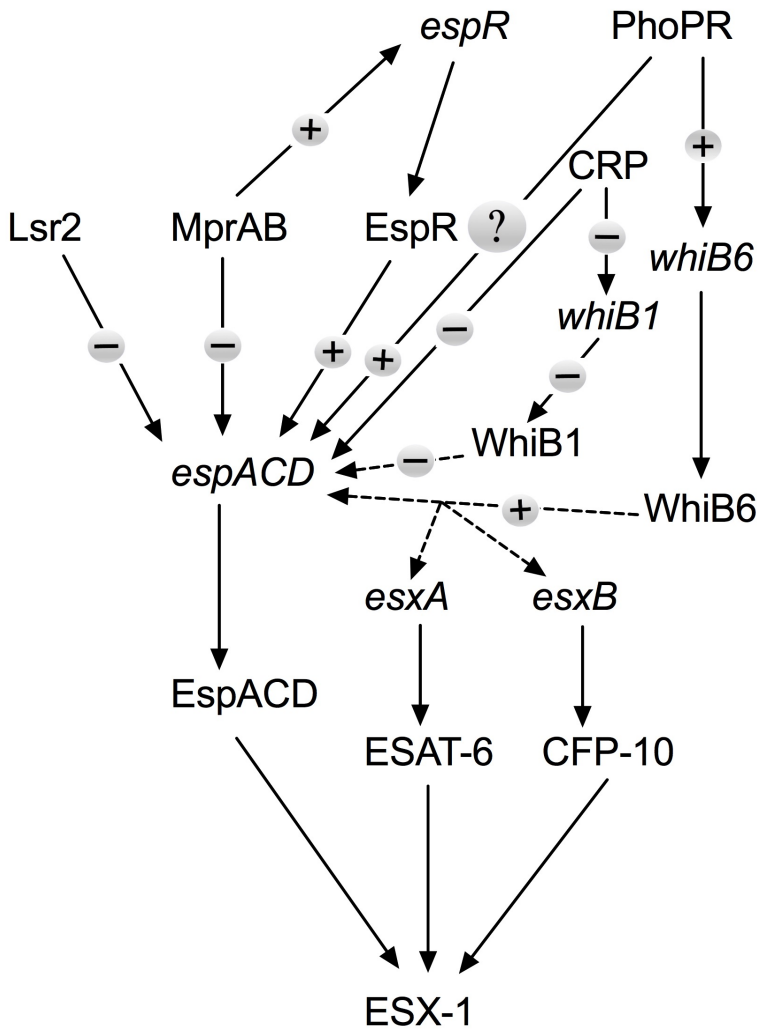


Figure 1.4: The genetic regulation of the ESAT-6 secretion system (ESX-1) in *M. tuberculosis*. Bold arrows indicate confirmed regulatory effects, while dotted and dashed arrows refer to the possibility of regulation. The (+) and (-) symbols refer to the up-regulation and down-regulation of gene expression respectively, while the (?) refers to the presence of a regulatory protein that mediates regulation but is not yet identified.

1.9 Overview of *M. tuberculosis* sigma factors

The ability of *Mtb* to face the harsh environments of hosts' macrophage cells during persistent infections requires precise regulation of the expression of genes that encode for enzymes of metabolic pathways and virulence factors. Important players in this regulation are alternative sigma (Sig) factors that respond to signals encountered in the host (Sachdeva *et al.,* 2010).

The holoenzyme of RNA polymerase (RNAP) is essential at the level of gene transcription. RNAP consists of four core subunits ($\alpha_2\beta\beta'\omega$), and the fifth is a dissociable sigma (σ) factor subunit (Figure 1.5). However, only the holoenzyme can initiate gene transcription. Sigma factors are crucial for promoter recognition and specificity (Gruber and Gross, 2003). Association and dissociation of sigma factors with core RNAP responds to cellular signals, thereby redirecting transcription (Vergne et al., 2004). In addition to sigma factors in directing transcription through interaction with RNAP core enzyme, many of them have been found play a role as targets for a variety of transcription factors, involving both DNA-binding activators and anti-sigma factors (Dove *et al.,* 2003).

Due to the importance of gene transcription for cell viability, all bacteria have essential sigma factors, which are required for the transcription of housekeeping genes in exponentially growing cells; and alternative sigma factors, which are environmentally responsive regulators that direct RNAP to specific classes of promoters in response to extracellular environmental conditions. At least one essential sigma factor is present in all bacterial cells (Ishihama, 2000). *Mycobacterium tuberculosis* encodes thirteen different sigma factors designated as SigA to SigM, each one has its own promoter specificity, to control the transcription of subsets of genes (Manganelli *et al.,* 2004).

1.10 The structure and the classification of sigma factors

Two distinct groups of sigma factors can be classified according to either Sig^{70} or Sig^{54} in *E. coli*. However, only Sig^{70} types are present in mycobacteria. Sig^{70} consists of four conserved regions, each one can be divided into sub-regions (Gruber and Gross, 2003). Region 1 is sub-divided into 1.1 and 1.2 sub-regions

and located at the *N*-terminal region of the protein, the presence of negatively-charged amino acids in sub-region 1.1 prevents the free sigma factors binding to DNA (Dombroski *et al.,* 1993). Region 2 has four sub-regions, 2.3 is required for forming the transcription bubble, while the sub-region 2.4 is essential for recognizing the –10 promoter element, a hexameric sequence located at 10 bp upstream of the transcription start point (Panaghie *et al.,* 2000). Three sub-regions are present in region 3, and an α-helix in sub-region 3.0 has been implicated in detecting the extended -10 promoter box, an extended -10 motif (TGN motif) located directly upstream of the -10 element of promoters (Sanderson *et al.,* 2003). The final region has 4.1 and 4.2 sub-regions, which is required for recognition of the - 35 promoter element, a hexameric sequence located about 35 bp upstream of the TSP; region 4.2 is also needed for partnering with many transcription factors. Simultaneously, all of the sigma factor regions contribute to different extents in binding with RNAP (Dove *et al.,* 2003; Gruber and Gross, 2003). In 1993, Ross et al. defined an AT-rich UP recognition element, usually extending between 41 and 61 bp upstream of the TSP, which can bind with the *C*-terminal domain of the α subunit of RNAP leading to increase promoter activity (Figure 1.5) (Ross *et al.,* 1993).

The family of Sig^{70} in *Mtb* can be classified according to their structure and physiological function into four groups: Group 1 includes essential or principal sigma factors, which contain the four-conserved regions, and is represented by SigA. Group 2 is known as primary factor-like sigma factors. Although, they are extremely similar to primary SigA factors, they lack most of region 1, and usually they are non-essential under normal physiological growth conditions. This group is known as SigB. Group 3 sigma factors also lack region 1, but it is the most heterogeneous compared to the essential SigA. These are known as alternative sigma factors and are represented by SigF. The other sigma factors can be classified under the group 4, which only has the conserved regions 2 and 4, and are described as extracytoplasmic function (ECF) sigma factors according to their roles in response to stress conditions (Wösten, 1998).

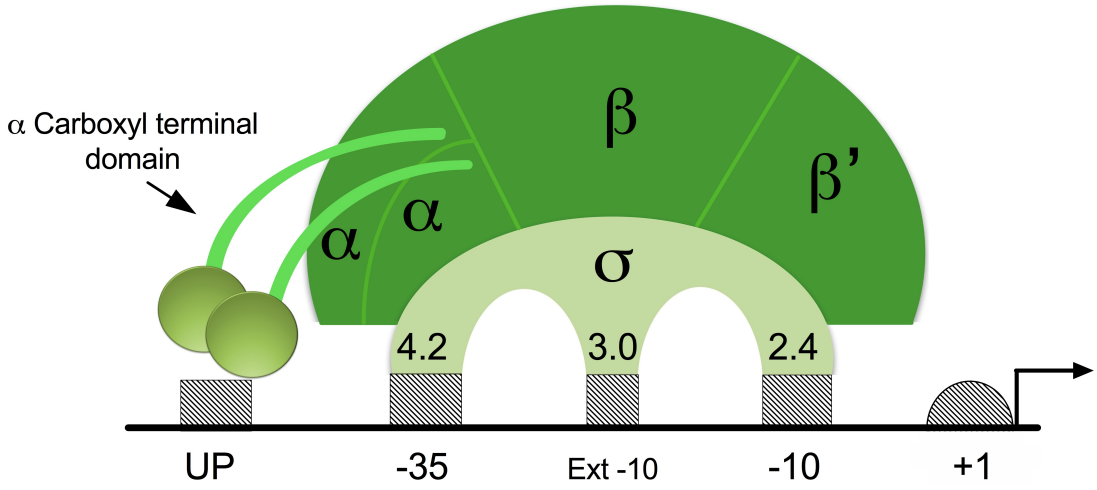


Figure 1.5: Promoter recognition elements by the holoenzyme of RNA polymerase adapted from (Sachdeva et al., 2010). The figure shows the subunits of holo-RNAP ($\alpha_2\beta\beta'\sigma$), and regions of Sig^{70} factor (2.4, 3.0 and 4.2) that interact with the DNA elements promoters. The +1 represents the transcription start point; -10, Ext -10 and -35 refer to -10 element, extended -10 element and -35 element respectively; while the UP is the up recognition element.

1.11 The physiological function of SigA factors

SigA is an essential protein in both *Mtb* and *M. smegmatis* (Gomez *et al.,* 1998; Manganelli *et al.,* 2004). Expression of *sigA* is constitutive and the amount of SigA is similar in *Mtb* H37Rv grown under a range of conditions. However, it has been reported that the expression of *sigA* in some clinical isolates is increased in human macrophages and in the lungs of mice after infection. In contrast, expression decreased when the cell growth reached stationary-phase and in hypoxic conditions (Volpe et al. 2006). The up-regulation of *sigA* after phagocytosis seems to be to improve intercellular growth of *Mtb* during superoxide stress, indicating the role of SigA in controlling the expression of virulence genes. Down-regulation of *sigA* during hypoxia and in stationary phase, might reflect the lower energy obtainable by bacterial cells, which leads to a decrease in the total mRNA pool, causing a drop in *sigA* transcription (Manganelli *et al.,* 1999; Wu *et al.,* 2004).

In addition to SigA regulating the transcription of housekeeping genes, there are several examples of its role in host-pathogen interactions often by regulating virulence in *Mycobacterium* species. During adaptation to hypoxia and in response to gasotransmitters such as NO and carbon monoxide, the two component Dos system is essential. DosR is believed to activate transcription by binding at a minimum of two sites in target promoters. One of these sites interferes with SigA -35 elements, suggesting that the DosR-SigA interaction is essential to activate transcription. Two point mutations E154A and E178A in DosR cause a severe defect in *Mtb* growth under hypoxia. Although the mutated DosR still bound DNA, the interaction with SigA was prevented completely, which was proved by *in vitro* and *in vivo* assays. This suggests a role of DosR-SigA complex in driving bacterial survival in hypoxic environments (Gautam *et al.,* 2014).

A Point mutation (R515H) in the *C*-terminal domain of SigA (sub-region 4.2), which has been identified to interact with the transcription regulator WhiB3 in *M. bovis,* caused attenuation of virulence in a guinea pig model of infection. This mutation prevents the interaction of SigA and WhiB3. It has also been reported that the inactivation of the *whiB3* gene in *M. bovis* caused a similar

attenuation; indicating, that preventing of SigA-WhiB3 interactions accounted for attenuation. However, different animal models did not show a similar effect for the same mutation. Also inactivation of *Mtb* WhiB3 caused only partial attenuation. It is possible that additional transcription factors compensate for the lack of WhiB3 function in *Mtb* due to differences in genetic backgrounds between *Mtb* and *M. bovis* (Steyn *et al.,* 2002).

Another study showed that WhiB7 interacts with region 4 of SigA. The residue R515 of SigA was indispensable for binding with WhiB7. WhiB7 is the only *Mtb* Wbl protein that has an AT-hook domain at its *C*-terminus, this domain is not essential for SigA binding, but crucial for DNA-binding at the minor grooves of AT-rich DNA. A point mutation E63D in the middle domain of WhiB7 had a similar effect as R515 mutation in SigA by preventing interaction of WhiB7 and SigA. Preventing WhiB7-SigA interaction leads to loss of WhiB7 function and antibiotic susceptibility (Burian et al., 2013).

A recent study confirmed the interaction of all *Mtb* Wbl proteins (except WhiB5) with SigA by *in vitro* and *in vivo* assays. This was achieved using *in vitro* pull down assays, and by using mycobacterial protein fragment complementation (M-PFC) in *M. smegmatis*, and the two hybrid BACTH system in *E. coli*. As shown previously, R515 of SigA is necessary for interaction with WhiB3 and WhiB7. This was tested with the rest of the Wbl proteins using the BACTH and M-PFC systems. Mutation of R515 clearly decreased interaction with WhiB3, 4, 6 and WhiB7, suggesting that all bind to SigA via a similar interface. However, the mutated SigA still interacted with WhiB1 and WhiB2 as strongly as the wild type, suggesting that WhiB1 and WhiB2 might target other sites in SigA for interaction (Feng et al., 2016). To date no studies have considered the role of the iron-sulfur clusters of Wbl proteins in interaction with SigA.

1.12 Aims of the study

As mentioned previously in this chapter, *Mtb* has the ability to "hide" from the immune system and enter into a dormant state and stay alive intracellularly for decades. Different redox agents are produced by macrophages such as ROI and RNI in response to *Mtb* infection. *Mycobacterium tuberculosis* respond to these agents by several regulator systems which lead to changes in its metabolism and expression of virulence systems by reprogram expression of a large number of its genes. Among these systems that responds to redox agents are DosR/T/S and Wbl proteins. WhiB1 is encoded by an essential gene and is the focus of this study. WhiB1 anchors an iron-sulfur cluster and is able to interact with DNA. However, losing the cluster leads to an induction of the DNA-binding activity. In addition, the reaction of WhiB1 with NO also switches the protein to a DNA-binding form. The WhiB1 iron-sulfur cluster does not react with O_2 but is very sensitive to NO leading to the suggestion that WhiB1 is a NO sensor (Smith *et al.* 2010). The *Mycobacterium tuberculosis sigA* gene is also essential and encodes a protein which can bind WhiB1, but the role of the iron-sulfur cluster in this interaction is unknown. According to these facts, a hypothesis of this study is: WhiB1 is a transcription regulatory factor that can bind SigA, and this binding is governed by response to redox agents via the cluster. Thus, this hypothesis was tested in this project by designing a set of biochemical and biophysical experiments. In addition, structural studies by NMR were also carried out to investigate the structure of WhiB1.

Chapter two

Material and methods

2.0 Materials and methods

2.1 Bacterial strains

The bacterial strains used in this study are listed in Table 2.1

Table 2.1: Strains of bacteria used in this study

Strain	Description	Reference
JRG2357	*E. coli* DH5α Δ*lacZYA-argF, lacZΔM15*	Invitrogen
JRG3196	*E. coli* DH5α / pDB55	Lab collection
JRG4968	*E. coli* BTH101 (strain for two hybrid BACTH system, geno type is: F-, *cya-99, araD139, galE15, galK16, rpsL1, hsdR2, mcrA1, mcrB1;* Str^R	Lab collection
JRG4972	*E. coli* DH5α / pUT18c-zip	Lab collection
JRG4974	*E. coli* DH5α / pKT25-zip	Lab collection
JRG 5946	*E. coli* DH5α / pGS2183	Lab collection
JRG5302	*E. coli BL21* λ(DE3) lysogen containing a chromosomal copy of the T7 RNAP under the control of the *lacUV5* promoter induced by IPTG	Invitrogen
JRG5386	*E. coli* BTH101 / pGS1669 and pGS1672	Lab collection

JRG5387	*E. coli* BTH101 / pGS1671 and pGS1673	Lab collection
JRG6161	*Mycobacterium smegmatis* MC^2155	Lab collection
JRG6769	*E. coli BL21* λ(DE3)/ pGS2500	This work
JRG6787	*E. coli* DH5α / pMyNT	Addgene
JRG6798	*Mycobacterium smegmatis* MC^2155 / pGS2524	This work
JRG6823	*Mycobacterium smegmatis* MC^2155 / pGS2542	This work
JRG6824	*Mycobacterium smegmatis* MC^2155 / pGS2543	This work
JRG6848	*E. coli* DH5α / pGS2524	This work
JRG6849	*E. coli* DH5α / pGS2560	This work
JRG6850	*E. coli* DH5α / pGS2561	This work
JRG6851	*E. coli* DH5α / pGS2563	This work
JRG6852	*E. coli* DH5α / pGS2564	This work
JRG6853	*E. coli* DH5α/ pGS2565	This work
JRG6854	*E. coli* DH5α / pGS2566	This work
JRG6855	*E. coli* DH5α / pGS2567	This work
JRG6856	*E. coli* DH5α / pGS2568	This work
JRG6857	*E. coli* BL21 λ(DE3) / pGS2566 and pGS2560	This work
JRG6858	*E. coli* BL21 λ(DE3) / pGS2565 and pGS2560	This work
JRG6859	*E. coli* BL21 λ(DE3) / pGS2183 and pGS2560	This work
JRG6862	*E. coli* BTH101 / pGS2567 and pGS2568	This work

JRG6863	*E. coli* DH5α / pGS2569	This work
JRG6864	*E. coli* DH5α / pGS2570	This work
JRG6865	*E. coli* DH5α / pGS2571	This work
JRG6866	*E. coli* DH5α / pGS2572	This work
JRG6867	*E. coli* DH5α / pGS2573	This work
JRG6868	*E. coli* DH5α / pGS2574	This work
JRG6869	*E. coli* DH5α / pGS2575	This work
JRG6870	*E. coli* DH5α / pGS2576	This work
JRG6871	*E. coli* DH5α / pGS2577	This work
JRG6872	*E. coli* DH5α / pGS2578	This work
JRG6873	JRG4968 (BTH101) / pGS2567 and pGS2579	This work

2.2 Plasmids

The plasmids used in this study are listed in Table 2.2.

Table 2.2: Plasmids used in this study

Plasmid	Description	Reference
pET21a	His_6-tag overexpression plasmid, Amp^R	Novagen
pET28a	His_6-tag overexpression plasmid, Kan^R	Novagen
pDB55	pT7-7 –*nifS*, Amp^R	Lab collection
pGS1669	pUT18, a two hybrid BACTH system vector, which encodes the T18 fragment downstream the MCS, Amp^R	Lab collection
pGS1671	pUT18c-zip, a two hybrid BACTH system control vector, which leucine zipper of GCN4 is genetically fused in frame to T18 frag.	Lab collection
pGS1672	pKT25, a two hybrid BACTH system vector, encodes the T25 fragment upstream the MCS, Kan^R	Lab collection
pGS1673	pKT25-zip, a two hybrid BACTH system control vector, which leucine zipper of GCN4 is genetically fused in frame to T25 frag.	Lab collection
pGS2183	pET28a-*sigA*, Kan^R	Lab collection
pGS2500	pET28a-*whiB1*, Kan^R	This work
pGS2521	Mycobacterial His_6-tag overexpression plasmid, Hygromycin-B^R	Addgen
pGS2524	pMyNT-*whiB1*, Hygromycin-B^R	This work
pGS2542	pMyNT-*whiB2*, Hygromycin-B^R	This work
PGS2543	pMyNT-*whiB3*, Hygromycin-B^R	This work
pGS2560	pET21a-*whiB1*, Amp^R	This work

pGS2561	pET21a-*whiB1*, with C9A mutation Amp^R	This work
pGS2563	pET21a-*whiB1*, with C40A mutation, Amp^R	This work
pGS2564	pET21a-*whiB1*, with C46A mutation, Amp^R	This work
pGS2565	pET28a-$sigA_{A1075\text{-}A1587}$, Kan^R	This work
pGS2566	pET28a-$sigA_{G1338\text{-}A1587}$, Kan^R	This work
pGS2567	pGS1669-with $sigA_{G1338\text{-}A1587}$, Amp^R	This work
pGS2568	pGS1672-with *whiB1,* Kan^R	This work
pGS2569	pET21a-*whiB1*, with D2R mutation, Amp^R	This work
pGS2570	pET21a-*whiB1*, with E12K mutation, Amp^R	This work
pGS2571	pET21a-*whiB1*, with D13R mutation, Amp^R	This work
pGS2572	pET21a-*whiB1*, with E15K mutation, Amp^R	This work
pGS2573	pET21a-*whiB1*, with D32R mutation, Amp^R	This work
pGS2574	pET21a-*whiB1*, with E45K mutation, Amp^R	This work
pGS2575	pET21a-*whiB1*, with D56R mutation, Amp^R	This work
pGS2576	pET21a-*whiB1*, with E65K mutation, Amp^R	This work
pGS2577	pET21a-*whiB1*, with D66R mutation, Amp^R	This work
pGS2578	pET21a-*whiB1*, with E67K mutation, Amp^R	This work
pGS2579	pGS1672-*whiB1,* with C9W and D11R Kan^R	This work

2.3 Culture media and growth conditions

2.3.1 Rich medium

Miller's Luria Broth (LB) medium was used to culture *E. coli.* The composition of this medium was:

Tryptone	10 g
Sodium chloride	10 g
Yeast extract	5 g
H_2O	To 1 litre

For LB agar, 15 g/l of bacteriological agar was added to the same components of LB.

2.3.2 *Escherichia coli* minimal medium

E. coli was grown in minimal medium for protein ^{15}N labelling prepared as descried by Cai et al., (1998) with some modifications

1.	KH_2PO_4	13.0 g
	K_2HPO_4	10.0 g
	Na_2HPO_4	9.0 g
	K_2SO_4	2.4 g
	$^{15}NH_4Cl$	1.0 g
	Deionized water	970 ml

The pH was adjusted to 7.4 before sterilization by autoclaving

2. trace elements "Sterilized by filtration" 10 ml

Trace elements	Per 100 ml
$FeSO_4$ ($7H_2O$)	0.60 g
$CaCl_2$ ($2H_2O$)	0.60 g
$MnCl_2$ ($4H_2O$)	0.12 g
$CoCl_2$ ($6H_2O$)	0.08 g
$ZnSO_4$ ($7H_2O$)	0.07 g
$CuCl_2$ ($2H_2O$)	0.03
H_3BO_3	2 mg

3. Thiamine	"Sterilized by filtration"	30 mg
4. 1M $MgCl_2$	"Sterilized by filtration"	10 ml
5. D-Glucose	"Sterilized by filtration"	2.0 g

2.3.3 The 7H9 complete medium

The 7H9 complete medium medium was used to grow *M. smegmatis* in primary culture. Medium was prepared according to the manufacturer's instructions.

Middlebrook 7H9 Medium powder (Difco, 271310)	4.7 g
Water	900 ml
ADS solution :	100 ml
2% glucose	
0.2% BSA	
0.8% NaCl	
Glycerol 50%	4 ml
Tween 80 20%	2.5 ml

The powder of Middlebrook 7H9 medium was dissolved in water and autoclaved first. The medium was left to cool to 45°C before adding the filter sterilized ADS solution, and glycerol or Tween.

Preparation of agar plates was done by adding 19 g/l of 7H10 Agar (Difco, 262710) instead of 7H9 medium, with the same additions.

2.3.4 The 7H9 expression medium

The 7H9 medium for protein expression in *M. smegmatis* was prepared in the same way as the 7H9 complete medium but without adding BSA.

2.3.5 *Mycobacterium smegmatis* minimal medium

Minimal medium of *M. smegmatis* was used to grow the bacteria for overexpression and ^{13}C and ^{15}N labelling of WhiB1 for NMR experiments. The recipe was modified from 7H9 medium as follows.

Glycerol-$^{13}C_3$	2.0 g	Sterilized by filtration
D-Glucose-$^{13}C_6$	2.0 g	
$^{15}NH_4Cl$ or $(^{15}NH_4)_2SO_4$	2.5 g	Sterilized by autoclave
Na_2HPO_4	2.5 g	
KH_2PO_4	1.0 g	
NaCl	0.1 g	
$MgSO_4$	0.005 g	
$CaCl_2$	0.0005 g	
$ZnSO_4$	0.001 g	
$CuSO_4$	0.001 g	
$Fe_2(SO_4)_3$	0.04 g	
Pyridoxine	0.001 g	
Thiamine	0.001 g	
Biotin	0.0005 g	
Deionized H_2O to a final volume of	1 l	

2.3.6 Growth media used for the bacterial BACTH two-hybrid system

2.3.6.1 MacConkey/maltose medium

MacConkey/maltose agar plates were prepared by dissolving 40 g/l of MacConkey agar (Difco 216830) in distilled water, before autoclaving and cooling to ~50°C. A stock solution of 20% maltose was sterilized by filtration. A 1% final concentration of maltose and the appropirate antibiotics were added to the prepared autoclaved MacConkey medium before pouring the plates. IPTG (final concentration 0.5 mM) was also added for inducing the expression of two hybrid proteins.

2.3.6.2 Maltose minimal medium/M63

A 5x of M63 minimal medium was prepared as following:

$(NH_4)_2SO_4$	10 g
KH_2PO_4	68 g
$FeCl_3$	0.5 mg
$FeSO_4.7H_2O$	2.5 mg
Vitamin B1	5 mg
Deionized H_2O	to a final volume of 1 l

The pH was adjusted to 7.0 with KOH and then autoclaved.

For preparing M63/maltose plates, 15 g of bacteriological agar was dissolved in 800 ml H_2O and autoclaved. Then 200 ml of autoclaved 5x M63 was added. The appropriate antibiotics, 0.2-0.4% maltose and 0.5 mM IPTG were also added before pouring the plates.

2.3.6.3 LB/X-gal medium

The LB agar medium was prepared as in Section 2.3.1, then the medium was left to cool to about 45°C, 40 µg/ml of the X-gal (Stock solution of X-gal was prepared 20 mg/ml in dimethyl formamide), 0.5 mM IPTG and appropriate antibiotics were added before pouring the plates.

2.3.7 Bacterial growth supplements

Media (broth and agar) was supplemented with appropriate antibiotics such as Kanamycin 35 µg/ml, Ampicillin 100 µg/ml, Streptomycin 100 µg/ml and Hygromycin B 100 µg/ml. To induce the expression of proteins 1.0 or 0.5 mM of IPTG was used. A 100 µg/l final concentration of ferric ammonium citrate or 0.5 mg/l of ferric chloride were also added to LB when needed.

2.3.8 Bacterial growth conditions

Primary cultures were grown from single colony or glycerol stocks for *E. coli* or *M. smegmatis*. The incubation times were 18 h for *E. coli*, 72 h for *M. smegmatis*.

Escherichia coli DH5α was cultured in LB medium at 37°C with shaking at 250 rpm; *E. coli* BL21 λ(DE3) was also cultured in LB medium at 37°C and 250 rpm until the OD_{600} of culture were reached to 0.65-0.75, then for protein expression the cultures were induced with 0.5-1 mM IPTG and incubated at 18°C for 12-18 h with shaking at 220 rpm. *Mycobacterium smegmatis* was cultured on 7H10 agar or in 7H9 broth medium at 37°C. Broth cultures were agitated at 220-250 rpm for protein expression. When the OD_{600} reached between 0.6 and 1.0, cultures were induced with 0.2% acetamide and incubated at 37°C for 12 h.

2.3.9 Bacterial growth measurement

Escherichia coli growth was estimated in broth media by reading the OD at 600 nm. For *M. smegmatis*, OD_{600} was also measured and the broth was supplemented with 2.0 ml/l glycerol and 0.5 ml/l Tween 80 to prevent the colonies clumping.

2.3.10 Bacterial strain storage

Glycerol stocks of bacterial strains were prepared for long-term storage as follows, cell pellets of 5 ml overnight cultures were suspended in 1.25 ml LB and 1 ml of 80% ($^{v}/_{v}$) glycerol with appropriate antibiotics. Then 1 ml was transferred into eppendorf tubes and stored at –70°C. Bacterial strains cultured on agar plates were also used for short-term storage.

2.3.11 Preparation of *E. coli* chemically competent cells

A single colony of a specific strain was inoculated in 10 ml of LB medium without antibiotics and incubated overnight at 37°C with shaking. A 1 ml aliquot from this culture was used to inoculate 100 ml of pre-warmed LB medium in a 250 ml flask, which was then incubated at 37°C until the OD_{600} reached to 0.5 (about 2 h). The culture was cooled in ice for 10 min, and the cells were

harvested at 4000 x*g*, 4°C for 5 min. Then, the supernatant was discarded and the cells were re-suspended gently in 30 ml of a chilled TFB1 buffer (see below) and kept on ice for 90 min. The cells were harvested again under the same conditions as above. The supernatant was discarded, and the cells were re-suspended in 4 ml of a chilled TFB2 buffer (see below). Aliquots of 200 µl were transferred into eppendorf tubes and stored at –70°C (QIAexpressionist, Qiagen).

TFB1 (pH 5.8)		**TFB2 (pH 6.8 with KOH)**	
RbCl	100 mM	MOPS	10 mM
$MnCl_2$	50 mM	RbCl	10 mM
Potassium acetate	30 mM	$CaCl_2$	75 mM
$CaCl_2$	10 mM	glycerol	15%
Glycerol	15%		

Buffers were sterilized by filtration.

2.3.12 Transformation of chemically competent cells

Frozen competent cells (Section 2.3.11) were thawed on ice. A 100 µl of the cell suspension was mixed with (0.1 ng-1 µg) pre-chilled DNA plasmid or ligation mixture, after that the mixture was kept on ice for 20 min. The cells were subjected to heat shock in a 42°C water bath for 90 sec and then chilled on ice for 2 min. Then 400 µl of LB was added to the cells and incubated for 1 h at 37°C with shaking. The culture was then harvested for 1 min at 4000 rpm. The supernatant was discarded and the pellet was re-suspended in 100 µl of LB and spread on LB-agar plates containing appropriate antibiotics and incubated at 37°C.

2.3.13 Preparation of electrocompetent cells

Primary cultures (5 ml) of *E. coli* and *M. smegmatis* were grown at 37°C with shaking (Section 2.3.8). Fifty ml of LB or complete 7H9 medium in 250 ml flasks was inoculated with 500 µl of overnight *E. coli* culture or 300 µl of *M. smegmatis*. For *E. coli* strains the new culture was incubated for 2-2.5 h at 37°C with shaking at 250 rpm, the OD_{600} reached to 0.6-0.65; while, for *M. smegmatis* the incubation time was overnight at 37°C and 220 rpm, the OD_{600}

usually reached between 0.6 and 1.0. The cultures were incubated on ice (for *E. coli* 30 min and 1-2 h for *M. smegmatis*), and then poured into 50 ml Falcon tubes, and then centrifuged at 2790 *xg* for 10 min at 4°C. The pellet was re-suspended with 50 ml of 10% ($^{v}/_{v}$) ice-cold glycerol and washed 3 times for 10 min each. After the last washing step, the bacterial pellet was re-suspended with 1000 µl of 10% ($^{v}/_{v}$) ice-cold glycerol. Aliquots (100 µl) of the cell suspensions were dispensed and flash-frozen in liquid nitrogen or dried ice before storage at -80°C.

2.3.14 Transformation of electrocompetent cells

Transformation was done by thawing aliquots of electrocompetent cells (prepared as in Section 2.3.13) on ice, and then plasmid DNA (~0.5 µg) was mixed with competent cells. After that, the mixture was transferred to pre-cooled electroporation cuvette on ice. The mixture was electroporated (one time at 1.8 kV for *E. coli* strains; and three times at 2.5 kV without in-between cooling for *M. smegmatis*) using a Hybaid Cell Shock unit. Immediately, 0.75 ml of LB or 7H9 complete medium was added and mixed with electroporated cells; the cell suspension was incubated at 37°C with shaking (1.0 h for *E. coli*, and 4.0 h for *M. smegmatis*). After incubation, the cells were then collected (1.0 min at 4000 rpm in a benchtop centrifuge). The supernatant was discarded and the pellet was re-suspended in 100 µl of broth medium and spread on LB (for *E. coli*) or 7H10 agar plates (for *M. smegmatis*) containing appropriate antibiotics and incubated at 37°C.

2.4 Nucleic acid methods

2.4.1 DNA storage

Genomic DNA and purified plasmids were stored in either deionized water or in EB buffer, (10 mM Tris-HCl, pH 8.5), at -20°C.

2.4.2 Plasmid purification

Qiagen QIAprep® Spin Miniprep Kit was used for purification of plasmid DNA from cultures of *E. coli* according to the manufacturer's instruction.

To extract plasmids from *M. smegmatis,* five to ten colonies were picked from 7H9 complete agar plates and washed with 10% ($^{v}/_{v}$) ice-cold glycerol, then mixed with electrocompetent *E. coli* DH5α and electroporated one time at 1.8 kV (Section 2.3.12). Plasmids were transferred from *M. smegmatis* to *E. coli*. After overnight incubation, *E. coli* colonies appeared but not *M. smegmatis.* A colony of *E. coli* was transferred to 5 ml LB broth with appropriate antibiotics and left to grow overnight, plasmids were extracted from cell pellets according to the manufacturer's instruction of Qiagen QIAprep® Spin Miniprep Kit.

2.4.3 Measurement of DNA concentration

The measurement of DNA concentration was done using the NanoPhotometer® P-300 Spectrophotometer with the double-stranded DNA program, dH_2O was used as a blank.

2.4.4 Agarose gel electrophoresis

For visualizing DNA fragments, agarose gel electrophoresis was done. Agarose (1% $^{w}/_{v}$) was dissolved in 1X TAE buffer by heating in a microwave oven. Before casting and after cooling, GelRed solution (Biotium) was added to the gel (1 µl/10 ml). DNA loading dye (5X), provided by Fermentas, was mixed with DNA samples, and then the samples were loaded on the gel. HyperLadder™ 1kb, supplied by Bioline, was used for estimation of DNA size. TAE buffer (1X) was used to electrophorese the agarose gels at 100 V for 1 h. Gels were viewed using a photodocumentation system.

10X TAE buffer

Tris	48.4 g
EDTA (0.5M, pH 8)	20 ml
Glacial acetic acid	11.42 ml
dH_2O	to 1 litre

2.4.5 DNA Gel extraction

DNA bands were cut from agarose gels, and extracted by using Qiagen QIAquick Gel Extraction Kit based on the manufacturer`s instruction.

2.4.6 Primer design

DNA primers were designed for cloning and site-directed mutagenesis. Primers were purchased from Eurofins MWG Synthesis.

2.4.7 Polymerase Chain Reaction (PCR)

The typical PCR contained:

DNA template (50-100 ng/µl)	2.0 µl
Forward primer (10 pmol/µl)	2.0 µl
Revierse primer (10 pmol/µl)	2.0 µl
PCR master mix (Thermo Scientific)	25 µl
dH_2O	19 µl

The PCR reaction profile was 30 cycles :

Initial denaturation	95°C - 5 min
Denaturation	95°C - 30 sec
Annealing	5 degree below the Tm°C of each primer - 30 sec
Extension	72°C - 1 min/1 kbp
Final extension	72°C - 10 min

Samples were separated and visualized by agarose gel electrophoresis.

2.4.8 DNA digestion by restriction endonuclease

A typical reaction was as follows, according to the New England BioLabs instructions.

Restriction enzyme	10 units
DNA	1 µg
10X NEBuffer	5 µl (1X)
Total reaction volume	50 µl
Incubation time	1 h, 37 °C

2.4.9 DNA ligation

Ligation reactions were according to the New England BioLabs instructions. T4 DNA Ligase was used in combination with 10X buffer. Ligation was usually performed in 20 µl reaction volumes as below, using a molar ratio of 1:3 vector to insert. The reaction was incubated at 16°C overnight. Next day the reaction was chilled on ice and used to transform into competent cells.

10X T4 DNA ligase buffer	2 µl
DNA	1:3 molar ratio vectors to insert
T4 DNA ligase	10 units
Nuclease-free water	to 20 µl

2.4.10 DNA purification

The Qiagen QIAquick PCR purification kit was used for DNA purification, according to the manufacturer`s instructions.

2.4.11 DNA sequencing

DNA sequencing was carried out by Core Genomic Facility in Medical School at the University of Sheffield using the appropriate oligonucleotide primers.

2.5 Protein methods

2.5.1 Protein concentration measurement

Estimation of protein concentration was achieved using Bradford reagent, supplied by Bio-Rad and according to the manufacturer`s (Bradford, 1976). Absorbance was measured at 595 nm using Biochrom WPA Lightwave II spectrophotometer. Protein concentration was also calculated using the Beer-Lambert law. A theoretical extinction coefficient factor was calculated using the ExPASy - ProtParam tool (http://web.expasy.org/protparam/), and the absorbance was measured at 280 nm by NanoPhotometer® P-300 Spectrophotometer.

2.5.2 Denaturing gel electrophoresis (SDS-PAGE)

The gel was prepared as described by (Laemmli, 1970). Both resolving and stacking gel were prepared as shown below. The glass plate sandwich was assembled using a holder, and then placed on a rack to pour the gel. The resolving gel was poured firstly between the glass plates, and ¼ of the space was left for stacking gel. This space was filled with 50% isopropanol. Once the gel had polymerized, the stacking gel was poured after discarding the isopropanol and washing twice with water. A comb was inserted and the gel was left for 20 min. Once the gel had polymerized, the comb was removed, and the gel was transferred to the electrophoresis tank filled with 1X SDS running buffer. Loading dye was prepared and mixed with samples at 1:5 ratio, boiled for 5 min at 100 °C and loaded separately in wells along with a protein ladder. Polypeptides were separated by electrophoresis for 1 h at 200 V. After that, the glass plates were removed and the gel was stained with Commassie blue stain solution for 1 h. Then, the stain solution was removed and the gel destained overnight. Precast gradient polyacrylamide gels (4-12%) (NuPAGE, provided by Invitrogen) were also used.

	15% Resolving gel (ml)	**5% Stacking gel (ml)**
dH_2O	2.7	6.95
30% ($^{w}/_{v}$) Acrylamide	3.8	1.7
3 M Tris-HCl (pH 8.3)	0.95	-
0.5 M Tris-HCl (pH 6.8)	-	1.25
10% ($^{w}/_{v}$) TEMED	0.02	0.02
10% ($^{w}/_{v}$) SDS	0.1	0.1

5X SDS loading dye		**1X SDS running buffer**	
Tris-HCl pH 6.8	0.225 M	Tris	3 g
Glycerol ($^{v}/_{v}$)	50%	SDS	1 g
SDS ($^{w}/_{v}$)	5%	Glycine	14.4 g
bromophenol blue ($^{w}/_{v}$)	0.05%	dH_2O to 1 litre	

2-Mercaptoethanol 0.25 M

	Commassie Blue stain	Destain
Acetic acid	100 ml	100 ml
Methanol	400 ml	400 ml
Coomassie Brilliant Blue	1.15 g/l	-
dH_2O to 1 litre		

2.5.3 Protein overproduction

2.5.3.1 Overproduction of WhiB1 in *E. coli*

The coding region of *whiB1,* which had been amplified from *M. tuberculosis* H37Rv using the primers Wb1F (5'-ATATATCATATGGATTGGCGCCACA AGGCG-3') and Wb1R (5'-ATATATGTCGACTCAGACCCCGGTACGGGC TTT-3'), was digested and ligated (Section 2.4.8, and 2.4.9) into pET28a (to engineer WhiB1 with *N*-terminal His_6-Thrombin-tag), and into pET21a (to engineer a non-tagged WhiB1) between NdeI and SalI sites respectively, which are underlined in the above primers. The integrity of insert and the success of ligation were confirmed by DNA sequencing (Section 2.4.11). The plasmids pGS2500 and pGS2560 were purified (Section 2.4.2**)** and transformed separately (Section 2.3.12) into *E. coli* JRG5302 (Table 2.1). Cultures for overproduction were grown in a 2 litre flask containing 500 ml of LB broth and kanamycin or ampicillin at 37°C with shaking until the OD_{600} reached 0.6-0.7 (Section 2.3.7), cultures then were induced with 1 mM IPTG and incubated at 18°C for 12-18 h with shaking 220 rpm. The cell pellets were harvested at room temperature by centrifugation at 17696 *xg* for 20 min and stored at -20°C until used for purification steps.

2.5.3.2 Overproduction of WhiB1 in *M. smegmatis*

The *whiB1* was amplified from *Mtb* H37Rv genomic DNA using the primers MSWb1F (5'-TATATACCATGGATTGGCGCCACAAGGCG-3') and MSWb1R (5'-TATATAAAGCTTTCAGACCCCGGTACGGGCTTT-3'), digested and ligated (Section 2.4.8, and 2.4.9) into pMyNT between NcoI and HindIII sites

respectively, which are underlined in the above primers. The integrity of insert and the success of ligation were confirmed by DNA sequencing (Section 2.4.11). The plasmid pGS2524 was purified (Section 2.4.2**)** and transformed (Section 2.3.14) into *M. smegmatis* MC^2155 JRG6798 (Table 2.1). Cultures for overproduction were grown in a 2 litre flask containing 500 ml of 7H9 broth (Section 2.3.3) and hygromycin B at 37°C with shaking until the OD_{600} reached 0.6-1.0 (Section 2.3.8), cultures then were induced with 0.2% acetamide and incubated at 37°C for 12 h with shaking 200 rpm, to induce the expression of WhiB1 with *N*-terminal His_6-TEV-tag. The cells were harvested at room temperature by centrifugation at 17696 *xg* for 30 min and stored at -20°C until used for cell lysate preparation.

2.5.3.3 Overproduction of SigA in *E. coli*

SigA was typically overproduced after transforming *E. coli* JRG5302 with pGS2183, which contains the coding region of *Mtb sigA* ligated into pET28a between NdeI and HindIII restriction sites (Section 2.3.12). SigA overproduction was done by culturing the host strain in a 2 litre flask containing LB broth and kanamycin at 37°C with shaking 250 rpm until the OD reached to 0.6-0.7. IPTG (1 mM) was added in order to induce the overexpression of SigA with *N*-terminal Thrombin tag, which was done at 18°C with 220 rpm for 12-18 h (Section 2.3.8). Cells were harvested by centrifugation at 17696 *xg* in room temperature for 20 min, and kept at -20°C until used for cell lysate preparation.

2.5.3.4 Overproduction of the complex WhiB1 and SigA

Co-expression of SigA and WhiB1 in *E. coli* JRG5302 was achieved by transformation with pGS2560 (Section 2.3.12). Then, electrocompetent cells of the transformed strain were prepared (Section 2.3.13), this was done by growing the strain in LB broth in the presence of ampicillin at 37°C with 250 rpm for 2-2.5 h (usually the OD reached 0.6-0.7), then the freshly prepared electrocompetent cells were transformed again with pGS2183 (Section 2.3.14). LB broth (0.75 ml) was added and the cells were incubated for 1 h at 37°C with shaking at 250 rpm. Cells were plated on LB agar containing ampicillin and kanamycin. Overproduction of the complex was done by culturing the host

strain *E. coli* JRG6859 carrying both pGS2560 and pGS2183 in a 2 litre flask containing 500 ml of LB broth and both ampicillin and kanamycin at 37°C with shaking at 250 rpm until the OD reached to 0.6-0.7. To induce both non-tagged WhiB1 and SigA with *N*-terminal His tag, IPTG was added to final concentration 1 mM, and the culture incubated at 18°C with shaking at 220 rpm for 12-18 h. Cells were harvested by centrifugation at 17696 *xg* at room temperature for 20 min, and kept at -20°C until used for cell lysate preparation.

2.5.3.5 Overproduction of the WhiB1-SigA$_{M359-D528}$, and WhiB1-SigA$_{A447-D528}$ complexes

Overproduction of the complex of untagged WhiB1 and untagged SigA$_{M359-D528}$ or untagged WhiB1 and His$_6$-tagged SigA$_{A447-D528}$ was achieved by amplifying the corresponding coding regions of *sigA* from pGS2183. For SigA$_{M359-D528}$ (SigA$_{C170}$) production, SigAc170F: (5'-ATATATAAGCTTCCGACCAGGCCC GCACCATCCG-3') and SigACTDR: (5'-ATATATGTCGACTCAGTCCAGGT AGTCGCGCAGG-3') were used, the amplified product was digested and ligated (Sections 2.4.8, and 2.4.9) into pET28a between NcoI and HindIII sites, which are underlined in the above primers. For SigA$_{A447-D528}$ (SigA$_{C82}$), SigAc82F (5'-ATATATCATATGGCCGTCGACGCGGTGTCCTTCA-3') and SigACTDR were used; the amplified product was also digested and ligated (Section 2.4.8, and 2.4.9) into pET28a between NdeI and HindIII sites, which are underlined in the above primers. The integrity of inserts and the success of ligation was confirmed by DNA sequencing (Section 2.4.11). The resulting plasmids pGS2565 and pGS2566 were purified (Section 2.4.2**)** and co-transformed individually with pGS2560 into JRG5302 *E. coli BL21* λ(DE3) creating JRG6858 and JRG6857 respectively as mentioned in Section 2.3.14, and Table 2.1.

Overproduction of the proteins was achieved by culturing the host strain carrying (pGS2560 and pGS2565) or (pGS2560 and pGS2566) in a 2 litre flask containing LB broth and both ampicillin and kanamycin at 37°C with shaking at 250 rpm until the OD reached to 0.6-0.7. To induce both non-tagged WhiB1 and non-tagged SigA$_{C172}$ complex or the complex of non-tagged WhiB1 and SigA$_{C82}$ with *N*-terminal His tag at the same time, IPTG was added to final

concentration 1 mM, and the culture was incubated at 18°C with shaking at 220 rpm for 12-18 h. Cells were harvested by centrifugation at 17696 *xg* at room temperature for 20 min, and kept at -20°C until needed.

2.5.4 Production of cell-free extracts

Frozen bacterial cells were suspended in 500 mM NaCl, 50 mM Tris pH 8.0 (or 25 mM NaH_2PO_4, pH 7.4), and then disrupted either by sonication 3 cycles each 20 sec for *E. coli,* or 9 cycles each 25 sec for *M. smegmatis.* Cells were sonicated on ice using the maximum amplitude value by Soniprep 150 Plus sonicator. Then, the cell lysate was separated from cell debris by centrifugation at 45,995 *xg* for 20 min at 4°C using Avanti J-251 centrifuge. The supernatant was used for purification immediately.

2.5.5 Protein purification

2.5.5.1 Purification by affinity chromatography

WhiB1, SigA, the complex of WhiB1-SigA, and the complex of WhiB1- $SigA_{A447-D528}$ (WhiB1:$SigA_{C82}$), which all had an *N*-terminal His_6-tag, were purified using 5 ml HiTrap™ chelating or HisTrap™ columns (GE Healthcare). The purification was carried out using the AKTA Pure machine; 50 ml of cell lysate was injected in to the loading loop, and the program of purification was set according to the instructions of the manufacturer. The purification buffers were prepared as shown below. The eluted fractions were collected and the protein concentrations were measured. Samples from each fraction were analyzed by SDS-PAGE to determine the purity of each fraction. The protein samples were stored at -20°C.

	Binding buffer A (pH 8.0)	Elution buffer B (pH 8.0)
Tris	50 mM	50 mM
NaCl	0.5 M	0.5 M
Imidazole	-	0.5 M
dH_2O to 1 litre		

2.5.5.2 Purification of untagged WhiB1-SigA$_{M359-D528}$ complex

Purification of the untagged complex of WhiB1-SigA$_{M359-D528}$ was carried out via three different steps. Anion exchange chromatography was the first step; this was done by using 5 ml HiTrap^{TM} Q HP column (GE Healthcare). Cell-free extracts were prepared in binding buffer (50 mM MES pH 6.5, 0.1 M NaCl) (Section 2.5.4), then 50 ml of cell lysate was applied to the column, and the program of purification was set according to the instructions of manufacturer. Elution buffer (50 mM MES pH 6.5, 1 M NaCl) was used to elute the protein. After identifying fractions containing the WhiB1-SigA$_{M359-D528}$ complex by SDS-PAGE, the protein was concentrated to 2 ml by Vivaspin column (Mwt cut-off 10,000) (GE Healthcare). The second step of purification was carried out via Superdex 200 gel filtration in 50 mM Tris pH 8.0 containing 0.5 M NaCl. SDS-PAGE was again used to identify the fractions which contained the WhiB1-SigA$_{M359-D528}$ complex. Complex containing fractions were diluted 5 times with 50 mM Tris pH 8.0 without salt, this was done to reduce the salt concentration of the protein to allow binding to heparin column beads in the third step of purification. Binding buffer was 50 mM Tris pH 8.0 containing 0.1 M NaCl, while the buffer of elution contained 1 M NaCl. The purity of the eluted protein was examined by SDS-PAGE and the protein sample was stored at -20°C until needed.

2.5.6 Estimation of protein molecular mass by size-exclusion chromatography

HiLoad 16/600 Superdex 200 column was used to estimate the molecular mass and determine the oligomeric state of proteins. The column was equilibrated with 50 mM Tris pH 8.0 containing 0.5 mM NaCl, and the program

of purification was set according to the instructions of the manufacturer. Protein molecular mass was calculated by comparing the data with a standard curve of molecules of known molecular weight, which were Blue dextran, Thyroglobulin, Conalbumin, Carbonic anhydrase, and Aprotinin, Ferritin1, Aldolase1, Ovalbumin, and Ribonuclease A (RNase A).

2.5.7 Estimation of WhiB1 sulfhydryl group reactivity

Cysteine alkylation by iodoacetamide (IAA) coupled with LC-MS mass spectrometry analysis has been used to determine the redox state of cysteine residues in apo-WhiB1 after losing the cluster. Reaction of IAA with a cysteine-thiol causes an increase in molecular mass of 57 Da. Reducing the WhiB1 (~100 μM) first with DTT followed by a reaction with IAA was used as a control. DTT was added to final concentration 5 mM with incubation for 30 min at room temperature, while IAA was added to final concentration 10 mM with incubation at room temperature for 1 h in the dark. Removing excess IAA was done by dialysis against 50 mM Tris pH 8.0 containing 0.5 M NaCl overnight.

2.5.8 Scanning UV Visible spectroscopy

The absorption spectra of proteins were obtained using Cary 50 Bio UV-visible with Hellma® quartz cuvette (10 mm), which had a screw cap lid to maintain anaerobic conditions. Scanning was between wavelengths 200-800 nm. Protein buffers were used as a blank to calibrate the instrument. Injection of reagents was achieved by using Hamilton gastight syringe through the Hellma® silicon seal of the lid.

2.5.9 Circular dichroism (CD) spectroscopy

Jasco J810 spectropolarimeter was used to characterise the secondary structure of WhiB1. Protein was prepared (30 μM) in 20 mM NaH_2PO_4, pH 7.4 containing 0.1 M NaCl. Holo-WhiB1 was kept anaerobic by closing the screw cap lid of the Hellma® quartz cuvette.

2.5.10 Liquid chromatography mass spectrometry (LC-MS)

Liquid chromatography mass spectrometry (LC-MS) was carried out by Mr. Simon Thorpe (The University of Sheffield Faculty of Science Mass Spectrometry Centre), using Agilent 1260 Infinity (LC instrument) and Agilent 6530 Q-ToF (MS instrument). The protein sample were diluted to 3 μM final concentration with the aqueous solution 0.1% ($^{v}/_{v}$) acetonitrile, 1% ($^{v}/_{v}$) formic acid and loaded onto Phenomenex Aeris Widepore 3.6u XB-C18 (50 mm x 2.1 mm) column. Elution of bound proteins was achieved by applying a linear gradient from 5% to 95% ($^{v/}_{v}$) acetonitrile, 0.1% ($^{v/}_{v}$) formic acid at a flow rate of (0.4 ml/min) for 10 min. The eluent was continuously introduced into the mass spectrometer using a positive mode electrospray ionization (ESI). For processing the LC peak spectra, compass data analysis using a maximum entropy v1.3 was used. Mass spectrometry analysis was used to verify the molecular weight of the purified proteins.

2.5.11 Electrospray ionization mass spectrometry (ESI-MS)

Electrospray ionization mass spectrometry (ESI-MS) was carried out by Dr. Jason Crack, University of East Anglia. The buffer of WhiB1 was exchanged using Zeba spin desalting columns (Thermo scientific) pre-equilibrated with 250 mM ammonium acetate pH 8.0. WhiB1 sample (~20 μM cluster) was directly injected (0.3 ml/h) into the ESI source of a Bruker micrOTOF-QIII mass spectrometer (Bruker Daltonics, Coventry, UK) in the positive ion mode (Agilent Technologies, California, USA). Full mass spectra (m/z 500-3500) were recorded for 5 min with parameters as follows: dry gas flow 4 L/min, nebuliser gas pressure 0.8 Bar, dry gas 180°C, capillary voltage 2750 to 4500 V, offset 500 V, ion energy 5 eV, collision RF 200 Vpp, collision cell energy 10 eV. Where necessary the isCID was increased from 0 to 140 eV, as required.

2.5.12 Site-directed mutagenesis (SDM)

QuikChange® II XL Site-Directed Mutagenesis Kit (Stratagene) was carried out for site-directed mutagenesis according to the manufacturer's instructions.

2.6 Iron- sulfur methods

2.6.1 Determination of the concentration of Holo-WhiB1

In order to determine the concentration of holo-WhiB1, Bradford assay and measurement by UV light at 280 nm both were used as described in Section (2.5.1). For Bradford measurements three repeats with different volumes of fully reconstituted [4Fe-4S] WhiB1 were measured and the average was taken. The protein was then treated with 10 mM EDTA in the presence of 1 mM TCEP-HCl. The protein was left in room temperature for 21 h. UV-visible spectrum confirmed the removal of the iron-sulfur cluster. Serial dilutions of this protein were measured at absorbance 280 nm. The protein concentration was calculated using the extinction coefficient factor of WhiB1 (Section 2.5.1). The correction factor between the two methods was calculated, and by using this method the concentration of holo-WhiB1 was measured by the Bradford method and adjusted using the correction factor.

2.6.2 WhiB1 iron-sulfur cluster reconstitution reaction

WhiB1 was purified as mentioned in (Section 2.5.5). The UV-visible spectrum was taken for all purified fractions to determine which fractions contained protein with a high concentration of clusters. The protein was reduced immediately after purification with 1 mM DTT and used directly for reconstitution. All steps of reconstitution were carried out in anaerobic cabinet. All buffers and solutions were left in the cabinet for at least 24 h before use. The reaction was set up using two different methods.

For the method using sodium sulfide, protein was dialyzed firstly against the reconstitution buffer: 25 mM NaH_2PO_4, pH 7.4 containing 0.25 M NaCl, 10 mM DTT, 10-fold molar excess of $FeCl_3$ and Na_2S for 12 h at 25 °C.

The second method used the NifS protein and the reaction was set up using the following recipe:

Purified-WhiB1	1X
Ammonium ferrous sulphate	10 X protein concentration
L-cysteine	20 X protein concentration
DTT	10 mM
NifS (Section 2.6.2.1)	10 µg for each 200 µM of WhiB1

The reaction was left overnight in the anaerobic cabinet. Next day the reconstituted-WhiB1 was dialyzed against anaerobic buffer: 25 mM NaH_2PO_4 (pH 7.4) with 0.25 M NaCl for at least 24 h within the cabinet to remove non-integrated materials (Section 2.6.2.2). UV-visible spectroscopic analysis was carried out after each reconstitution reaction to determine the concentration of incorporated [4Fe-4S] cluster.

2.6.2.1 Expression and purification of the desulfurase enzyme (NifS)

NifS protein from *Azotobacter vinelandii* has a cysteine desulfurase activity. As a result, it was used to provide sulfide from L-cysteine for the *in vitro* reconstitution of WhiB1 iron-sulfur cluster. The plasmid pDB55, which encodes NifS, was provided by Dennis R. Dean (University of Virginia). Using the expression protocol reported by (Zheng et al., 1993), the protein was expressed in a low concentration. Changing the inducer by using 1 mM IPTG instead of 1% ($^{w}/_{v}$) lactose and induction for 8 h instead of 2 h, the protein expressed very well in the soluble fraction at 25°C.

NifS protein was firstly purified by anion exchange chromatography using 5 ml HiTrap™ Q HP column (GE Healthcare), the protein was eluted between 0.35 M and 0.5 M NaCl in 50 mM Tris pH 8.0. The purified fractions were collected and concentrated to 2 ml, before size-exclusion chromatography on Superdex 200 gel filtration column equilibrated with 50 mM Tris pH 8.0 containing 0.5 M NaCl. The protein eluted as a monomer, was aliquoted and stored at -80 °C

2.6.2.2 Purifying holo-WhiB1 after reconstitution

In order to remove the contaminants from reconstitution reactions such as the excess amounts of DTT, L-cysteine and iron salts (also removing the

imidazole), therefore allowing the use of holo-WhiB1 in further assays, the protein was dialyzed against 25 mM NaH_2PO_4 buffer (pH 7.4) containing 0.25 M NaCl. Dialysis was done within the anaerobic cabinet for 21 h at 25 °C. No precipitation of protein occurred during dialysis of the reconstituted protein. In contrast dialysis against 50 mM Tris buffer (pH 8.0) with 0.5 M of NaCl caused high precipitation of protein.

2.6.3 Estimation of the iron content of reconstituted-WhiB1

After protein reconstitution and clean up by dialysis, iron assays were done to measure the total amount of iron in WhiB1. Iron standard solution (BDH), final concentration of iron 18 mM, was used to create the standard curve. Holo-WhiB1 (100 µl) was mixed with 380 µl of 1% ($^{w}/_{v}$) final concentration of trichloricacetic acid (TCA) and boiled for 5 minutes. The sample was centrifuged at maximum speed in a bench-top centrifuge for 2 min. The supernatant was taken and mixed with 400 µl of saturated sodium acetate, 30 µl of 10% ($^{w}/_{v}$) of Bathophenothroline sulphonic acid and 90 µl of 20% ($^{w}/_{v}$) of freshly prepared sodium ascorbate. After the addition of sodium ascorbate, the sample colour changed to red. The sample was centrifuged for 5 minutes and the absorbance measured at 535 nm. The reading of absorbance was compared with the standard curve and the iron concentration was calculated. Following that, the concentration of cluster and the extinction coefficient factor were calculated.

2.6.4 Oxygen sensitivity of the WhiB1 iron-sulfur cluster

In order to examine the ability of the WhiB1 iron-sulfur cluster to react with O_2, the protein was reconstituted, and purified by dialysis (Sections 2.6.2 and 2.6.2.2). The UV-visible spectrum of holo-WhiB1 was recorded as mentioned in Section 2.5.8. Oxygen-saturated buffer (50 mM Tris, pH 7.4 containing 0.5 M NaCl, 5% ($^{v}/_{v}$) glycerol) was added to a final concentration of 110 µM, the UV-visible spectrum then was recorded at several time points. This assay was carried out with and without presence of reducing agents. Oxygen was also applied directly to the sample by transfer the holo-protein to a cuvette outside

the anaerobic cabinet and the lid of the cuvette was left open, and by bubbling the protein sample with air three times via a 100 µl Hamilton gastight syringe.

2.6.5 Nitric oxide sensitivity of the WhiB1 iron-sulfur cluster

The NO donor Proli NONOate (Cayman) was used to investigate the sensitivity of the iron-sulfur cluster of WhiB1 to NO. The Proli NONOate half-life is 1.8 seconds in 0.1 M phosphate buffer (pH 7.4).

Proli NONOate was prepared by dissolving the powder in 10 mM sodium hydroxide. The concentration was determined at 252 nm using the extinction coefficient factor of 8,400 M^{-1} cm^{-1}. In a pH-dependent manner Proli NONOate dissociates to NO and L-Proline; alkaline solutions of Proli NONOate were firstly mixed with buffer of assay in a Hamilton gastight syringe to initiate the release of NO and to reduce the effect of the basic pH. The solution then was directly injected to the protein sample through the Hellma® silicon seal of Hellma® quartz cuvette lid. All experiments were done under anaerobic conditions.

2.7 Protein-protein interaction

The interaction of WhiB1 and SigA was carried out *in vitro* and *in vivo* in *E. coli*, to determine the specificity of interaction.

2.7.1 *In vitro* interaction investigation of WhiB1 and SigA

The *in vitro* interaction between WhiB1 and SigA or SigA domains was achieved by co-expression of both proteins at the same time from two different recombinant plasmids co-transformed in the same *E. coli* cell (strain JRG5302, Table 2.1), and under the control of T7 promoter and IPTG induction (Sections 2.5.3.4 and 2.5.3.5). Pull-down assays were used to purify the complexes of (WhiB1-SigA) or (WhiB1-$SigA_{A447\text{-}D528}$) via affinity chromatography first and then by size-exclusion chromatography. Co-purification was done by binding the bait protein, the SigA or $SigA_{A447\text{-}D528,}$ which were engineered with *N*-terminal His_6 tags, on HiTrap™ chelating columns; the immobilized SigA binds the WhiB1,

which lacked a tag. Size-exclusion chromatography was carried out after that to confirm the specificity of interaction and to determine the size of the complexes (Section 2.5.5.1). WhiB1-SigA$_{M359-D528}$ complex was purified by ion-exchange chromatography and then by size-exclusion chromatography, because both proteins lacked tags (Section 2.5.5.2).

2.7.2 *In vivo* interaction investigation of WhiB1 and SigA$_{A447-D528}$

To detect the *in vivo* interaction between WhiB1 and SigA, Bacterial Adenylate Cyclase-based Two-Hybrid (BACTH) system was used. The genes encoding WhiB1 and SigA$_{A447-D528}$ were amplified via PCR. For WhiB1, pKT25-WB1 F: (5'-TATATATCTAGAGATGGATTGGCGCCACAAGGCG-3'), and pKT25-WB1 R: (5'-ATATATGGTACCTCAGACCCCGGTACGGGCTTT-3') primers were used; the amplified PCR product was then digested and ligated (Sections 2.4.8, and 2.4.9) into pGS1672 plasmid between XbaI and KpnI sites, which are underlined in the above primers, producing the plasmid construct pGS2568. While, for SigA$_{A447-D528,}$ pUT18-SigA F: (5'-ATATATAAGCTTGA TGGCCGTCGACGCGGTGTCC-3'), pUT18-SigA R: (5'-ATATATGGTACCC GGTCCAGGTAGTCGCGCAGGAC-3') were used; the amplified PCR product was then digested and ligated (Sections 2.4.8, and 2.4.9) into pGS1669 plasmid between HindIII and KpnI sites, which are underlined in the above primers, producing the plasmid pGS2567. As a result, WhiB1 and SigA were sub-cloned into pKT25 and pUT18 vectors respectively as in-frame fusions at the *C*-terminal end of T25, and the *N*-terminal end of T18. The recombinant plasmids were propagated in *E. coli* K12 (XL1-Blue). Plasmid DNA was purified by Qiagen miniprep kit (Section 2.4.2). The plasmids pGS2568 and pGS2567 encoding the T25-WhiB1 and SigA $_{A447-D528}$-T18 hybrid proteins were co-transformed (Section 2.3.14) into electrocompetent *E. coli* JRG4968 cells (Table 2.1). Transformed cells were plated on selective medium or indicator medium plates (Section 2.3.6), and incubated at 30°C for 24 to 36 h.

Interaction of WhiB1 and SigA$_{C82}$ was measured in aerobic cultures of JRG6862 grown in 5 ml LB broth containing ampicillin (200 μg/ml), kanamycin (25 μg/ml) and streptomycin (25 μg/ml) in 25 ml Sterilin tubes at 37°C for 16 h with rocking; where indicated, cultures were supplemented with dipyridyl (0.5

mM final concentration). Anaerobic cultures were grown in LKB MES (tryptone, 10 g/l; yeast extract, 5 g/l; KCl, 6.4 g/l buffered with 100 mM MES (2-(*N*-morpholino) ethanesulfonic acid)) pH 5.5 supplemented with IPTG (100 μg/ml) and either fumarate or nitrite (1.4 mM) as indicated. Growth under the latter conditions exposes the culture to NO, generated chemically from acidified nitrite (Tiso and Schechter, 2015). Aerobic and anaerobic control cultures, consisting of *E. coli* BTH101 transformed with pKT25 (pGS1672) and pUT18 (pGS1669) (JRG5386), or with pKT25-zip (pGS1673) and pUT18c-zip (pGS1671) (JRG5387) were also analyzed. The former strain contains only the BACTH vectors and hence acts as a negative control; the latter strain expresses fusions to the leucine zipper region of GCN4 thereby acting as a positive control (Karimova et al., 1998). β-Galactosidase activities were measured according to the method of (Miller, 1972). A minimum of three independent cultures were analyzed for each condition.

Chapter three

Isolation and biochemical characterization of *M. tuberculosis* WhiB1

3.0 Isolation and biochemical characterization of *M. tuberculosis* WhiB1

3.1 Introduction

The Wbl (WhiB-like) family of proteins are only found in actinomycetes (Soliveri et al., 2000). Wbl-proteins are small in size, varying in length from 87 to 130 amino acid residues, and possess four conserved cysteine residues within the *N*-terminal region and a predicted DNA-binding motif towards the *C*-terminal (Soliveri et al., 2000; den Hengst and Buttner, 2008). The cysteine residues are essential for anchoring an iron-sulfur cluster (Jakimowicz et al., 2005; Raghunand and Bishai, 2006a; Smith et al., 2012).

Any of the seven Wbl proteins of *M. tuberculosis* isolated after aerobic overproduction in *E. coli* exhibited a mixture of [2Fe-2S]-, [4Fe-4S]- and apo-Wbl forms. In contrast, after anaerobic *in vitro* reconstitution, only a [4Fe-4S] cluster was bound (Walters et al., 2006; Alam et al., 2009; Smith et al., 2010; Chawla et al., 2012; Burian et al., 2013). The [4Fe-4S]- forms react very quickly with NO but not with O_2 (Chawla et al., 2012; Singh et al., 2007; Smith et al., 2010). Similar observations were made with *Streptomyces coelicolor* WhiD1 (Crack et al., 2010). This suggests a possible function as NO sensors, and indeed the [4Fe-4S]-WhiB1 was found to bind with DNA only when it had reacted with NO (Smith et al., 2010). The variation of iron-sulfur cluster forms of recombinant Wbl proteins expressed in *E. coli* raises questions about the true nature of the proteins' cluster in mycobacteria. All the previous studies of Wbl proteins were carried out using *E. coli* expression systems for the protein overproduction. Here, *M. tuberculosis* WhiB1 was expressed in *M. smegmatis* and the biophysical and the biochemical characteristics of the isolated protein was compared to that purified from *E. coli*.

3.2 Overproduction, purification, biochemical and biophysical characterization of recombinant WhiB1 produced in *E. coli*

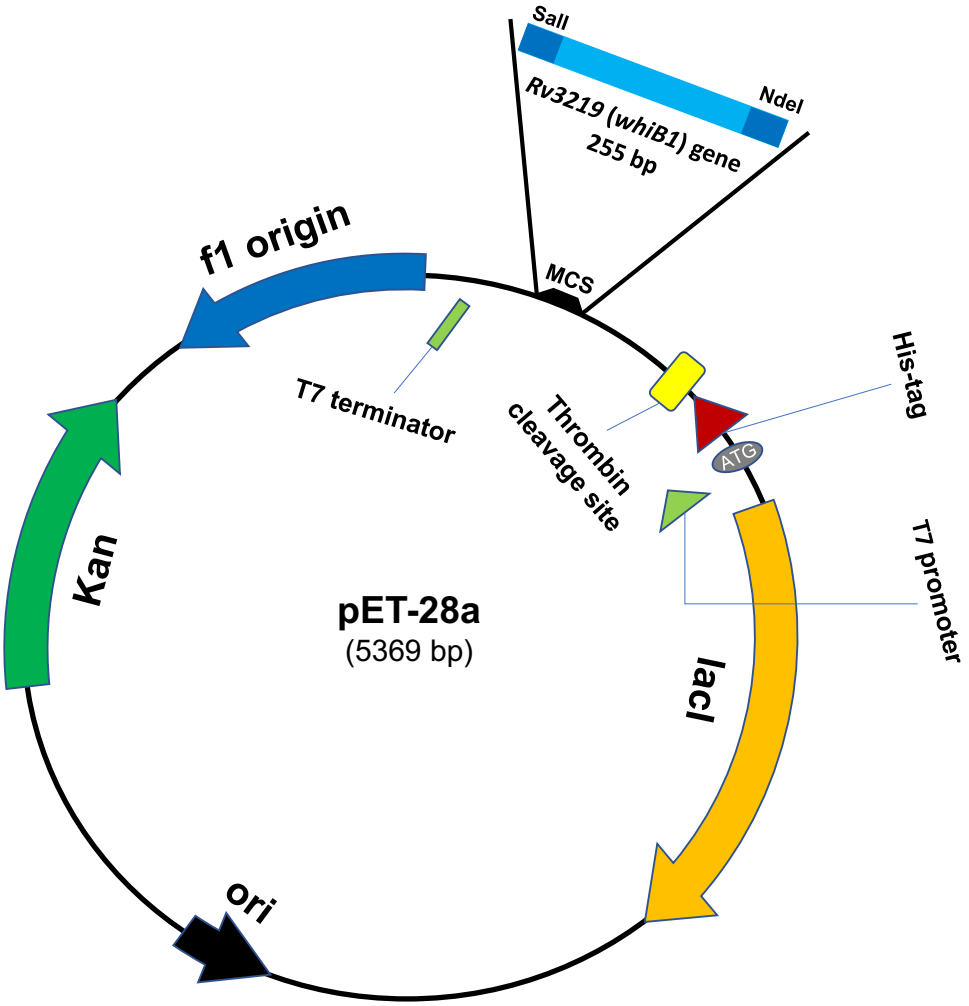
3.2.1 WhiB1^{Ec} overproduction and isolation

The WhiB1^{Ec} overproduction strain *E. coli* JRG6769 was created by transforming chemically competent *E. coli* JRG5302 with pGS250, encoding WhiB1 with an *N*-terminal His_6-tag and a linker containing a Thrombin cleavage site (His_6-ThCS-WhiB1^{Ec}) (Figure 3.1).

To find the optimum WhiB1^{Ec} overproduction conditions, cultures (50 ml) of *E. coli* JRG6769 were induced with a final concentration of 1 mM IPTG and incubated at 25°C or 18°C for 6 h, 12 h or 18 h (Section 2.5.3.1). Cell samples from each condition were taken for analysis by SDS-PAGE. All conditions resulted species at ~11 kDa corresponding to the WhiB1^{Ec} protein. However, the best condition of WhiB1^{Ec} overproduction was at 18°C for 18 h, which was chosen for WhiB1^{Ec} overproduction at large scale (1-4 L of LB broth culture).

For WhiB1^{Ec} isolation, cell-free extract was prepared by re-suspending cell pellets in binding buffer consisting of 500 mM NaCl, 50 mM Tris pH 8.0, and then sonicated according to Section 2.5.4. Subsequent protein purification was as described in Section 2.5.5.1. The elution profile from the HiTrap chelating column showed three peaks, the first peak after the imidazole gradient was formed from non-specifically bound materials, usually representing proteins with high histidine contents, while the second and the third peaks contained WhiB1^{Ec} protein (Figure 3.2A). Fractions of elution profile were all analysed by SDS-PAGE which confirmed the presence of WhiB1^{Ec} (Figure 3.2B).

A



B

MGSSHHHHHHSSGLVPR↓GSH- WhiB1

His_6-tag Thrombin cleavage site

Figure 3.1: Schematic diagram of pGS2500. [A] The *whiB1* open reading frame (255 bp) was amplified by PCR from genomic DNA of *M. tuberculosis* with engineered NdeI/SalI restriction sites, and ligated into pET28a to create pGS2500. **[B]** The plasmid encodes a His_6-ThCS-$WhiB1^{Ec}$ fusion protein. The arrow marks the peptide bond susceptible to cleavage by Thrombin.

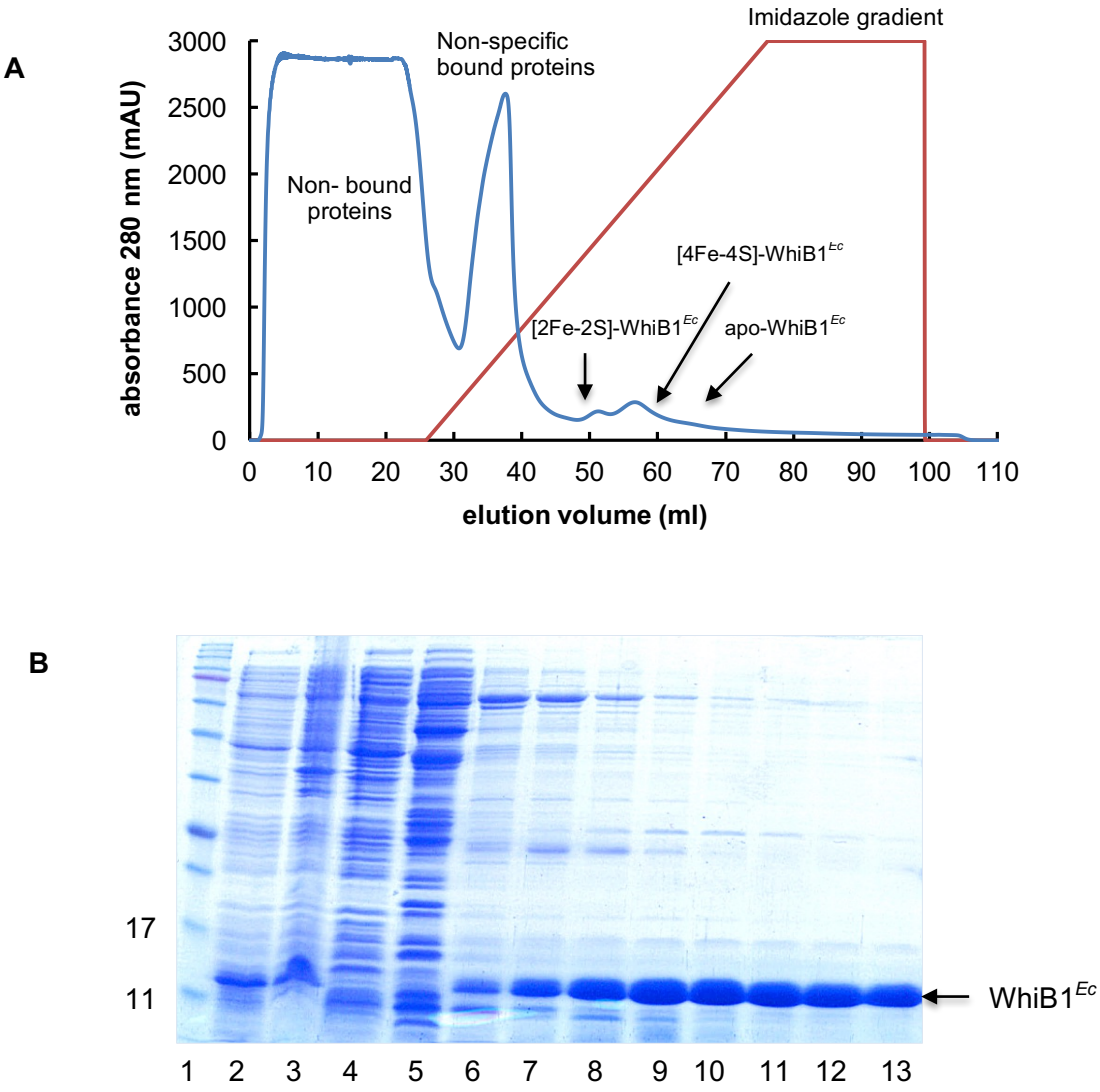


Figure 3.2: [A] Elution profile of HiTrap chelating column of *E. coli* cell lysate containing overproduced WhiB1^{Ec}. The trace shows protein absorbance at 280 nm (blue line) of injected cell lysate, and the gradient concentration of imidazole from 0-0.5 M (red line). Eluted WhiB1^{Ec} appears as two peaks between ~50-60 ml, the first peak represents the [2Fe-2S]-form, while the second represents WhiB1^{Ec} with [4Fe-4S] cluster form; apo-WhiB1^{Ec} is eluted at the position between 60-70 ml (Section 3.2.2). **[B] Coomassie stained SDS-PAGE gel of isolated fraction in the elution profile of [A].** Lane 1, protein ladder (sizes in kDa are indicated); lane 2 and 3, cell-free extract and cell debris respectively of induced WhiB1^{Ec}; fractions of non-bound proteins spread between 0-25 ml, all mixed together and loaded in lane 4; lane 5, non-specific bound proteins, usually contain high poly Histidine content, eluted at 35 ml; lanes 6-13, fractions of WhiB1^{Ec} eluted between 48-70 ml, as indicated in **[A]**.

3.2.2 UV-visible spectroscopic characterization of purified WhiB1^{Ec}

Fractions from HiTrap chelating chromatography containing WhiB1^{Ec} protein were coloured (Figure 3.2A and B). WhiB1^{Ec} containing fractions at the lower imidazole concentration were yellow, those at higher imidazole concentrations were brown. At still higher imidazole concentration, colourless fractions were associated with apo-WhiB1$^{Ec.}$

The UV-visible spectrum of WhiB1^{Ec} eluting first showed absorbance at 325, 420, 465, and between ~550-600 nm (Figure 3.3, black line). This absorption spectrum is usually representative of [2Fe-2S] cluster-containing proteins (Khoroshilova et al., 1997). In contrast, the spectrum of WhiB1^{Ec} eluting second showed only a single feature at 420 nm (Figure 3.3, red line), which is a typical of proteins containing [4Fe-4S] clusters (Crack et al., 2009). The colourless WhiB1^{Ec} sample did not show any detectable feature except the one at 280 nm, suggesting the sample did not contain an iron-sulfur cluster, i.e. an apo-form (Figure 3.3, blue line). This result indicated the existence of three different forms of WhiB1^{Ec} when the protein was overproduced in *E. coli* and purified under aerobic conditions. Both cluster forms of WhiB1^{Ec} degraded upon prolonged exposure to air. Thus, to fully understand the ability of WhiB1^{Ec} to anchor an iron-sulfur cluster, what final form it can take, and the cluster modifications upon exposure to air, an anaerobic *in vitro* reconstitution of the iron-sulfur cluster was carried out (Section 3.3.1). This was followed by treatment of the reconstituted sample of WhiB1^{Ec} with O_2-saturated buffer and UV-visible spectroscopic analysis to examine the formation of different cluster species during aerobic degradation (Section 3.3.4).

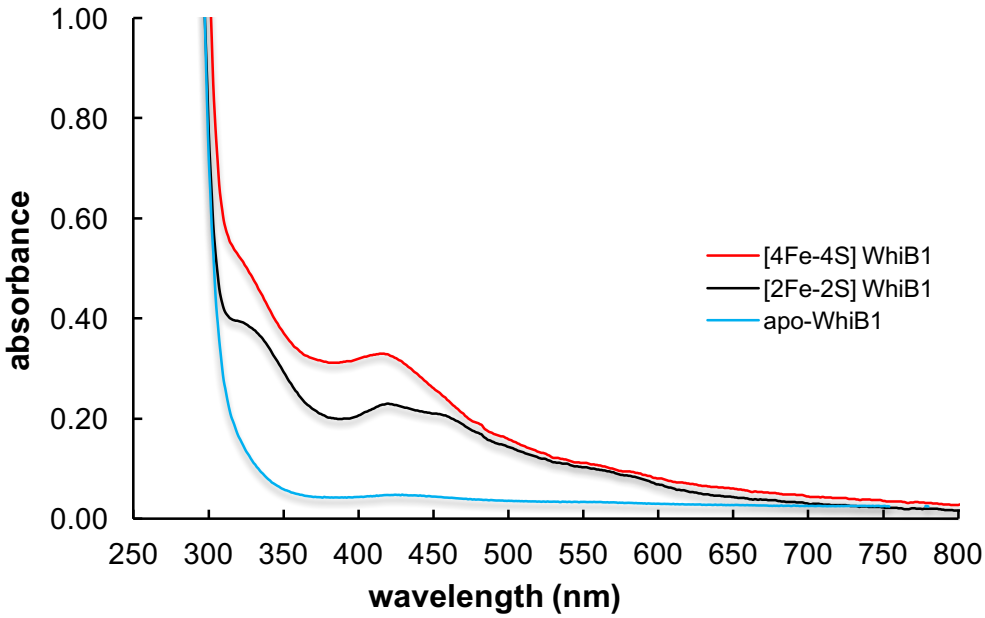


Figure 3.3: UV-visible spectroscopic analysis of WhiB1^{Ec} forms eluted during HiTrap chelating chromatography. The [4Fe-4S]-WhiB1^{Ec} form (red trace) shows only one peak at 420 nm. Black trace shows shoulders at 325, 420, 460, and ~550 nm, these features indicate the presence of [2Fe-2S] clusters in WhiB1^{Ec}. Blue trace represents the apo-form of WhiB1^{Ec}, where there are no any detectable features in the visible region of the spectrum. The order of elution from HiTrap chelating column was: [2Fe-2S]-, [4Fe-4S]-, and apo-WhiB1^{Ec} (Figure 3.2A).

3.3 Properties of the WhiB1^{Ec} iron-sulfur clusters

3.3.1 Holo-WhiB1^{Ec} can be reconstituted *in vitro*

To investigate the ability of apo-WhiB1^{Ec} to re-anchor an iron-sulfur cluster, the *in vitro* reconstitution reaction was carried out as described in Section 2.6.2. UV-visible spectra were recorded at several time points during the course of the reconstitution reaction (Figure 3.4A). The spectra showed the appearance of a peak at 420 nm as cluster formation progressed. This suggested the assembly of a [4Fe-4S] cluster represents the holo- form of WhiB1^{Ec}. The spectra did not indicate formation of [2Fe-2S] clusters during the reaction. The data also suggested that 7 to 9 h of incubation time under anaerobic conditions was sufficient for WhiB1^{Ec} to be fully occupied with [4Fe-4S] clusters (Figure 3.4A). Increased absorbance at 280 nm was also caused by the formation of iron-sulfur clusters. However, the absorbance at 280 nm was significantly higher than might be expected based on the proportion of cluster formed within the protein. Therefore, the components of the reconstitution reaction were incubated under anaerobic conditions without the presence of WhiB1^{Ec} protein for about 9 h. UV-visible spectra showed two features, a peak at 285 nm and a shoulder at 420 nm (Figure 3.4B), suggesting the formation of iron-complexes in solution. However, these contaminants were very labile and were eliminated when the reconstituted holo-protein was dialyzed under anaerobic conditions (Section 2.6.2.2; Figure 3.5).

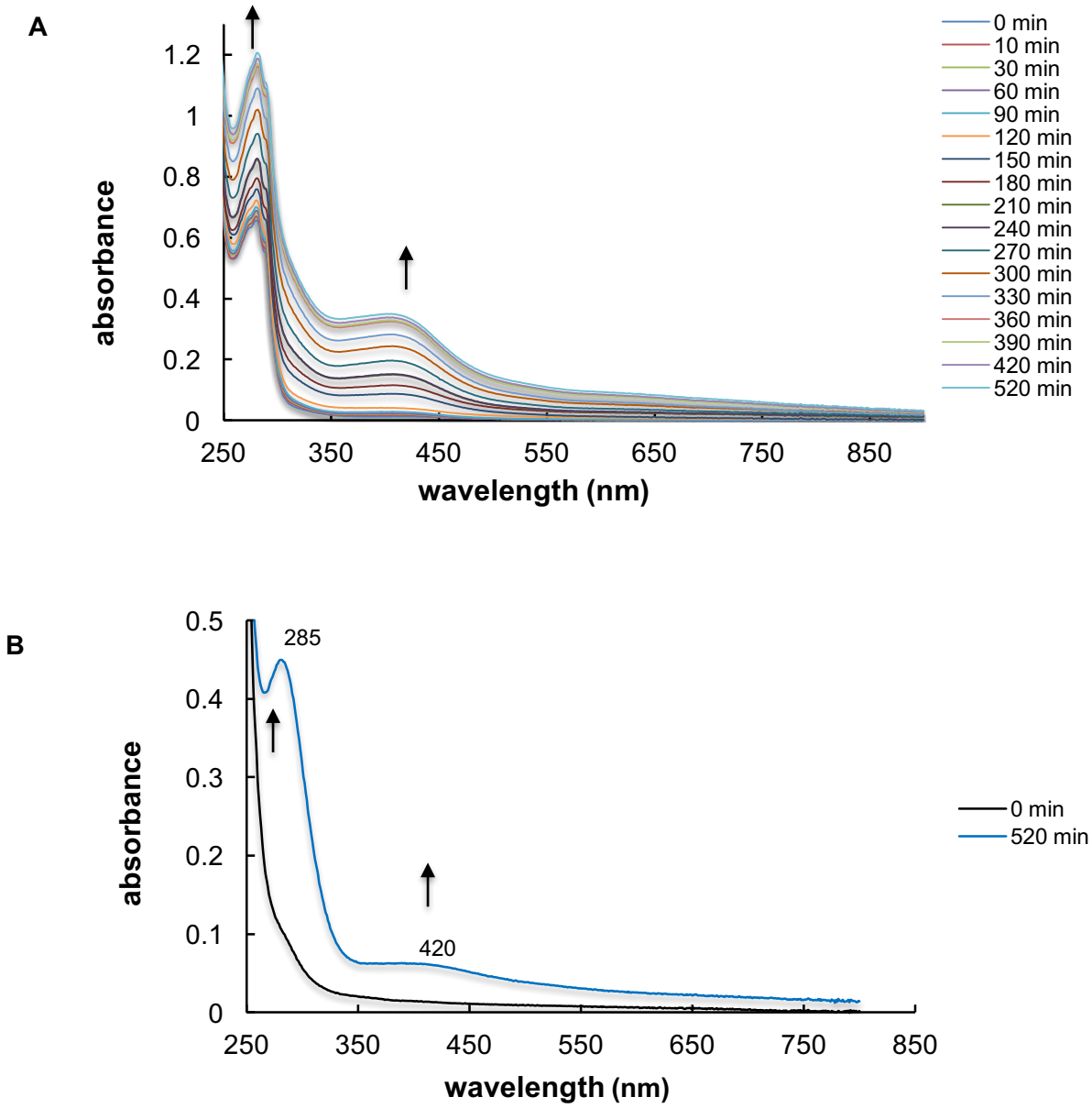


Figure 3.4: Iron-sulfur cluster reconstitution of WhiB1^{Ec}. [A] UV-visible spectra of iron-sulfur cluster formation in WhiB1^{Ec} over ~9 h. Reaction was carried out under anaerobic conditions in closed screw capped quartz cuvette containing WhiB1^{Ec}, 40 μM; NifS, 2 μg; $FeNH_4(SO_4)_2$, 0.4 mM; L-cysteine, 0.2 mM and 10 mM of DTT. Spectra were recorded at several time points and cluster occupation was achieved after ~7h of incubation at 25°C. **[B] UV-Visible spectrum before (black trace) and after (blue trace) of reconstitution reaction in the absence of WhiB1^{Ec}.** The sample comprised of the same reaction contents as **[A]** under similar reaction conditions, but in the absence of WhiB1^{Ec} protein.

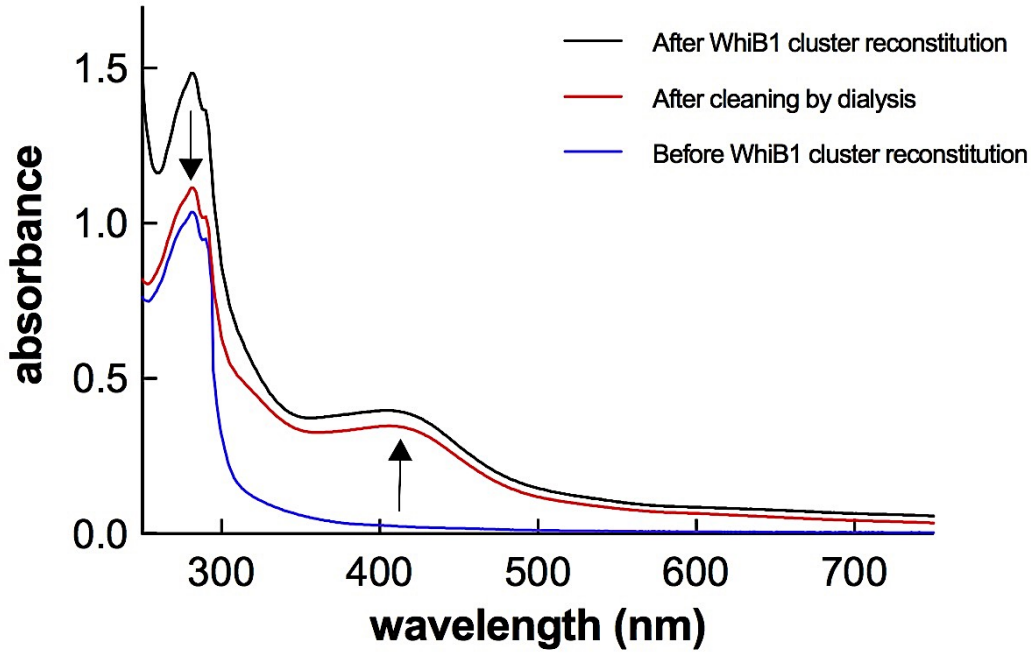


Figure 3.5: Purification of [4Fe-4S]-WhiB1^{Ec} form from unincorporated contaminants of reconstitution reactions. The UV-visible spectra show: apo-WhiB1^{Ec} (blue trace) before the reconstitution reaction; reconstituted WhiB1^{Ec} (black trace); and reconstituted WhiB1^{Ec} after dialysis against 25 mM of NaH_2PO_4 buffer (pH 7.4) containing 0.25 M NaCl inside the anaerobic cabinet at room temperature (red trace). After dialysis under anaerobic condition the absorbance at 420 nm of the reconstituted WhiB1^{Ec} was relatively unchanged; but the enhanced absorbance at 280 nm attributed to non-specific components of the reconstitution reaction was diminished to a level commensurate with absorbance contributed by the cluster.

3.3.2 Protein concentration measurement of [4Fe-4S]-WhiB1^{Ec}

The [4Fe-4S]-WhiB1^{Ec} concentration was determined as described in Section 2.6.1. The concentration of the protein was estimated by Bradford assay and measurement by UV light at 280 nm as described in Section 2.5.1. However, there are some drawbacks in both methods, which can affect the accuracy of calculating the concentration of WhiB1^{Ec}. This has implications for other analytical experiments in which specific concentrations of WhiB1^{Ec} are needed. The Bradford method can give overestimation or underestimation of concentration for some proteins. Although the measurement of protein concentration by $A_{280\ nm}$ is potentially more accurate than the Bradford method, the presence of iron-sulfur clusters in Wbl proteins can severely affect the accuracy of these measurements. This is because the iron-sulfur clusters can absorb the UV light at 280 nm (Depamphilis et al., 1974), and even when the extinction coefficients of iron-sulfur clusters at 280 nm are known, it is difficult to determine the protein concentration by this method due to the type of iron-sulfur present and the level of cluster occupation.

To resolve these problems, the Bradford method of protein estimation, using bovine serum albumin as standard, was initially used to calculate holo-WhiB1^{Ec} concentration. Then, O_2-saturated buffer, supplemented with chelating (10 mM EDTA) and reducing (1 mM TCEP-HCl) agents, was added to remove the iron-sulfur cluster of WhiB1^{Ec} (Section 2.6.4). When, the peak at 420 nm was lost (Figure 3.7A), apo- WhiB1^{Ec} concentration was measured according to the Beer-Lambert Law at 280 nm, using the WhiB1^{Ec} extinction coefficient factor calculated via the Expasy tool ($\varepsilon_{280\ nm}$ = 16,500 M^{-1} cm^{-1}). The correction factor between the two methods was measured (0.79), which was required for accurate measurement of protein concentration. For further calculations of the [4Fe-4S] WhiB1^{Ec} of the Bradford method was applied and the reading was adjusted using the above correction factor.

3.3.3 Holo-WhiB1^{Ec} has a [4Fe-4S] cluster

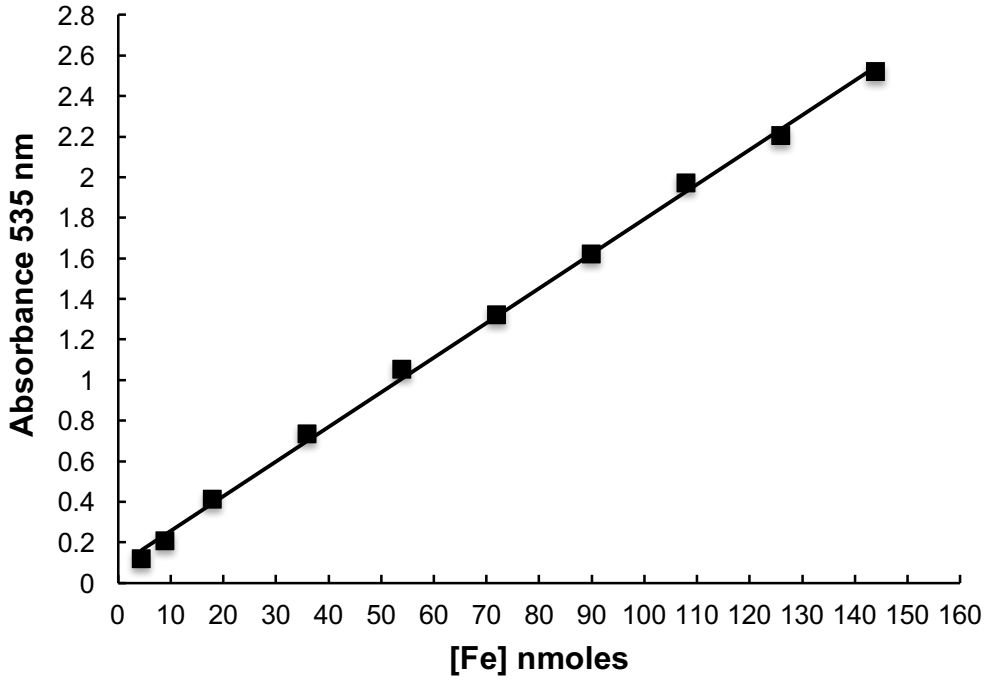
The brown colour of reconstituted holo-WhiB1^{Ec} and its UV-visible spectrum were consistent with the presence of a [4Fe-4S] cluster. To confirm this, iron content was measured as described in Section 2.6.3. Reconstituted WhiB1^{Ec} was prepared and purified as in Sections 2.6.2 and 2.6.2.2 respectively. Protein concentration was measured as previously mentioned (Section 2.6.1). The assay was callibrated using a standard iron solution (BDH) (Figure 3.6A), and the ratio of iron to protein was calculated for three different preparations (Figure 3.6B). The iron content of holo-WhiB1^{Ec} was ~3.8 iron atoms per protein. Hence, it was concluded that WhiB1^{Ec} is a protein containing a [4Fe-4S] cluster after anaerobic reconstitution.

3.3.4 Holo-WhiB1^{Ec} is relatively insensitive to oxygen

As previously shown, aerobic purification of WhiB1^{Ec} resulted in three different forms of the protein, apo-, [2Fe-2S]-, and [4Fe-4S] forms. The iron-containing forms of WhiB1^{Ec} lost their clusters upon prolonged exposure to air. To investigate the protein's sensitivity to O_2, reconstituted [4Fe-4S] WhiB1^{Ec} was treated with O_2-saturated buffer as previously described (Section 2.6.4). The reconstituted WhiB1^{Ec} was prepared and purified first as described (Sections 2.6.2 and 2.6.2.2). The [4Fe-4S] cluster concentration after reconstitution was ~11 μM, which then was treated with O_2-saturated buffer equal to a final O_2 concentration of 110 μM. The cluster break down rate was extremely slow; however, the holo-protein changed to the apo-form after nearly 5h of incubation (Figure 3.7A). The [4Fe-4S] WhiB1^{Ec} was also exposed to air directly by bubbling the protein with air three times using a 100 µl Hamilton gastight syringe without adding buffer or any reducing agent. The cluster concentration dropped by about a half in first 50 minutes (Figure 3.7B). Then, the protein sample started to precipitate after losing the cluster, suggesting the possible formation of some intermolecular disulfide bridges between the molecules of apo-WhiB1^{Ec}.

The sensitivity of [4Fe-4S]-WhiB1^{Ec} to air with and without the presence of reducing agents was studied (Figure 3.8). The reason for including reducing

A



B

WhiB1^{Ec} (nmoles)	Fe (nmoles)	Ratio of iron to WhiB1^{Ec}
4.33	16	3.70
6.26	24	3.83
8.26	31	3.75

Figure 3.6: Measurement of the iron content of reconstituted WhiB1^{Ec} protein. [A] Different concentrations of iron standard solution (BDH) were measured and plotted against the absorbance at 535 nm. **[B]** Ratio of iron to WhiB1^{Ec}. The iron content of the WhiB1^{Ec} protein samples was calculated, then divided by the protein concentration of each sample to calculate the ratio.

agents in this assay was to prevent the possible formation of inter-molecular disulfide bonds when the cysteine residues are released from the cluster, which can lead to protein precipitation.

The WhiB1^{Ec} [4Fe-4S] cluster was extremely stable when exposed to O_2 in the presence of DTT. Protein lost less than 10% of the cluster during ~16 hours after exposure to 110 µM of O_2 in the presence of DTT. DTT then was replaced with TCEP-HCl, which resists air oxidation unlike DTT (Han and Han, 1994). About 60% of the cluster was dissociated within 4 h of exposure to 110 µM of O_2. However, in the presence of no reducing agent, WhiB1^{Ec} lost more than 90% of cluster in ~4 h, suggesting the role of reducing agent in stabilizing the cluster. Removing the His_6-ThCS tag from the protein was carried out by treatment with Thrombin protease (from human plasma, Sigma), where 10 units of protease was used for each 100 nmole of WhiB1^{Ec} with overnight incubation at 4°C (Appendix 1). Treatment of the reconstituted, non-tagged WhiB1^{Ec} with O_2-saturated buffer did not modify the cluster's sensitivity to O_2.

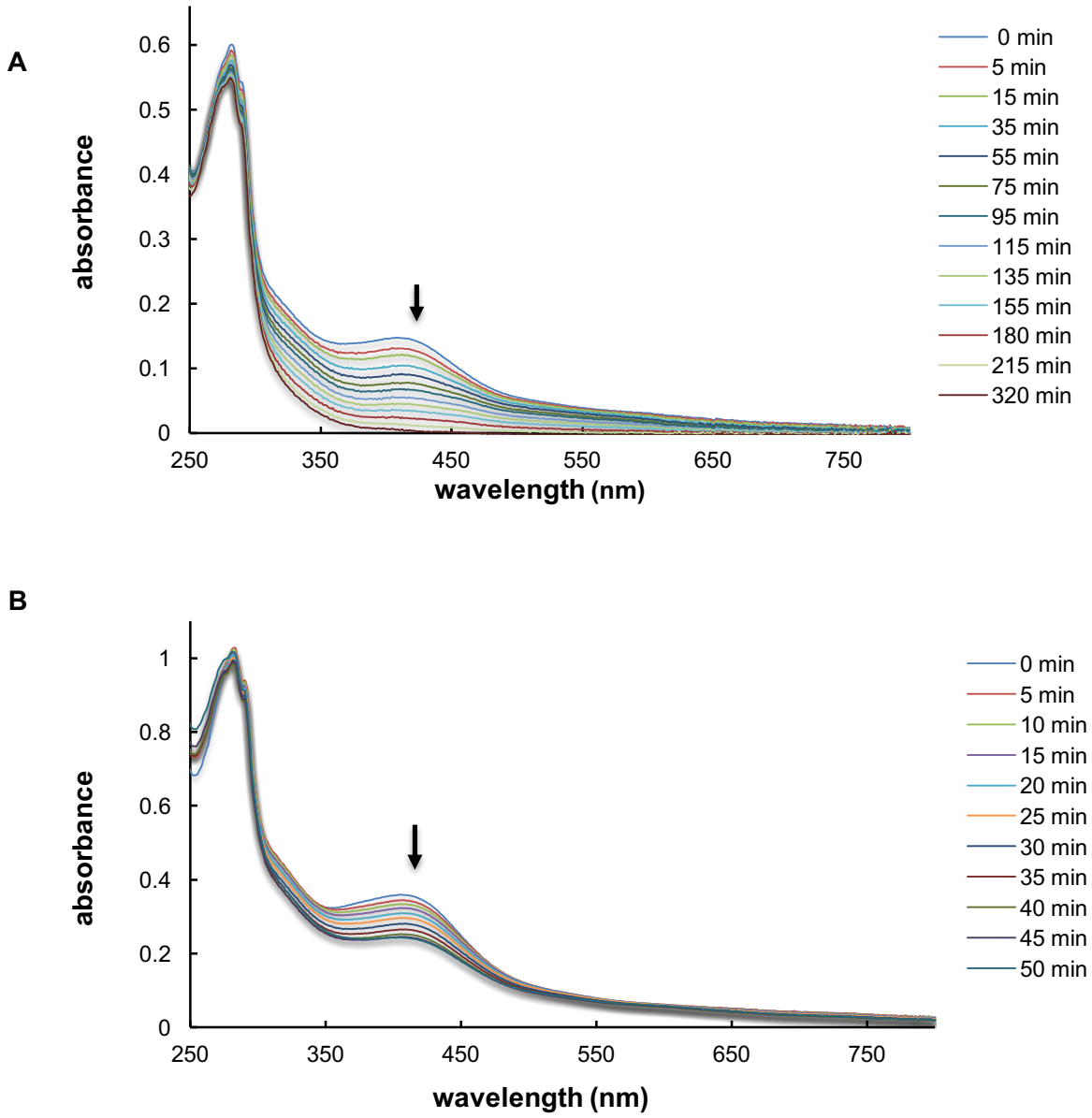


Figure 3.7: Oxygen insensitivity of [4Fe-4S]-WhiB1^{Ec}. [A] UV-visible spectra of reconstituted WhiB1^{Ec} after exposure to air. Air-saturated buffer (50 mM Tris, pH 7.4 containing 0.5 M NaCl, and 5% glycerol) with a final concentration 110 μM of O_2 was added to ~11 μM of [4Fe-4S] cluster and incubated at 25°C. Spectra were taken at the time points indicated in the figure. **[B] UV-Visible specta of [4Fe-4S]-WhiB1^{Ec} showing cluster breakdown after direct exposure to air.** Arrows indicate the spectra direction within the period of incubation.

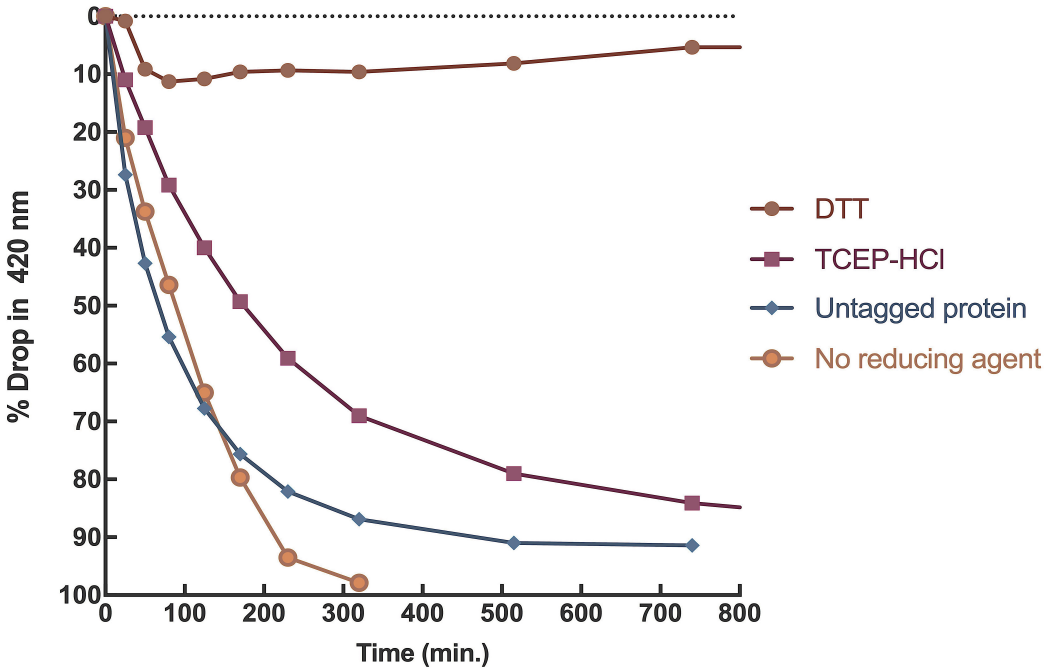


Figure 3.8: Iron-sulfur cluster degradation rates of reconstituted [4Fe-4S]-WhiB1^{Ec} under aerobic conditions in the presence and absence of reducing agents and after removing the His$_6$-ThCS-tag. The [4Fe-4S]-WhiB1^{Ec} was exposed to air-saturated buffer (110 µM final concentration of O_2), 1 mM of DTT or TCEP-HCl was used to reduce WhiB1^{Ec}. The percentage of cluster degradation of each condition was calculated along the incubation time. Untagged reconstituted WhiB1^{Ec} was treated with O_2-saturated buffer in the absence of reducing agents.

3.3.5 The cysteine status of apo-WhiB1^{Ec}

Alkylation of cysteine residues by using iodoacetamide (IAA) coupled with analysis by Liquid chromatography mass spectrometry (LC-MS), which was carried out as described in Section 2.5.10 by Mr. Simon Thorpe (The University of Sheffield Faculty of Science Mass Spectrometry Centre), was used to determine the redox status of cysteine residues after cluster break down (Section 2.5.7). Cysteine-thiol reaction with iodoacetamide lead to increase the protein molecular weight of 57 Da. Reducing WhiB1^{Ec} first with DTT followed by a reaction with IAA caused an increase in the molecular mass of WhiB1^{Ec} by 232 Da, suggesting that the four-cysteine residues of WhiB1^{Ec} were in a reduced form (Figure 3.9, blue line). LC-MS analysis of alkylated native apo-WhiB1^{Ec} with IAA showed three main peaks of different masses; 11477 Da, equivalent to apo-WhiB1^{Ec} with two intra-molecular disulfide bonds; and 11593 Da, corresponding to apo-WhiB1^{Ec} with one intra-molecular disulfide bridge and two IAA modified thiol; and 11,709 Da, indicating that all four cysteine residues were in the reduced form and reacted with IAA.

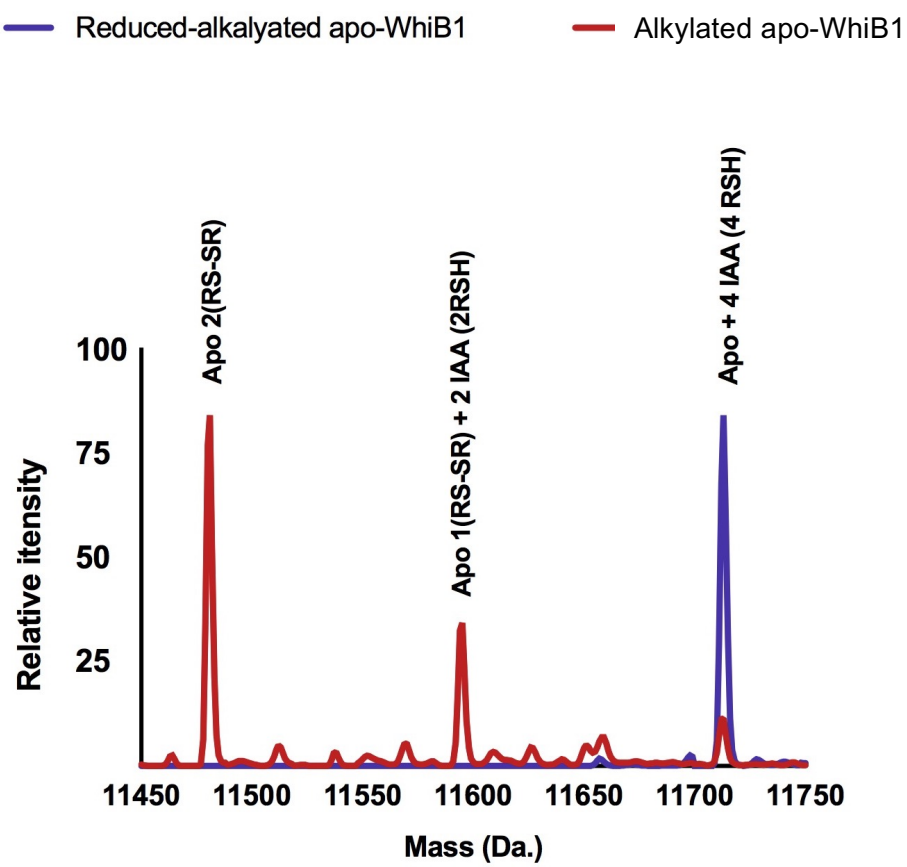


Figure 3.9 LC-MS spectra of alkylated apo-WhiB1^{Ec} after iron-sulfur cluster degradation. Reconstituted WhiB1^{Ec} (100 μM) was prepared and cleaned as previously described, then treated with air-saturated buffer without the presence of TCEP-HCl, and left in aerobic condition until the protein completely lost the cluster as judged by the UV-visible spectrum. Reducing the protein with 5 mM DTT for 30 min, followed by a reaction with 10 mM IAA in the dark for 1 h was done for preparing the control (blue line). The native apo-WhiB1^{Ec} was also incubated with 10 mM IAA, but without treatment with the reducing agent. Following alkylation, WhiB1^{Ec} was cleaned by dialysis as described in Section 2.5.7. Three main peaks with different masses are highlighted (bold fonts). Mass spectra were obtained by Mr. Simon Thorpe.

3.4 Overproduction, purification, biochemical and biophysical characterization of WhiB1^{Ms} overexpressed in *M. smegmatis*

3.4.1 WhiB1^{Ms} overproduction and isolation

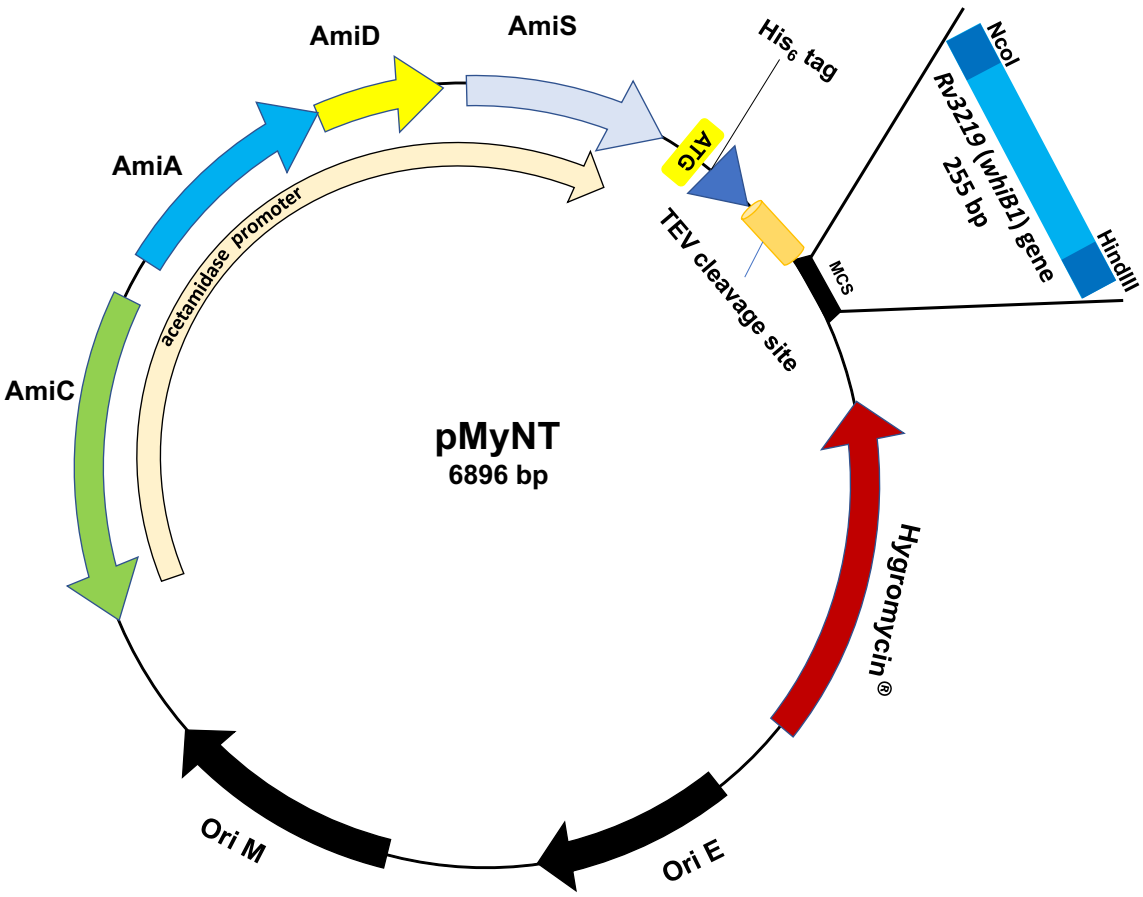
Mycobacterium smegmatis has a WhiB1 protein that is 96% identical to that of *M. tuberculosis* (Figure 3.10). Therefore, expression of recombinant *M. tuberculosis* WhiB1 in *M. smegmatis* instead of *E. coli* was carried out to further characterize the iron-sulfur cluster of the protein. This is the first time that a Wbl protein has been expressed, overproduced and isolated from a mycobacterium host.

WhiB1^{Ms} overproduction was done by transforming electro-competent *M. smegmatis* MC^{2}155 cells with pGS2524 (Table 2.1 and 2.2), which encodes the recombinant *N*-terminal His$_6$ tag-TEV protease cut site-WhiB1 (His$_6$-TEVCS-WhiB1^{Ms}) under the control of the acetamidase promoter as descried in Section 2.3.14 (Figure 3.11). Cultures of *M. smegmatis* JRG6798 were induced, once the O.D. reached ~1.0, with 0.2% acetamide, and then incubated for protein expression a further 12 h at 37°C with shaking 250 rpm (Section 2.5.3.2). The cell-free extracts contained a high level of WhiB1^{Ms} in the soluble fraction (Figure 3.12A, lane 3). Therefore, WhiB1^{Ms} purification was carried out by affinity chromatography as mentioned in Section 2.5.5.1. Cell lysate of *M. smegmatis* was prepared by re-suspending cell pellets in binding buffer 500 mM NaCl, 50 mM Tris pH 8.0, and then sonicated (Section 2.5.4). WhiB1^{Ms} was eluted from HiTrap chelating column in only one form with a dark brown colour, the purification profile showed only a single peak for WhiB1^{Ms} (Figure 3.12 B), which was unlike the purification profile of WhiB1^{Ec} overproduced in *E. coli* (Figure 3.2A). Fractions were analysed by SDS-PAGE which confirmed the presence of WhiB1^{Ms} protein (Figure 3-12A, lane 4).

Overproduction and isolation of WhiB2^{Ms} and WhiB3^{Ms} were also achieved similarly as WhiB1^{Ms} using pGS2542 and pGS2543 (Table 2.2), which encode His$_6$-TEVCS-WhiB2^{Ms} and His$_6$-TEVCS-WhiB3^{Ms} respectively. WhiB2 and WhiB3 are not the focus of this study and only studied for comparison purposes with some results of WhiB1^{Ms}. *Mycobacterium smegmatis* MC^{2}155 JRG6823 and JRG6824 (Table 2.1) were used for overproducing WhiB2 and WhiB3 respectively.

Figure 3.10: Sequence alignment of WhiB1 proteins from *M. smegmatis* MC^{2}155 and *M. tuberculosis* H37Rv. The sequences of *M. smegmatis* MC^{2}155 *whiB1* were obtained from NCBI website; https://www.ncbi.nlm.nih.gov/protein/AFP38350.1, and sequence of *M. tuberculosis H37Rv whiB1 from* https://www.ncbi.nlm.nih.gov/protein/NP_217735.1, and aligned by multalin tool (Corpet, 1988). Residues that are identical in both proteins (*), residues with similar properties (.) are indicated.

A



B

MK<u>HHHHHH</u>PSAG<u>ENLYFQ</u>↓G<u>A-WhiB1</u>

His_6-tag TEV cleavage site

Figure 3.11: Schematic diagram of pGS2524 expression plasmid. [A] The *whiB1* open reading frame (255 bp) was amplified by PCR from genomic DNA of *M. tuberculosis* with engineered NcoI/HindIII restriction sites, and ligated into pMyNT to create pGS2524. **[B]** The plasmid encodes a His_6-TEVCS-$WhiB1^{Ms}$ fusion protein.

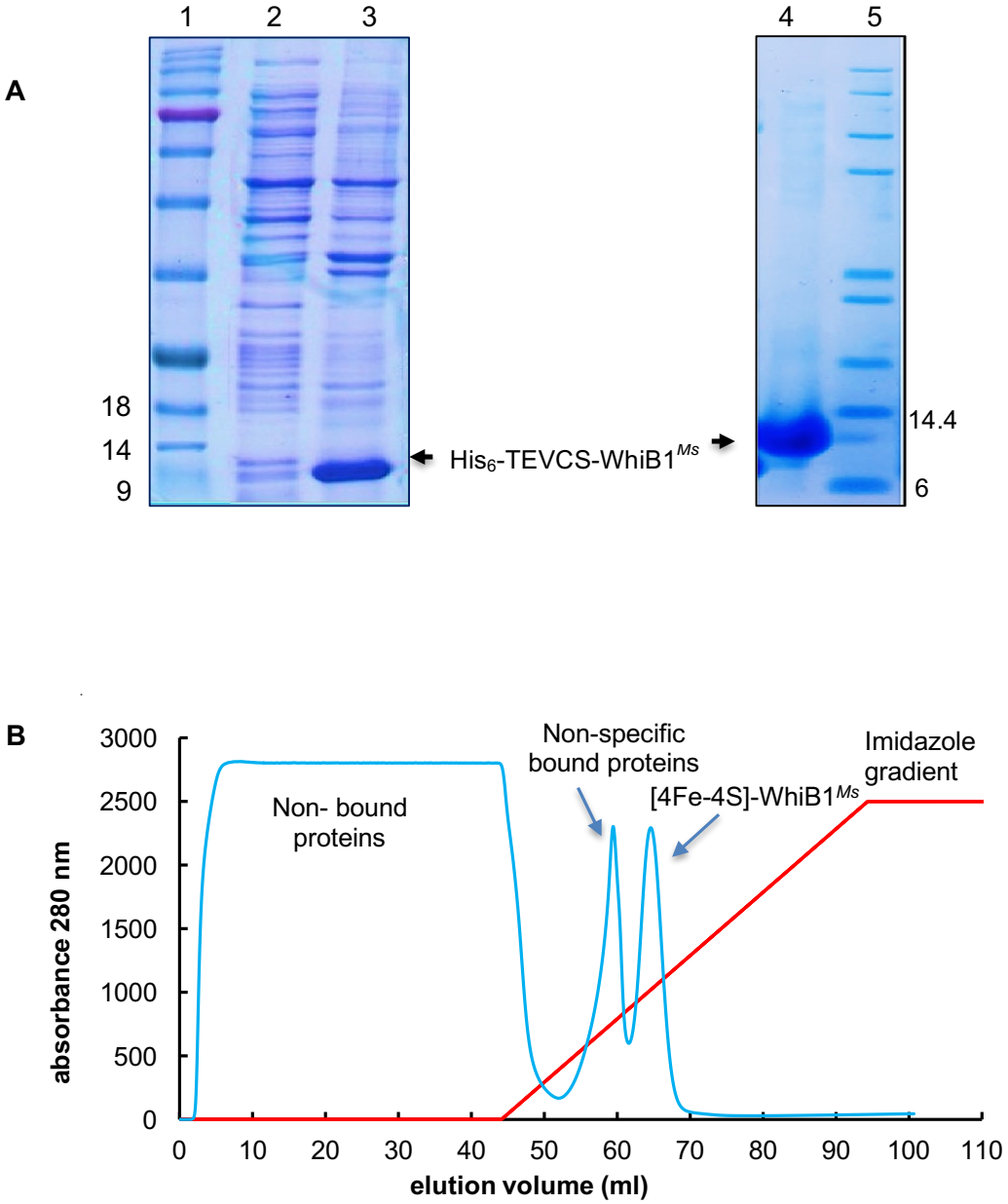


Figure 3.12 WhiB1^{Ms} overproduction and isolation from a mycobacterial host. [A] Coomassie stained SDS-PAGE gel of samples from WhiB1^{Ms} overproduction and isolation. Lane 1, protein ladder (sizes in kDa are indicated); lanes 2 and 3, cell debris and cel-free extract respectively from induced cultures of *M. smegmatis* JRG6798 with acetamide; lane 4, isolated WhiB1^{Ms} by a nickel-charged column; lane 5, protein ladder. **[B] Elution profile from HiTrap chelating chromatography of cell lysate obtained from *M. smegmatis* JRG6798.** The chart shows protein absorbance at 280 nm (blue line), and the imidazole gradient from 0-0.5 M (red line). Eluted [4Fe-4S] WhiB1^{Ms} appears as only one peak between ~62-70 ml.

3.4.2 WhiB1^{Ms} has a stable [4Fe-4S] cluster

The UV-visible spectroscopy trace of isolated WhiB1^{Ms} by HiTrap chelating column, showed only one peak at 420 nm (Figure 3.13). The brown colour of the protein and the trace of UV-visible spectrum indicated the presence of [4Fe-4S]-WhiB1^{Ms} form. The total iron content of WhiB1^{Ms} was also measured as described in Section 2.6.3. The average molar ratio of iron to protein in three repeats (3.79, 3.95, and 3.88) was 3.87 of iron to each a monomer molecule of WhiB1^{Ms}. Moreover, the total protein mass including the iron-sulfur cluster was measured by native electrospray ionization mass spectrometry (EIS-MS), in a collaboration with Dr. Jason Crack, University of East Anglia as described in Section 2.5.11. The molecular mass of holo-protein was 11987 Da., which indicated the presence of an additional ~335 Da - the mass of a [4Fe-4S] cluster – compared to the molecular mass of apo-WhiB1^{Ms} 2(RS-SR) (11632 Da.) (Figure 3.21). These results confirmed that WhiB1^{Ms}, which was overproduced in *M. smegmatis,* binds only a [4Fe-4S] cluster, unlike the WhiB1 overproduced in *E. coli* which was isolated in three different forms (Section 3.2.2).

To test the stability of [4Fe-4S] cluster of WhiB1^{Ms} upon addition of O_2 to final concentration of 110 μM (Section 2.6.4). UV-visible spectra were obtained before and after O_2 exposure (Figure 3.13). After 15 h of incubation under these conditions no change was noticed at 420 nm, unlike the spectral changes of WhiB1^{Ec} (Figure 3.7). This result showed that the iron-sulfur cluster of WhiB1^{Ms} behaved differently to that of WhiB1^{Ec} in similar aerobic environments.

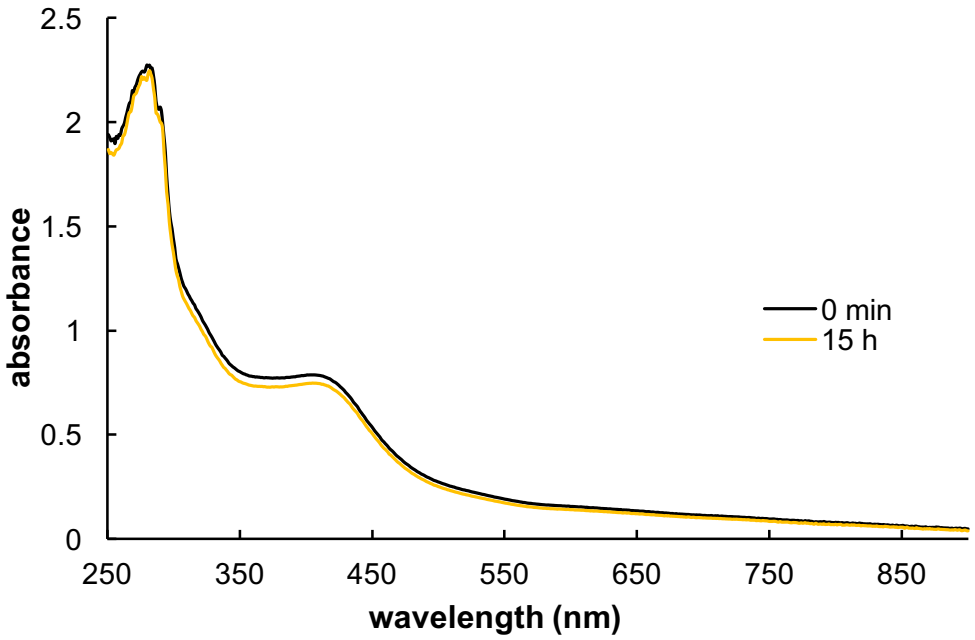


Figure 3.13: The [4Fe-4S] cluster stability of isolated WhiB1^{Ms}. UV-visible spectra WhiB1^{Ms} (44μM cluster) before (black trace), and after 15 h of incubation (yellow trace) in the presence of 110 μM final concentration of O_2.

3.4.3 Circular dichroism (CD) spectroscopy properties of WhiB1

To better understand the reasons behind the differences in cluster stability between WhiB1 overproduced in *E. coli* and that overproduced in *M. smegmatis,* CD spectra were obtained (Section 2.5.9). The WhiB1 samples were prepared in 20 mM NaH_2PO_4, pH 7.4 containing 0.1 M NaCl, where buffer exchange was achieved using 0.5 ml Zeba™ spin desalting columns (Thermo Fisher) pre-equilibrated with the mentioned buffer. For comparing the secondary structure composition of the protein, absorption in the far UV (240 nm and below) was taken for each sample of 30 μM of apo-, holo-WhiB1^{Ec}, and holo-WhiB1^{Ms} (Figure 3.14). The far UV-CD spectra showed a clear difference in negative molar ellipticity between ~204-220 nm of apo- and holo-WhiB1^{Ec} (region of random coil and α-helix), or between ~208-222 nm of holo-WhiB1^{Ms}. These data suggested that the presence of [4Fe-4S] cluster in WhiB1 changes the structure of the protein. Holo-WhiB1^{Ms} showed a different trace with an increased proportion of α-helices compared with the reconstituted WhiB1^{Ec}, The near UV-visible CD spectrum of WhiB1^{Ms} was also studied, this was done by preparing the protein sample in the same buffer mentioned above but with a higher cluster concentration (295 μM). The result was compared with the spectrum of WhiB1^{Ec}, which was previously obtained by Dr. Jason Crack (Smith et al., 2010). The near UV-CD spectrum 260-320 nm represents the region of aromatic amino acids, which is used to study protein tertiary structure (Kelly et al., 2005). The spectrum of WhiB1^{Ms} in this region showed an increase of the positive molar ellipticity at 292 nm compared with WhiB1^{Ec} (Figure 3.15), which could be due to the presence of Tyr 16 and Phe 17 amino acids next to each other in the TEV cleavage site (Figure 3.11B); taking into consideration that Tyrosine is absent in WhiB1 (Figure 3.1B). In contrast, the visible CD spectrum shows positive features at 427 nm and 515 nm, like the WhiB1^{Ec}, indicating some similarity in the iron-sulfur cluster environment of both WhiB^{Ec} and WhiB1^{Ms}.

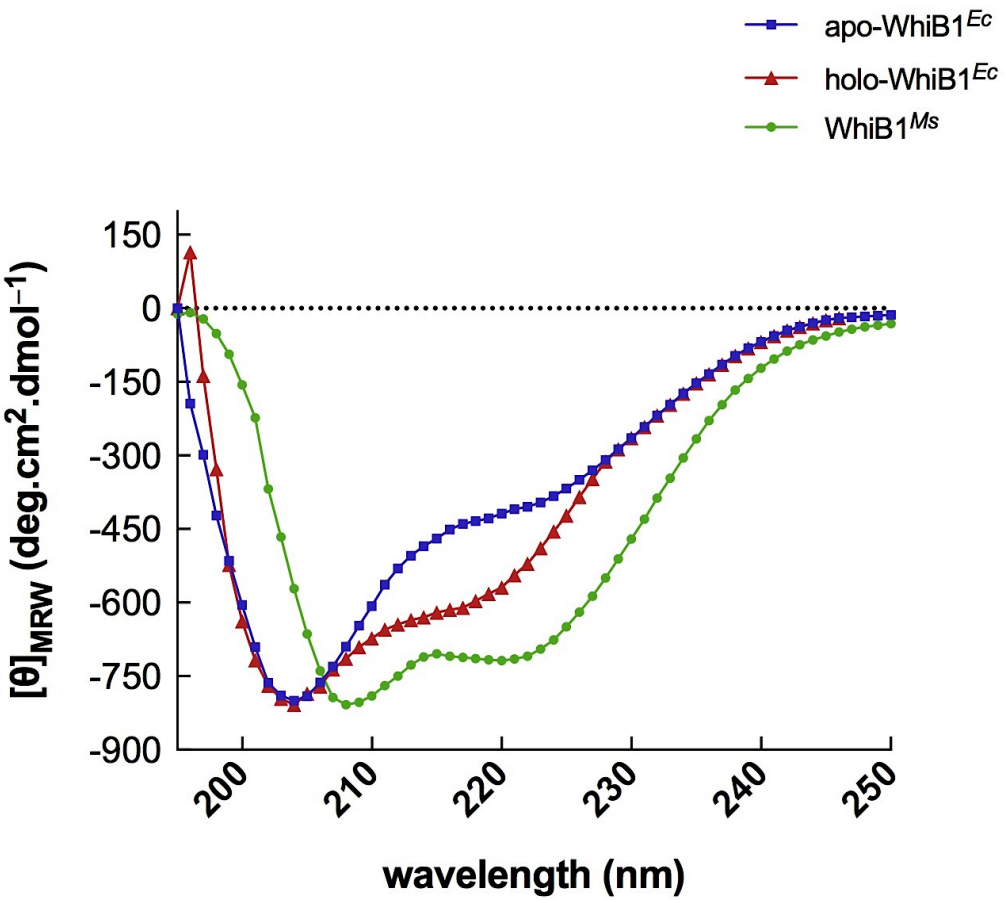


Figure 3.14 Far-UV CD analysis of apo-, holo-WhiB1^{Ec} and WhiB1^{Ms}. The spectra were recorded using 30 μM of each WhiB1 sample in 20 mM NaH_2PO_4, pH 7.4 containing 0.1 M NaCl at 25°C. The trace of apo-WhiB1^{Ec} (■) shows a feature at 204 nm, while both traces of holo-WhiB1^{Ec} spectrum (▲) WhiB1^{Ms} (●) show shoulder or feature at 218-222 nm. WhiB1^{Ms} shows another feature at 208 nm.

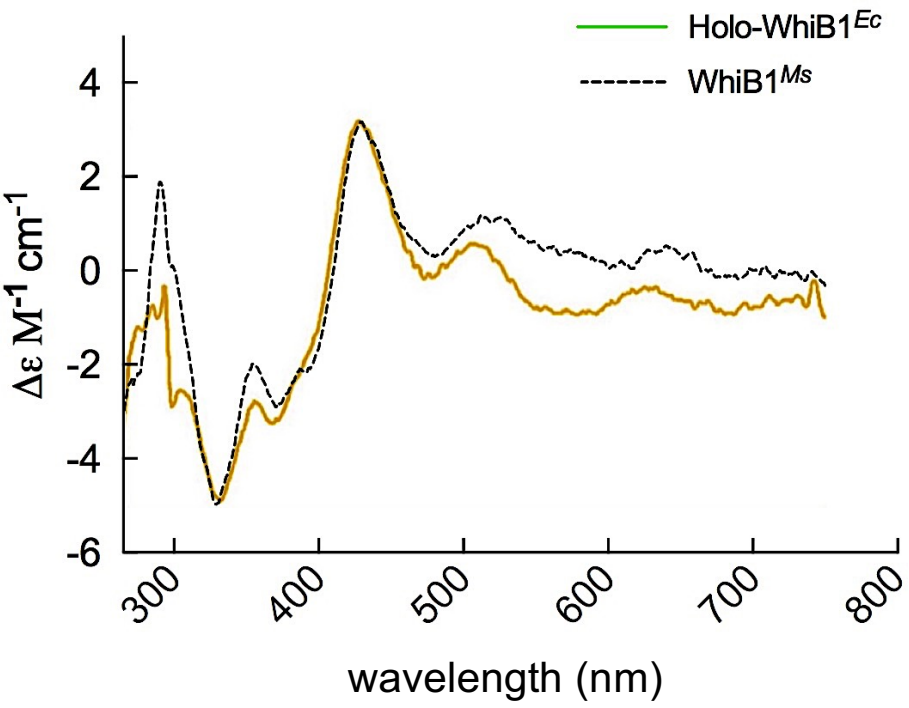


Figure 3.15: Near UV-visible CD analysis of holo-WhiB1^{Ec} and WhiB1^{Ms} proteins. The spectrum of [4Fe-4S]-WhiB1^{Ms} (295 μM cluster concentration) (dashed line) was compared with the 196 μM cluster of [4Fe-4S]-WhiB1^{Ec} spectrum (orange line) from Smith et al., (2010).

3.4.4 The His_6-TEVCS tag enhances cluster occupancy of $WhiB1^{Ms}$

It has been shown that the [4fe-4S] cluster of $WhiB1^{Ms}$ is more stable and this protein is more ordered than the previously characterized $WhiB1^{Ec}$ protein. As well as being produced in *M. smegmatis*, $WhiB1^{Ms}$ has His_6-TEV tag rather than the His_6-Thrombin tag present in $WhiB1^{Ec}$. To establish whether the enhanced cluster stability and structural ordering were due to the His_6-TEV tag or the host, holo-$WhiB1^{Ms}$ was treated with TEV protease at a ratio of 5:100 ($^{w}/_{w}$) of $WhiB1^{Ms}$ substrate. The reaction was carried out in 25 mM NaH_2PO_4 (pH 7.5) with 0.1 M NaCl at 20°C for overnight. Complete digestion was unsuccessful even when the concentration of TEV protease was increased 4 fold (Figure 3.16 lane 2). This suggested that the TEV protease recognition site was inaccessible. Treatment of $WhiB1^{Ms}$ with 1.5 M guanidine HCl for 30 min on ice resulted in the release of the iron-sulfur cluster. Subsequent protein re-folding of apo-$WhiB1^{Ms}$ in the same buffer mentioned above permitted removal of the His-tag by TEV protease treatment (Figure 3.16 lanes 3 to 5). For cleaning the un-tagged $WhiB1^{Ms}$ from cleaved His-tags and the protease, the reaction mixture was injected through nickel-charged Hi-Trap chelating column. The cleaved His-tags, and the TEV protease, which had a His tag as well, bound to the matrix; while, cleaved $WhiB1^{Ms}$ without a His-tag passed straight through. Anaerobic reconstitution of the cleaved, refolded apo-$WhiB1^{Ms}$ resulted in acquisition of a [4Fe-4S] cluster which showed a similar level of O_2 sensitivity to that of holo-$WhiB1^{Ec}$ (Figure 3.17A). However, when the $WhiB1^{Ms}$ protein was denatured, refolded and then reconstituted without removing the His_6-TEV tag, the iron-sulfur cluster was again stable in air (Figure 3.17B; 3.27). Moreover, far UV-CD spectrum of cleaved, refolded apo-$WhiB1^{Ms}$ showed a similar trace to that of apo-$WhiB1^{Ec}$ (Figure 3.18); but if the $WhiB1^{Ms}$ protein was refolded and reconstituted without removing the His_6-TEV tag, the CD-spectral features were similar to that of the original $WhiB1^{Ms}$ spectrum (Figure 3.15). These data indicate that the protein was refolded correctly after denaturation by guanidine HCl, and that the His_6-TEV-tag is responsible for stabilizing the WhiB1 [4Fe-4S] cluster possibly due to its role in changing the secondary structure of the protein.

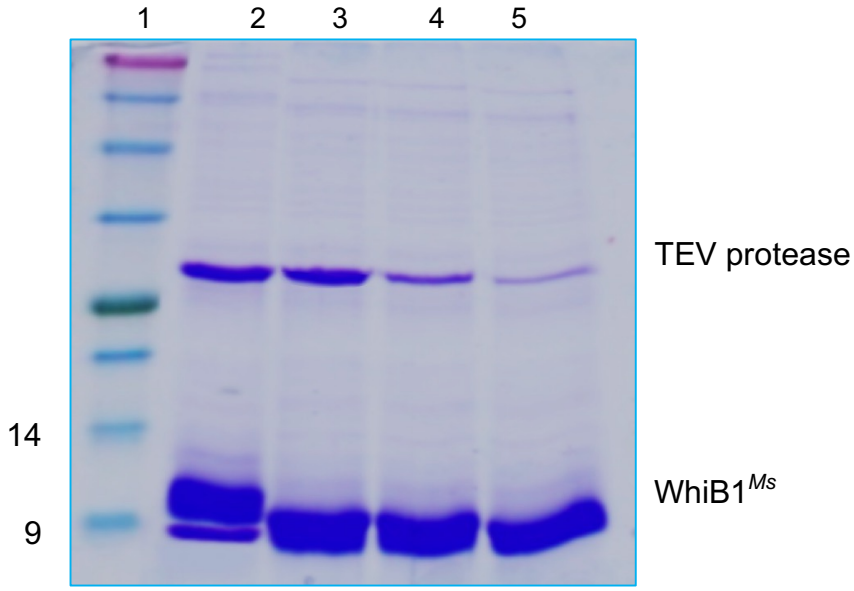


Figure 3.16: Coomassie blue stained SDS-PAGE gel of WhiB1^{Ms} treated with TEV protease. Lane 1, protein ladder (sizes in kDa are indicated); lane 2, holo-WhiB1^{Ms} treated with 20% TEV protease; lanes 3, 4, and 5, refolded apo-WhiB1^{Ms} digested with TEV protease at a ratio of 20, 10, 5:100 ($^{w}/_{w}$) respectively. Digestion was carried out in 25 mM NaH_2PO_4 (pH 7.5) with 0.1 M NaCl at 20°C overnight.

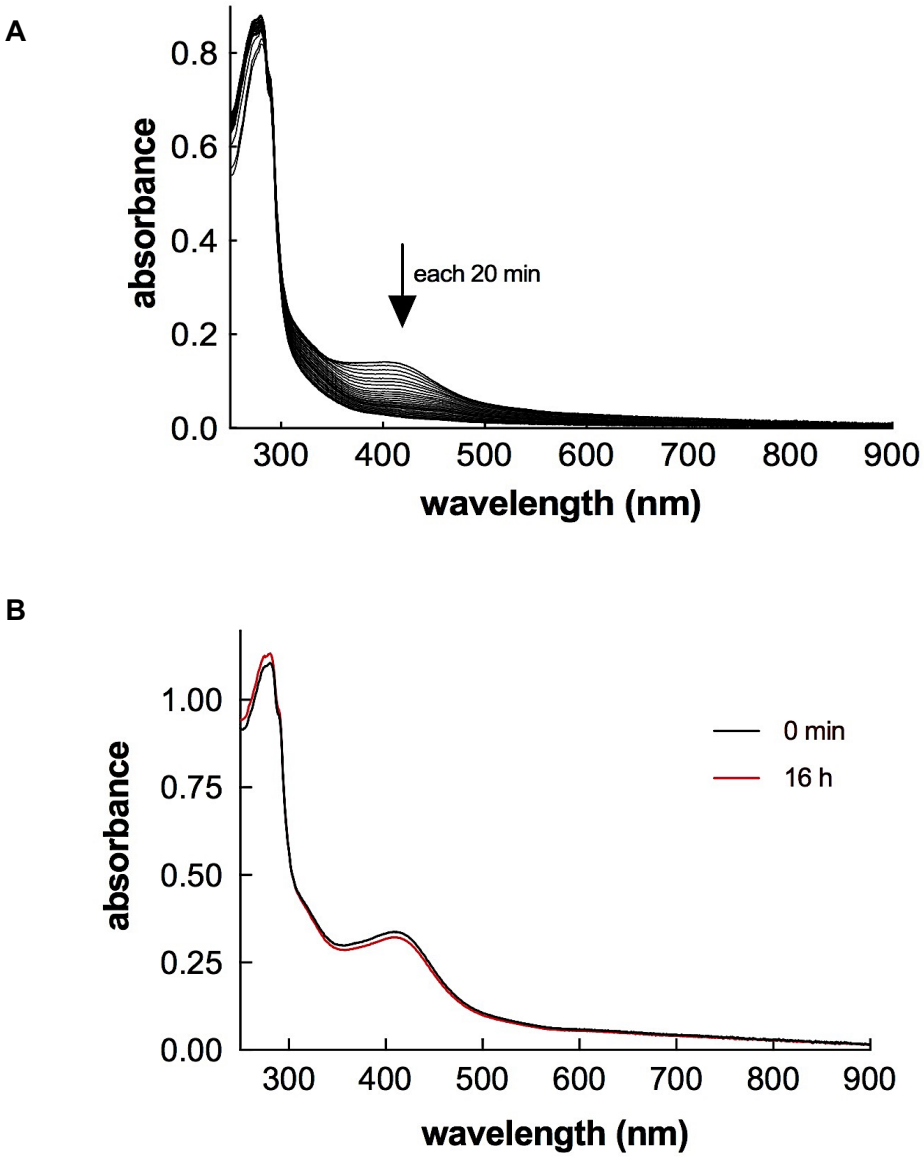


Figure 3.17: The iron-sulfur cluster stability to O_2 of refolded-reconstituted WhiB1^{Ms} in the presence or absence the His$_6$-TEV tag. [A] UV-visible spectrum of refolded-reconstituted untagged WhiB1^{Ms}, the cluster was broken down after adding O_2-saturated buffer with a final concentration 110 μM of O_2. **[B]** UV-visible spectrum of refolded-reconstituted WhiB1^{Ms} (but it was not treated with TEV protease) shows the stability of the cluster in the presence of 110 μM of O_2. Arrow indicates the spectra direction within the period of incubation.

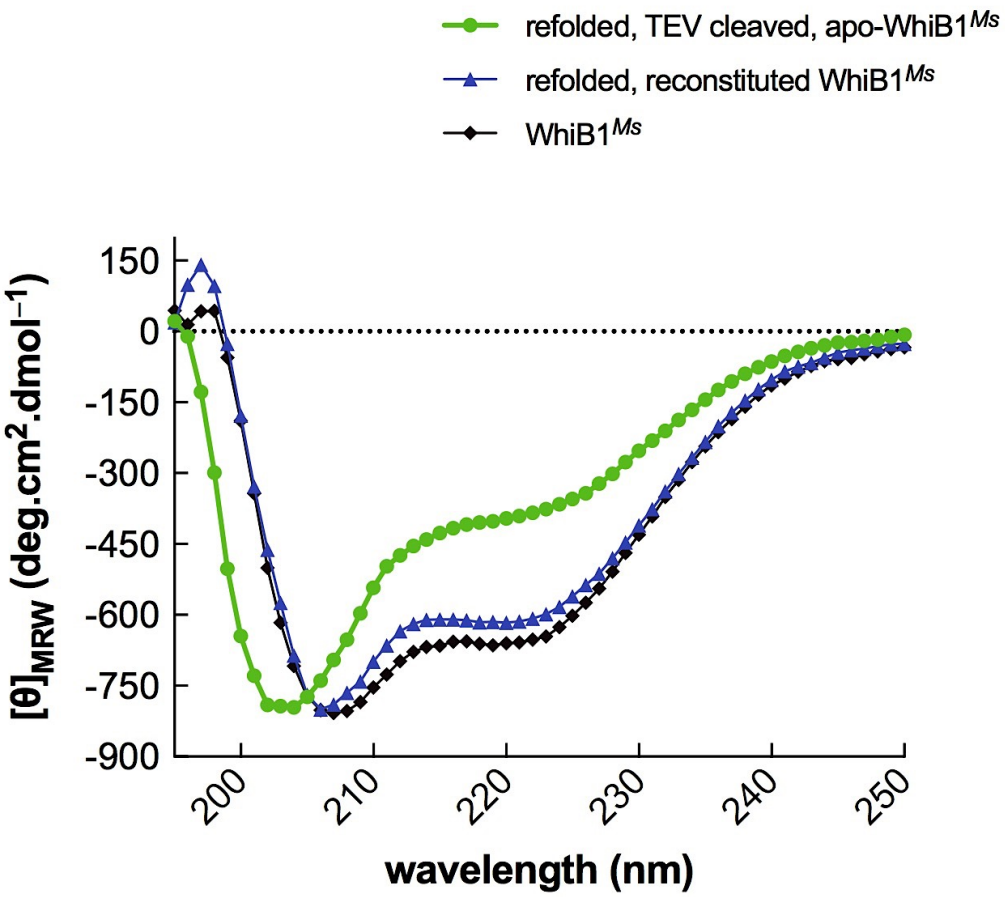


Figure 3.18: Secondary structural properties of apo- and holo-WhiB1^{Ms} with and without His$_6$-TEV tag. The spectra were recorded using 30 μM of each WhiB1 sample in 20 mM NaH_2PO_4, pH 7.4 containing 0.1 M NaCl at 25°C. The Far-UV CD analysis of WhiB1^{Ms} (■); the refolded, His-TEV tag cleaved, apo-WhiB1^{Ms} (●); and the refolded, reconstituted, non-cleaved WhiB1^{Ms} (▲) are indicated.

3.4.5 The iron-sulfur cluster is essential for folding WhiB1

As shown previously in Figure 3.14, the presence of the cluster in WhiB1^{Ec} changed the CD spectral properties of the protein in comparison to the apo-form; however, the secondary structure of holo-WhiB1^{Ms} was significantly different in comparison to holo-WhiB1^{Ec} as judged from CD spectra. Moreover, the presence of TEV tag in the protein appears to be responsible for stabilizing the iron-sulfur cluster as shown in Figure 3.17. These data raised a question, whether the TEV tag alone is responsible for stabilizing WhiB1 structure or the cluster is also needed. Apo- and holo-WhiB1^{Ms} protein were compared using 1D proton NMR spectroscopy. For apo-WhiB1^{Ms} preparation, holo-protein was treated with 1.5 M guanidine HCl and incubated on ice for 30 min. WhiB1 refolding was achieved as mentioned in the previous section. The 1D proton NMR spectra of 0.4 mM of apo- and holo-WhiB1^{Ms}, prepared in 25 mM NaH_2PO_4 (pH 7.0), containing 0.25 M NaCl, were obtained in collaboration with Prof. Michael Williamson and Mrs. Andrea Hounslow (University of Sheffield, Department of Molecular Biology and Biotechnology, NMR facility).

1D ^{1}H NMR is a powerful way to qualitatively judge whether a protein is folded or not. If a protein is well folded, the signals in the spectrum will be sharp, narrow and well dispersed and cover a large area of chemical shifts, unlike an unfolded protein which gives a poor signal dispersion with broader peaks. This is mainly because the environment of protons in an unfolded protein are more similar than in a folded one where each proton is constrained in a specific conformation, and its chemical environment depends on its location in the protein tertiary structure (Kwan et al., 2011; Page et al., 2005).

The spectrum of holo-WhiB1^{Ms} showed more sharp peaks spanning a wide range of chemical shifts, especially in the amide area, from about 7 to 10 ppm in comparison to apo-WhiB1^{Ms} (Figure 3.19). These data suggest that the iron-sulfur cluster is crucial for the folding of WhiB1.

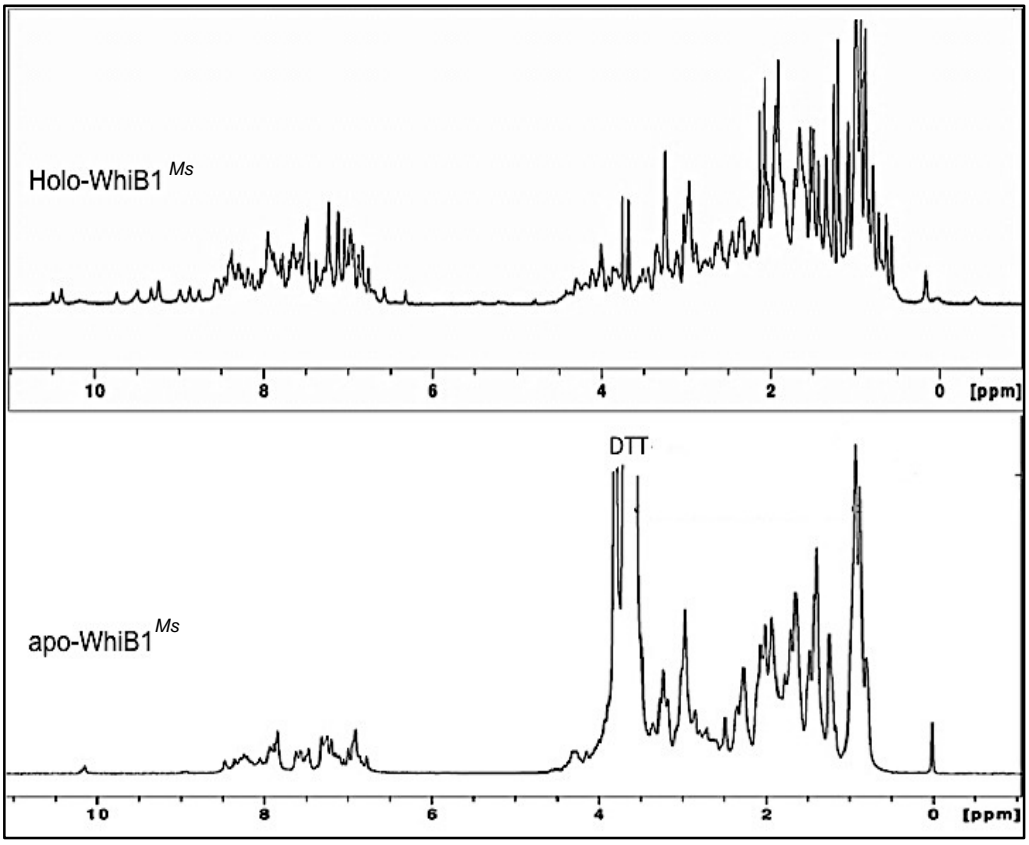


Figure 3.19: The iron-sulfur cluster is crucial for folding WhiB1^{Ms}. 1D ^{1}H-NMR spectra of 0.4 mM apo- and holo-WhiB1^{Ms} prepared in 25 mM NaH_2PO_4 (pH 7.0), containing 0.25 M NaCl were recorded at 25°C.

3.4.6 Liquid chromatography mass spectrometry (LC-MS) study of WhiB1^{Ms}

Some [4Fe-4S] cluster containing regulatory proteins, like *E. coli* FNR and *Psedumonas putida* ANR, were reported to form a cysteine persulfide-liganded [2Fe-2S] cluster upon exposure to O_2 (Zhang et al., 2012; Ibrahim et al., 2015). To investigate the ability of [4Fe-4S] WhiB1^{Ms} to retain sulfide bridges after cluster dissociation, LC-MS analysis was used. LC-MS analysis was carried out by Mr. Simon Thorpe (The University of Sheffield Faculty of Science Mass Spectrometry Centre) as described previously (Section 2.5.10). Apo-WhiB1^{Ms} exhibited one main peak at a molecular weight of 11,636 Da, revealing of the presence of four thiol groups in the protein (Figure 3.20A). This result suggested that WhiB1 lost iron and sulfur atoms of the cluster without sulfur retention. The mass trace was different however, to the WhiB2^{Ms}, which showed an apo-form with two intra-molecular disulfide bonds (peak at a molecular mass of 12868 Da), in addition to other peaks at +32, +64, +96 and 128 mass units compared to apo-WhiB2^{Ms}, indicating the ability of WhiB2^{Ms} to retain cluster sulfide after cluster dissociation (Figure 3.20B). WhiB2 ^{Ms} overproduction and purification were achieved exactly as described for WhiB1^{Ms} in section 3.4.1.

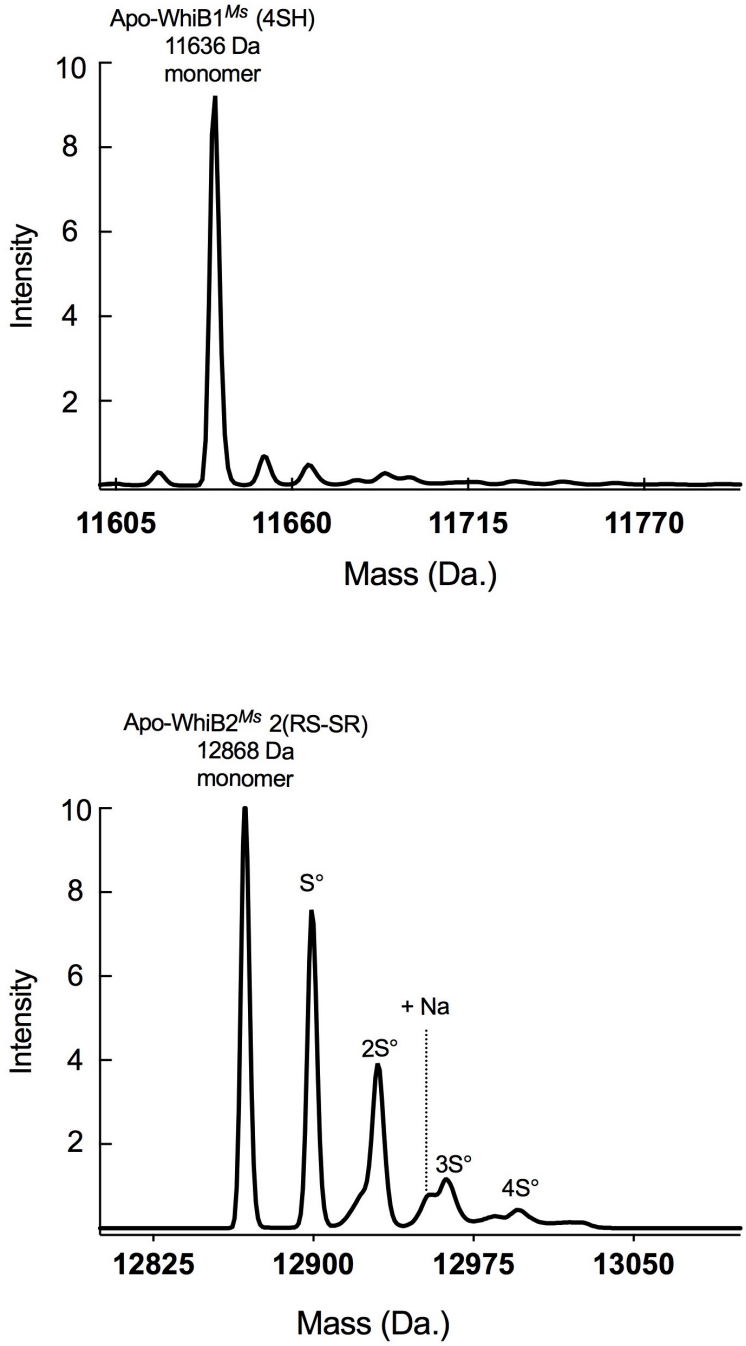


Figure 3.20: LC-MS analysis of Holo-WhiB1^{Ms} and holo-WhiB2^{Ms}. Protein samples (50 µM cluster), in 50 mM Tris, pH 8.0 containing 0.5 M NaCl were applied in the presence of 0.1% formic acid to LC instrument for 10 min before mass analysis. **[A]** The peak at 11636 Da represents the monomer apo-WhiB1^{Ms} in its reduced form. **[B]** The peak at 12868 Da corresponds to the monomer WhiB2^{Ms} containing two disulfide bonds, peaks indicated with S°, 2S°, 3S°, 4S° represent the addition of one, two, three and four sulfur atoms retained by apo-WhiB2^{Ms}.

3.4.7 WhiB1^{Ms} cluster degradation is initiated by loss of sulfur atom

Electrospray ionization mass spectrometry (ESI-MS) is a powerful procedure to study the oxidation states and the stoichiometry of proteins containing metal cofactors such as iron-sulfur clusters. Moreover, even the labile forms of these clusters, which can be degraded quickly, can be studied (Johnson et al., 2000; Yu et al., 1993). In addition, it provides an excellent tool to examine metalloproteins in their native conformations, where the proteins compact structure with their metal cofactors are retained without exposure to denaturation conditions (Troxler et al., 1999; Veenstra et al., 1998).

In-source collision-induced dissociation (isCID) was carried out using electrospray ionization mass spectrometry by Dr. Jason Crack (Section 2.5.11), to study the fragmentation pathway of WhiB1^{Ms} [4Fe-4S] cluster. WhiB1^{Ms} (~20 μM cluster) was prepared in ammonium acetate, and injected through the capillary tube of the ESI part maintained at 2.7-4.5 kV, in positive ion mode. The nebuliser gas was set at 0.8 Bar, with a gas temperature of 180°C, and 4 L/min flow rate. The emitted ions in the gaseous phase entered the inlet inner capillary tube. WhiB1^{Ms} [4Fe-4S] cluster fragmentation was achieved by increasing the voltage (from 0 to 150 eV in 16 cycles) between the capillary exit and the hexapole skimmer of the spectrometer.

The data confirmed that the [4Fe-4S] form of WhiB1 was the predominant species present at the beginning of the experiment at 0 eV (Figure 3.21, the yellow line). However, after changing the capillary exit voltages, a series of peaks at different mass units lower than holo-WhiB1^{Ms} started to appear after each cycle. Analysis of each species suggested that the sequence of intermediates in the degradation of the WhiB1^{Ms} [4Fe-4S] was initiated by the loss of a sulfur atom as follows:

[4Fe-4S] → [4Fe-3S] → [3Fe-3S] → [2Fe-2S] → apo;

or [4Fe-4S] → [4Fe-3S] → [3Fe-3S] → [3Fe-2S] → [3FeS] → apo.

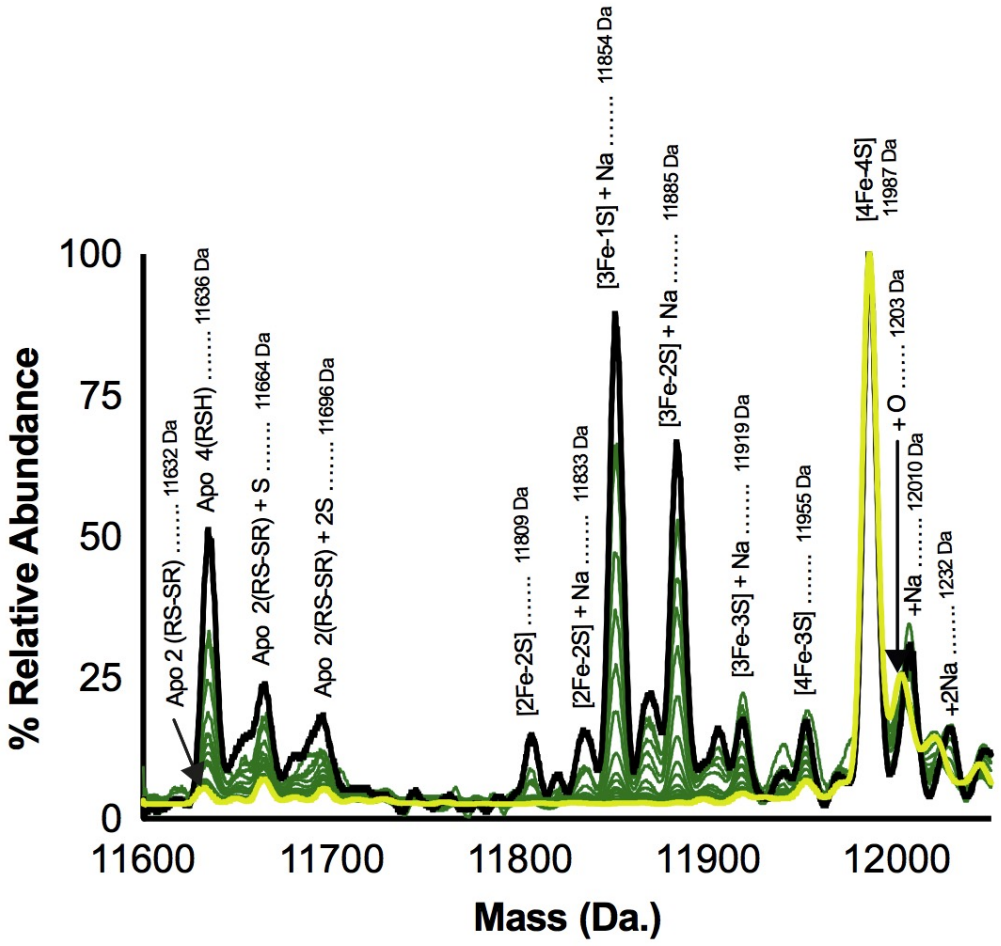


Figure 3.21: Electrospray ionization mass spectrometry (ESI-MS) analysis of WhiB1^{Ms}. The yellow line represents the intact [4Fe-4S]-WhiB1^{Ms}, before changing the capillary exit voltage, with a molecular mass of 11987 Da. The in-source collision-induced dissociation (isCID) voltage was increased from 0 to 150 eV, in the order of (0, 25, 50, 75, 85, 95, 100, 105, 110, 115, 120, 130, 140, 150 eV); where, the yellow line represents the first run at 0 eV, while the black line represent the last run at 150 eV. The WhiB1^{Ms} species that correspond to the masses detected are shown above the major peaks.

3.4.8 The *in vitro* reaction of holo-WhiB1^{Ms} with NO is not influenced by the presence of His$_6$-TEV-tag

Smith et al. (2010) reported that the [4Fe-4S] cluster of WhiB1, which was overproduced in *E. coli*, reacts very quickly with eight molecules of NO to produce a nitrosylated cluster. Thus, the sensitivity of WhiB1^{Ms} to NO was analysed to investigate the possible effect of His$_6$-TEV-tag on the cluster behavior when it reacts with NO. The reaction was carried out using the NO donor (Proli NONOate) as described in Section 2.6.5. WhiB1^{Ms} (44 μM cluster) was exposed under anaerobic conditions to 0.3 mM of Proli NONOate, which is equivalent to 0.6 mM of NO, for 1 h. After the incubation, the UV-visible spectrum of the protein showed an increase in absorbance at 360 nm combined with a decrease in the absorbance of 420 nm (the [4Fe-4S] absorbance) (Figure 3.22A). These changes suggest the formation of a nitrosylated iron-sulfur cluster complex in WhiB1 protein. To test how quick the reaction of WhiB1^{Ms} [4Fe-4S] cluster with NO, the absorbance at 360 nm was measured over time (Figure 3.22B); this was achieved by repeating the same reaction above, where the proli NONOate was injected into the protein sample in a sealed Hellma 10 mm cuvette, which had a screw-top lid to maintain anaerobic conditions, and the absorbance over time was read immediately after injection. Results showed that the change in the absorbance happened instantly during the period of Proli NONOate injection (Figure 3.21 B), suggested that the reaction was extremely rapid.

To estimate the ratio of reacted NO per [4Fe-4S] cluster of WhiB1^{Ms}, holo-WhiB1^{Ms} was incubated with increased amounts of Proli NONOate under anaerobic condition in a sealed cuvette as described in Section 2.6.5. Increased concentration of Proli NONOate (5.6-207 μM) (stock concentration 2.8 mM in 10 mM NaOH, equivalent to 5.6 mM of NO) were introduced to holo-WhiB1^{Ms} (44 μM cluster) in 35 additions, each one contained 11.2 μM of NO, except the last two additions where the concentration was doubled (Figure 3.23). Each addition of Proli NONOate was mixed firstly with (50 mM NaH_2PO_4, 100 mM NaCl, pH 7.4), and then directly injected into the protein sample via a gastight syringe (Hamilton). These mixtures were then incubated at 25°C for 3

min prior the spectroscopic measurements. The spectra showed an increase after each addition at 360 nm whilst simultaneously decreasing at 420 nm as shown in Figure 3.23A. These observations indicated the formation of nitrosylated iron-sulfur complex (Cruz-Ramos et al., 2002). Cumulative change in absorbance at 360 nm was calculated after each Proli NONOate addition, taking in to account the dilution factor. This was plotted against the NO:WhiB1[4Fe-4S] ratio (Figure 3.23B). This showed that the reaction ended at a [NO]/[FeS] ratio of 8.0. Suggesting the formation of an octa-nitrosylated iron-sulfur complex.

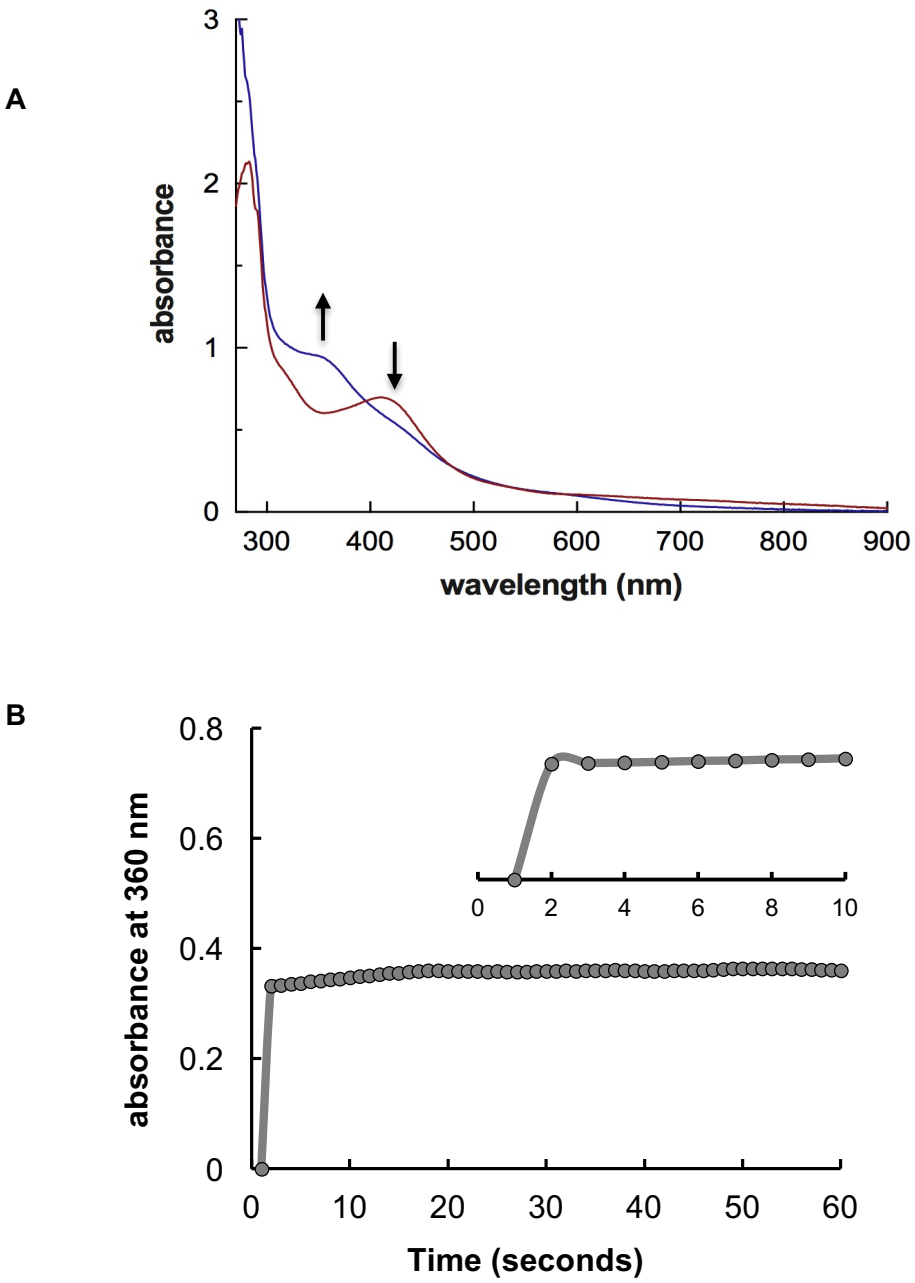


Figure 3.22: Reaction of WhiB1^{Ms} with NO. [A] UV-visible spectra of WhiB1^{Ms} after reaction with NO. WhiB1^{Ms} (44 μM cluster) before (red line) and after (blue line) adding 300 μM Proli NONOate (equivalent to 600 μM of NO) for 1 h were recorded; arrows indicate the spectra change direction. Proli NONOate was mixed firstly with assay buffer (50 mM NaH_2PO_4, 100 mM NaCl, pH 7.4), and then directly injected into the protein sample **[B]** The absorbance change at 360 nm over time after adding the NO donor (600 μM of NO) into WhiB1^{Ms} (44 μM cluster). The protein containing buffer was 50 mM NaH_2PO_4, pH 7.4 containing 500 mM NaCl and 5% glycerol.

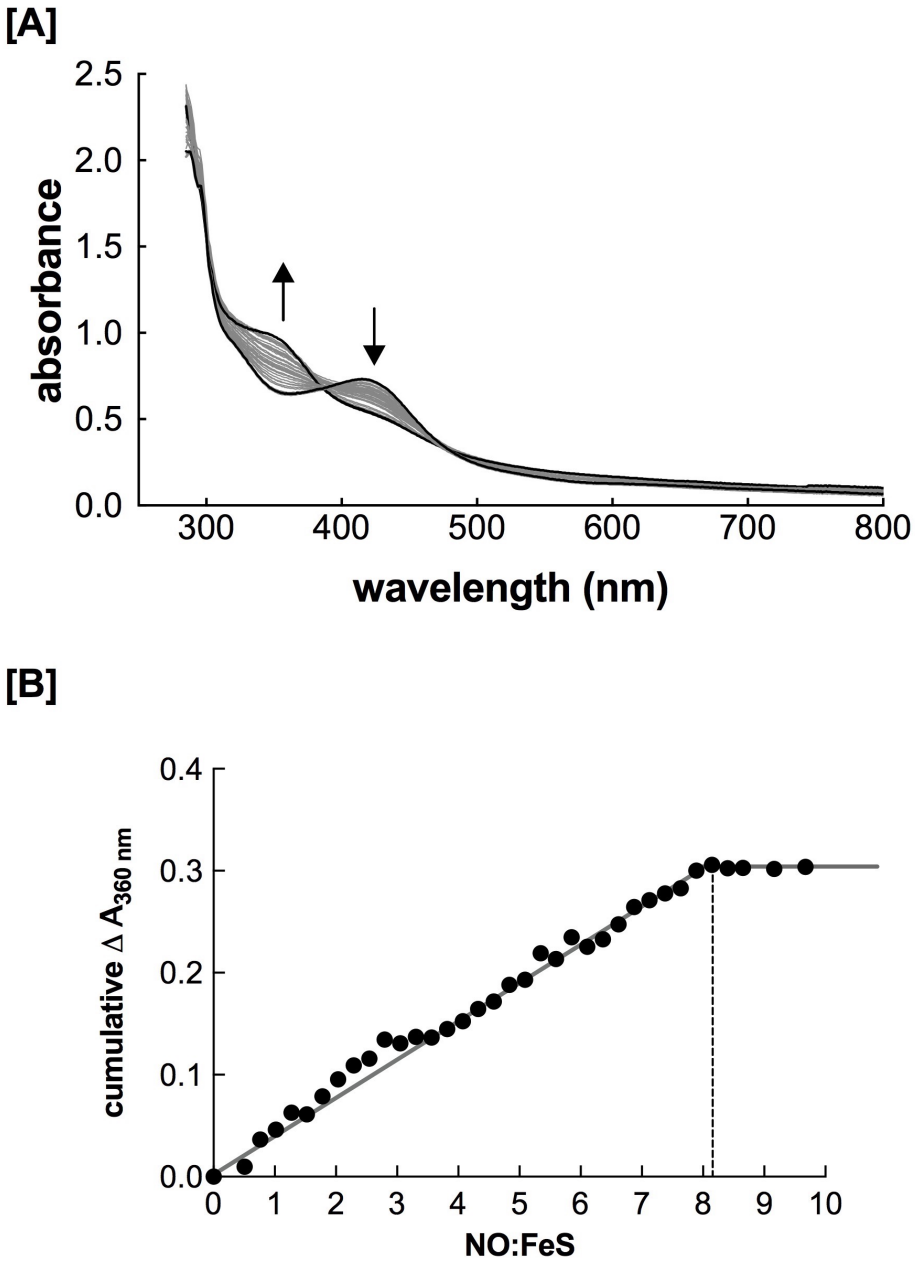


Figure 3.23: Stoichiometry of the reaction between WhiB1^{Ms} [4Fe-4S] and NO. [A] UV-visible spectra following NO reaction with WhiB1^{Ms}. Increasing amounts (5.6-207 μM) of Proli NONOate (equivalent to 11.2-414 μM of NO) were introduced to WhiB1^{Ms} (44 μM cluster) in 35 additions. The arrows refer to the direction of change in the absorbance. Each spectrum was taken 3 min after NO addition. **[B]** The cumulative changes in absorbance at 360 nm were plotted versus the NO:[4Fe-4S] ratio. The dashed line indicates the point at which no further change in absorbance was observed.

3.5 Gel filtration analysis of WhiB1

Size exclusion chromatography was used to determine the oligomeric state of apo- and holo-WhiB1^{Ec} and WhiB1^{Ms} as described in Section 2.5.6. The HiLoad 16/600 Superdex column was calibrated using a group of known molecular weight proteins (Figure 3.24 A). The buffer of 50 mM Tris-HCl (pH 8.0) containing 0.5 M NaCl was left inside an anaerobic cabinet for 24 h whilst being stirrerd. It had then degassed under vacuum for 1 h, before 10 mM of DTT was added, and the buffer immediately used for column equilibration. This helped to maintain low O_2 conditions inside the column, and prevented the slow cluster dissociation of WhiB1^{Ec}; however, the same conditions were also used for both apo-WhiB1^{Ec} and WhiB1^{Ms}. Each sample (0.4 mM, 2 ml) was applied into the column, the elution profile of protein trace of each sample showed a similar peak position at 92 ml (Figure 3.25). UV-visible spectrum of eluted holo-WhiB1^{Ec} showed a strong peak at 420 nm, suggesting that the protein was retained the cluster. To estimate the molecular weight of WhiB1, the standard curve of calibration proteins (Figure 3.24B) was used. An estimate of ~11 kDa was obtained, which suggest that WhiB1 is a monomer (the actual molecular weight of apo-WhiB1^{Ec} is 11.48 kDa, and WhiB1^{Ms} is 11.63 kDa) in the presence or absence of its [4Fe-4S] cluster.

A

Standard	Mwt (kDa)	Kav value
Aprotinin	6.5	0.78
Ribonuclease	13.7	0.66
Carbonic anhydrase	29	0.56
Ovalbumin	43	0.46
Conalbumin	75	0.39
Aldolase	158	0.24
Ferritin	440	0.1
Thyroglobulin	669	0.04

B

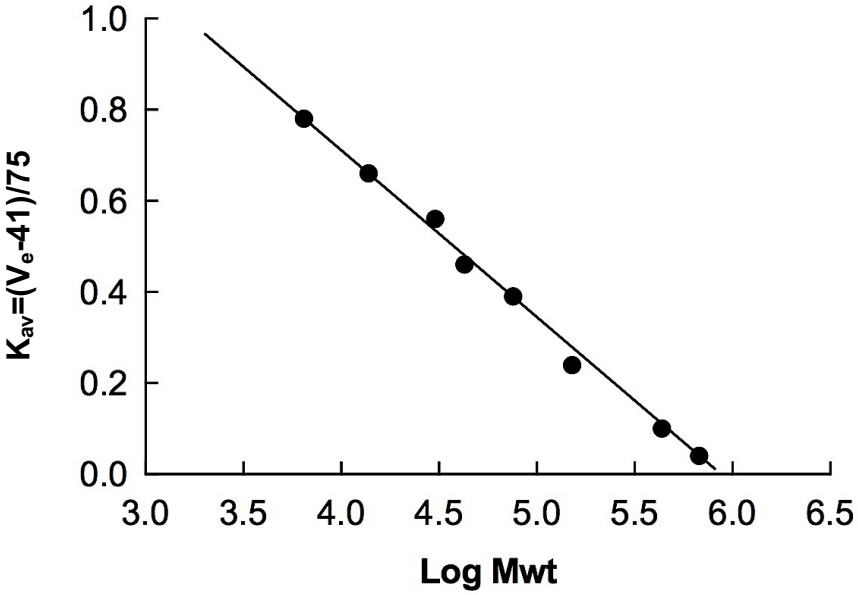


Figure 3.24: Gel filtration standard curve. [A] List of known molecular weight (Mwt) molecules used to create the calibration curve. [B] The calibration curve. The K_{av} value [$K_{av} = (V_e - V_0)/(V_t - V_0)$] was plotted against log molecular weight (Mwt) of the protein in table [A]. The WhiB1 K_{av} equal to 0.68.

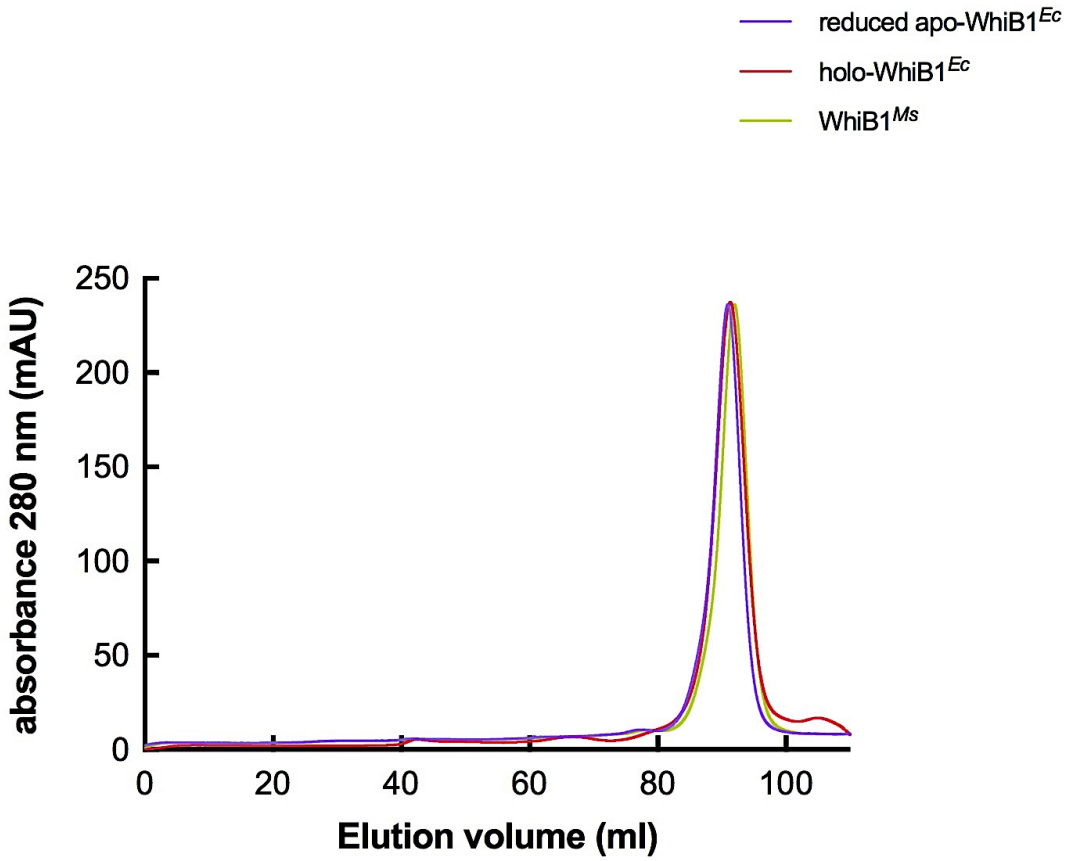


Figure 3.25: Size-exclusion chromatography profile of apo- and holo-WhiB1^{Ec} or WhiB1^{Ms}. WhiB1 (0.4 mM) was eluted as a single peak at 92 ml (**K_{av}** equal to 0.68) from HiLoad 16/600 Superdex column pre-equilibrated with anaerobic buffer of 50 mM Tris-HCl (pH 8.0) containing 0.5 M NaCl and 10 mM DTT.

3.6 Discussion

In vitro biochemical and biophysical characteristics of *M. tuberculosis* WhiB1 iron-sulfur clusters have been studied in this chapter. WhiB1 overproduction and isolation were achieved for the first time using a *M. smegmatis* expression system in addition to *E. coli*. This was done to determine the nature of the WhiB1 cluster in a mycobacterial host, to study its sensitivity to O_2 and NO, in addition to its role in structuring the protein.

WhiB1 overproduced in *E. coli* with a significantly shorter *N*-terminal tag than that used in previous studies (WhiB1^{Ec}) (Smith et al., 2010) adopted three different protein forms [4Fe-4S]- [2Fe-2S]- and apo-forms (Figure 3.26A and B); as judged by UV-visible spectra (Figure 3.3). In addition, the *in vitro* reconstitution of WhiB1^{Ec} iron-sulfur cluster under anaerobic conditions indicated that the [4Fe-4S] cluster was the final form present in WhiB1^{Ec}, which was confirmed by estimation of total iron content (Figure 3.6). These results are consistent with previous for WhiB1 (Alam et al., 2009; Smith et al., 2010), and for other *Mtb* Wbl proteins such, WhiB3 (Singh et al., 2007), WhiB4 (Alam et al., 2007), and WhiB7 (Burian et al., 2013). Interestingly, the [2Fe-2S] WhiB1^{Ec} form could not be observed by UV-visible spectra during the WhiB1^{Ec} [4Fe-4S] degradation in the presence of O_2 (Figure 3.7A), or even during the [4Fe-4S] cluster reconstitution (Figure 3.4A). According, to these data, it has been suggested initially that the incomplete acquisition of iron could be a reason for the appearance of the [2Fe-2S] WhiB1^{Ec} (Smith et al., 2010). Supplementation of overproduction LB medium with 0.5 mg/l ferric chloride, resulted in all eluted fractions from nickel charged column as [4Fe-4S]- or [2Fe-2S]- forms without any apo-form (Appendix 2, Figure S.2A); increasing the concentration of iron source to 5 mg/l did not change this result, indicating that the presence of apo-form is possibly due to iron starvation or cluster degradation, but the [2Fe-2S] WhiB1 was not formed because of *in vitro* cluster redox degradation or *in vivo* iron starvation.

The results of the current study suggest that the [2Fe-2S] cluster form observed during WhiB1^{Ec} protein isolation might arise from the presence of different iron-sulfur cluster biogenesis system(s), where the scaffold and

coordination of iron-sulfur cluster systems could be different in *E. coli* compared to *Mycobacterium sp.,* which might lead to this variation in the cluster forms (Py and Barras, 2010).

The sensitivity and the stability of WhiB1^{Ec} [4Fe-4S] cluster to O_2 in the presence or absence of reducing agents has been also investigated. The WhiB1^{Ec} [4Fe-4S] cluster was only degraded after prolonged incubation in air (Figure 3.7). Similar degradation rates were observed for *M. tuberculosis* WhiB3, where 75% of the WhiB3 [4Fe-4S] cluster was lost within 180 min after introducing air bubbles through the protein for 2 min, and this converted the WhiB3 [4Fe-4S] cluster to a [2Fe-2S] (Singh et al., 2007). In contrast, the *Streptomyces coelicolor* [4Fe-4S] WhiD (the homolog of *Mtb* WhiB3) reacted similarly to WhiB1^{Ec} with air-mediated degradation proceeding without a [2Fe-2S] intermediate (Crack et al., 2009).

Smith et al., (2010), reported that the [4Fe-4S] cluster of WhiB1 was not degraded after 2 h of incubation with O_2-saturated buffer in the presence of DTT. Result of this chapter confirmed that the DTT is the main reason behind the cluster stability of reconstituted [4Fe-4S] WhiB1^{Ec}; cluster degradation over 16 h in the presence of O_2 and DTT showed less than 10% cluster loss (Figure 3.8). The same underestimation rate of the cluster degradation might be behind the stability of *Mtb* WhiB4 cluster, which was reported as a stable over 48 h of air oxidation in the presence of DTT by Alam et al., (2007). Similar observations were made by Jakimowicz et al., (2005), for *Streptomyces coelicolor* [4Fe-4S] WhiD, and they argued that in the absence of DTT the cluster degradation rate was much faster, and cluster conversion intermediates were not detected by absorption spectroscopy. It was also reported that the stabilizing effect of DTT on the [4Fe-4S] cluster of a non-Wbl protein, the *Bacillus subtilis* NsrR, was due to the ability of DTT to play a role as an exogenous ligand which can bind with the cluster during the cluster degradation (Yukl et al., 2008). Perhaps DTT interacts with holo-WhiB1^{Ec} iron or sulfur atoms through its thiol groups leading to stabilization of the cluster. Accordingly, when a non-thiol reducing agent (TCEP-HCl) was used (Figure 3.8), cluster stability was lowered. Replacement of a cysteine ligand by a DTT ligand is shown in Figure 3.26D.

WhiB1 overproduction in *M. smegmatis* (WhiB1^{Ms}) with an *N*-terminal His_6-TEV tag, rather than His_6-Thrombin tag (WhiB1^{Ec}) showed that only an O_2-insensitive [4Fe-4S] form was isolated (Figure 3.13). Results also showed that the secondary structural properties of holo-WhiB1^{Ms} were more ordered than the characterized apo- and holo-WhiB1^{Ec} (Figure 3.14); however, the visible CD-spectra properties of the [4Fe-4S] cluster region in both holo-WhiB1^{Ms} and holo-WhiB1^{Ec} were similar (Figure 3.15), indicating that the cluster environment is similar in both proteins. Complete removal of the WhiB1^{Ms} His-TEV-tag by TEV protease treatment was not possible when the protein held a [4Fe-4S] cluster, but this was possible for the apo-protein (Figure 3.16). Reconstituted untagged WhiB1^{Ms} lost cluster stability when it was exposed to O_2; in contrast to the reconstituted His_6-TEV-tagged form (Figure 3.17; 3.27). In addition, removing the tag changed the far UV-CD spectral properties, which became similar to WhiB1^{Ec} (Figure 3.18). All these data provide strong biochemical and biophysical evidence that the His-TEV-tag of WhiB1^{Ms} contributes to the stability of the cluster under aerobic conditions. Thus, this study suggests that, *in vivo,* the properties of the WhiB1 cluster could be modified by a natural partner protein(s), and the TEV-tag site could mimic these interactions.

Burian et al., (2013), reported that the WhiB7 protein when co-expressed with RpoV protein in *E. coli*, WhiB7 protein was eluted from nickel-charged column as a complex with a brown colour and showed only a single peak at 420 nm of the UV-visible spectrum, indicating, that the interaction between [4Fe-4S]-WhiB7 and RpoV stabilized the iron-sulfur cluster of WhiB7.

When WhiB2 was overproduced and isolated with the presence of His_6-TEV-tag using the *M. smegmatis* expression, two cluster forms [4Fe-4S] and [2Fe-2S] were found in the isolated protein, suggesting that the presence of TEV cleavage tag did not stabilize the WhiB2^{Ms} cluster (Appendix 4). Indicating, that the TEV cleavage site possibly has a unique interaction site within WhiB1.

In vitro reconstitution of isolated His_6-TEV-WhiB2^{Ms} produced only [4Fe-4S] form, and after protein treatment with 110 µM O_2 in the absence of reducing agents, the [4Fe-4S] cluster absorbance at 420 nm declined by half within ~4 h (Appendix 4, Figure S.4A); followed by an increase in the protein absorbance

to the original level but with the characteristics of [2Fe-2S] spectrum (Appendix 4, Figure S.4A). The [2Fe-2S] form was more stable than the WhiB2^{Ms} [4Fe-4S] form. WhiD *Streptomyces coelicolor* showed a similar spectrum alteration during [4Fe-4S] cluster degradation after O_2 exposure (Jakimowicz et al., 2005). In contrast, WhiB1^{Ms} did not show a [2Fe-2S] form even after TEV cleaved, reconstituted WhiB1^{Ms} was exposed to O_2 (Figure 3.17A).

Moreover, LC-MS analysis of WhiB1 showed that both WhiB1^{Ec} (Figure 3.9) and WhiB1^{Ms} (Figure 3.20A) did not retain sulfur adducts upon cluster degradation and were mainly found in reduced forms or as oxidized intra-molecular S-S bridge forms. In contrast, WhiB2^{Ms} LC-MS analysis (Figure 3.20B) showed the presence of up to four extra sulfur adducts attached to the apo-form after cluster dissociation; suggesting that when the WhiB2^{Ms} [4Fe-4S] converted to the [2Fe-2S] form, the protein firstly lost two iron atoms and then cysteine residues gained extra one or two sulfur atoms from the original broken cluster to form cysteine persulfide ligands (Zhang et al., 2012; Ibrahim et al., 2015), as follows:

$[4Fe\text{-}4S\ (CysS)_4] \rightarrow [2Fe\text{-}2S\ (CysS)_3\ (CysSS)_1] \rightarrow apo\ (CysSS)_{0,\ 1,\ 2,\ 3\ or\ 4}$
or $[4Fe\text{-}4S\ (CysS)_4] \rightarrow [2Fe\text{-}2S\ (CysS)_2\ (CysSS)_2] \rightarrow apo\ (CysSS)_{0,1,2,3\ or\ 4}$

In contrast, the isCID-MS analysis of WhiB1^{Ms} [4Fe-4S] cluster (Figure 3.21; 3.28) showed that the cluster dissociation pathway started by losing the sulfur atoms first then iron as follows:

$[4Fe\text{-}4S] \rightarrow [4Fe\text{-}3S] \rightarrow [3Fe\text{-}3S] \rightarrow [2Fe\text{-}2S] \rightarrow apo;$
or $[4Fe\text{-}4S] \rightarrow [4Fe\text{-}3S] \rightarrow [3Fe\text{-}3S] \rightarrow [3Fe\text{-}2S] \rightarrow [3FeS] \rightarrow apo.$

It is concluded that WhiB1 in the absence of a partner protein/stabilizing tag under aerobic conditions, loses its [4Fe-4S] cluster without forming intermediate cluster forms, unlike the WhiB2 protein; suggesting different possible functions of each *Mycobacterium* Wbl proteins.

Proton 1D NMR spectrum of holo-WhiB1^{Ms} showed that the presence of the [4Fe-4S] cluster was essential for structuring the protein (Figure 3.19), CD spectra of holo-WhiB1^{Ec} also showed that the presence of the cluster changed

the spectral properties compared to apo-WhiB1^{Ec} (Figure 3.14). Similar results were reported for *Mtb* apo-reduced or oxidized forms of WhiB1, WhiB2, WhiB3, WhiB4 and WhiB7 which exhibit a large proportion of disordered structure (Garg et al., 2007; Alam et al., 2007, 2009). However, the presence of the cluster in *Streptomyces coelicolor* WhiD changed the secondary structural properties similarly to that observed for WhiB1 (Crack et al., 2009). Reconstituted [4Fe-4S] WhiB2^{Ms} and WhiB3^{MS} (both with His-TEV-tag with O_2-responsive clusters) under anaerobic conditions showed CD-spectra features dominant at the region of α-helices (208 nm and 222 nm) (Appendix 5).

Several iron-sulfur cluster containing proteins were reported to act as a NO sensor proteins, where the reaction of the clusters with NO led to changes in protein conformation, and altered the DNA-binding properties (Cruz-Ramos et al., 2002; Ding and Demple, 2000; Smith et al., 2010; Tucker et al., 2008). The [4Fe-4S]-WhiB1 protein was found to react with NO very rapidly (milliseconds), and the reaction led to an octa-nitrosylated cluster formation (Dimerization of Roussin's red ester cluster) (Crack et al., 2010; Smith et al., 2010). Similar findings were reported in *Streptomyces coelicolor* WhiD protein (Crack et al., 2010). This study confirmed that [4Fe-4S]-WhiB1^{Ms} reacts very quickly with eight molecules of NO, and the TEV-tag does not modify cluster behavior towards NO reaction (Figure 3.22, 3.26 E and F, 3.29).

Size exclusion chromatography analysis of reduced apo- and holo-WhiB1^{Ec}, and holo- WhiB1^{Ms} showed that WhiB1 was eluted from gel filtration column at a molecular weight of ~11 kDa (Figure 3.25), indicating that the holo- and apo-WhiB1 are monomers; holo-WhiB2^{Ms} and holo-WhiB3^{Ms} also were eluted from the column as monomer molecules (Appendix 3). Similar results are reported in holo-WhiD (Crack et al., 2009).

Results presented in this chapter indicate that WhiB1 is indeed a protein that has a [4Fe-4S] cluster, which is not sensitive to O_2 but very sensitive to NO. The biochemical and biophysical assays show clearly that cluster is essential for folding the protein, and the presence of TEV tag changes the stability of this cluster. Thus, studying WhiB1 cluster stability and role when WhiB1 interacts

with SigA, and the structure properties of holo-WhiB1 will be the subject of the next two chapters.

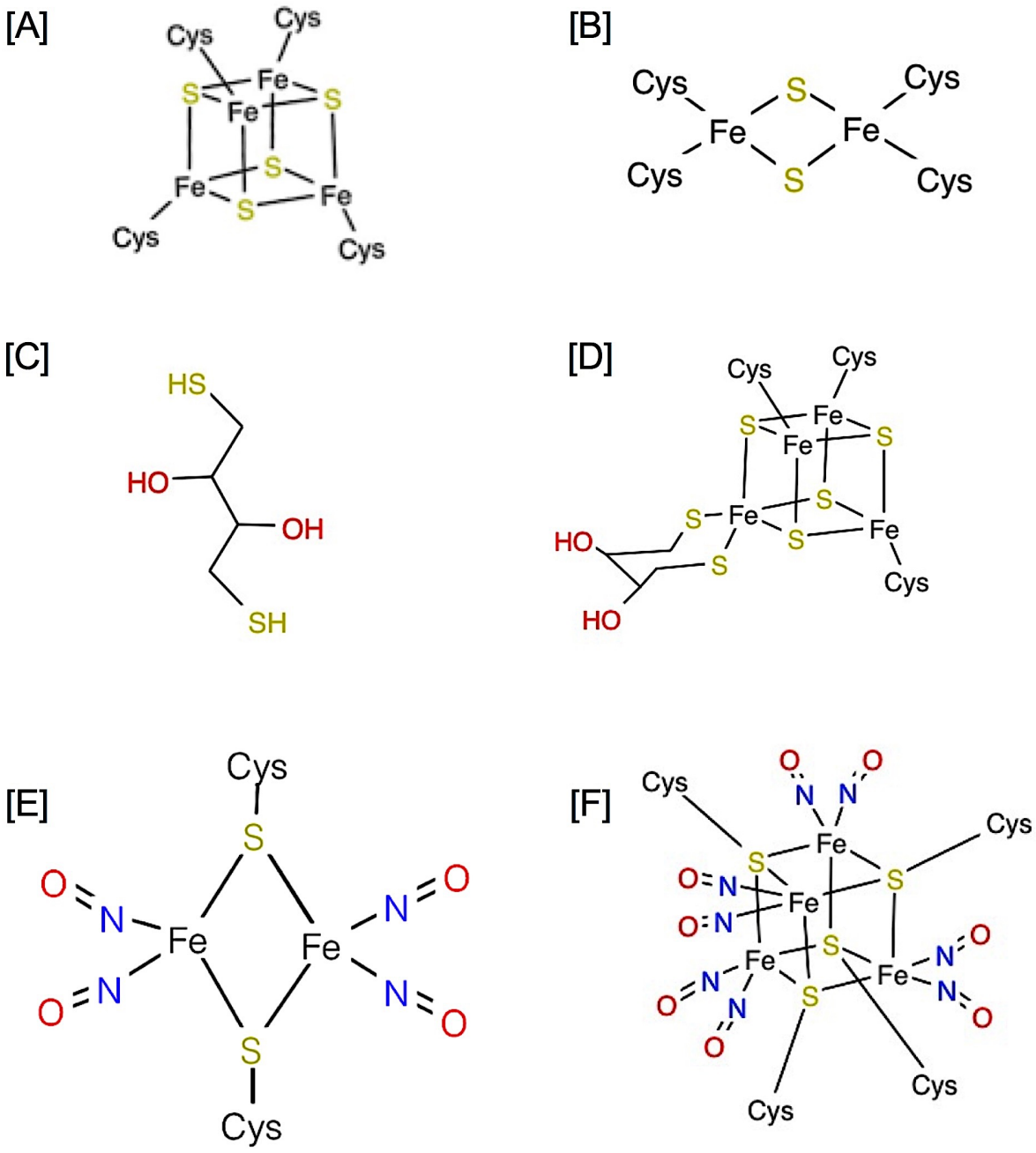


Figure 3.26: Schematic representation of biological FeS clusters and molecules used in their characterization. [A] and **[B]**, [4Fe-4S] cluster and [2Fe-2S] cluster respectively liganded by four cysteine residues; **[C]** Dithiothreitol (DTT); **[D]** [4Fe-4S] cluster liganded by three cysteine residues and one exogenous DTT ligand, adapted from (Hinckley and Frey, 2006); **[E]** Roussin's red ester cluster with bridging cysteine residues; **[F]** Octa-nitrosylated [4Fe-4S] cluster. All were generated via online CORINA Interactive Demo tools (https://www.mn-am.com/online_demos/corina_demo_interactive).

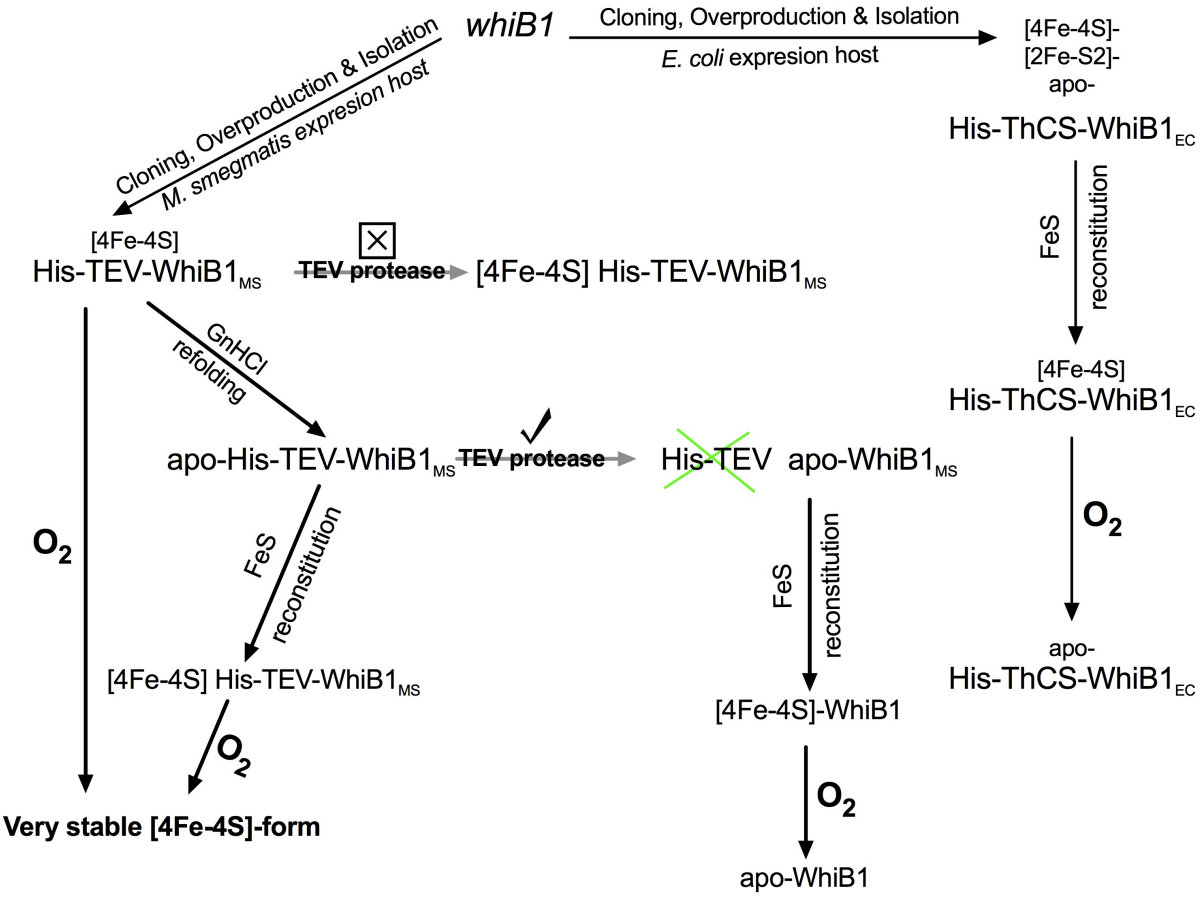


Figure 3.27: Various forms of WhiB1 analyzed in this work and their sensitivity to O_2. Overproduction and isolation of WhiB1^{Ec} in *E. coli* resulted three different forms of apo-, [2Fe-2S]-, and [4Fe-4S]-WhiB1. The [2Fe-2S] and [4Fe-4S] forms can be degraded in the presence of O_2 to apo-form, which in turn under anaerobic conditions can be reconstituted again to [4Fe-4S]-form. WhiB1^{Ms} overproduced and isolated in *M. smegmatis* has only [4Fe-4S] cluster, which is not degraded in the presence of O_2. Removing the TEV-tag in the presence of the cluster by TEV protease is not possible and can only be digested by removing the cluster after refolding the protein. The TEV cleaved apo-WhiB1 can anchor the [4Fe-4S] cluster after a cluster reconstitution reaction, the resulted untagged holo-WhiB1 however can lose the cluster after a prolong incubation in the presence of O_2 similarly as holo-WhiB1^{Ec}. In contrast, refolding the protein without cleaving the TEV tag and then reconstitute the cluster resulted a very stable iron-sulfur cluster in the protein similarly as the native WhiB1$^{Ms.}$

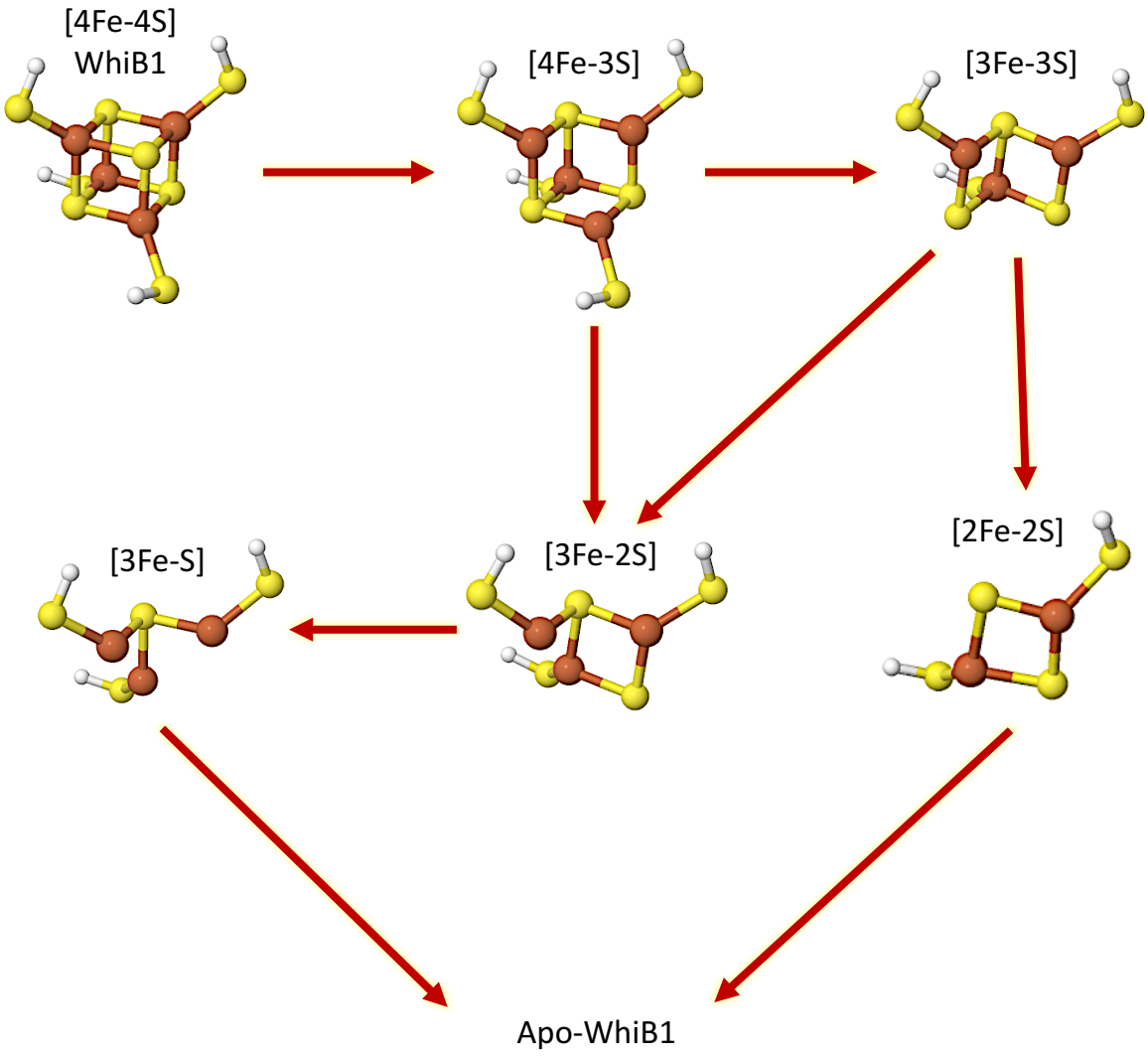


Figure 3.28: Cluster dissociation pathway of WhiB1^{Ms} after changing the capillary voltage exist of isCID-mass spectrometry. The 3D structure models of iron-sulfur clusters were generated via online CORINA Interactive Demo tools (https://www.mn-am.com/online_demos/corina_demo_interactive).

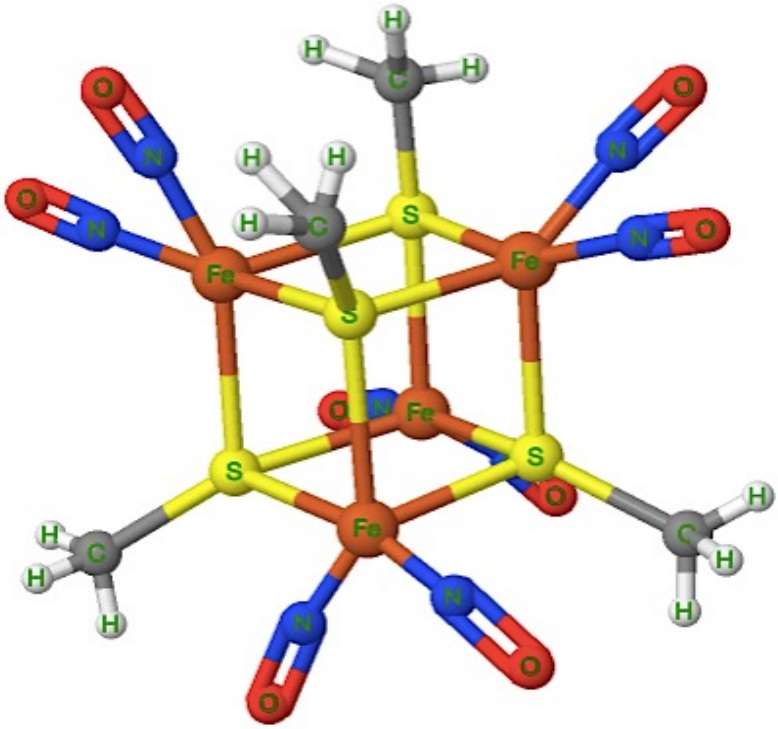


Figure 3.29: Putative model of octa-nitrosylated cluster resulted from the reaction of [4Fe-4S] WhiB1^{Ms} with NO. The model was generated via online CORINA Interactive Demo tools (https://www.mn-am.com/ online_demos/corina _demo _interactive).

Chapter four

Structure-function relationships of WhiB1:SigA complex

4.0 Structure-function relationships of WhiB1:SigA complex

4.1 Introduction

Mycobacterium tuberculosis encodes thirteen different sigma (σ) factors designated as SigA (σ^A) to SigM (σ^M) (Manganelli et al., 2004). SigA, a member of the Sig^{70} (σ^{70}) family, is the essential sigma factor in both *M. tuberculosis* and *M. smegmatis* (Gomez et al., 1998; Manganelli et al., 2004), other sigma factors are non-essential and are only produced in response to specific stress conditions (Wösten, 1998). The Sig^{70} family share four designated regions (1 to 4), each divided in to sub-regions, some of which (4.2, 2.3, 2.4 and 2.5) are required to recognize elements (-35, -10, and extended -10) of targeted promoters (Dove et al., 2003; Gruber and Gross, 2003; Panaghie et al., 2000; Sanderson et al., 2003).

In Mycobacteria, Wbl proteins were found to interact with SigA. For example, SigA protein of *M. bovis* and *M. tuberculosis* interact with WhiB3. Moreover, it was reported that the *M. bovis* SigA R515H mutant was attenuated due to impaired interaction between SigA and WhiB3 (Steyn et al., 2002). Similar observations were made by Burian et al., 2013, for *M. tuberculosis* WhiB7 interacting with region 4 of SigA via R515. A recent study showed that all *M. tuberculosis* Wbl proteins except WhiB5 interact with SigA, and WhiB3, 4, 6, and 7 all interact via the SigA R515 residue, but WhiB1 and WhiB2 do not, suggesting a different binding site is involved (Feng et al., 2016). However, none of the above reports studied the role of the iron-sulfur cluster of the Wbl proteins in interaction with SigA. Moreover, the binding interface between WhiB1 and SigA is not known. Here, *M. tuberculosis* WhiB1 and SigA binding domains were defined, the biochemical and biophysical characteristics of WhiB1 iron-sulfur cluster in the complex were also studied, particularly the reaction of [4Fe-4S]-WhiB1:SigA with NO.

4.2 Isolation, biochemical and biophysical characteristics of the WhiB1:SigA complex

4.2.1 Overproduction and purification of the WhiB1:SigA complex

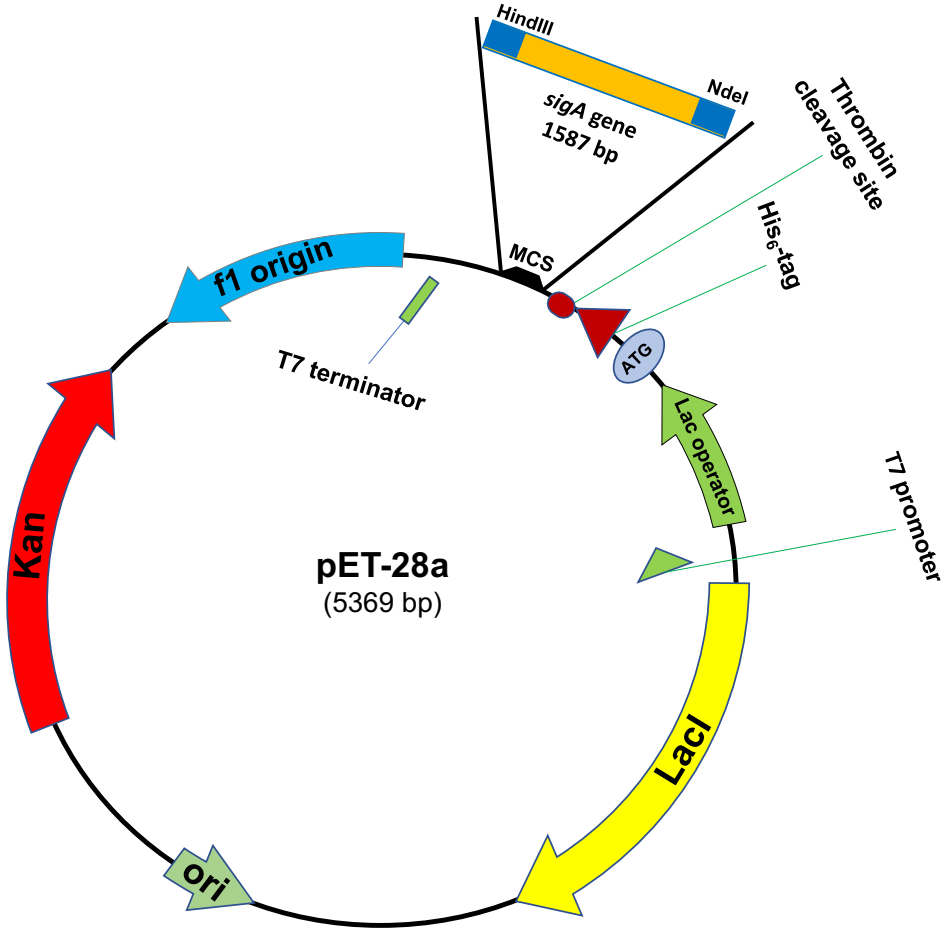
WhiB1:SigA complex overproduction was achieved using *E. coli* JRG6859. This strain was created by transforming *E. coli* JRG5302 with pGS2183, encoding SigA with an *N*-terminal His_6-tag and a linker containing a Thrombin cleavage site (His_6-ThCS-SigA), and pGS2560, encoding WhiB1 without a tag (Section 2.3.14; Figure 4.1 and 4.2). Previous studies of Wbl:SigA complexes were done using His-tagged Wbl proteins and non-tagged SigA (Burian et al., 2013; Feng et al., 2016), and it was reported that the untagged SigA was able to bind nickel-charged column beads resulting in non-specific SigA binding along with specific binding of Wbl proteins. Non-specific SigA binding complicated the interpretation of the data obtained (Feng et al., 2016). By tagging SigA and knowing that non-tagged WhiB1 cannot bind nickel-charged beads (Section 3.4.4) SigA:WhiB1 interaction were more easily detected.

Overproduction of the WhiB1:SigA complex was achieved by growing *E. coli* JRG6859 in LB broth, and the cultures were induced with a final concentration of 1 mM IPTG with incubation for 12-18 h at 18°C and an agitation of 250 rpm (Section 2.5.3.4).

For WhiB1:SigA complex isolation, cell-free extract was prepared by re-suspending cell pellets in binding buffer consisting of 500 mM NaCl, 50 mM Tris pH 8.0, and then sonicated according to Section 2.5.4. Subsequent protein purification was as described in Section 2.5.5.1.

The elution profile from the HiTrap chelating column showed only two peaks, the first peak after the imidazole gradient was formed from non-specifically bound materials, usually representing proteins with high histidine contents, while the second peak contained brown coloured WhiB1:SigA complex protein (Figure 4.3A). Fractions were analyzed by SDS-PAGE gel, which confirmed the presence of the complex (Figure 4.3B).

A



B

MGSSHHHHHHSSGLVPR↓GSH- SigA

His_6-tag Thrombin cleavage site

Figure 4.1: Schematic diagram of pGS2183. [A] The *sigA* open reading frame (1587 bp) was amplified by PCR from genomic DNA of *M. tuberculosis* with engineered NdeI/HindIII restriction sites, and ligated into pET28a to create pGS2183, this was done by Dr. Roger Buxton (NIMR Mill Hill, London). **[B]** The plasmid encodes a His_6-ThCS-SigA fusion protein. The arrow marks the DNA coding for the peptide bond susceptible to cleavage by Thrombin.

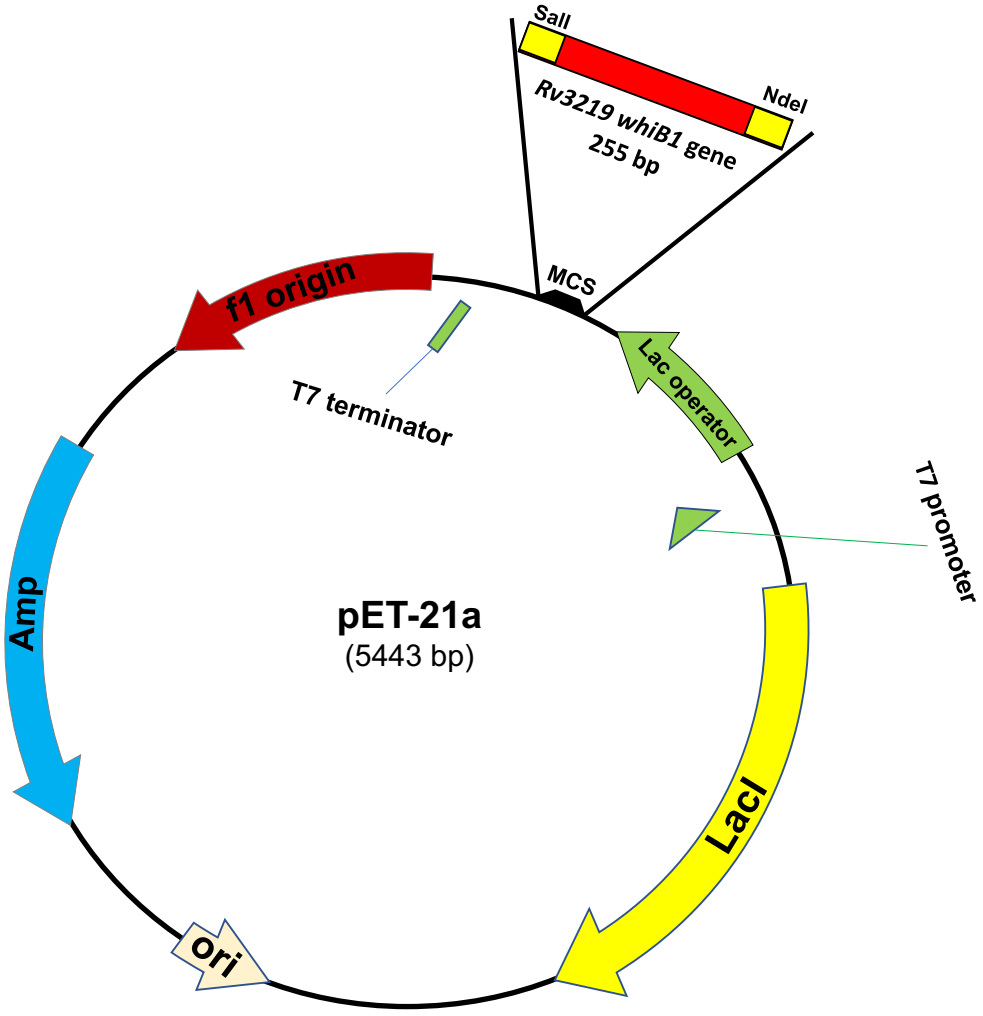


Figure 4.2: Schematic diagram of pGS2560. [A] The *whiB1* open reading frame (255 bp) was amplified by PCR from genomic DNA of *M. tuberculosis* with engineered NdeI/SalI restriction sites, and ligated into pET21a to create pGS2560. The plasmid encodes a non-tagged WhiB1.

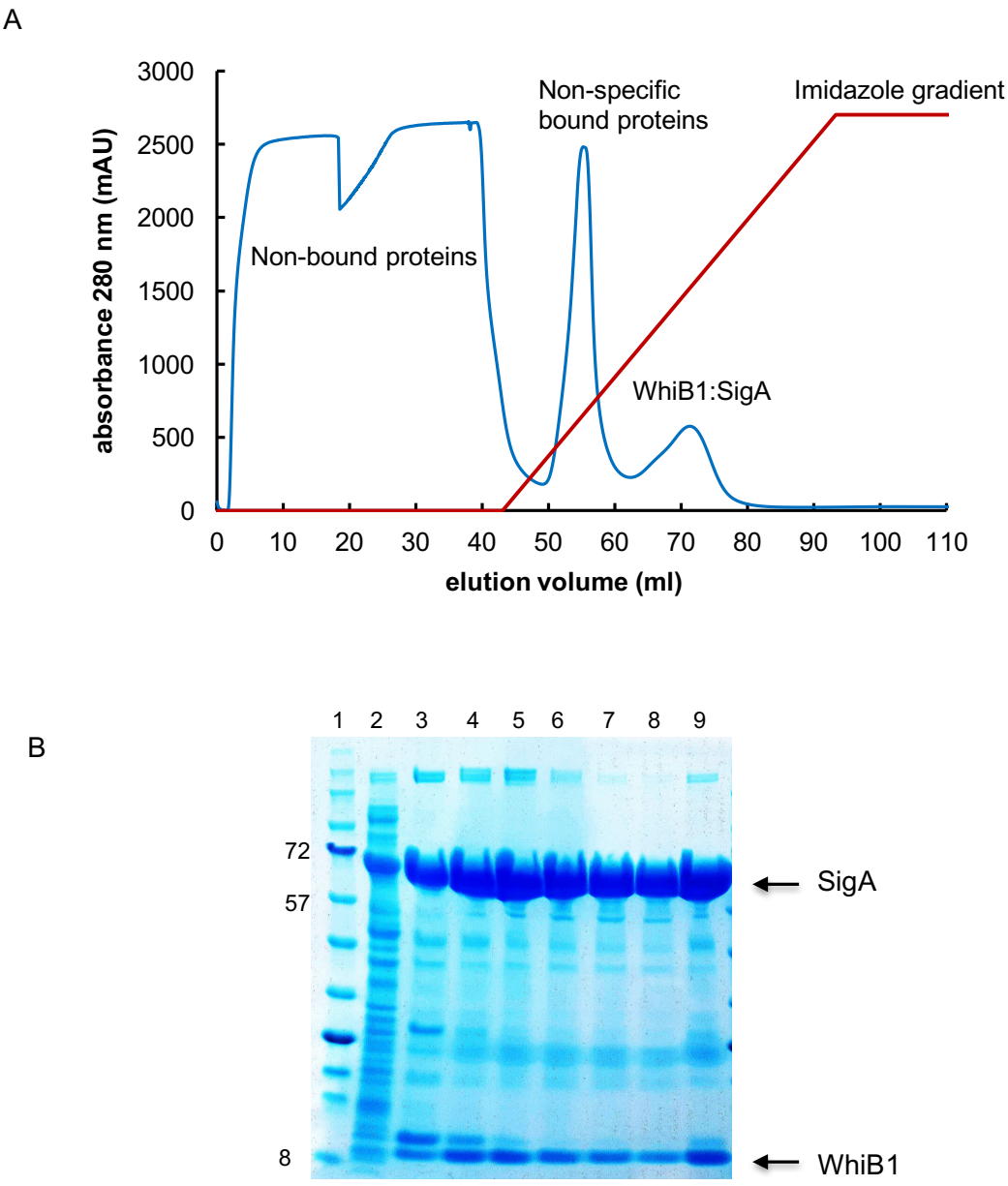


Figure 4.3: [A] Elution profile of HiTrap chelating column of *E. coli* cell lysate containing overproduced WhiB1 and SigA. The trace shows protein absorbance at 280 nm (blue line) of injected cell lysate, and the gradient concentration of imidazole from 0-0.5 M (red line). Eluted WhiB1:SigA complex appears as a single peak between ~62-78 ml. **[B] Coomassie stained SDS-PAGE gel of isolated fractions from the elution profile of [A].** Lane 1, protein ladder (sizes in kDa are indicated); lane 2 cell free-extract of induced WhiB1 and SigA proteins; lanes 3-9, fractions of WhiB1: SigA complex eluted between ~62-78 ml, as indicated in **[A]**.

4.2.2 Size exclusion chromatography analysis of WhiB1:SigA complex

Because SigA was engineered with *N*-terminal His_6-tag and WhiB1 without tag, it was assumed that after the complex isolation by nickel-charged HiTrap chelating column, some molecules of free SigA protein could also bind on column beads and elute with fractions of WhiB1:SigA complex. Thus, size exclusion chromatography analysis was carried out to get a pure sample of the complex for further experiments, and to test the stability of the complex after passage through the column. HiLoad 16/600 Superdex 200 gel filtration column pre-equilibrated with 50 mM Tris (pH 8.0) containing 0.5 M NaCl aerobic buffer was used (Section 2.5.6). The isolated WhiB1:SigA (0.29 mM) protein sample from HiTrap chelating column was applied directly to the column with 1.5 ml/min flow rate. The elution profile showed two peaks at 42 ml and 54 ml (Figure 4.4A). The protein concentration in the first peak was about 10x less compared to the second peak as judged by Bradford assays (Section 2.5.1); however, the first peak showed a high UV absorbance as indicated in the elution profile trace. SDS-PAGE gel of the eluted fractions showed that the first peak represented free SigA protein (lane 4, inset, Figure 4.4A), while the second peak was WhiB1:SigA complex (lanes 5 and 6, inset, Figure 4.4A). UV-visible spectra showed that the first peak absorbed the light at 260 nm, while the second had two features at 280 nm and 420 nm (Figure 4.4A). It appears that the free SigA protein was enriched with nucleic acids, supported by agarose gel electrophoresis and staining for nucleic acids, where the sample of protein showed a DNA smearing (lane 2, inset, Figure 4.4B). The presence of a feature at 420 nm in the second eluted peak indicates that the WhiB1:SigA sample had an iron-sulfur cluster. However, the sample of the complex visualised on the agarose gel showed a very faint smearing (lane 3, inset, Figure 4.4B), suggesting that there is less DNA bound to SigA:WhiB1 complex.

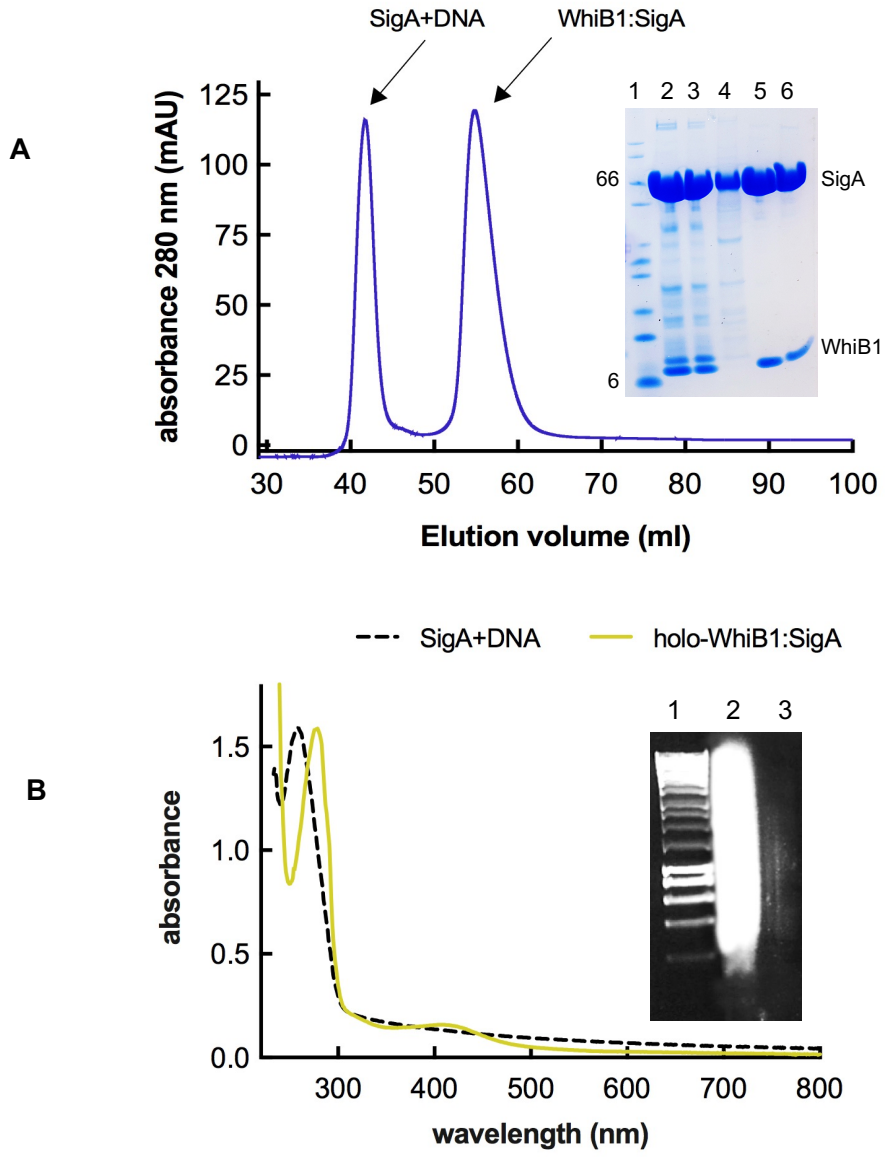


Figure 4.4: [A] Size-exclusion chromatography of isolated WhiB1:SigA complex. Free SigA protein was eluted as a single peak at 42 ml**,** while the complex of WhiB1:SigA was eluted at 54 ml from HiLoad 16/600 Superdex column pre-equilibrated with aerobic buffer of 50 mM Tris-HCl (pH 8.0) containing 0.5 M NaCl. ***Inset***, Coomassie stained SDS-PAGE gel of eluted fractions of **[A].** Lane 1, protein ladder (sizes in kDa are indicated); lanes 2-3, WhiB1:SigA samples loaded onto gel filtration column, lane 4 fraction of free SigA eluted at 42 ml position; lanes 5-6, fractions of WhiB1:SigA complex eluted at 54 ml position, as shown in **[A]**. **[B] UV-visible spectra of free SigA and WhiB1:SigA complex.** ***Inset,*** gel red stained agarose gel of eluted fractions in the elution profile of **[A].** Lane 1, DNA ladder; lane 2 and 3, free SigA, and WhiB1:SigA proteins respectively of the elution profile of **[A].**

4.2.3 Purification of the WhiB1:SigA complex from nucleic acids

To investigate whether the binding of WhiB1 with SigA was mediated by DNA, as well as to rid the sample of nucleic acids, two methods of purification were carried out. The isolated complex was treated first with Benzonase nuclease (Sigma-Aldrich) (1.0 unit $A_{260\ nm}$ of 1.0 for 30 min at 37°C, final volume of reaction 1.0 ml), in reaction buffer of 50 mM Tris (pH 8.0) containing 0.5 M NaCl. Then, the treated sample was loaded directly into HiLoad 16/600 Superdex gel filtration column pre-equilibrated with aerobic 50 mM Tris-HCl (pH 8.0) containing 0.5 M NaCl (Figure 4.5). The second method was by HiTrap Heparin chromatography; where, the eluted protein complex from the nickel-charged HiTrap column was diluted with 50 mM Tris, pH 8.0 containing 0.1 M NaCl to a final concentration of 0.25 M NaCl. The low NaCl containing WhiB1:SigA sample was then applied to the Heparin column (Figure 4.6A), and the eluted protein was loaded directly onto the gel filtration column and all eluted fractions were analyzed by SDS-PAGE gel (Figure 4.6B). The elution profiles after the treatments showed the absence of the peak associated with free SigA bound to DNA (Figure 4.5).

In addition, the WhiB1:SigA complex was still present (Figure 4.5), suggesting that complex integrity was not adversely influenced by the presence of the nuclease, and that WhiB1:SigA binding is not mediated through DNA. A similar observation was noticed when 1M salt was present in the elution buffer, where the absorbance of first peak at 42 ml, representing SigA:DNA, was significantly smaller (Figure 4.5, black trace), while the WhiB1:SigA complex was at the same elution position; suggesting release of contaminating nucleic acids from SigA without complex dissociation due to high NaCl concentration.

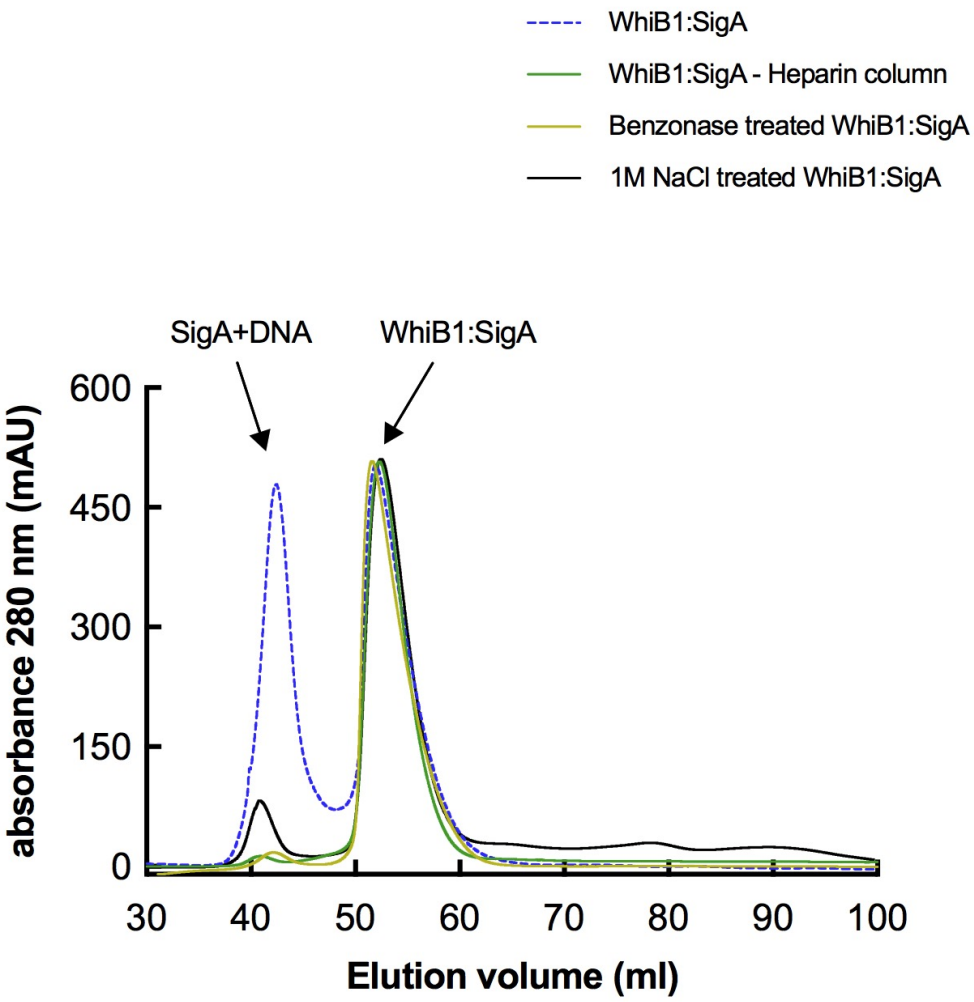


Figure 4.5: WhiB1:SigA complex purification from contaminating DNA. HiLoad 16/600 Superdex gel filtration column pre-equilibrated with buffer of 50 mM Tris-HCl (pH 8.0) containing 0.5 M NaCl was used. Protein (-----) (0.29 mM), which was initially isolated using nickel-charged HiTrap chelating column was directly injected onto gel filtration column. Two eluted peak positions at 42 ml (SigA+DNA) and at 54 ml (WhiB1:SigA) were detected. The contaminant DNA in free SigA peak was removed when the sample was treated prior the separation by gel filtration column with Benzonase (—) or further purified by Heparin chromatography (—); DNA contamination was also decreased after eluting the WhiB1:SigA complex in the presence of 1M NaCl instead of 0.5 M.

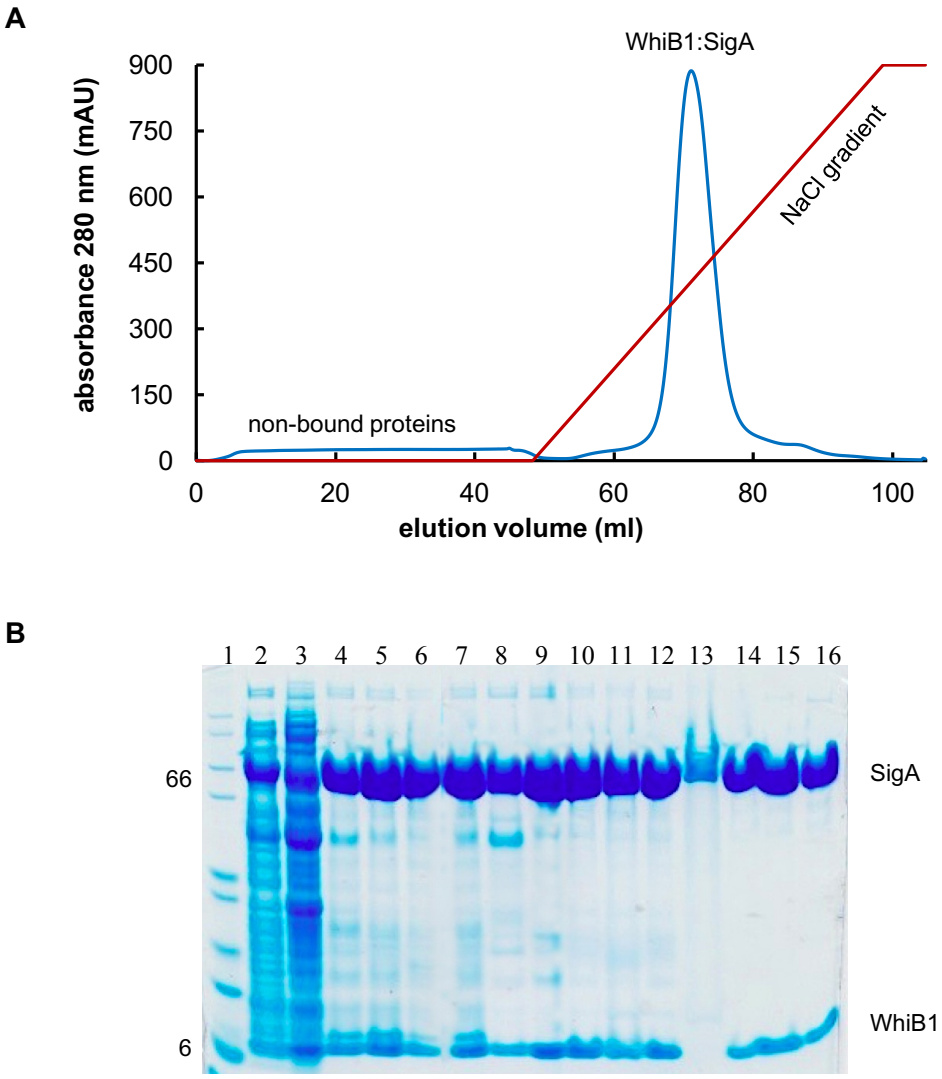


Figure 4.6: [A] Elution profile from HiTrap Heparin column of injected WhiB1:SigA complex. The trace shows protein absorbance at 280 nm (blue line) of injected WhiB1:SigA complex and the concentration of NaCl from 0-0.5 M (red line), in the presence of 50 mM Tris (pH 8.0). Eluted WhiB1:SigA complex appears as a single peak between ~65-78 ml. **[B] Coomassie stained SDS-PAGE gel of isolated fractions of WhiB1:SigA protein.** Lane 1, protein ladder (sizes in kDa are indicated); lanes 2 and 3, cell debris and cell free-extract of induced WhiB1:SigA; lanes 4-6, fractions of WhiB1:SigA complex eluted from nickel-charged column, which were then mixed (lane 7) and loaded into Heparin column; lanes 8-11, fractions of WhiB1:SigA complex eluted between ~65-78 ml, as indicated in **[A]**; which were mixed together (lane 12) and injected into gel filtration column (Figure 4.5); lane 13, and lanes 14-16, represent free SigA and WhiB1:SigA complex respectively separated by gel filtration column in Figure 4.5.

4.2.4 UV-visible spectroscopy of the WhiB1:SigA complex is consistent with the presence of a stable [4Fe-4S] cluster

The UV-visible spectrum of WhiB1:SigA, which was first isolated by HiTrap chelating chromatography, then HiTrap Heparin chromatography, and finally by HiLoad 16/600 Superdex gel filtration chromatography, exhibited a peak at 420 nm (Figure 4.7A). The brown colour of the sample and the UV-visible spectrum indicated the presence of [4Fe-4S] cluster in the WhiB1:SigA complex.

To test the stability of [4Fe-4S] cluster of the complex upon addition of O_2-saturated buffer to final concentration of 110 μM (Section 2.6.4), the protein was concentrated first using Vivaspin 30 kDa MW cut off filtration unit (Sigma-Aldrich) to yield, 70 μM final cluster concentration. The UV-visible spectra were obtained before and after O_2 exposure (Figure 4.7B). After 8 h of incubation under these conditions no change was noticed at 420 nm. Indeed, it was observed that the complex of WhiB1:SigA did not precipitate or lose the cluster even when the sample was left at 4ºC for 1 month; suggesting the high stability of the complex. These data indicate that the iron-sulfur cluster of WhiB1 in complex with SigA protein behaved differently to that of WhiB1^{Ec} (Figure 3.7), but similarly to WhiB1^{Ms} with His_6-TEV tag (Figure 3.13). This result suggests that the presence of the partner protein SigA changed the iron-sulfur cluster stability of WhiB1 protein.

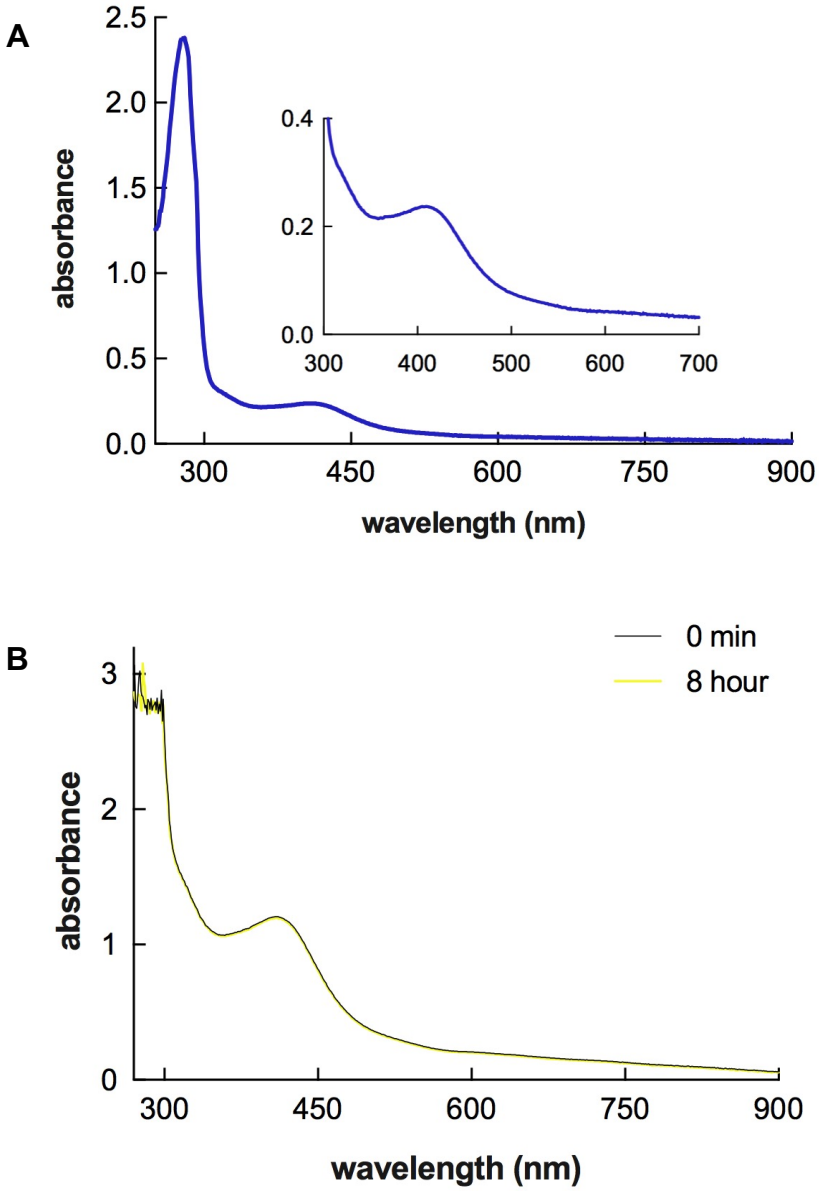


Figure 4.7: UV-visible spectroscopy of the WhiB1:SigA complex. [A] The spectrum of holo-WhiB1:SigA protein shows a peak at 420 nm. **[B] The [4Fe-4S] cluster stability of isolated WhiB1:SigA.** UV-visible spectra of the complex (70 μM cluster) before (black trace), and after 8 h of incubation (yellow trace) in the presence of 110 μM final concentration of O_2.

4.2.5 The WhiB1 iron-sulfur cluster is essential for interaction with SigA

Gaskell et al. (2007) reported that the activity of SigM in *Streptomyces coelicolor* is under the control of anti-sigma factor RsmA, which binds a [2Fe-2S] cluster; and, this type of RsmA:SigM interaction is governed by the presence of the iron-sulfur cluster. It was concluded, that abolishing the interaction was activated by losing the RsmA [2Fe-2S] cluster in response to an as yet undiscovered signal. Thus, to investigate whether the iron-sulfur cluster of WhiB1 is essential for interaction with SigA, site-directed mutagenesis was carried out using pGS2560, which encoded a non-tagged WhiB1, as a template (Section 2.5.12). The Cys 9, Cys 40, and Cys 46 of WhiB1 were individually replaced by alanine leading to pGS2561, pGS2563, and pGS2564 respectively; (WhiB1 Cys37Ala mutant was excluded from this investigation because an extra unintended mutation (Ala7Val) was found after sequencing the mutated pGS2562 construct). Smith et al. (2012) reported that all four cysteine residues in WhiB1 are essential for anchoring the [4Fe-4S] cluster. As a result any Cys mutation should cause loss of the cluster. Plasmids of pGS2561, pGS2563, and pGS2564 were individually co-transformed with pGS2183, encoding SigA with an *N*-terminal His_6-tag (Figure 4.1), into the electro-competent *E. coli* JRG5302. Overproduction and purification of the corresponding WhiB1:SigA complexes were as described in Section 4.2.1. Results showed that none of the three WhiB1 Cys mutants was co-purified with SigA by nickel-charged HiTrap chelating chromatography in contrast to wild type WhiB1. The SDS-PAGE gel analysis showed that WhiB1 eluted along with the non-specific binding proteins at low concentration of imidazole, while only SigA protein was eluted by high imidazole concentration (Figure 4.8). These data suggest that the WhiB1 iron-sulfur cluster is essential for WhiB1-SigA interaction, and complex disassembly could be induced in a response to specific stimulus/stimuli that cause cluster degradation or modifications or possibly under iron starvation conditions.

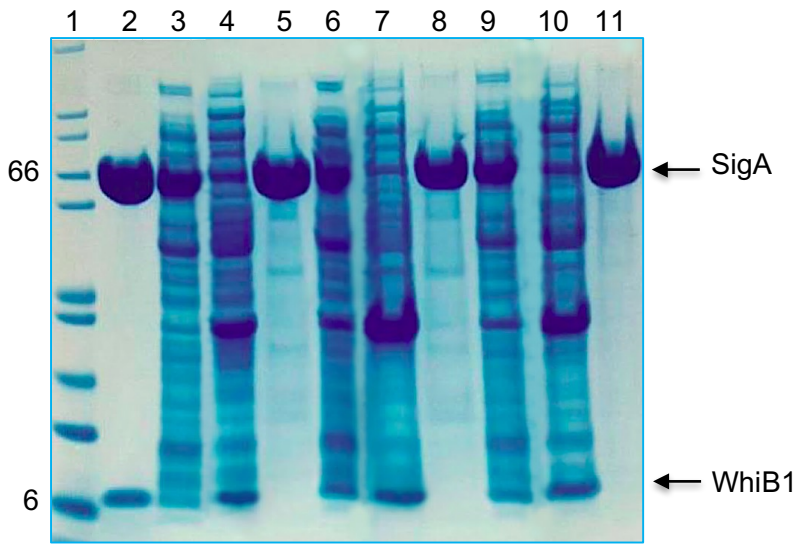


Figure 4.8: Coomassie stained SDS-PAGE gel analysis of His_6-SigA:(Cys mutants) WhiB1 complexes. Lane 1, protein ladder (sizes in kDa are indicated); lane 2, isolated His_6-SigA:WhiB1; lane 3, cell-free extract of *E. coli* expressing His_6-SigA and WhiB1(Cys9Ala); lane 4, flow through from nickel-charged column after application of the cell-free extract shown in lane 3 by low concentration of imidazole; lane 5, eluate from the nickel affinity column loaded with the cell-free extract shown in lane 3 by high concentration of imidazole; lanes 6-8, and lanes 9-11, represent the same order of lanes 3-5 contents, but the cell free-extract of each group belong to *E. coli* expressing His_6-SigA:WhiB1(Cys40Ala) and His_6-SigA:WhiB1(Cys46Ala) respectively instead of His_6-SigA:WhiB1(Cys9Ala).

4.2.6 LC-MS study of the [4Fe-4S]-WhiB1:SigA complex

As found in the previous chapter, WhiB1^{Ms} cannot retain sulfur after cluster dissociation in contrast to WhiB2^{Ms} which does (Section 3.4.6; Figure 3.20). To investigate the ability of [4Fe-4S]-WhiB1 in the complex with SigA to form sulfide bridges after cluster dissociation and complex separation, LC-MS analysis was carried out by Mr. Simon Thorpe (The University of Sheffield Faculty of Science Mass Spectrometry Centre) as described previously (Section 2.5.10). Two peaks were detected, indicating the separation of the complex. Apo-WhiB1 showed a single peak at a molecular weight of 9,318 Da, revealing of the presence of four thiol groups in the apo- protein; while, SigA protein exhibited one main peak at a molecular weight of 59,933 Da (Figure 4.9). These data suggest that WhiB1 lost the iron-sulfur cluster without sulfur retention, and the complex was not stable under these conditions.

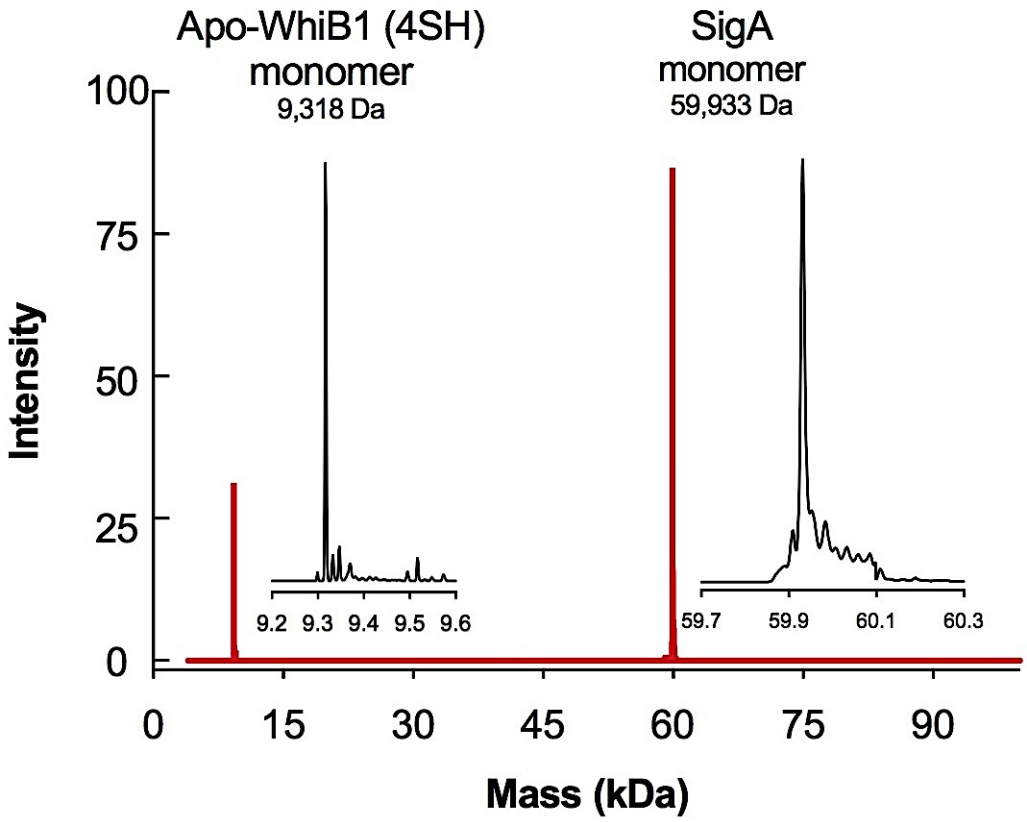


Figure 4.9: LC-MS analysis of the holo-WhiB1:SigA complex. Protein sample (~50 μM cluster), in 50 mM Tris, pH 8.0 containing 0.5 M NaCl were applied in the presence of 0.1% formic acid to LC instrument for 10 min before mass analysis. The peak at 9,318 Da represents the monomer apo-WhiB1 in its reduced form, and the peak at 59,933 Da represents the monomer SigA protein. Mass spectrum was obtained by Mr. Simon Thorpe.

4.2.7 The presence of the His_6-TEV tag in $WhiB1^{Ms}$ prevents interaction with SigA

To investigate the ability of [4Fe-4S]-$WhiB1^{Ms}$, which was overproduced in *M. smegmatis* (Section 3.4.1), to bind with SigA, and to study the influence of the His_6-TEV tag on the binding, SigA overproduction was carried out by transforming electro-competent *E. coli* JRG5302 with pGS2183 (Figure 4.1). Free SigA protein overproduction and isolation conditions were similar to that of the WhiB1:SigA complex (Section 2.5.3.3; 2.5.5.1). SigA protein after isolation by nickel-charged HiTrap chelating chromatography was also passed through HiTrap Heparin column and HiLoad 16/600 Superdex gel filtration column, this was done to free the protein of nucleic acids. $WhiB1^{Ms}$ was mixed with SigA at a molar ratio of 1:1 (0.5 mM), in 25 mM NaH_2PO_4 (pH 7.4) containing 0.2 M NaCl for 1 h on ice or at room temperature before separation by gel filtration column. Results showed that no interaction between $WhiB1^{Ms}$ and SigA was detected, and each one was separated and eluted at its individual elution position (Figure 4.10A). SDS-PAGE gel analysis of eluted fractions confirmed this interpretation (Figure 4.10B). Feng et al. (2016) reported that His_6-ThCS-WhiB1 can interact with non-tagged SigA when both were co-expressed in *E. coli.* However, results here showed that the binding also happened when His_6-ThCS-SigA co-expressed with non-tagged WhiB1 (Section 4.2.1), suggesting that the presence of *N*-terminal His_6-ThCS tag in either SigA or WhiB1 has no effect on WhiB1:SigA interaction. However, this interaction did not occur when *N*-terminal His_6-TEV-[4Fe-4S] $WhiB1^{Ms}$ was mixed with *N*-terminal His_6-ThCS-SigA, suggesting that the presence of TEV-tag prevents WhiB1:SigA interaction. As a result, it was concluded that the presence of the TEV cleavage site in the $WhiB1^{Ms}$ tag could mimic the presence of SigA in changing the conformation of WhiB1 protein, consistent with the presence of the TEV tag or SigA leading to stabilization the iron-sulfur cluster of WhiB1.

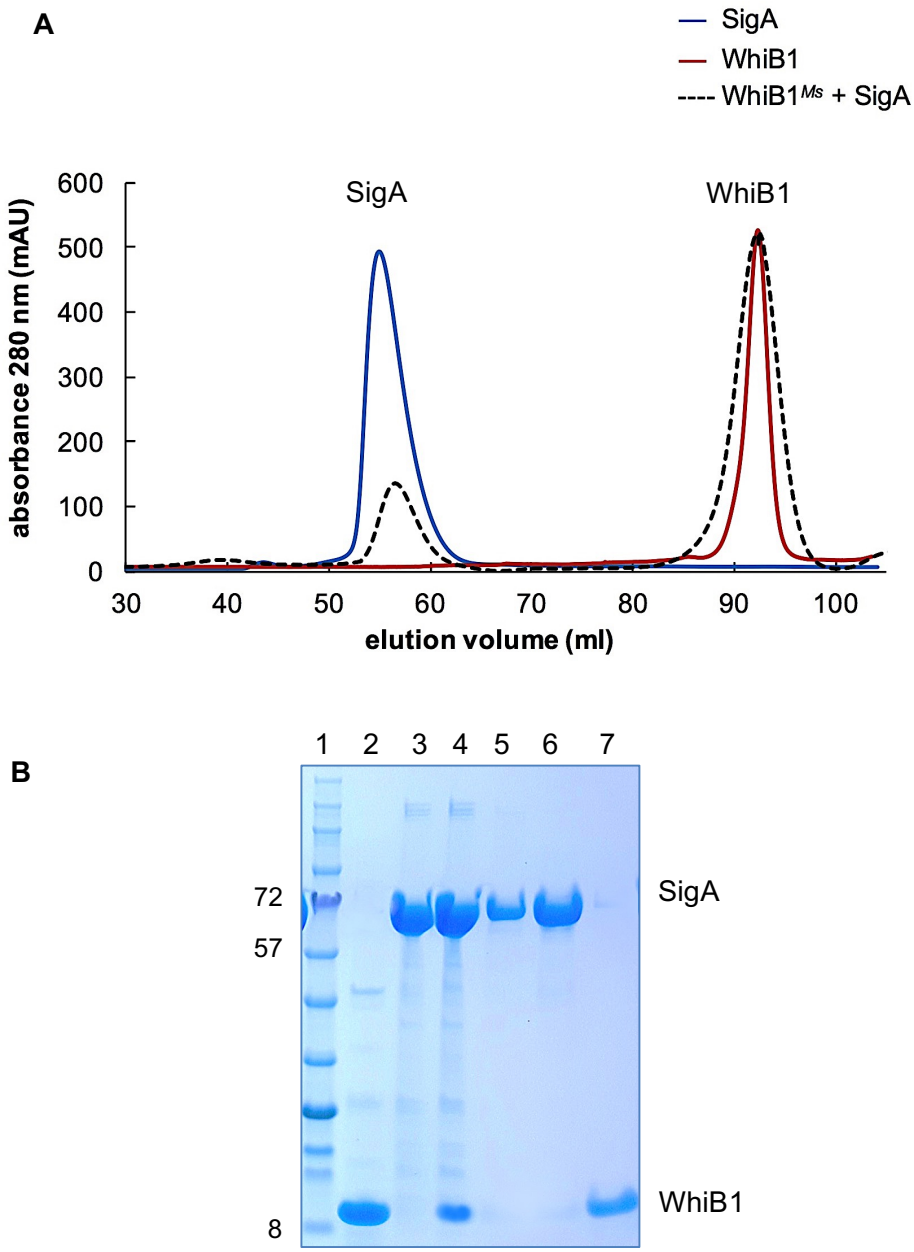


Figure 4.10: *In vitro* analysis of WhiB1^{Ms} interaction with SigA. [A] Holo-His$_6$-TEV-WhiB1^{Ms} and His-ThCS-SigA were isolated separately and then mixed at a molar ratio of 1:1 (0.5 mM) before separation by size exclusion chromatography (**----**). Reaction was carried out in 25 mM NaH_2PO_4 (pH 7.4) containing 0.2 M NaCl. The elution positions of free WhiB1 (—) and free SigA (—) before mixing are indicated. **[B]** SDS-PAGE gel analysis of eluted fractions of Holo-His$_6$-TEV-WhiB1^{Ms} and His-ThCS-SigA mixture in Figure **[A]**. Lane 1, protein ladder (sizes in kDa are indicated); lane 2, His$_6$-TEV-WhiB1^{Ms}; lane 3, His-ThCS-SigA; lane 4, the loaded mixture of WhiB1^{Ms} and SigA; lanes 5 and 6, fractions eluted between 52-60 ml; fractions eluted between 90-95 ml, were mixed and loaded at lane 7. Fractions positions are shown in **[A].**

4.3 The sensitivity of the holo-WhiB1:SigA complex to NO

4.3.1 The [4Fe-4S]-WhiB1:SigA complex reacts very quickly with eight molecules of NO

As shown in Section 3.4.8, WhiB1^{Ms} reacts very quickly with NO via its iron-sulfur cluster forming an octa-nitrosylated iron-sulfur complex. To investigate whether [4Fe-4S]-WhiB1:SigA complex reacts with NO, Proli NONOate was used as a NO donor (Section 2.6.5). The [4Fe-4S]-WhiB1:SigA complex was prepared by purifying the sample through nickel-charged HiTrap chelating column as a first step then by HiTrap Heparin column and after that by HiLoad 16/600 Superdex gel filtration to isolate the complex from free SigA and any contaminating DNA (Secction 4.2.3). Holo-WhiB1:SigA (~ 36 µM) was exposed to NO under anaerobic conditions in a sealed cuvette with screw-top lid, by injecting 0.26 mM Proli NONOate (0.52 mM NO) in to the protein sample, and incubating for 30 min. UV-visible spectra showed an increase the absorbance at 355-360 nm, suggesting the formation of a nitrosylated iron-sulfur cluster in WhiB1:SigA complex (Figure 4.11A). To test the reaction speed, the change in the absorbance at 355 nm was measured over time after Holo-WhiB1:SigA (36 µM cluster) was exposed to 0.26 mM of Proli NONOate (Figure 4.11B). The reaction was extremely rapid similar to that of WhiB1^{Ms} (Figure 3.21B). Nitric oxide titration was also carried out to estimate the ratio of NO per [4Fe-4S]-WhiB1:SigA. Increased concentrations of Proli NONOate (1.05-47.25 µM equivalent to 2.1-94.5 µM NO, Proli NONOate stock solution 380 µM in 10 mM NaOH) were injected into the Holo-WhiB1:SigA (10 µM cluster) in 43 additions, each one contained 2.1 µM of NO, except the last three additions where the concentration was doubled (Figure 4.12). Each addition was mixed first with (50 mM NaH_2PO_4, 100 mM NaCl, pH 7.4) to neutralize the sodium hydroxide, and directly injected in to the protein. The reaction was left ~1.5-3 min at 25ºC before reading the spectrum. The spectroscopic result showed an increase at 355-360 nm with a decrease at 420 nm after each addition with an isosbestic point at 390 nm. These data suggest the formation of a nitrosylated iron-sulfur complex (Crack et al., 2010; Cruz-Ramos et al., 2002; Smith et al., 2010). Cumulative change in absorbance at 355 nm after each Proli NONOate addition was calculated (dilution factors were accounted for) and plotted against the NO:[4Fe-4S] ratio, which showed that the reaction ended at a ratio

of ~7.6 (Figure 4.12B), suggesting the formation of an octa-nitrosylated iron-sulfur complex as seen for WhiB1^{Ms} with His_6-TEV tag (Figure 3.22).

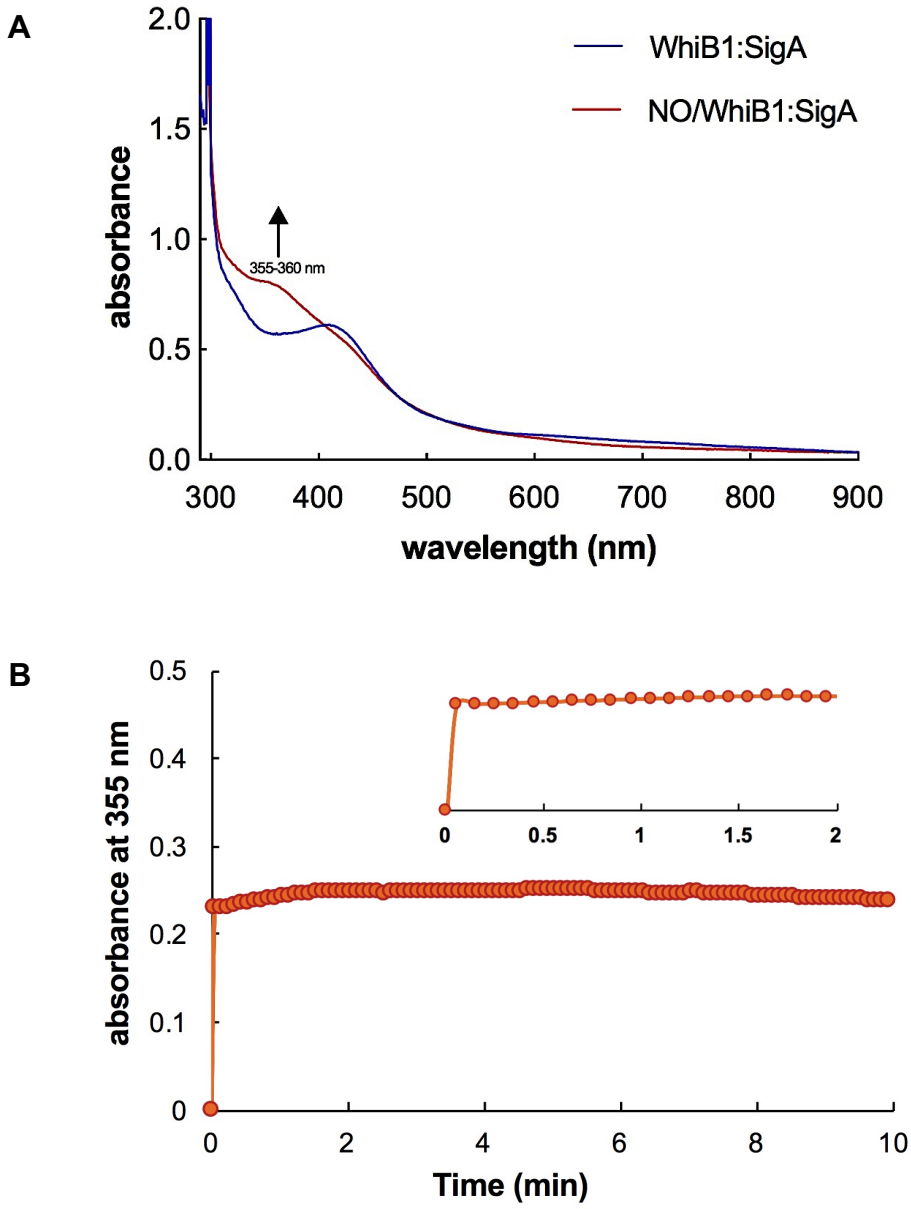


Figure 4.11: The sensitivity of holo-WhiB1:SigA to NO. [A] UV-visible spectra of WhiB1:SigA complex reaction with NO. The spectra of holo-WhiB1:SigA (~36 μM cluster) before (blue line) and after (red line) adding 260 μM Proli NONOate (equivalent to 520 μM of NO) after 30 min of incubation the spectrum was recorded; arrow indicates the direction of spectral change. Proli NONOate was mixed firstly with (50 mM NaH_2PO_4, 100 mM NaCl, pH 7.4), and then directly injected into the protein sample **[B]** The absorbance change at 355 nm plotted over time after adding the NO donor (520 μM of NO) into Holo-WhiB1:SigA (~ 36 μM cluster). Buffer of protein was 50 mM Tris, 500 mM NaCl, pH 7.4.

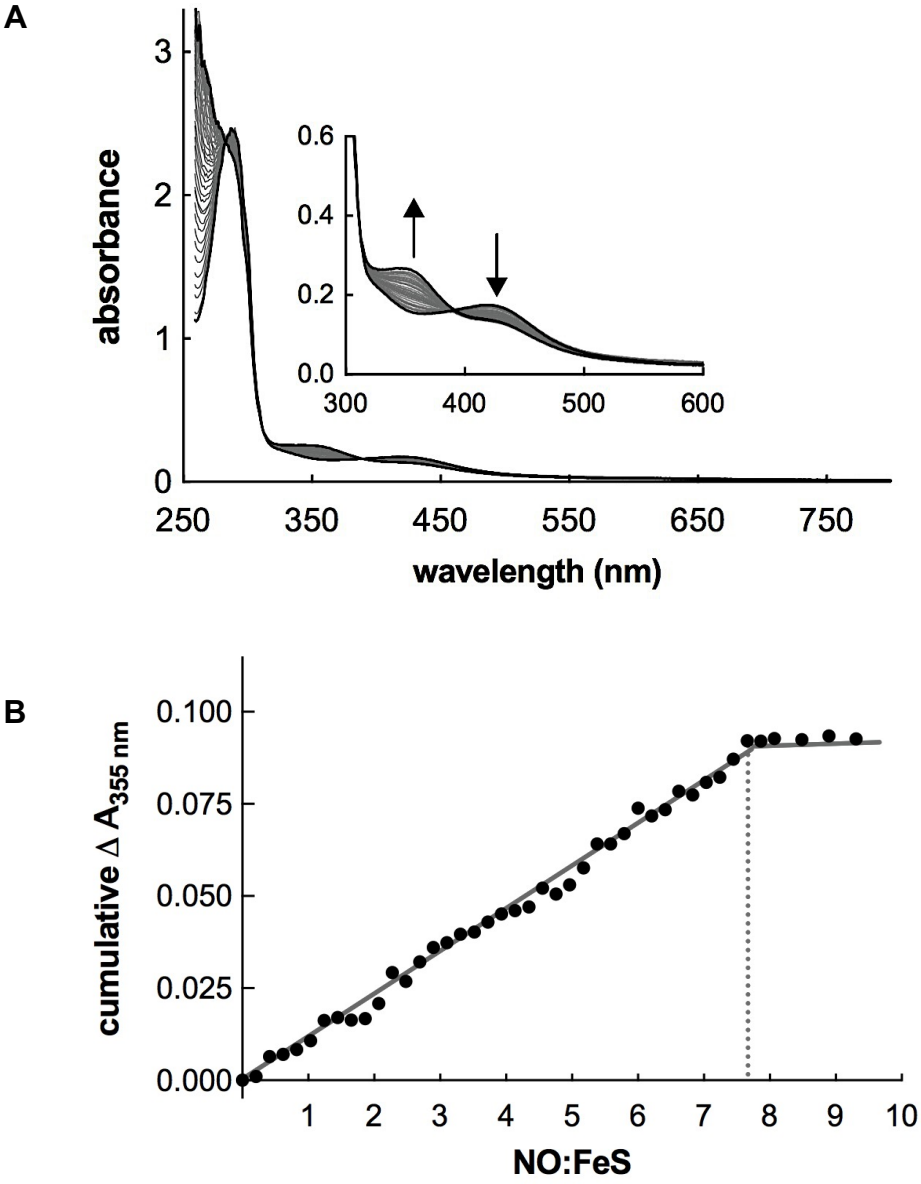


Figure 4.12: Stoichiometry of the reaction between [4Fe-4S]-WhiB1:SigA and NO. [A] UV-visible spectra following NO reaction with [4Fe-4S]-WhiB1:SigA. Increasing amounts (1.05-47.25 μM) of Proli NONOate (equivalent to 2.1-94.5 μM of NO) were introduced to holo-WhiB1:SigA (10 μM cluster) in 43 additions. The arrows refer to the direction of change in the absorbance. Each spectrum was taken 1.5 min after NO addition. **[B]** The cumulative changes in absorbance at 355 nm were plotted versus the NO:[4Fe-4S] ratio. The dashed line indicates the point at which no further change in absorbance was observed.

4.3.2 Exposure of the holo-WhiB1:SigA complex to NO results in the formation of a nitrosylated complex and release of WhiB1

To study the WhiB1:SigA complex state after exposure to NO by size exclusion chromatography, the sensitivity of nitrosylated WhiB1:SigA complex to O_2 was tested first. This test was designed to establish whether the nitrosylated WhiB1:SigA iron-sulfur cluster was stable under the aerobic conditions used for gel filtration. The complex was isolated as described previously (Sections 4.2.1, 4.2.3), and then was nitrosylated using Proli NONOate (Section 4.3.1). The nitrosylated sample was exposed to air directly by bubbling the protein sample with air three times using a 100 µl Hamilton gas tight syringe with opening of the lid of the cuvette and without any reducing agent. The degradation rate of nitrosylated WhiB1:SigA complex was monitored by UV-visible spectra; each spectrum was recorded every 30 min at 25ºC (Figure 4.13). The cluster degradation rate was extremely slow; after 1 h of incubation an absorbance decrease of ~10% at 355 nm was detected. According to this, gel filtration analysis could be carried out to study the oligomeric state of the WhiB1:SigA complex after iron-sulfur cluster nitrosylation. Nevertheless, the gel filtration buffer (50 mM Tris (pH 8.0) containing 500 mM NaCl) was left inside an anaerobic work station overnight whilst being stirred, after that the buffer was degassed under vacuum for 1 h, and used for equilibration of the gel filtration column. This was done to maintain a low level of O_2 inside the column to prevent the cluster degradation. Holo-WhiB1:SigA (1 ml of 40 µM cluster) was nitrosylated by Proli NONOate solution (400 µM) (equivalent to 800 µM of NO), then incubated inside an anaerobic cabinet for 1 h prior injecting the sample on to the gel filtration column. The flow rate was set (2 ml/min) for 50 min. The elution profile of nitrosylated-WhiB1:SigA showed three different peaks at 42 ml, 56 ml, and 93 ml (Figure 4.14A). SDS-PAGE analysis of eluted fractions showed that the WhiB1:SigA complex, free SigA, and free WhiB1 were eluted at the above positons, respectively (Figure 4.14A). UV-visible spectra showed that the complex of WhiB1:SigA was nitrosylated, as a clear peak at ~360 nm was present in the UV-visible spectrum (Figure 4.14B).; while, the released WhiB1 was quite diluted, but after concentrating the sample using a Vivaspin column, the spectrum showed that the WhiB1 was in the apo-form (Figure 4.14C).

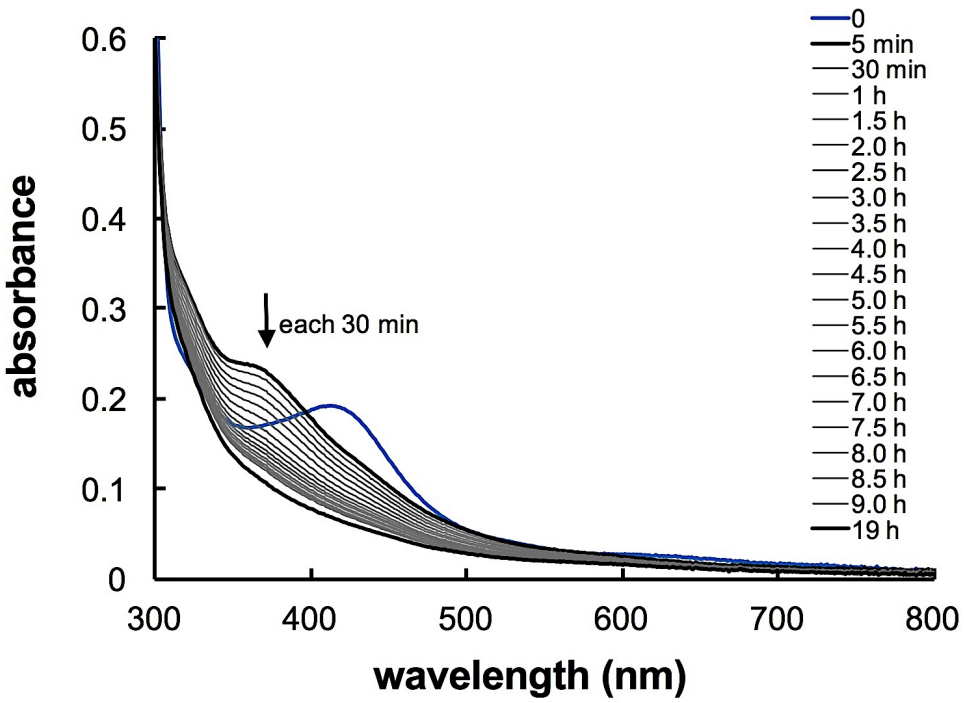


Figure 4.13: The O_2 sensitivity of the nitrosylated iron-sulfur cluster of the WhiB1:SigA complex. Holo-WhiB1:SigA complex (11 µM) in 50 mM Tris (pH 7.4) containing 500 mM NaCl, was treated with 90 µM Proli NONOate (equivalent to 180 µM of NO). Then the nitrosylated sample was exposed to air directly by bubbling the protein sample with air three times using a 100 µl Hamilton gas tight syringe. UV-visible spectra were recoded every 30 min in 25ºC. The arrow indicates the direction of change in absorbance.

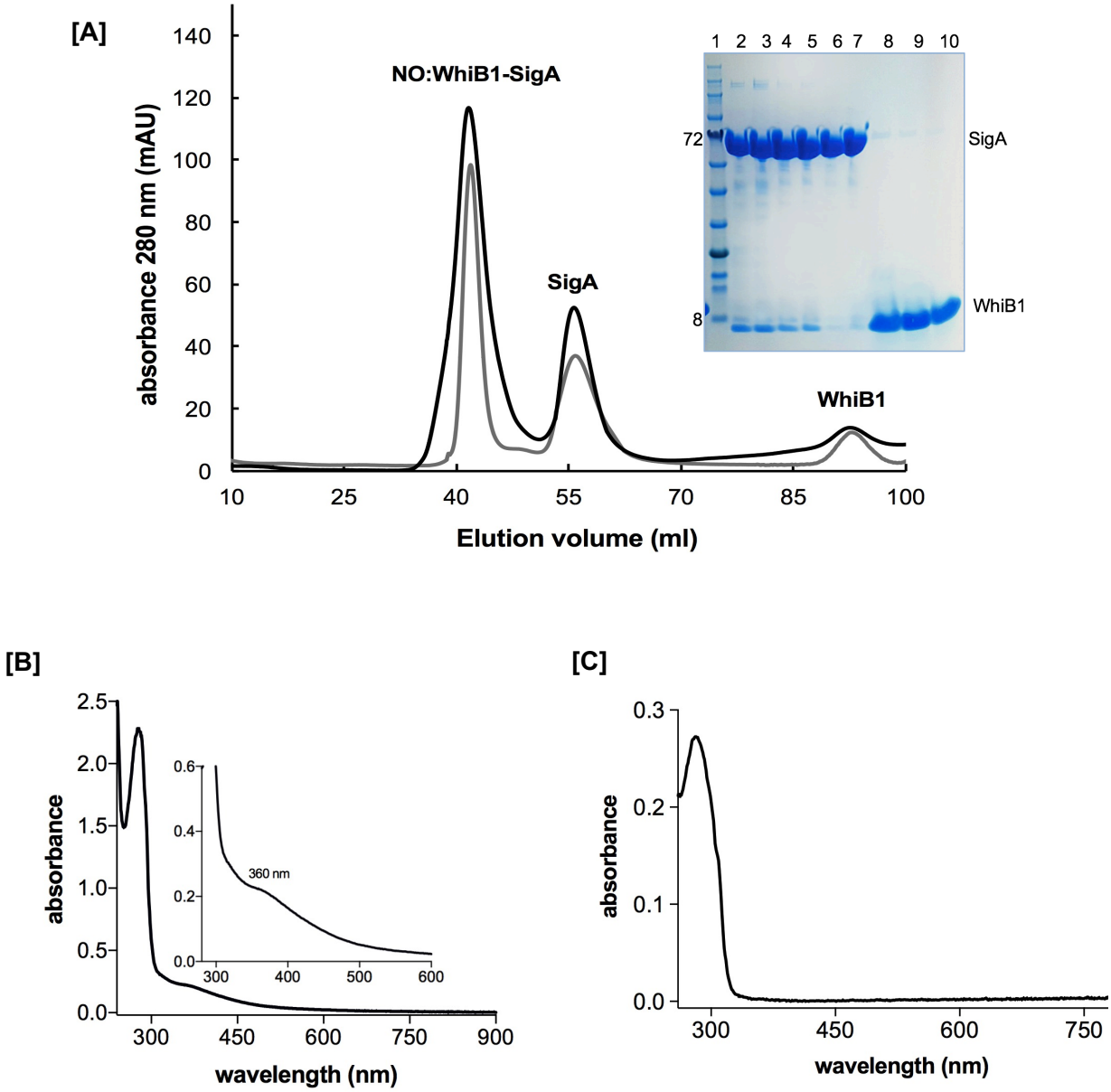


Figure 4.14: Size-exclusion chromatography analysis of nitrosylated WhiB1:SigA complex. [A] Holo-WhiB1:SigA complex (40 µM cluster) was treated with 400 µM of Proli NONOate (800 µM of NO) before applying the sample to a gel filtration column pre-equilibrated with anaerobic-degassed buffer of 50 mM Tris (pH 8.0) containing 500 mM NaCl. Nitrosylated WhiB1:SigA protein was eluted at 42 ml**,** while free SigA, and free WhiB1 were eluted at 56 ml and 93 ml respectively. ***Inset:*** **SDS-PAGE gel of ~20 µg of each eluted fraction in the elution profile.** Lane 1, protein ladder (sizes in kDa are indicated); lanes 2, nitrosylated WhiB1:SigA sample loaded into gel filtration column; lanes 3-5, nitrosylated WhiB1:SigA eluted at ~42 ml position; lanes 6-7, free SigA eluted at 56 ml position; lanes 8-10, free WhiB1 eluted at ~93 ml; as shown in the elution profile. The UV-visible spectra of nitrosylated WhiB1:SigA eluted at ~42 ml and free WhiB1 eluted at ~93 ml are shown in **[B]** and **[C]** respectively.

4.4 The holo-WhiB1:SigA$_{M359-D528}$ reaction with NO

4.4.1 Introduction

The observations that the WhiB1 iron-sulfur cluster was required for interaction with SigA (Section 4.2.5), and that the holo-WhiB1:SigA complex disassembled after exposure to NO (Section 4.3.2) suggested that WhiB1 might act as an anti-SigA factor, with NO as the external signal that could initiate complex separation. To begin mapping the interaction interface, different region of the SigA protein were tested for WhiB1 binding.

Thermus thermophilus Sig^{70} crystal structure showed that the protein is composed of four domains: *N*-terminal domains 1 and 2 (ND1,2), linker domain (LD) and *C*-terminal domain (CD). The ND2, LD, and CD contain the terminus of region 2 (sub region 2.5) and all of regions 3 and 4 (the -35 hexamer binding region) as shown in Figure 4.31 (Vassylyev et al., 2002). *Mycobacterium tuberculosis* WhiB7 and WhiB3 interact with SigA, and ND1, which composes two-thirds of SigA, was not essential for this binding; whereas WhiB3 was found to interact with *M. tuberculosis* SigA$_{369–528}$ (SigA$_{C160}$) (Steyn et al., 2002), and WhiB7 with *M. smegmatis* SigA$_{297–466}$, which is identical (99.6%) to *M. tuberculosis* SigA$_{359–528}$ (SigA $_{C170}$) (Burian et al., 2013).

To test whether WhiB1 also binds with SigA$_{C170}$ (the last 170 amino acid at the *C*-terminal domain (residues 359–528) of SigA, where the *N*-terminal domain 1 is not present), a new expression plasmid containing the *M. tuberculosis* *sigA*$_{A1075-A1587}$ (*sigA*$_{C170}$) coding region was made, and the WhiB1:SigA$_{C170}$ binding was investigated.

4.4.2 WhiB1:SigA$_{C170}$ overproduction and purification

For WhiB1:SigA$_{C170}$ complex overproduction, *E. coli* JRG6858 was used. This strain was created by transforming electro-competent *E. coli* JRG5302 with pGS2565 (Figure 4.15) and pGS2560 (Figure 4.2) that encode SigA$_{C170}$ and WhiB1 respectively, both without tags (Section 2.3.14). WhiB1:SigA$_{C170}$ complex overproduction was carried out by growing *E. coli* JRG6858 in LB broth with induction by 1 mM IPTG and incubation for 12-18 h at 18°C with an agitation speed of 250 rpm (Section 2.5.3.5).

For isolation of the WhiB1:$SigA_{C170}$, cell-free extract was prepared by re-suspending cell pellets in binding buffer consisting of 100 mM NaCl, 50 mM MES (pH 6.5) and then sonicated according to Section 2.5.4. Subsequent protein purification was done via three different purification steps (Section 2.5.5.2). The first step was by anion exchange chromatography using HiTrap Q HP column (GE Healthcare). The elution profile (A_{280}), the brown colour of fractions containing the complex and SDS-PAGE analysis helped to identify the fractions containing the WhiB1:$SigA_{C170}$ complex (Figure 4.16). The second step of purification was achieved using a HiLoad 16/600 Superdex gel filtration column pre-equilibrated with 50 mM Tris (pH 8.0) containing 0.5 M NaCl buffer (Figure 4.17A). Fractions from the HiTrap Q HP column, were collected and concentrated by Vivaspin unit with a 10 kDa MW cut off (Sigma-Aldrich), and then the sample was applied to the gel filtration column. Fractions from the column were then analyzed by SDS-PAGE to identify those containing WhiB1:$SigA_{C170}$ complex (Figure 4.17B).

The fractions from the gel filtration column containing the WhiB1:$SigA_{C170}$ complex, also contained contaminating *E. coli* proteins, espicially at ~40 kDa (Figure 4.17B). Therefore, another step of purification was needed. Post-gel filtration WhiB1:$SigA_{C170}$ was diluted with binding buffer (50 mM Tris pH 8.0, containing 0.1 M NaCl) to reduce the NaCl to a final concentration of 0.2 M. This sample was then applied to a HiTrap Heparin column (Figure 4.18A). The SDS-PAGE analysis of the complex eluted showed an un-tagged WhiB1:$SigA_{C170}$ complex essentially free of contaminants (Figure 4.18B).

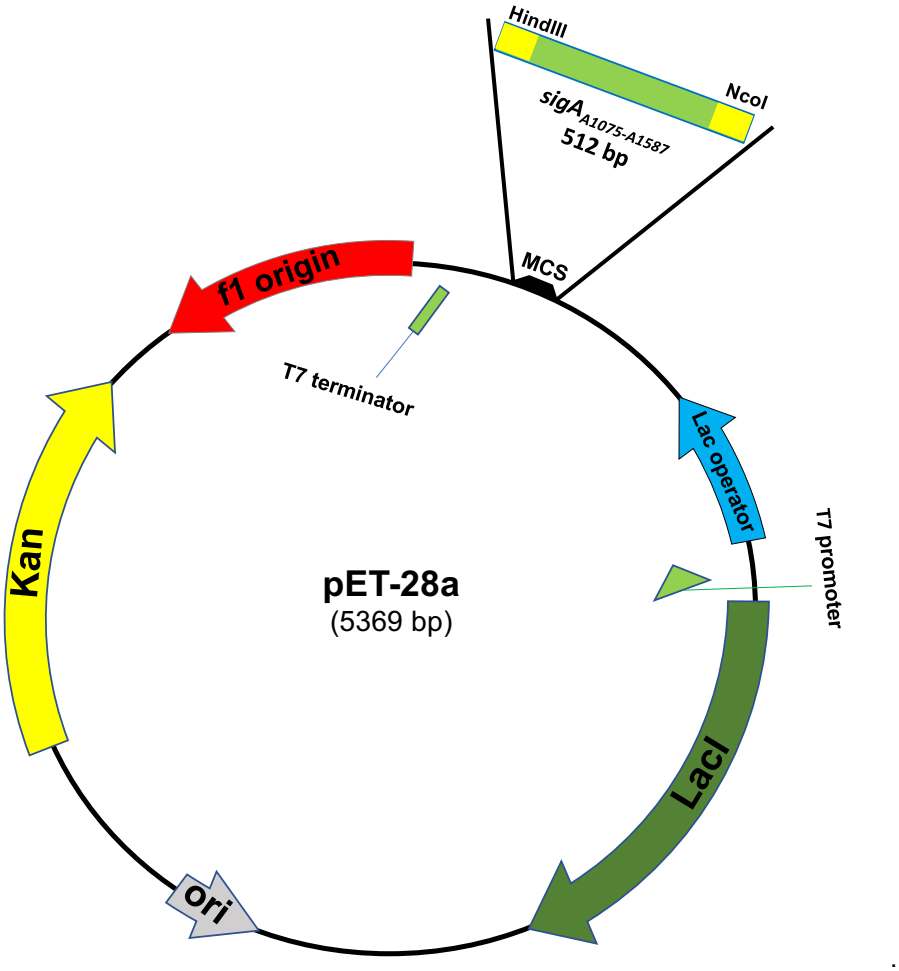


Figure 4.15: Schematic of pGS2565. [A] A portion (512 bp) of the *sigA* open reading frame ($sigA_{A1075\text{-}A1587}$) was amplified by PCR from genomic DNA of *M. tuberculosis* with engineered NcoI/HindIII restriction sites, and ligated into pET28a to create pGS2565. The plasmid encodes an un-tagged $SigA_{C170}$.

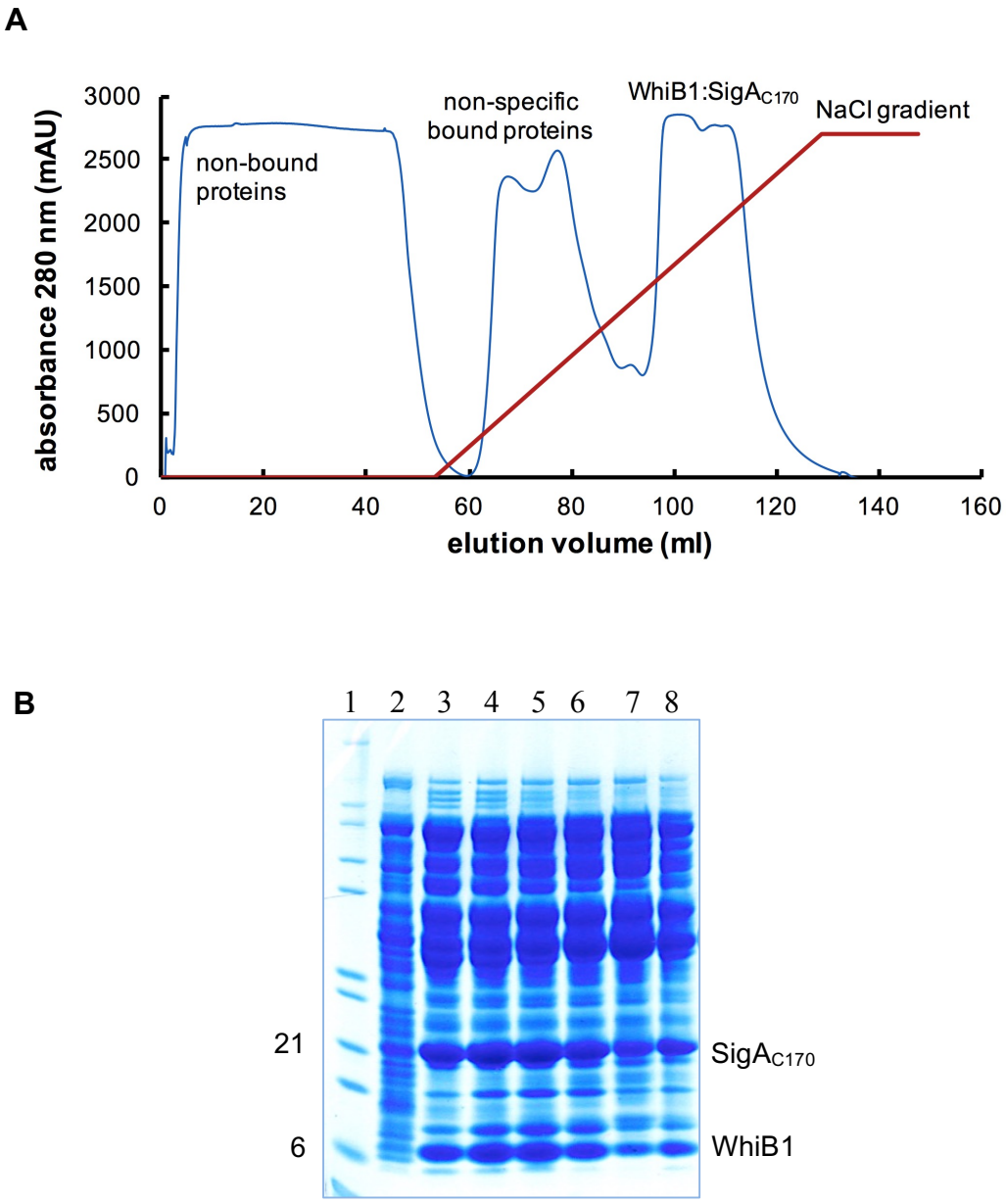


Figure 4.16: Isolation of un-tagged WhiB1:SigA$_{C170}$ complex I. [A] Elution profile of HiTrap Q HP column of *E. coli* cell lysate containing overproduced WhiB1 and SigA$_{C170}$. The trace shows protein absorbance at 280 nm (blue line), and the gradient concentration of NaCl from 0-0.5 M (red line). WhiB1:SigA$_{C170}$ complex eluted with other *E. coli* proteins between ~100-120 ml. **[B] Coomassie-stained SDS-PAGE of fractions from the separation shown in [A].** Lane 1, protein ladder (sizes in kDa are indicated); lane 2 cell-free extract of induced WhiB1:SigA$_{C170}$; lanes 3-8, fractions of eluate between ~100-120 ml which contain the complex, as indicated in **[A]**.

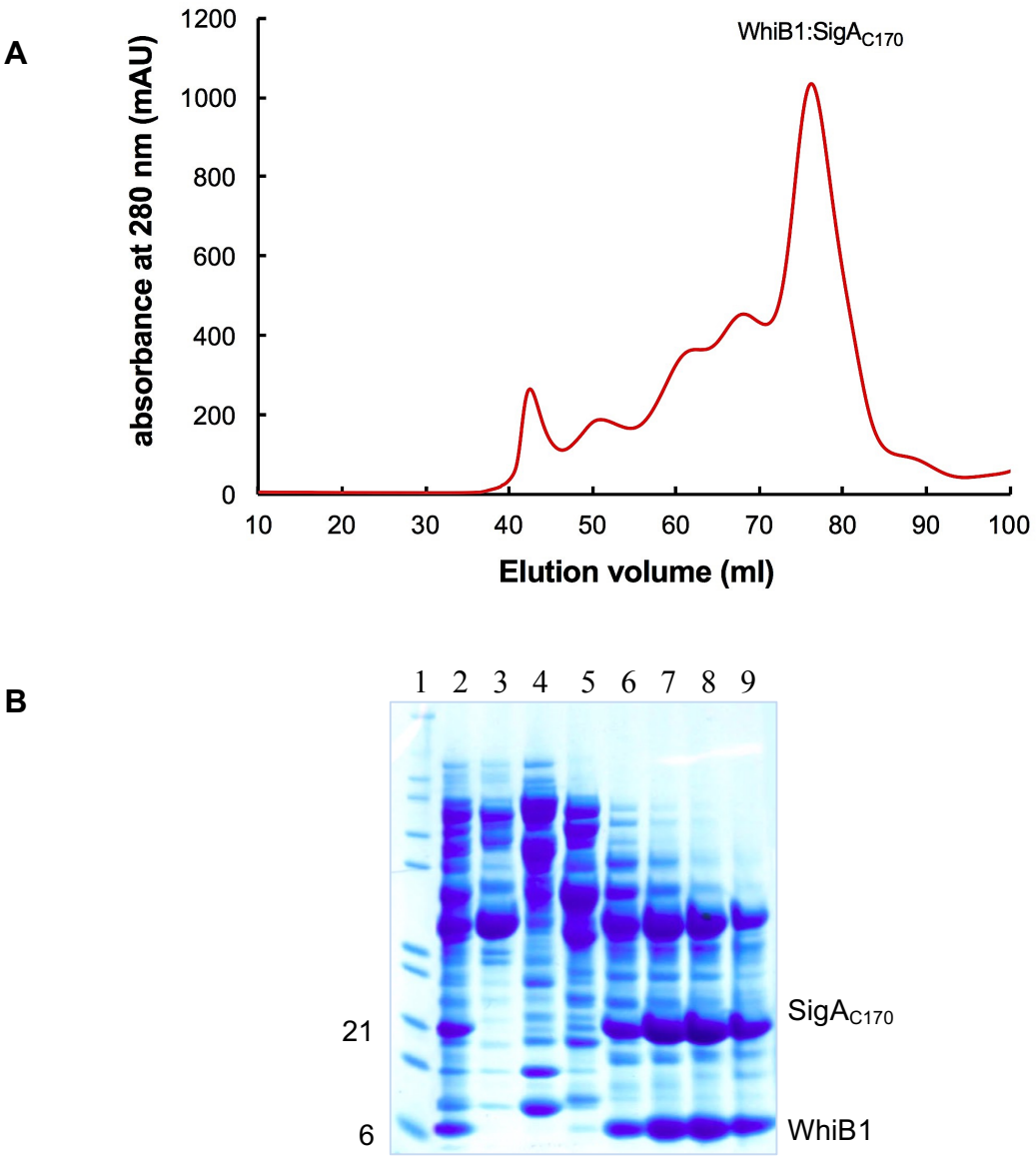


Figure 4.17: Isolation of un-tagged WhiB1:SigA$_{C170}$ complex II. [A] Elution profile from HiLoad 16/600 Superdex gel filtration column of sample eluted from HiTrap Q HP column (Figure 4.16). The trace shows protein absorbance at 280 nm (red line). WhiB1:SigA$_{C170}$ complex eluted with other *E. coli* proteins between ~72-84 ml. **[B] Coomassie stained SDS-PAGE of fractions from the elution profile of [A].** Lane 1, protein ladder (sizes in kDa are indicated); lane 2 loaded protein sample containing WhiB1:SigA$_{C170}$ purified by HiTrap Q HP column (Figure 4.16); lanes 3, 4, and 5, fractions eluted at 50, 62, and 68 ml respectively; Lanes 6-9, fractions eluted at ~72-84 ml, as indicated in **[A]**.

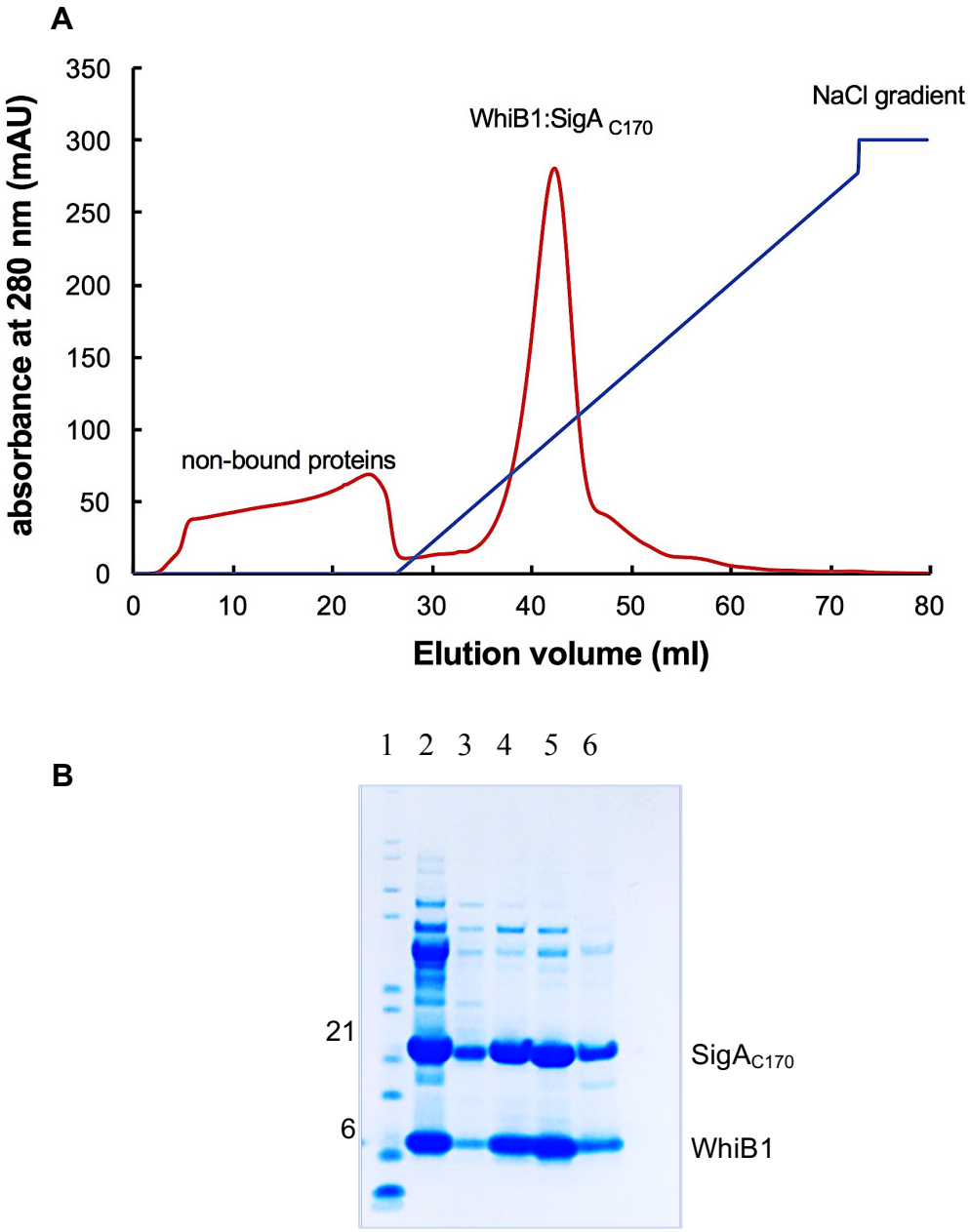


Figure 4.18: Isolation of un-tagged WhiB1:SigA$_{C170}$ complex III. [A] Elution profile from HiTrap Heparin HP chromatography of sample eluted from the gel filtration step (Figure 4.17). The trace shows protein absorbance at 280 nm (red line), and the gradient of NaCl from 0-0.5 M (blue line). WhiB1:SigA$_{C170}$ complex eluted between ~38-47 ml. **[B] Coomassie stained SDS-PAGE of fractions from the elution profile shown in [A].** Lane 1, protein ladder (sizes in kDa are indicated); lane 2 loaded protein sample containing WhiB1:SigA$_{C170}$ purified by gel filtration (Figure 4.17); lanes 3-6, fractions of WhiB1:SigA$_{C170}$ complex eluted between ~38-47 ml as indicated in **[A]**.

4.4.3 Gel filtration analysis of WhiB1:SigA$_{C170}$ complex

Previously, the WhiB1:SigA complex, and free SigA eluted anomalously during separation by gel filtration compared to their molecular weights (Figure 4.5). As a result, it was not possible to know the stoichiometry of the WhiB1:SigA complex. To attempt to answer this question the WhiB1:SigA$_{C170}$ complex was analyzed by gel filtration to determine its size and define its oligomeric state. WhiB1:SigA$_{C170}$ was purified (Section 4.4.2) and concentrated using a Vivaspin unit with a 10 kDa MW cut off to ~1.0 mM. It was then loaded onto a gel filtration column pre-equilibrated with 50 mM Tris (pH 8.0) containing 0.5 M NaCl. Protein was eluted as a single peak at ~77.5 ml (Figure 4.19). This suggested that WhiB1:SigA$_{C170}$ was a heterodimer with a molecular weight of ~31.5 kDa; calculated molecular weight of [4Fe-4S] WhiB1: SigA$_{C170}$ is 29 kDa.

4.4.4 UV-visible spectroscopic characterization of the WhiB1:SigA$_{C170}$ complex

To investigate the WhiB1:SigA$_{C170}$ complex oligomeric state after nitrosylating the iron-sulfur cluster, UV-visible spectra were obtained. The WhiB1:SigA$_{C170}$ complex after three steps of purification (Section 4.4.2) under aerobic conditions still showed a clear single feature at 420 nm (Figure 4.20, black line). This indicates the stability of [4Fe-4S] cluster. Treating the WhiB1:SigA$_{C170}$ complex (~24 µM cluster) with Proli NONOate (300 µM) (equivalent to 600 µM of NO) resulted in nitrosylation of the WhiB1:SigA$_{C170}$ complex. A peak appeared at ~360 nm accompanied by decreasing absorbance at 420 nm, and formation of two isosbestic points at 390 and 470 nm (Figure 4.20, blue line). The holo- and nitrosylated- forms of the WhiB1:SigA$_{C170}$ complex were exposed separately to O_2 by bubbling the protein samples with air three times using a 100 µl Hamilton gas tight syringe. The holo-WhiB1:SigA$_{C170}$ complex after 5 h of incubation did not show any absorbance change at 420 nm (Figure 4.20, yellow line); while, the nitrosylated WhiB1:SigA$_{C170}$ complex absorbance at 360 nm decreased by ~5% after 1 h of incubation (Figure 4.20, red line), indicating the stability of the WhiB1 iron-sulfur cluster in the complex to air.

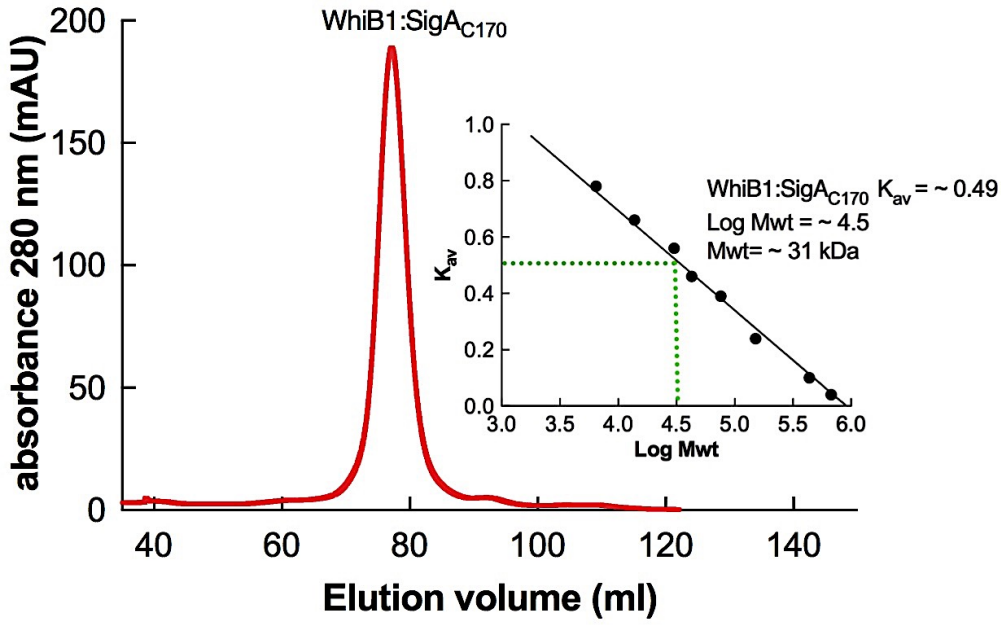


Figure 4.19: Size-exclusion chromatography analysis of holo-WhiB1:SigA$_{C170}$. The un-tagged holo-WhiB1:SigA$_{C170}$ complex (1.0 mM) was eluted from HiLoad 16/600 Superdex gel filtration column pre-equilibrated with buffer of 50 mM Tris (pH 8.0) containing 0.5 M NaCl as a single peak at ~77.5 ml (K_{av} equal to ~0.49). ***Inset***, the gel filtration calibration curve, the K_{av} value was plotted versus log [Mwt] of the standards proteins as shown in Figure 3.23.

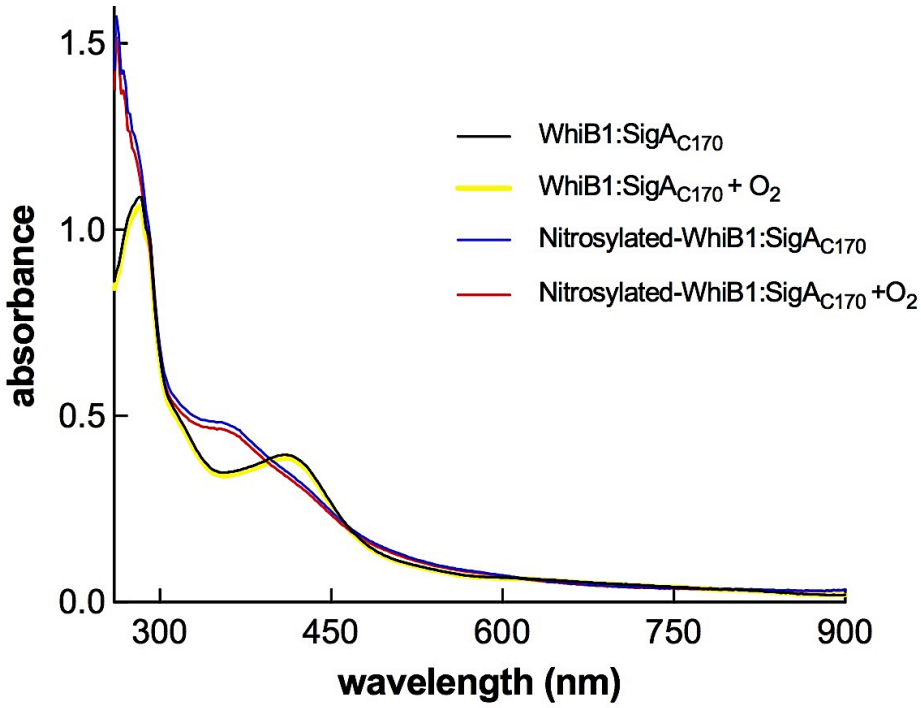


Figure 4.20: UV-visible spectroscopy of holo-WhiB1:SigA$_{C170}$. The spectrum of holo-WhiB1:SigA$_{C170}$ protein (~24 μM cluster) shows a peak at 420 nm (black line). Protein buffer was 50 mM Tris (pH 8.0) containing ~0.17 M NaCl. After bubbling the protein sample with air three times using a 100 μl Hamilton gas tight syringe, no changes to the original spectrum were detected (yellow line) after 5 h of incubation. Treating the protein with Proli NONOate (300 μM) (equivalent to 600 μM of NO) resulted a peak at ~360 nm with decreasing absorbance at 420 nm forming two isosbestic points at 390 and 470 nm (blue line). Bubbling the nitrosylated sample with air resulted in decreased absorbance at ~360 nm by ~5% within 1 h (red line).

4.4.5 Nitric oxide triggers WhiB1 release from WhiB1:SigA$_{C170}$ complex

As shown previously (Section 4.3.2), when the holo-WhiB1:SigA complex was exposed to NO, a partial separation of the nitrosylated complex occurred (Figure 4.14). Therefore, the WhiB1:SigA$_{C170}$ complex was nitrosylated under anaerobic conditions using Proli NONOate. Holo-WhiB1:SigA$_{\mathbf{C170}}$ complex (50 µM cluster), in 50 mM Tris (pH 7.4) containing 500 mM NaCl and 0.1 mM DTT was exposed to ~500 µM of Proli NONOate (equivalent to 1 mM of NO) and incubated in anaerobic workstation for 1 h before analysis by gel filtration. The elution profile from the HiLoad 16/600 Superdex column pre-equilibrated with anaerobic-degassed (50 mM Tris (pH 8.0) containing 500 mM NaCl), showed two peaks at 80 ml and ~90 ml (Figure 4.21A). The SDS-PAGE analysis showed that almost all the released SigA$_{C170}$ was eluted in the first peak, while the released WhiB1 was eluted at ~90 ml (Figure 4.21B). However, the released WhiB1 was not eluted exactly at the usual position of 93 ml (Figure 4.21, orange line), and the SDS-PAGE showed also faint species equivalent to SigA$_{C170}$ present in fractions containing WhiB1 (lanes 9-12, Figure 4.21B). The same was noticed with fractions of SigA$_{C170}$, where the SDS-PAGE showed several faint species migrating faster than SigA$_{C170}$ that could be WhiB1 (lanes 3-9, Figure 4.21B). This suggests, some of the complex remains intact and eluted with released WhiB1 or SigA$_{C170}$. The UV-visible spectrum of eluted WhiB1 after concentrating the sample by a Vivaspin unit suggested that the protein was in the apo-form.

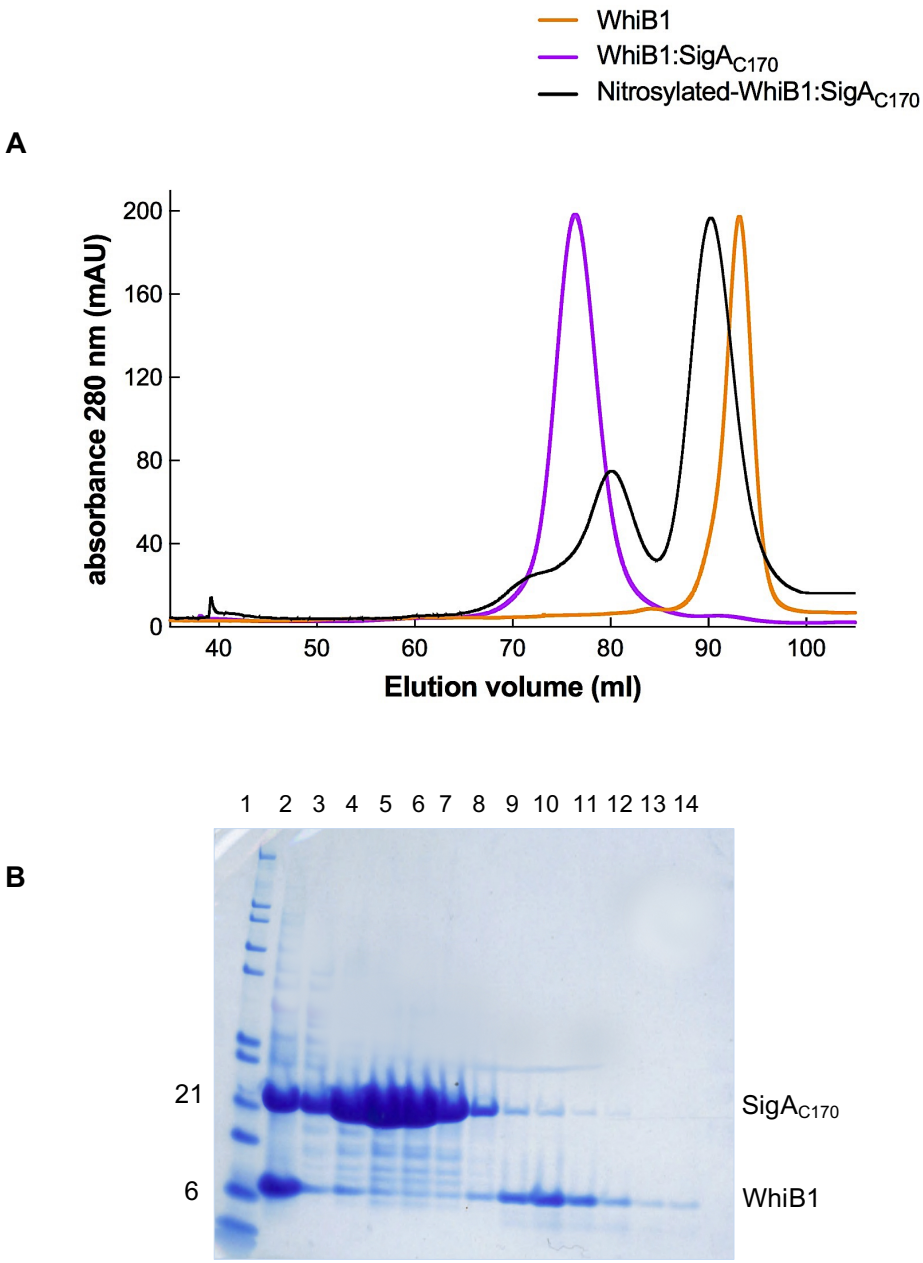


Figure 4.21: Size-exclusion chromatography analysis of nitrosylated WhiB1:SigA$_{C170}$ complex. [A] Holo-WhiB1:SigA$_{C170}$ complex (50 µM cluster), in 50 mM Tris (pH 7.4) containing 500 mM NaCl and 0.1 mM DTT, was treated with ~500 µM of Proli NONOate (1 mM of NO) before applying the sample into gel filtration column pre-equilibrated with anaerobic-degassed buffer of 50 mM Tris (pH 8.0) containing 500 mM NaCl. Free SigA$_{C170}$, and free WhiB1 were eluted at 80 ml and ~90 ml respectively. **[B] SDS-PAGE of fractions in the elution profile of Figure [A].** Lane 1, protein ladder (sizes in kDa are indicated); lanes 2, nitrosylated WhiB1:SigA$_{C170}$ sample loaded into gel filtration column; lanes 3-8, eluted fractions from ~72 -85 ml; lanes 9-14, eluted fractions at ~85-98 ml; as shown in Figure **[A]**.

4.5 Biochemical and biophysical characteristics of WhiB1:SigA$_{C82}$ complex

4.5.1 WhiB1 interacts with region 4.0 of SigA

Several transcription factors, including WhiB7 in *M. tuberculosis*, are reported to bind at region 4.0 of Sig^{70}, the essential region for recognition and binding at -35 promoter elements, (Burian et al., 2013; Dove et al., 2003). Thus, to investigate the ability of WhiB1 to bind this region of SigA, Bacterial Adenylate Cyclase Two-Hybrid System (BACTH) (Karimova et al., 1998) was carried out using *E. coli* JRG4968 (*E. coli* Δ*cya*) strain (Figure 4.22). The BACTH system is an *in vivo* assay for screening protein-protein interaction based on the reconstitution of complementary fragments (T25 and T18) of adenylate cyclase (CyaA) catalytic domain from *Bordetella pertussis*. The putative interacting proteins (X and Y) are genetically fused to T25 and T18. Association of X and Y results in functional complementation of the CyaA fragments, causing cAMP synthesis. Cyclic AMP induces the catabolite activator protein (CAP, also known as cAMP receptor protein, CRP) by effecting a conformational change in CAP, leading to DNA-binding at CAP dependent promoters. Some of these promoters control genes for maltose and lactose utilization and hence cAMP production (protein interaction) can be detected using selective or indicator media (Figure 4.22).

Interaction of WhiB1 with region 4.0 of SigA (SigA$_{C82}$, the *C-terminal* 82 amino acids of SigA, (residues 447–528)) was investigated using the BACTH system. This was achieved by transforming electro-competent BTH101 (JRG4968) strain of *E. coli* Δ*cya* cells with pGS2568 which encodes T25-WhiB1, *whiB1* gene was cloned into pGS1672 (pKT25) plasmid as an in-frame fusion with the *C*-terminal end of T25; and also with pGS2567 which encodes T18-SigA$_{C82}$, *sigA$_{C82}$* was cloned into pGS1669 (pUT18) as an in-frame fusion at the *N*-terminal end of T18 as described in 2.7.2 (Figure 4.23). Transformed cells were plated on selective media (Maltose minimal/X-gal medium, Section 2.3.6.2) or indicator media (MacConkey/maltose, and LB/X-gal medium; Sections 2.3.6.1, 2.3.6.3 respectively) plates, in the presence of IPTG for induction. For positive control pGS1671 and pGS1673 (Table 2.2) were used; and for the negative

control pGS2567 and pGS2579, that encodes T25-WhiB1 with C9W mutation in WhiB1 (WhiB1 D11R mutation is also encoded by this plasmid, this is unintended mutation) were used. The four cysteine residues of WhiB1 are essential for WhiB1:SigA interaction as described previously (Section 4.2.5). Two extra negative controls were also used, empty pUT18 co-transformed with pGS2568 (T25-WhiB1) plasmid; and empty pKT25 co-transformed with pGS2567 (T18- $SigA_{C82}$) into BTH101 cells. After incubation for 24-30 h at 30°C (for minimal medium the incubation time was 72 h), colonies on MacConkey medium were red, and blue on LB/X-gal, and on Maltose minimal/X-gal media, unlike the negative controls, where the colonies of bacteria were white with a yellow background on MacConkey plates, and white on LB/X-gal medium, while there was no growth on a minimal medium. These data provide a strong evidence that WhiB1 interacts with region 4.0 of SigA (Figure 4.23).

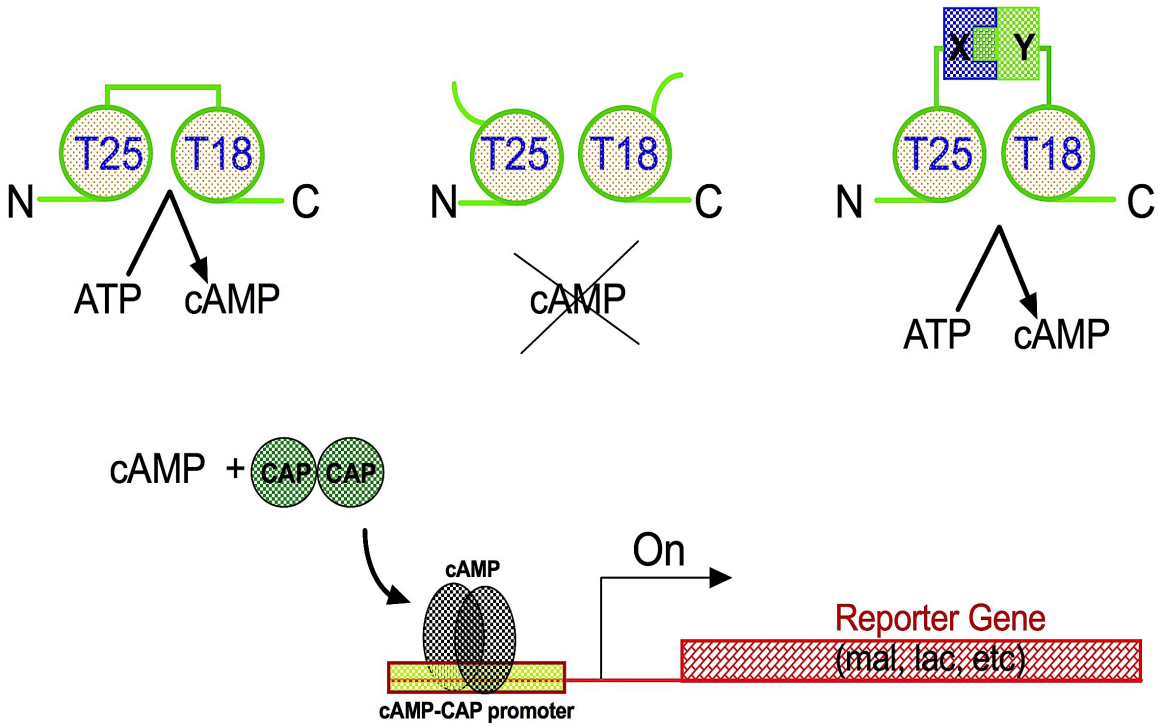


Figure 4.22 Schematic diagram of the two hybrid BACTH system adapted from Karimova et al. (1998). The catalytic domain fragments (T25 and T18) of *Bordetella pertussis* adenylate cyclase, when they are expressed as a full length protein (1-399 amino acid fused together) in *E. coli* can generate cAMP. However, if the two fragments are co-expressed separately (T25 $_{1\text{-}224}$; and T18 $_{224\text{-}399}$), no cAMP is produced. If interaction is mediated through two genetically fused interacting proteins (X and Y), the two fragments of the adenylate cyclase catalytic domain come close enough to restore cAMP synthesis. Cyclic AMP binds CAP (CRP) to permit binding at specific promoters in *E. coli* Δ*cya,* such as the *lac* operon, which leads to a transcription of the reporter gene (*lacZ*).

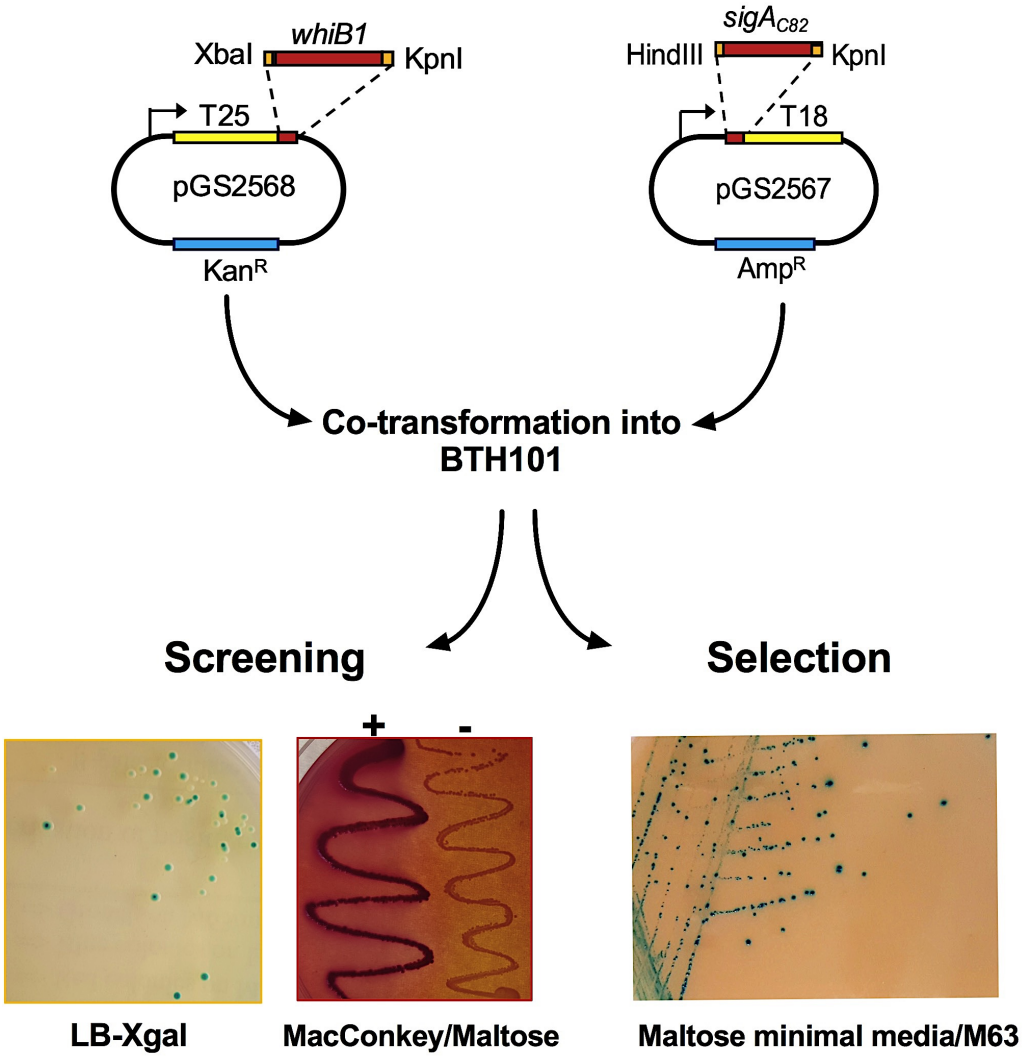


Figure 4.23: *In-vivo* interaction of WhiB1 and $SigA_{C82}$ using the BACTH system. The pGS2568 plasmid, which encodes T25-WhiB1, was created by ligating *whiB1* gene into pKT25 plasmid between XbaI and KpnI. While, pGS2567, encoding $SigA_{C82}$-T18, was produced by inserting $sigA_{C82}$ into pUT18 plasmid between HindII and KpnI. Both plasmids were co-transformed into *E. coli* Δ*cya* BTH101, and plated on MacConkey/Maltose or LB/Xgal media for screening; or on Maltose minimal medium/M63 plus Xgal for selection, all supplemented with Amp, Kan and IPTG. pGS2579 was used instead pGS2568 as a negative control, where the WhiB1 has a C9W mutation. Colonies on LB/Xgal and the minimal media were blue in colour, and red on MacConkey agar (+). White colonies with yellow background was the colour of negative control (-) on MacConkey plates.

4.5.2 Interaction between WhiB1 and $SigA_{C82}$ is disrupted by iron-starvation and exposure to NO *in vivo*

In vivo investigation of iron-starvation and NO effect on WhiB1:$SigA_{C82}$ complex integrity using the bacterial two-hybrid approach was carried out by Prof Jeffrey Green as described in Section 2.7.2. Significant interaction, as indicated by β-galactosidase activities ~20-fold greater than the negative control (JRG5386) and ~3-fold lower than the positive control (JRG5387), was detected in cultures of JRG6862, which expressed the T25-WhiB1 and T18-$SigA_{C82}$ fusions. *In vitro* analyses (Figure 4.8) indicated that the WhiB1 iron-sulfur cluster was required for formation of a complex with SigA. Restricting the availability of iron by supplementing culture medium with dipyridyl decreased β-galactosidase activity in the *E. coli* JRG6862 cultures, indicating that interaction of WhiB1 and $SigA_{C82}$ was impaired, consistent with the *in vitro* data. In contrast, the output from *E. coli* JRG5387, the leucine zipper control, was unaffected by iron-starvation (Figure 4.24). A major host response to bacterial infection, including the human response to an *M. tuberculosis* infection, is to restrict the access of the bacteria to sources of iron (Niederweis et al., 2015; Sritharan, 2016). This suggests that the WhiB1:SigA interaction could be disassembled in response to host mediated iron-starvation. The effect of NO on formation of the WhiB1:SigA complex was tested in anaerobic cultures grown in the presence of nitrite at pH 5.5. A significant reduction in β-galactosidase activity was observed compared to the control cultures (Figure 4.24). Thus, despite the presence of the endogenous NO sensing and detoxification systems of *E. coli*, the recombinant WhiB1 was able to respond, suggesting that it is sufficiently sensitive to act as an NO-responsive regulator in *M. tuberculosis*. It should also be noted that there was no significant difference in complex formation when aerobic cultures were compared to those grown under anaerobic (fumarate respiratory) conditions, suggesting that the WhiB1 iron-sulfur cluster was stable in the presence of O_2 *in vivo* (Figure 4.24). Despite the use of an heterologous host, the demonstration of disassembly of the WhiB1:SigA complex *in vivo* under conditions of iron-starvation and in the presence of NO, but not in response to O_2, is consistent with the *in vitro* analyses of the WhiB1:SigA complex and suggests a role for WhiB1 as an NO- and iron-responsive sensor-

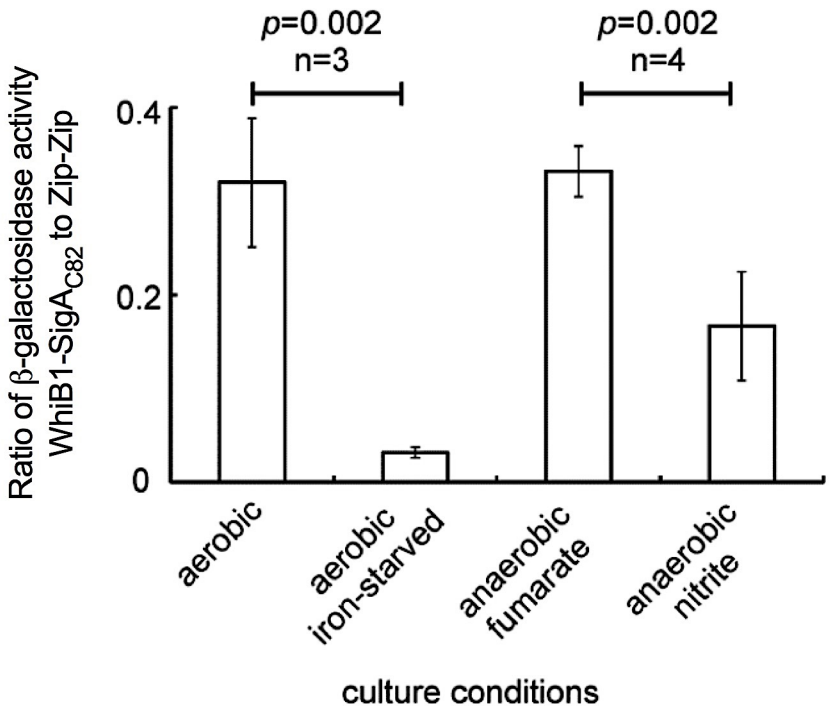


Figure 4.24: Bacterial two-hybrid analysis of interaction between *M. tuberculosis* WhiB1 and SigA$_{C82}$ fused to the T-25 and T-18 domains, respectively, of *Bordetella pertussis* adenylate cyclase in *E. coli* BTH101. Interaction between the fusion partners is manifest as enhanced production of β-galactosidase. Control cultures were of the same host transformed with plasmids encoding fusions to the GCN4 leucine zipper (Zip-Zip) (Karimova et al., 1998). The data are shown as the β-galactosidase activities obtained for WhiB1 and SigA$_{C82}$ divided by the those obtained for control (Zip-Zip) cultures grown under the indicated conditions. The chart shows the mean and standard deviation for independent cultures; significance (t-test) is indicated. This experiment was carried out by Prof Jeffrey Green.

regulator during *M. tuberculosis* infections.

4.5.3 Overproduction and purification of WhiB1:SigA$_{C82}$ complex

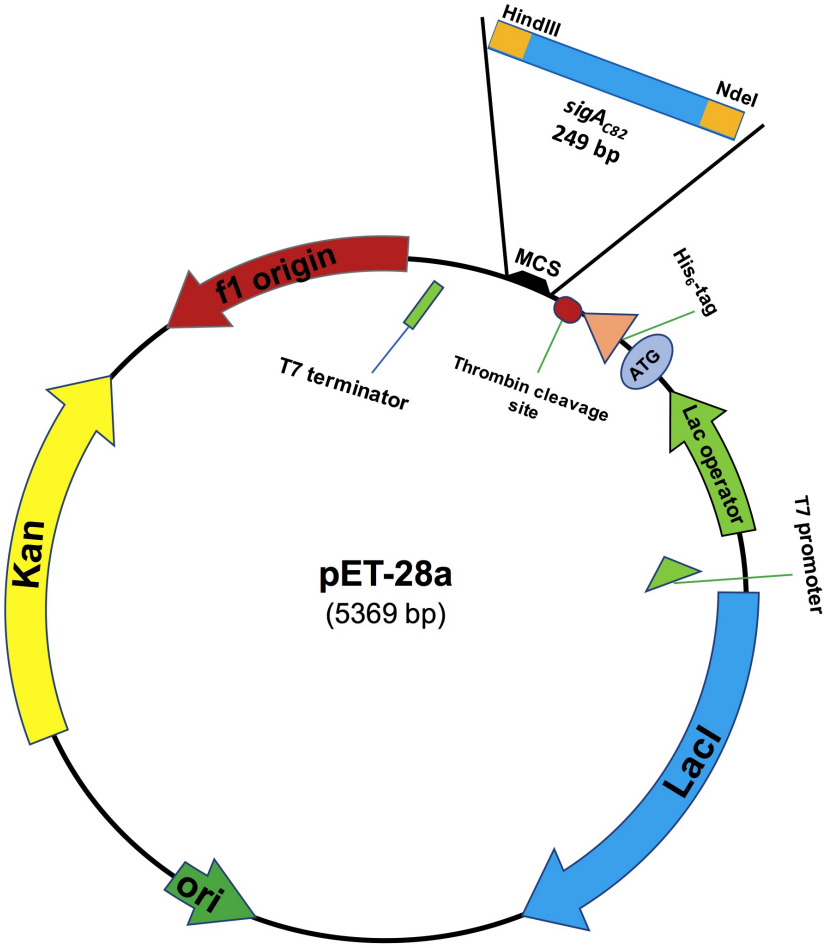
WhiB1:SigA$_{C82}$ complex overproduction was carried out as for WhiB1:SigA and WhiB1:SigA$_{C170}$ complexes (Section 4.2.1; 4.4.2). Where, *E. coli* JRG6857 was created by transforming electro-competent *E. coli* JRG5302 with pGS2566, encoding SigA$_{C82}$ with an *N*-terminal His$_6$-tag and a linker containing a Thrombin cleavage site (His$_6$-ThCS-SigA$_{C82}$), and pGS2560, encoding WhiB1 without a tag (Figure 4.2; 4.25). Transformation was done as described in Section 2.3.14.

Overproduction conditions for WhiB1:SigA$_{C82}$ complex were similar to those for WhiB1:SigA or WhiB1:SigA$_{C170}$. The *E. coli* JRG6857 cultures were grown until the $OD_{600\ nm}$ reached 0.6-0.7 and then induced with 1 mM final concentration of IPTG with a further incubation at 18°C for 12-18 h and shaking at 250 rpm (Section 2.5.3.5).

For protein isolation, affinity chromatography using a nickel-charged HiTrap chelating column was used. If the proteins of the complex interact, the non-tagged WhiB1 will bind His-tagged SigA$_{C82}$ which in turn will bind to the nickel-charged beads of the column.

The *E. coli* cell lysate containing over-produced WhiB1 and SigA$_{C82}$ was prepared by sonicating the cell pellets, suspended in binding buffer (500 mM NaCl, 50 mM Tris, pH 8.0) as described previously (Section 2.5.5.1). The elution profile from HiTrap chelating chromatography showed two peaks after applying a gradient concentration of imidazole from 0-0.5 M. The first peak was formed from non-specifically bound materials; while, the WhiB1:SigA$_{C82}$ complex was eluted as a brown-coloured protein in the second peak (Figure 4.26A). Fractions were analyzed by SDS-PAGE which confirmed the elution position of the complex (Figure 4.26B).

A



B

MGSSHHHHHHSSGLVPR↓GSH-SigA$_{C82}$

His$_6$-tag Thrombin cleavage site

Figure 4.25: Schematic diagram of pGS2566. [A] The *sigA*$_{C82}$ open reading frame (249 bp) was amplified by PCR from genomic DNA of *M. tuberculosis* with engineered NdeI/HindIII restriction sites, and ligated into pET28a to create pGS2566. **[B]** The plasmid encodes a His$_6$-ThCS-SigA$_{C82}$ fusion protein. The arrow marks the peptide bond susceptible to cleavage by Thrombin.

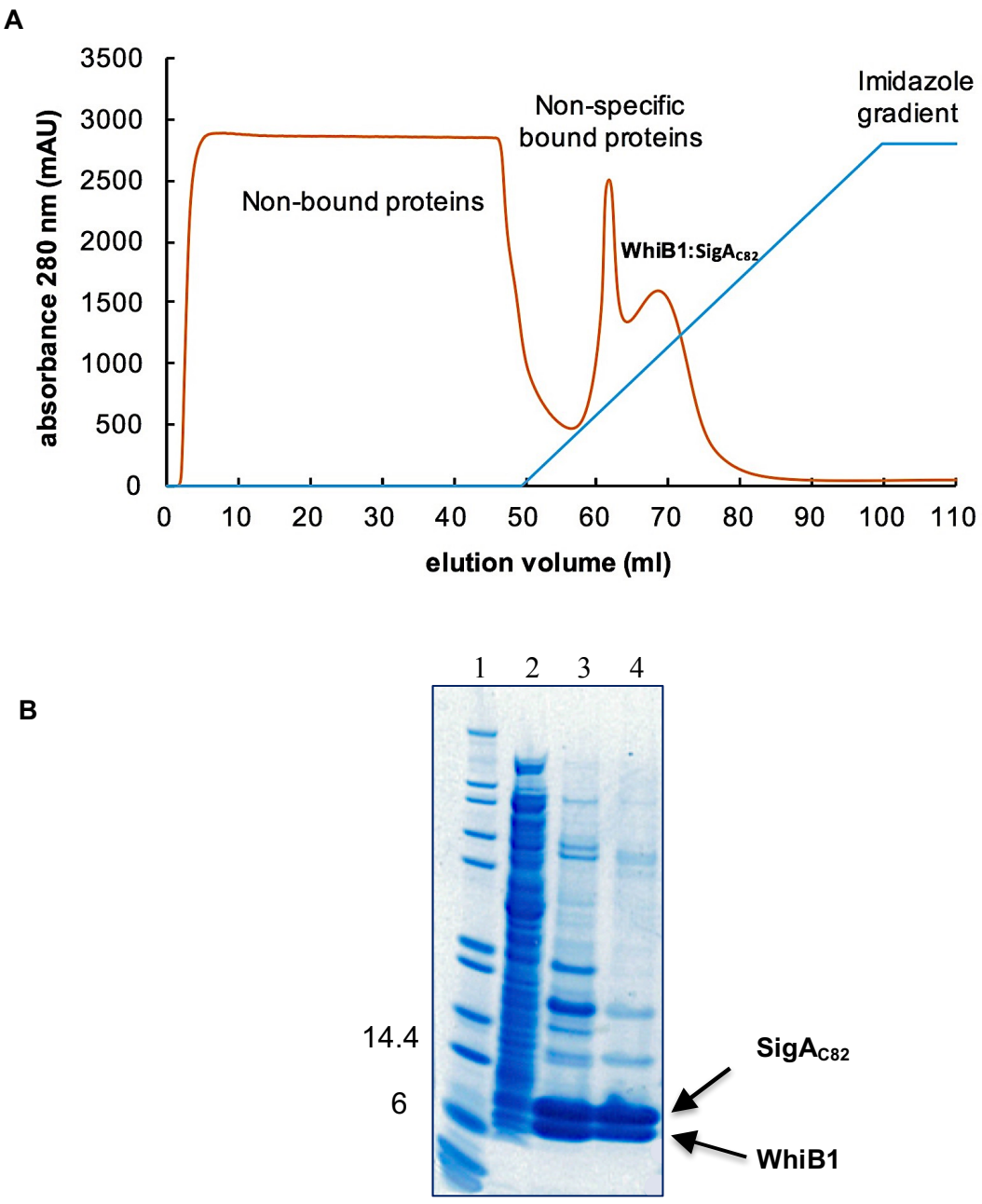


Figure 4.26: [A] Nickel-charged HiTrap chelating column elution profile of *E. coli* cell lysate containing overproduced WhiB1 and SigA$_{C82}$. The trace shows protein absorbance at 280 nm (red line), and the concentration of imidazole from 0-0.5 M (blue line). Eluted WhiB1:SigA$_{C82}$ complex appears between ~63-78 ml. **[B] Coomassie stained SDS-PAGE of isolated fractions from [A].** Lane 1, protein ladder (sizes in kDa are indicated); lane 2 cell-free extract of induced WhiB1:SigA$_{C82}$; lanes 3-4, fractions of WhiB1: SigA $_{C82}$ complex eluted between ~63-78 ml, as indicated in **[A]**.

4.5.4 Gel filtration analysis of WhiB1: $SigA_{C82}$ complex

Size exclusion gromatography of WhiB1:$SigA_{C82}$ was carried out for several purposes. During the isolation of the complex by affinity chromatography (Section 4.5.1) free $SigA_{C82}$ molecules, which can also bind the nickel-charged column, might be eluted along with the complex due to the presence of *N*-terminal His_6-tag in $SigA_{C82}$, unlike WhiB1 which was overproduced without a tag. Thus, HiLoad 16/600 Superdex 200 pg column was used as a second step of isolation to get a high purity sample for further experiments. In addition, gel filtration was useful for studying the integrity of WhiB1:$SigA_{C82}$ complex as well as the WhiB1 iron-sulfur cluster under aerobic conditions. Moreover, the analysis by size-exclusion chromatography provided information about the molecular mass of the complex and its oligomeric state. After injecting 0.5 mM of the complex onto the gel filtration column pre-equilibrated with 50 mM Tris (pH 8.0) containing 0.5 M NaCl, the complex was eluted at 86 ml, as confirmed by SDS-PAGE (Figure 4.27). This suggested that the molecular weight of protein was ~ 21 kDa, indicating that the complex is a heterodimer. In addition, the complex eluted with a brown colour and the UV-visible spectrum showed a peak at 420 nm (Figure 4.28), indicating the presence of [4Fe-4S] cluster which was stable during gel filtration for ~ 1 h (flow rate 1.5 ml/min). These data suggest that the [4Fe-4S] cluster of WhiB1:$SigA_{C82}$ complex under aerobic conditions was stable. The UV-visible spectra were also not changed after treating the protein sample with O_2-saturated buffer for 1 h, confirming complex and cluster integrity. The spectrum also suggested that the complex was a dimer protein, WhiB1:$SigA_{C82}$ complex concentration was estimated using the extinction coefficient factor (19480 M^{-1} cm^{-1}) at 280 nm, calculated by Expasy tool (Section 2.5.1). Then the WhiB1 concentration was estimated using the extinction coefficient (16500 M^{-1} cm^{-1}) at 420 nm of the [4Fe-4S] cluster (Sweeney and Rabinowitz, 1980). The ratio of WhiB1:$SigA_{C82}$ was 1:1, confirming the gel filtration chromatography result.

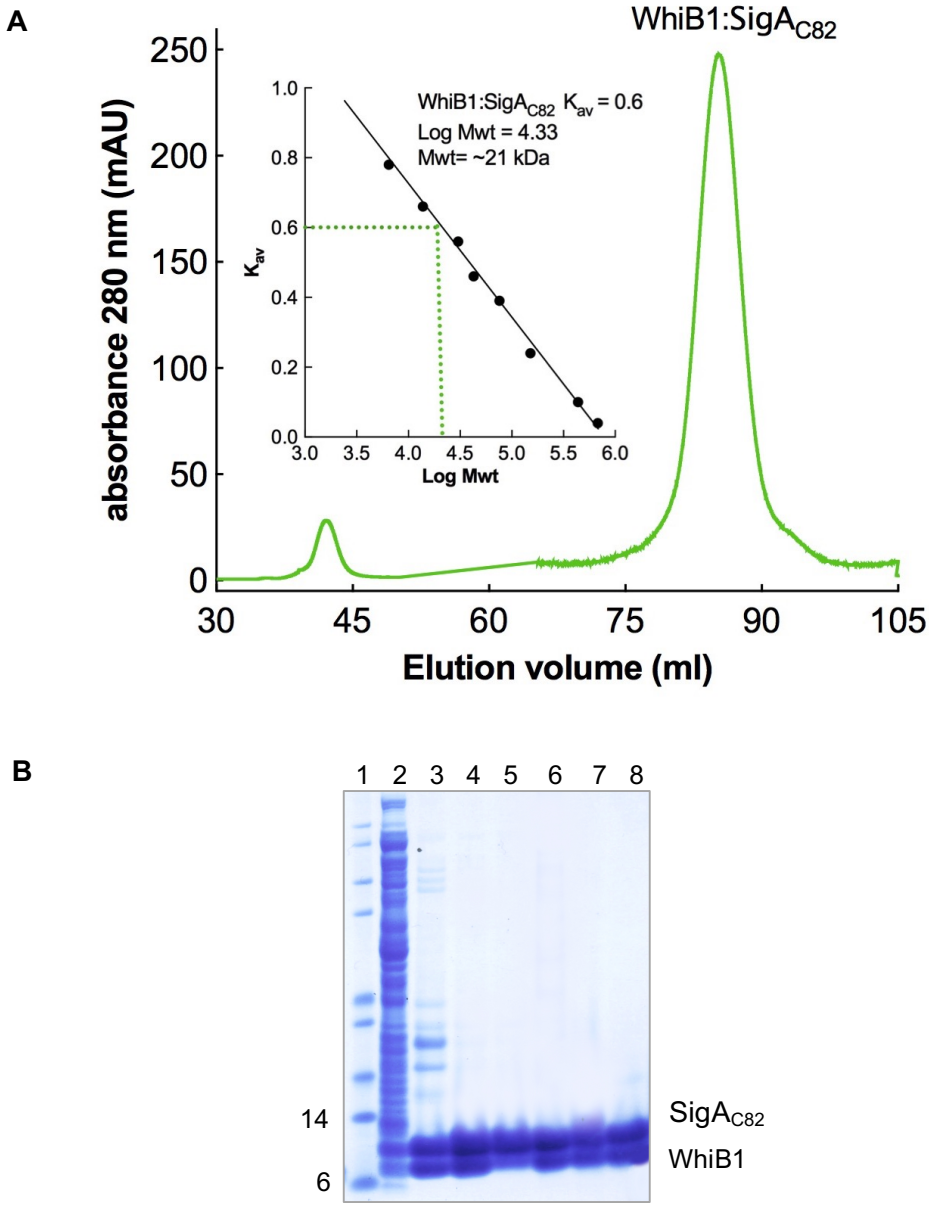


Figure 4.27: Gel filtration analysis of holo- WhiB1:SigA$_{C82}$. [A] The holo-WhiB1:SigA$_{C82}$ complex (0.5 mM) was injected onto HiLoad 16/600 Superdex gel filtration column pre-equilibrated with 50 mM Tris (pH 8.0) containing 0.5 M NaCl. The holo-complex was eluted as a single peak at 86 ml with a K_{av} equal to 0.6. ***Inset*, the gel filtration standarization curve.** The K_{av} value was plotted versus log [Mwt] of the standards proteins as shown in Figure 3.24. **[B] SDS-PAGE of fractions from the elution profile shown in [A].** Lane 1, protein ladder (sizes in kDa are indicated); lane 2, cell-free extract of induced WhiB1:SigA$_{C82}$; lane 3, sample loaded into gel filtration column, lanes 4-8, fractions from ~77 -90 ml as shown in **[A]**.

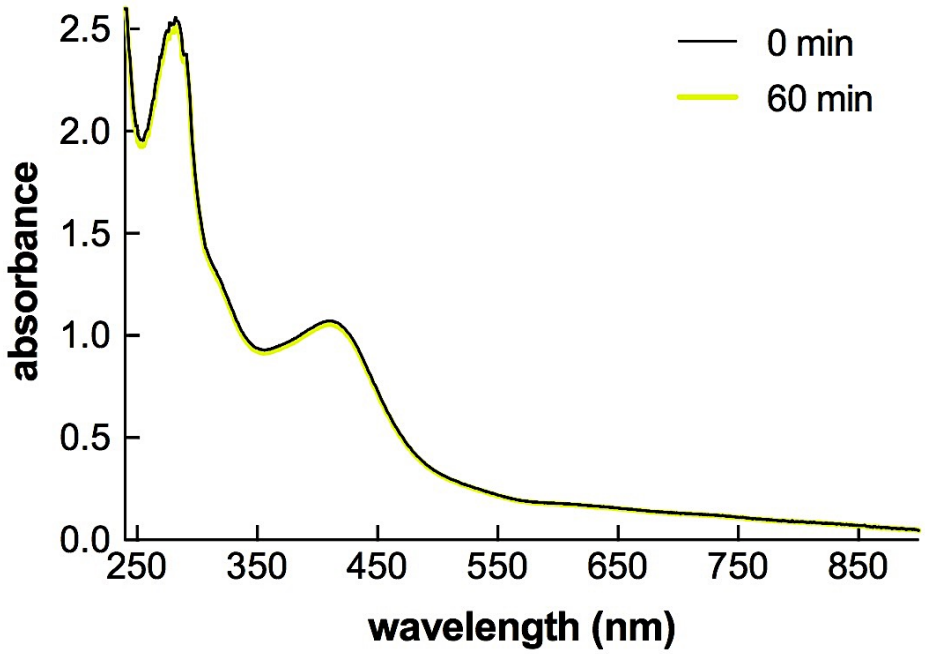


Figure 4.28: UV-visible spectra of holo-WhiB1:SigA$_{C82}$ complex. The spectrum of holo-WhiB1:SigA$_{C82}$ protein (~65 μM cluster), which was isolated using gel filtration column pre-equilibrated with 50 mm Tris (pH 8.0) containing 0.5 M NaCl, shows a peak at 420 nm (black line). After treating the protein sample with O_2-saturated buffer (final concentration of 110 μM), no changes in the spectrum were detected (yellow line) after 1 h of incubation.

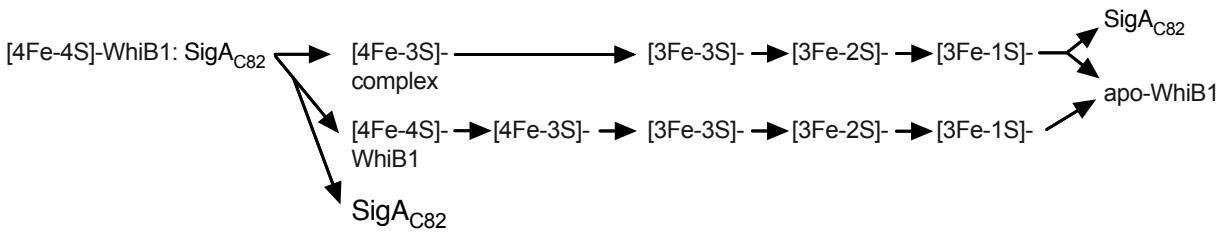
4.5.5 WhiB1:SigA$_{C82}$ complex iron-sulfur cluster degradation is initiated by loss sulfur atoms

In-source collision-induced dissociation (isCID) was carried out using electrospray ionization mass spectrometry in a collaboration with Dr. Jason Crack. This provided information on the pathway of WhiB1 cluster degradation when it is in complex with SigA$_{C82}$, and the consequences of that on the complex integrity. The protein buffer was exchanged using Zeba desalting columns (Thermo Scientific) pre-equilibrated with 250 mM ammonium acetate pH 8.0, and then the protein sample was diluted to ~36 μM (cluster). The complex then was injected directly (0.3 ml/h) into the ESI source of a Bruker micrOTOF-QIII mass spectrometer through the capillary tube of the ESI part, maintained at 2.7-4.5 kV in positive ion mode; collision radio frequency (RF) was fixed at 650 Vpp, and collision cell energy at 6 eV. The nebuliser gas was set at 0.8 Bar, with a gas temperature of 180°C, and 4 L/min flow rate, the quadrupole set to filter out incoming ions of ≤ 20 kDa. The emitted ions from [4Fe-4S]-WhiB1:SigA$_{C82}$ (ion energy was set at 4 eV) in the gaseous phase entered the inlet inner capillary tube; where, cluster fragmentation and complex separation were studied by increasing the voltage (from 0 to 200 eV in 21 cycles) between the capillary exit and the hexapole skimmer of the spectrometer.

Free [4Fe-4S]-WhiB1, free SigA$_{C82,}$ and the complex of [4Fe-4S]-WhiB1: SigA$_{C82}$ were present at 0 eV (Figure 4.29, green line), suggesting some complex separation happened during sample ionization. To exclude the incoming free WhiB1 and SigA$_{C82}$ species, Quad 1 of the mass spectrometer was set to filter out masses below 20 kDa (WhiB1: SigA$_{C82}$ complex has a distinct ionization profile >2000 m/z), this should significantly reduce incoming free WhiB1 and SigA$_{C82}$ ions versus the complex, if the complex can resist dissociation during transits. Under these conditions a significant reduction in the intensity of free WhiB1 and free SigA$_{C82}$ was observed (Figure 4.29, the black line), suggesting the complex was sufficiently stable to be studied by this technique.

The isCID data showed that WhiB1:$SigA_{C82}$ with the [4Fe-4S] cluster was the predominant species at the beginning of the experiment (0 eV), with a molecular weight of 21,191 Da (Figure 4.30C, black line). This confirmed that the WhiB1 present in a complex with SigA has a [4Fe-4S] cluster, and the complex is a heterodimer. Changing the capillary exit voltage resulted in the detection of a series of species at molecular weights lower than the [4Fe-4S]-WhiB1:$SigA_{C82}$ complex. These data indicated that the degradation of the [4Fe-4S] cluster in the WhiB1:$SigA_{C82}$ complex was initiated by loss of a sulfur atom whilst maintaining complex integrity. Following this different iron-sulfur clusters were formed in WhiB1, but not the apo-form whilst WhiB1 remained in complex with $SigA_{C82}$.

Four peaks between 21-21.2 kDa appeared after changing the voltage, suggesting the formation of [4Fe-4S]-, [4Fe-3S]-, [3Fe-2S]-, [3Fe-1S]-WhiB1:$SigA_{C82}$ forms (Figure 4.30B). Data also showed free $SigA_{C82}$ at 11,522 Da and free holo-WhiB1 were released after changing the capillary exit voltage. Moreover, cluster degradation of released holo-WhiB1 was quite similar to that seen for $WhiB1^{Ms}$ with His_6-TEV tag (Figure 3.21), where cluster fragmentation was also initiated by sulfur atom loss (Figure 4.30A). Interestingly, a [2Fe-2S] form was not identified, confirming the previous observation of not seeing the [2Fe-2S] spectrum in $WhiB1^{Ec}$ during the cluster reconstitution or degradation (Section 3.3.1, 3.3.4). According to these data it was concluded that the following pathways of cluster fragmentation and complex separation were occurring.



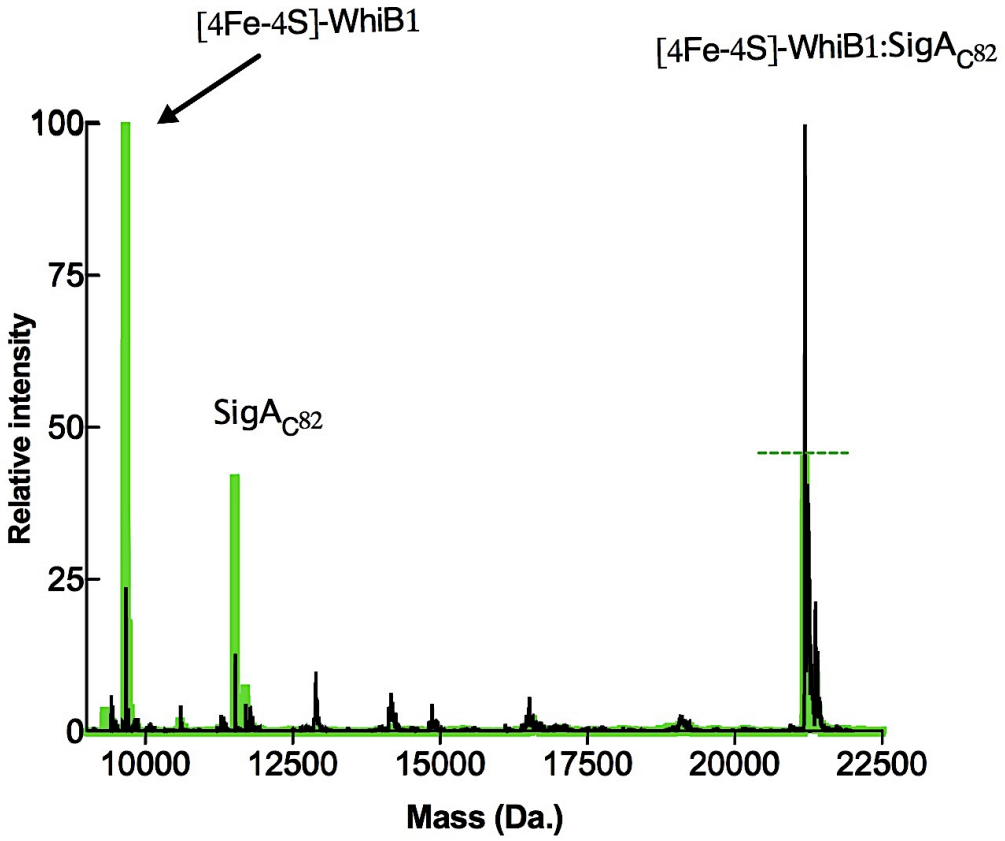


Figure 4.29: Mass-spectrometry analysis of [4Fe-4S]-WhiB1:SigA$_{C82}$ before and after a mass filter applied below 20 kDa. The complex of [4Fe-4S]-WhiB1:SigA$_{C82}$ (~36 μM cluster) in 250 mM ammonium acetate pH 8.0, was injected in to the mass spectrometer and the spectrum was read. Spectrum before filtration (green line) shows the [4Fe-4S]-WhiB1:SigA$_{C82}$ at 21,191 Da (the top intensity point ends at the dashed line ~45%); free SigA$_{C82}$ at 11,522 Da; and [4Fe-4S]-WhiB1 at 9,669 Da. The relative intensity decreased to less than 25% of free WhiB1 and SigA$_{C82}$ after applying the mass filtration (black line).

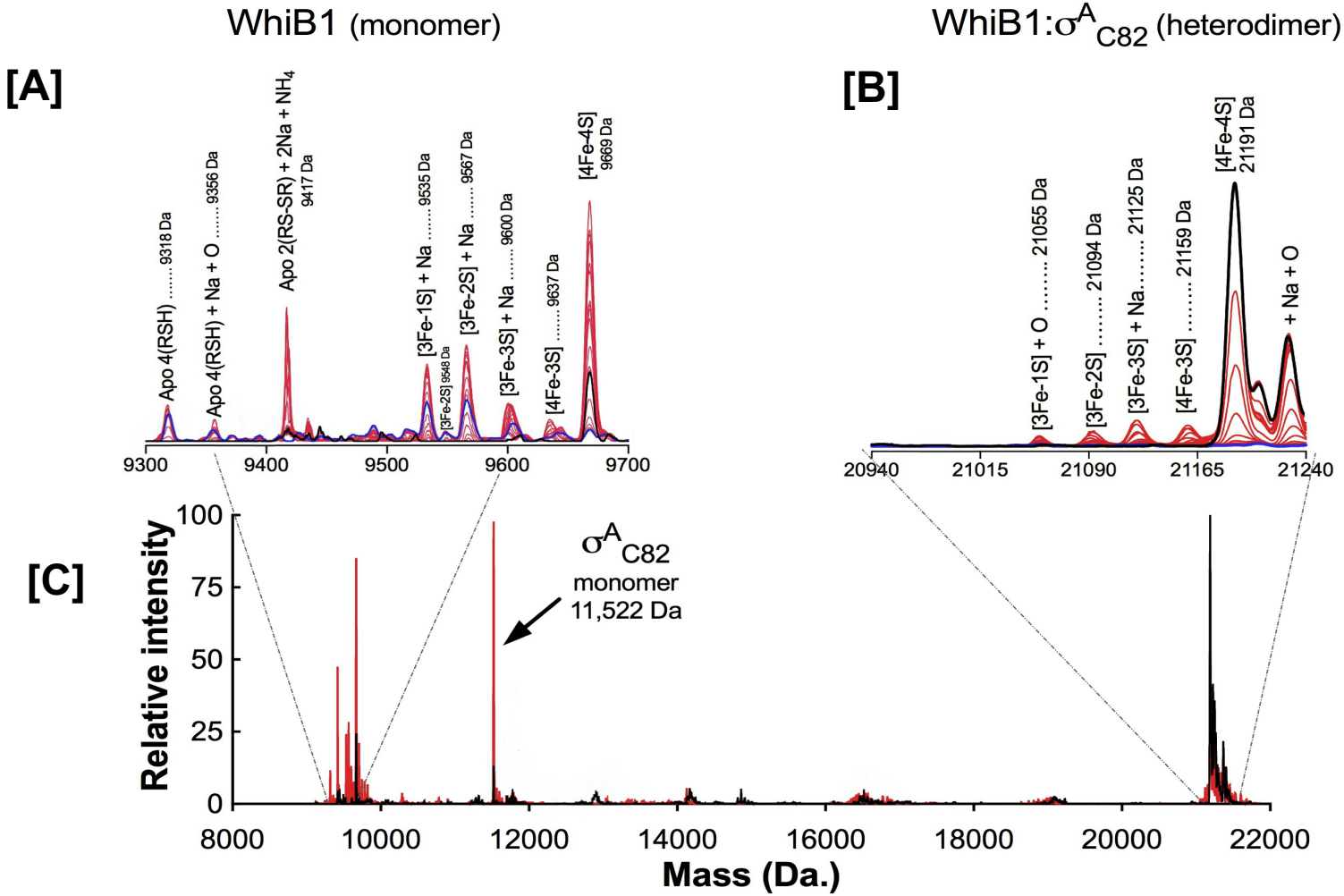


Figure 4.30: Electrospray ionization mass spectrometry (ESI-MS) analysis of WhiB1:SigA$_{C82}$. The black line in Figure **[C]** represents the intact [4Fe-4S]-WhiB1:SigA$_{C82}$ complex before changing the capillary exit voltage. The in-source collision-induced dissociation (isCID) voltage was increased from 0 to 200 eV, in the order of (0, 10, 20, 30 etc… 200 eV); where, the black line represents the first run at 0 eV, while the blue line represents the last run at 200 eV, red traces represent the runs in between. The WhiB1 species that correspond to the masses detected are shown above the major peaks in Figure **[A]**; while, the species of WhiB1:SigA$_{C82}$ are shown as in Figure **[B]**.

4.6 Discussion

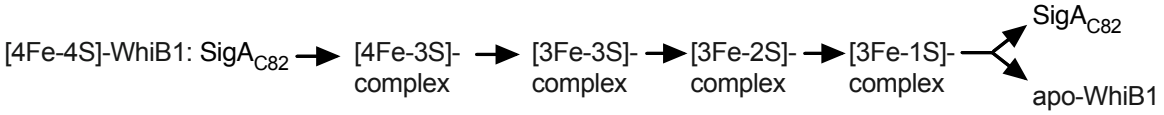
The biochemical and biophysical characteristics of the iron-sulfur cluster of WhiB1 in a complex with SigA have been investigated in this chapter. In addition, the binding interface between WhiB1 and SigA has been studied. Co-expression of WhiB1 and SigA (or SigA domains) in *E. coli* was carried out to identify the nature of the WhiB1 iron-sulfur cluster when the protein in a complex with SigA.

When SigA, (or $SigA_{C170}$, or $SigA_{C82}$) was overproduced in *E. coli* along with WhiB1, a complex was formed which could be isolated by chromatography, as has been observed for some other Wbl proteins (Burian et al., 2013; Feng et al., 2016; Steyn et al., 2002). The WhiB1:SigA complex was yellowish-brown in colour, as were WhiB1:$SigA_{C170}$ and WhiB1:$SigA_{C82}$ (Section 4.4.1; 4.5.2). The UV-visible spectra of these complexes showed a peak at 420 nm; revealing the presence of [4Fe-4S] cluster which was not degraded upon prolonged exposure to O_2 (Figure 4.7, 4.20, 4.28). The presence of a [4Fe-4S] cluster in WhiB1 when in complex with SigA was precisely identified by measuring the mass of the WhiB1:$SigA_{C82}$ complex using ESI-mass spectrometry (Figure 4.29, 4.30; black lines). This result is similar to that shown by Burian *et al.* (2013), who reported that WhiB7 in complex with SigA contained a [4Fe-4S] cluster. These data indicated that the partner protein SigA stabilizes the WhiB1 iron-sulfur cluster.

Although, both $WhiB1^{Ec}$ and the complex of WhiB1:SigA were expressed in *E. coli*, unlike the $WhiB1^{Ms}$ which was overproduced in *M. smegmatis*; the cluster stability of WhiB1 in the complex was quite similar to that seen in $WhiB1^{Ms}$ containing a His_6-TEV-tag but was different to $WhiB1^{Ec}$. This result supports the previous findings that the cluster integrity in $WhiB1^{Ms}$ was probably because of the TEV tag presence, and was not due to the *M. smegmatis* host (Section 3.4.4). Moreover, the *in vitro* mixing of His_6-TEV-$WhiB1^{Ms}$ with free SigA resulted in no interaction as judged by gel filtration (Section 4.2.6), suggesting that the TEV-tag in $WhiB1^{Ms}$ could mimic the interaction site(s) of SigA causing stabilization of the WhiB1 [4Fe-4S] cluster.

The presence of the iron-sulfur cluster seems essential for some proteins to bind a sigma factor but not crucial for others. For example, Kumar *et al.* (2012) reported that the non-Wbl protein, the *M. tuberculosis* anti-sigma factor RshA, which has a cysteine-ligated [2Fe-2S] cluster after protein isolation, and [4Fe-4S] after cluster *in vitro* reconstitution, was still able to interact with SigH after removing the capacity to bind an iron-sulfur cluster by mutation of two cysteine residues, indicating that the cluster was not essential. In contrast, the *Streptomyces coelicolor* anti-sigma factor RsmA required its [2Fe-2S] cluster for interaction with SigM (Gaskell et al., 2007). While, in *M. tuberculosis* WhiB7, which has four cysteine residues anchoring a [4Fe-4S] cluster, after Cys 48 or both Cys 45 and Cys 48 were mutated to alanines and by using BacterioMatch system, no growth was detected under selective conditions. Suggesting no WhiB7:SigA interaction occurred in the absence of the WhiB7 iron-sulfur cluster. However, pull down assays could not confirm the result due to instability of the WhiB7 variants with SigA or even alone (Burian et al., 2013). Three individual Cys mutations (C9A, C40A, and C46A) in WhiB1 resulted in the inability of WhiB1 to co-purifiy with SigA (Figure 4.8). In addition, the two hybrid BACTH system confirmed the inability of the WhiB1 C9W mutant to bind $SigA_{C82}$ (Section 4.5.1). Thus, it is concluded that the WhiB1:SigA interaction is governed by the presence of the iron-sulfur cluster, and the interaction could be modulated in a response to external signal that modifies or degrades the WhiB1 cluster. *In vivo* WhiB1 and $SigA_{C82}$ complex formation was also impaired under iron starvation conditions, this result is consistent with the *in vitro* data (Section 4.5.2, Figure 4.24).

The results of isCID mass spectrometry also showed that the degradation of [4Fe-4S] cluster of WhiB1:$SigA_{C82}$ complex was initiated by losing sulfur atoms, and the complex remained intact until WhiB1 completely lost the cluster (Figure 4.30 B); the cluster modification pathway is as follows:



This result is consistent with the previous finding of cluster dissociation in $WhiB1^{Ms}$, which was also initiated by sulfur loss (Figure 3.21). In addition, the

results of LC-MS analysis of the [4Fe-4S]-WhiB1:SigA complex showed that after cluster degradation and complex separation, the WhiB1 cysteine residues were not able to retain sulfur adducts (Figure 4.9). In contrast, preliminary data for WhiB2^{Ms} indicated that it lost iron atoms first and retained the sulfur to form a cysteine persulfide ligands to a [2Fe-2S] cluster (Figure 3.20B). These data confirm that WhiB1 loses sulfur atoms first and then iron but without a sulfur retention, verifying the data of isCID mass spectrometry (Figure 4.30).

Smith *et al.* (2010), reported that both apo- and nitrosylated-WhiB1 forms were able to bind the non-coding region of *whiB1*, but the holo-form did not. Here, a similar pattern of interaction with NO as seen previously in WhiB1^{Ms} (Figure 3.22, 3.23), where, the reaction of [4Fe-4S]-WhiB1:SigA with NO was extremely rapid, and titration showed that nearly eight molecules of NO reacted with each WhiB1 [4Fe-4S] cluster (Figure 4.11; 4.12). The NO effect on complex integrity was then analyzed by gel filtration and showed a partial separation of WhiB1 from the complex, along with the formation of a nitrosylated complex (Figure 4.14). The analysis of nitrosylated-WhiB1:SigA$_{C170}$ complex by gel filtration also showed release of WhiB1 with presence of some intact complex molecules as judged by SDS-PAGE (Figure 4.21). Experiments also showed that the degradative effect of O_2 on nitrosylated WhiB1:SigA was extremely slow (Figure 4.13) similar to the nitorsylated-WhiB1 behaviour reported by Smith *et al.* (2010). This excludes the possible side effect of O_2 - which might occur due to the incomplete anaerobic conditions of experiments - on nitrosylated complex integrity. According to these data, it is concluded that NO triggers the release of WhiB1 from the complex. However, the reason(s) for the partial separation of WhiB1 and the formation of a nitrosylated complex, as seen in the case of nitrosylated WhiB1:SigA complex, is not totally clear. Perhaps it could be due to the formation of different forms of nitrosylated iron-sulfur cluster, such as the presence of individual pairs of Roussin's red ester (REE), octa-nitrosylated iron-sulfur (dimerization of RRE), or even Roussin's black ester (hepta-nitrosylated iron-sulfur cluster, RBE) in each a monomer. Possibly some of these clusters might trigger complex separation, while others do not. This suggestion is supported by the findings made by Serrano *et al.* (2016), who reported that the *S. coelicolor* NsrR and the Wbl protein (WhiD)

when they reacted with NO, the RRE and Roussin's Black salt (RBS) related products were diagnosed; moreover, the RBE was suggested to be formed in both proteins due to replacing of the RBS sulfide ligands by one or more Cys thiolate group. It was also reported by Crack *et al.* (2016), that different nitrosylated species of *S. coelicolor* NsrR showed a different pattern in activating/deactivating the protein-DNA binding ability. The [4Fe-4S]-NsrR form was able to bind the non-coding area of *nsrR, hmpA1, hmpA2*. Approximately, 2.5 NO per protein cluster were enough to deactivate the DNA binding at *hmpA2,* in contrast 4 and 8 NO per cluster were needed to completely abolish the binding of *hmpA1* and *nsrR* promoters respectively. This indicates that the presence of different nitrosylated iron-sulfur species leads to different protein function. The same thing could be true in the case of WhiB1, where different nitrosylated iron-sulfur species of WhiB1 might carry out different genetic regulatory functions in *M. tuberculosis. In vivo* interaction of WhiB1 with *C*-terminal domain in *E. coli* exhibited a significant decrease in β-galactosidase activity under NO conditions. This result is also congruent with the *in vitro* data and suggest that WhiB1 is a NO-responsive regulatory protein (Section 4.5.2, Figure 4.24).

The binding interface between WhiB1 and SigA was also studied here. As mentioned previously, SigA proteins share four main domains (Figure 4.31), and in *M. tuberculosis* the *N*-terminal domain 1 comprises two thirds of the protein size (359 a.a), while the *N*-terminal domain 2, linker domain and *C*-terminal domain covers the rest (167 a.a) (Lonetto et al., 1992; Vassylyev et al., 2002) (Figure 4.32). BACTH system results indicated that WhiB1 binds the *C*-terminal domain consisting of region 4 of SigA (last 82 a.a) (Figure 4.32). This region of SigA is essential for binding -35 elements of promoters, and is also a target for several regulatory transcription factors, some of which consist of a sigma factor binding domain and a sensory domain to respond to external stimuli (Dove et al., 2003). These interactions between a regulatory (WhiB1) and a sigma factor (SigA) can either prevent the formation of holo-RNAP by sequestering the sigma factor from binding with core RNAP acting in this case as anti-sigma factor (Paget, 2015); or reprogram transcription by binding with sigma factor and DNA promoters simultaneously (Dove et al., 2003). The fact

that *M. bovis* is attenuated by the R515H mutation in region 4 of SigA, and similar attenuation was caused by *whiB3* deletion, strongly suggest a model in which the stimulation of unknown virulence gene(s) requires WhiB3 and SigA interaction (Collins et al., 1995; Steyn et al., 2002). Furthermore, WhiB7, which is the only *M. tuberculosis* Wbl that has an AT-hook domain, can bind also region 4 of SigA. The AT-hook binding motif, the abolishing of WhiB7:SigA interaction in the presence of the R515H mutation in SigA and the WhiB7-specific multi-drug sensitivity effect due to R515H mutation, which is similar to the effect of the AT-hook deletion, all suggest that WhiB7 acts through stabilizing the SigA–DNA binding, providing an extra selectivity for specific promoters under specific conditions (Burian et al., 2013). Thus, it seems that WhiB1 might act similarly to WhiB7 via its binding with SigA, and interaction with DNA. SigA is the major vegetative sigma factor and an essential protein critical for driving the expression of housekeeping genes in *M. tuberculosis* and *M. smegmatis* (Gomez et al., 1998; Manganelli et al., 2004). However, it was discovered that the expression of SigA is increased in human macrophages and or the lungs of mice after *M*. *tuberculosis* infection. This modulates the expression of essential genes as well as genes responsible for virulence, leading to enhance intercellular growth (Volpe et al., 2006; Wu et al., 2004). Hence, it is unlikely that *M. tuberculosis* Wbl proteins act as anti-SigA factors to completely abolish SigA-RNAP interaction due to the essentiality of SigA; instead, this type of binding might be essential for regulation of specific genes by stabilizing SigA-DNA interaction by providing additional promoter selectivity, especially since WhiB1 has a *C*-terminal domain rich in positive amino acids that are crucial for DNA-binding (Smith et al., 2012). The upregulation of SigA during infection of macrophages may also suggest that some SigA could interact with WhiB1 for inducing the expression of other alternative sigma factors which might be necessary for virulence or adaptation, while a basal level of free SigA is still available for regulating housekeeping genes. Indeed, *whiB1* was reported to be possibly under the direct control of SigF in *M. smegmatis* (Hümpel et al., 2010). Moreover, ChIP-Seq analysis of

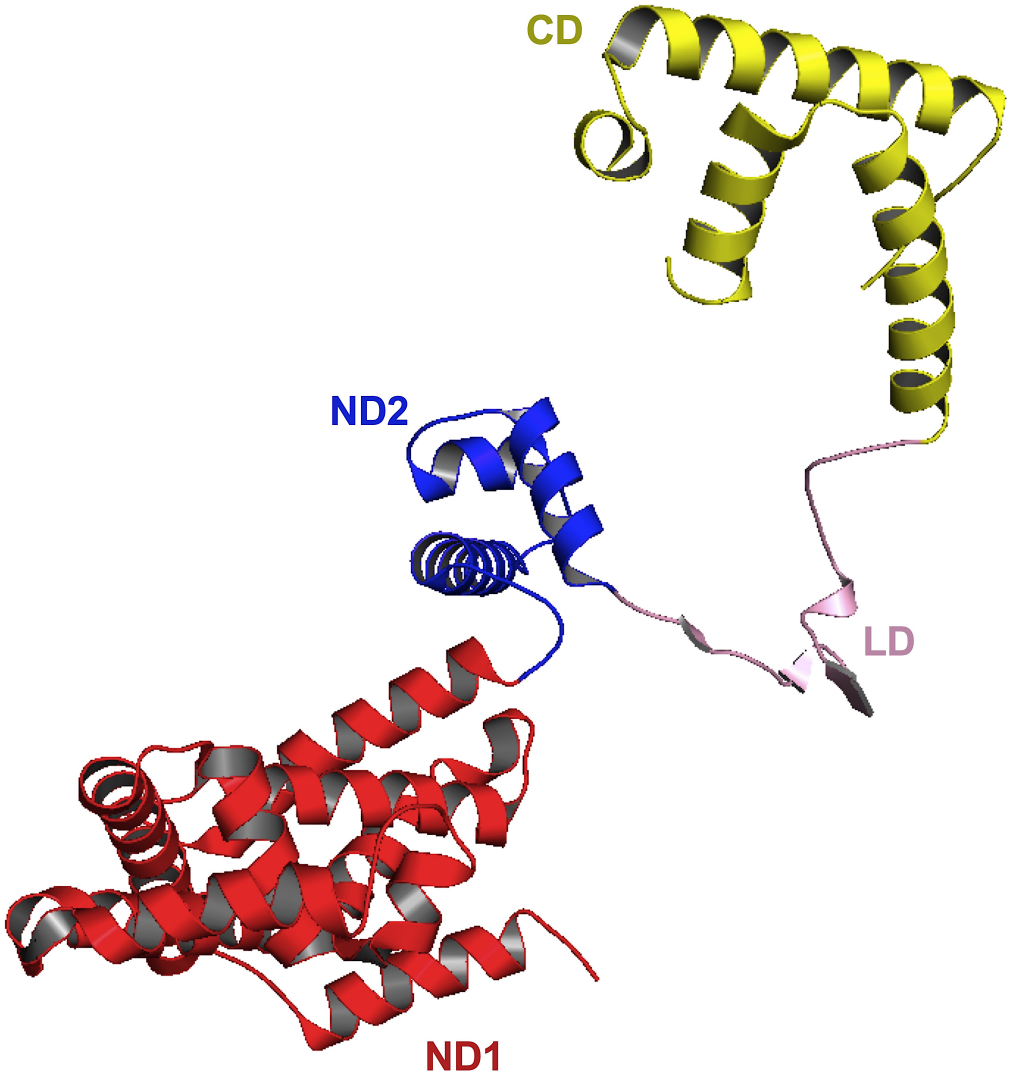


Figure 4.31: Ribbon diagram of *T. thermophilus* SigA structure. Three-dimensional structure of SigA (Vassylyev et al., 2002). The *N*-terminal domain (ND1), the *N*-terminal domain (ND2), the linker domain (LD), and *C*-terminal domain are indicated. The structure was obtained from protein data bank website (http://www.rcsb.org/pdb/explore.do?structureId=1iw7) using the PDB number 1IW7. The figure was generated here using Pymol software (DeLano, 2002).

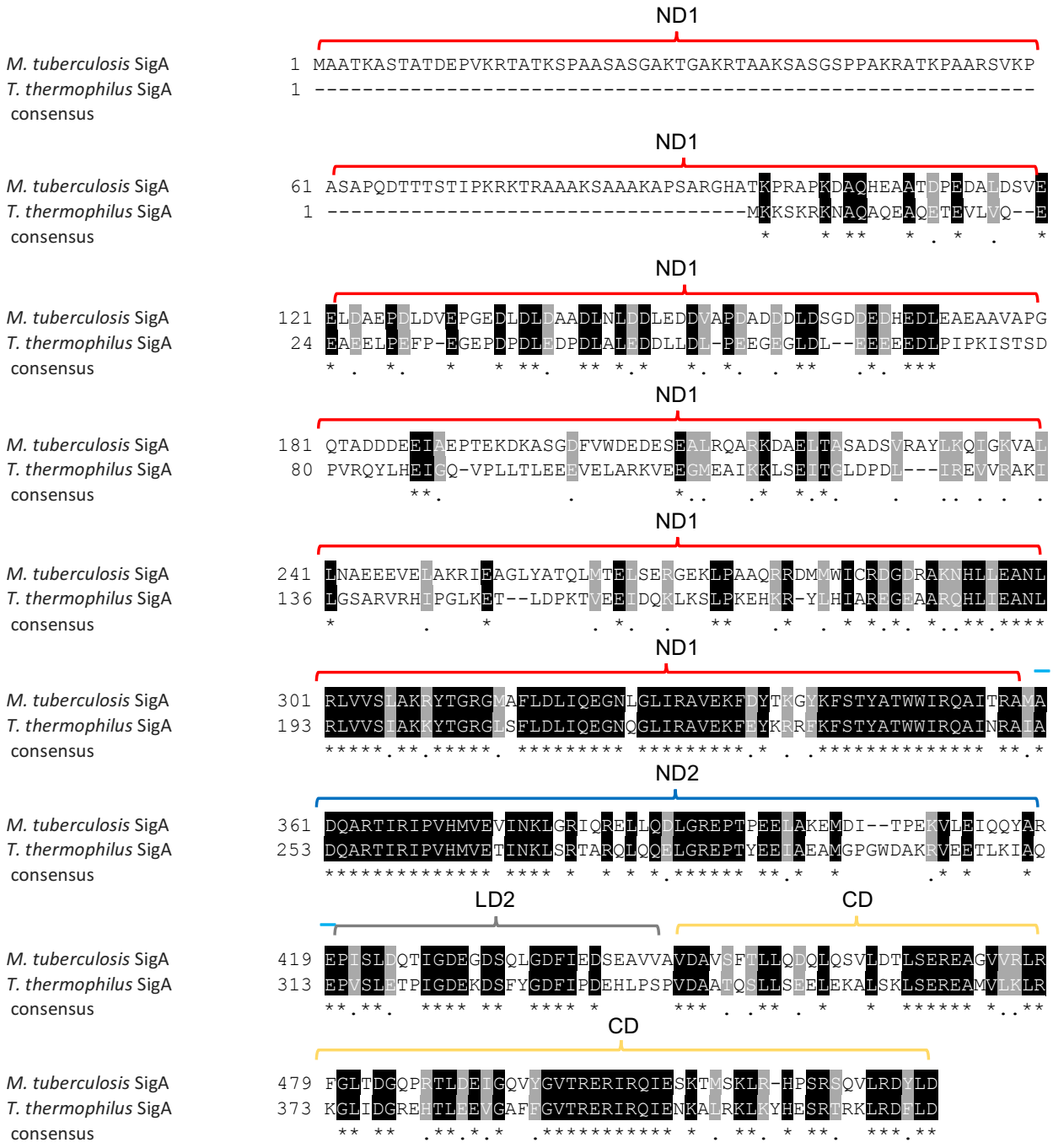


Figure 4.32: Sequence alignment of *M. tuberculosis* SigA and *T. thermophiles* SigA proteins. The sequences of *M. tuberculosis* SigA was obtained from NCBI website; https://www.ncbi.nlm.nih.gov/protein/ALB19907.1, and sequence of *T. thermophiles* SigA from https://www.ncbi.nlm.nih.gov/protein/81699349?report=genpept, and aligned by multalin tool (Corpet, 1988). Residues that are identical in both proteins (*), and residues with similar properties (.) are indicated.

TB data base (http://www.tbdb.org), showed the presence of a WhiB1 binding site within the non-coding region of *sigF*, suggesting the possible mutually dependent regulation role of gene expression of WhiB1 and SigF (Figure 4.33). Loss of *sigF* decreases virulence and sickness-associated tissue damage of *M. tuberculosis* in mice and guinea pigs, in addition to altering cell wall composition and effecting cell wall-associated proteins (Chen et al., 2000; Geiman et al., 2004; Karls et al., 2006; Williams et al., 2007). Loss *sigF* in *M. smegmatis* increases sensitivity to heat shock, acidic pH, oxidative stress, disables of carotenoid synthesis, and alters the cell wall permeability (Gebhard et al., 2008; Provvedi et al., 2008). These phenotypes indicates the important role of SigF in virulence, and a possible regulatory link between WhiB1 or WhiB1:SigA complex.

Nitric oxide production in activated macrophages is the first antimicrobial weapon against *M. tuberculosis*, which needs to adapt via switching on/off several gene transcription pathways (Yang et al., 2009). The *M. tuberculosis* WhiB1 [4Fe-4S] cluster was found to act as a specific NO sensor converting the WhiB1 from non-active DNA-binding form to an active form (Smith et al., 2010). Here it was found that the [4Fe-4S] cluster of WhiB1 is essential for WhiB1:SigA interaction, and NO can trigger the separation of the complex. This suggests that the cluster has a key role in regulating WhiB1:SigA formation/disassembly in a response to NO. Release of nitrosylated-WhiB1, the DAN-binding form, and the presence of NO-WhiB1:SigA complex adds another layer of complexity in the control of *M. tuberculosis* gene expression during the infection process. Further DNA-protein binding assays with the WhiB1:SigA complex in the presence and absence of NO should highlight the contributions of these processes. Moreover, the ability of apo-WhiB1 to bind DNA in addition to its ability to interact with GlgB, which is responsible for glycogen synthesis from glucose-1-phosphate, in addition to its possible function in the metabolism of trehalose (Chandra et al., 2011; Garg et al., 2009, 2007b), indicates the diverse roles of WhiB1 forms (Figure 4.33).

To conclude, *M. tuberculosis* WhiB1 interacts directly with the major sigma factor (SigA), region 4 of the *C*-terminal domain, and this binding is governed by the presence of WhiB1 iron-sulfur cluster; which is insensitive to O_2 but very

sensitive to NO. Cluster nitrosylation results in complex separation. The sensing of NO which causes complex separation, the DNA-binding ability of nitrosylated WhiB1, and the essentiality of WhiB1, all indicate the essential role of WhiB1 in controlling gene expression of *M. tuberculosis* during infection.

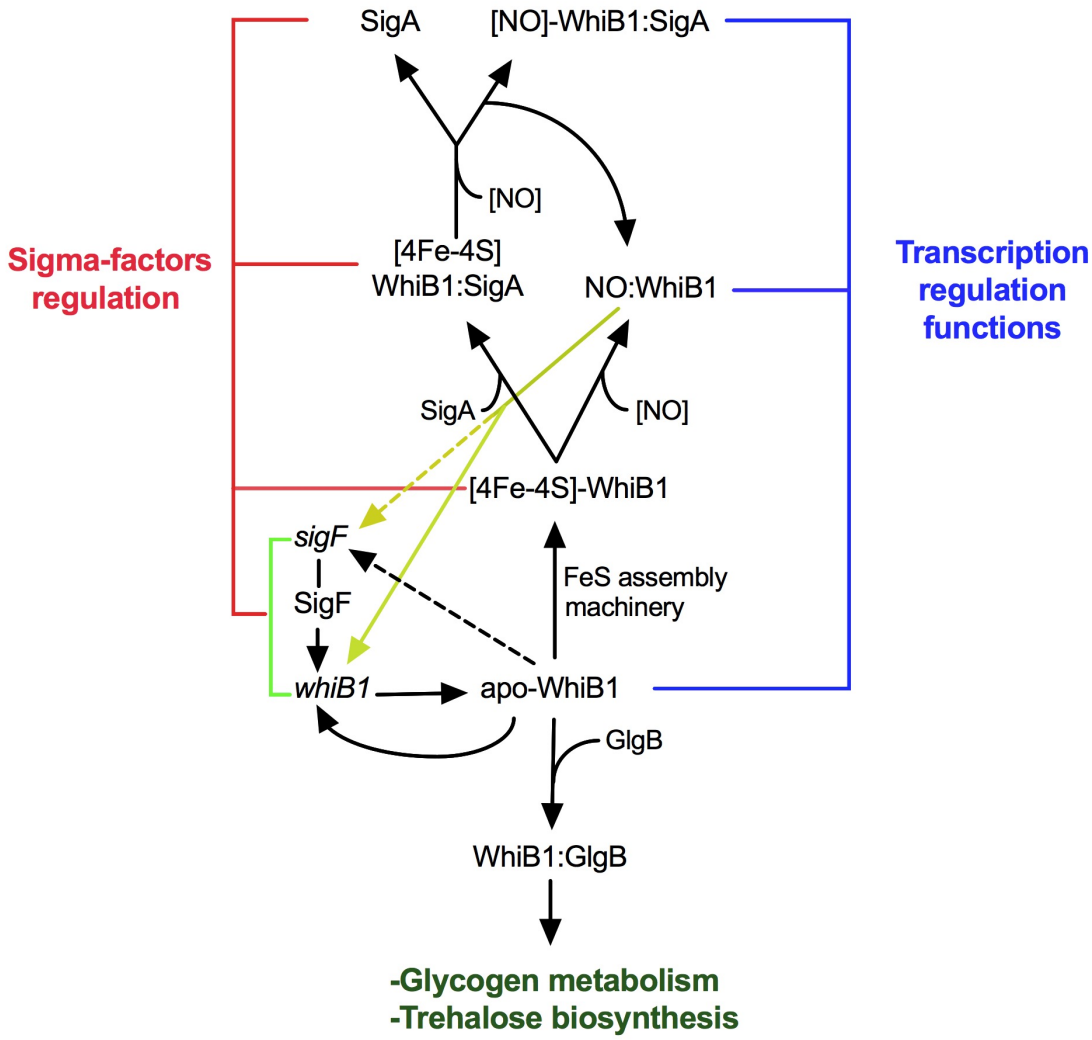


Figure 4.33: Formation of different WhiB1 forms and their possible regulatory roles. Bold arrows indicate confirmed pathways, while dashed arrows refer to possibility of regulation.

Chapter five

Structural characterization of *M. tuberculosis* WhiB1

5.0 Structural characterization of *M. tuberculosis* WhiB1

5.1 Introduction

Structural information for Wbl proteins was lacking at the start of this research and one of the main aims was to fill this gap using X-ray crystallography or NMR spectroscopy. One likely reason for the lack of structural information for proteins of this family is the fragility of iron-sulfur clusters under aerobic conditions, especially since the cluster is essential for WhiB1 folding (Section 3.4.5) (Alam et al., 2007, 2009; Crack et al., 2009; Garg et al., 2007a). However, the iron-sulfur cluster is not the only reason for the absence of a crystal structure. Several trials, carried out here, for crystalizing WhiB1 were not successful, even when the WhiB1 iron-sulfur cluster was insensitive to O_2; e.g. the His-TEV tag WhiB1 or when the WhiB1 was in a complex with SigA or SigA domains. This may suggest the presence of a flexible region of the protein prevents crystallization. In addition, application of NMR techniques to iron-sulfur containing proteins is problematic because Iron-sulfur cluster paramagnetism can confound the assignment of specific residues, especially those that anchor the cluster (Cheng and Markley, 1995; Piccioli and Turano, 2015). The cross-relaxation determination of proton-proton interaction (the Nuclear Overhauser Effect, or NOE) for residues close to the cluster is challenging due to the rapid relaxation of protons caused by the paramagnetic centre. This can seriously impair structural resolution near the cluster due to the importance of NOE data in protein structure determination (Cheng and Markley, 1995).

The aerobic stability of the iron-sulfur cluster of the TEV-tag-WhiB1 construct allowed the application of NMR structural studies to be attempted (Figure 5.1)

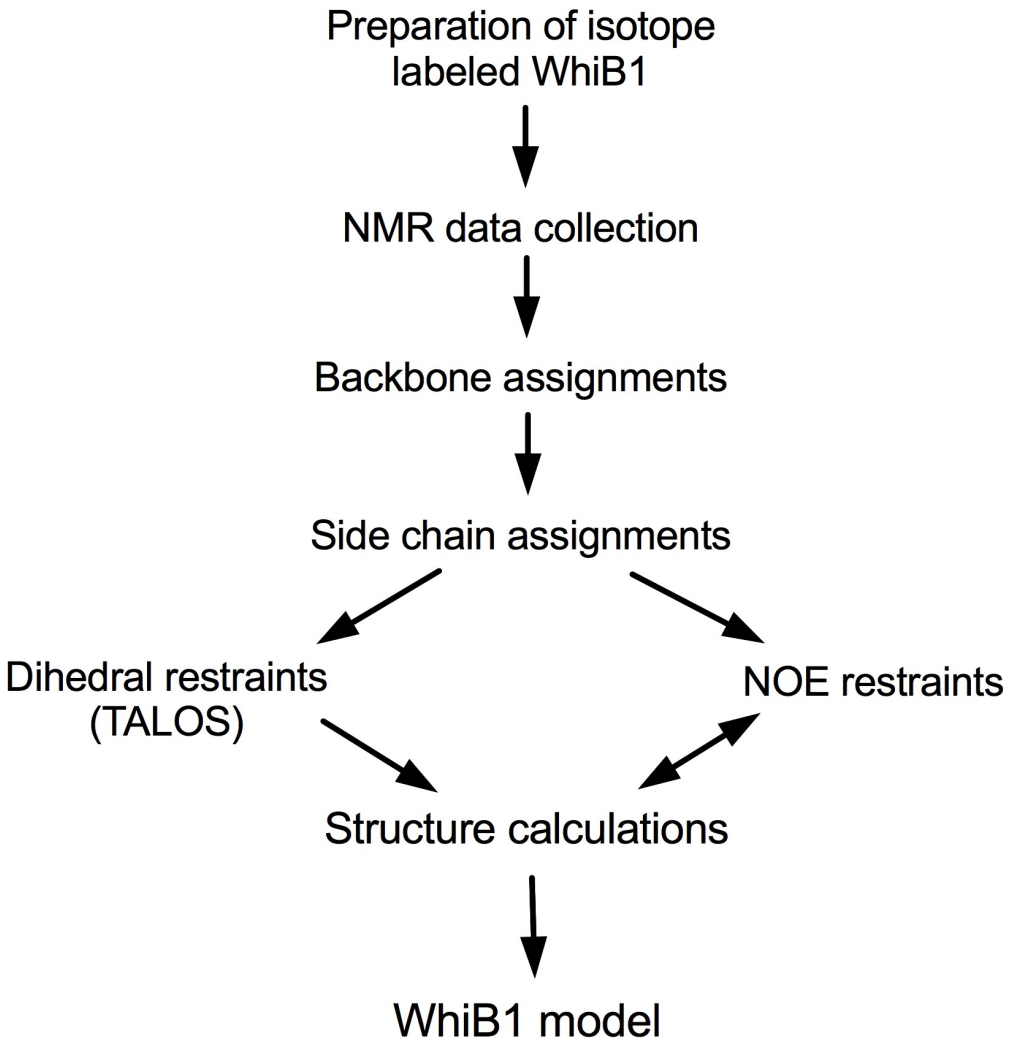


Figure 5.1: Flow chart showing the major steps taken to create a structural model of WhiB1.

5.2 General overview of applications of NMR to proteins

This section provides information to explain some basic concepts and facts about the NMR of proteins. NMR spectroscopy is a powerful method used to obtain information about the dynamics and structure of molecules at the level of atoms. In NMR, the nuclei of isotopes that have an odd number of protons and/or neutrons have a non-zero spin (Figure 5.2A). When an external magnetic field of the appropriate strength is applied, the magnetic moments of these nuclei will align parallel with the applied field (Figure 5.2B). However, when nuclei have an even number of protons and/or neutrons, they cannot be observed in NMR because they have zero-nuclear spin. Therefore, NMR experiments use active nuclei such as ^{13}C instead of ^{12}C to study proteins. When specific nuclei are irradiated with a short radio-frequency pulse, the magnetic field causes the nuclear spins to rotate away from the z axis towards the xy plane (Figure 5.2 B). The spins then precess around the z axis, generating a voltage in a coil of wire placed around the sample. This oscillating voltage is detected and can be translated into a spectrum which can provide information to identify the atom type, the chemical environment, etc. Samples under investigation are usually located in an NMR probe which is placed in the centre of the magnet. Irradiating the sample and receiving the emitted radio-frequency resonance from the sample are detected by this probe which contains a coil for this purpose (Rattle, 1995).

An NMR spectrometer is composed of three parts: a superconducting magnet, which provides the external magnetic field; an NMR console, which produces and regulates the radio-frequency pulses, is responsible for exciting the sample inside the probe and simultaneously for receiving and detecting the signals emitted from the sample; and finally, a computer terminal supported by various NMR software (Figure 5.2C). By using several NMR programs, protein structures can be studied based on structural restraints, such as the distance between proton nuclei (NOEs), chemical shifts, dihedral angles, etc., which are measured during different NMR experiments (Marion, 2013).

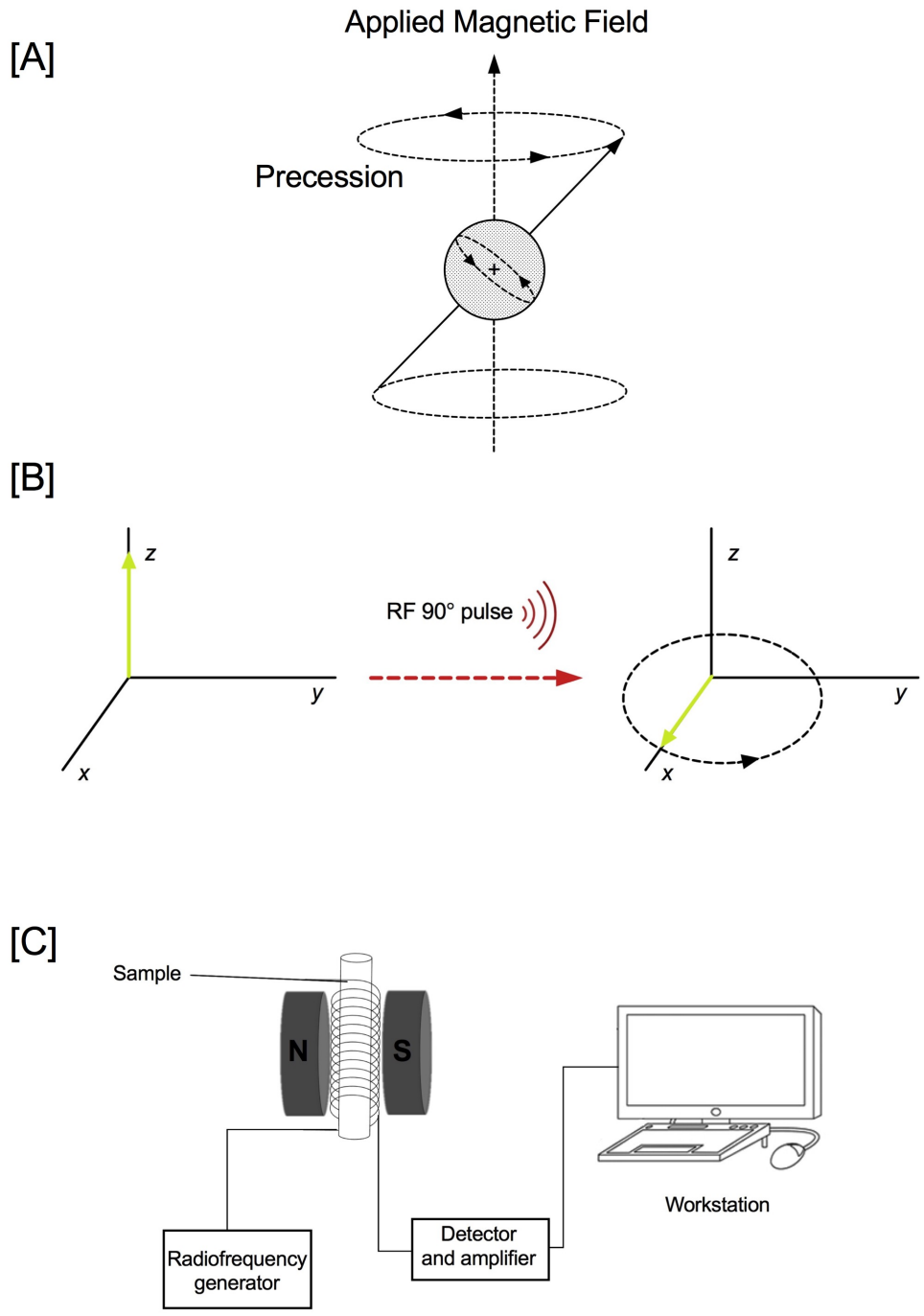


Figure 5.2: Basic concepts of active nuclei properties in an NMR spectrometer. **[A]** Nuclei, that have an odd number of proton and/or neutrons have a magnetic moment allowing for precessional motion under magnetic field. **[B]** Magnetization of nuclei can be rotated by a radio-frequency field (RF). Applied radiation by 90° will rotate nuclei magnetization along with the X-axis, once the RF has been stopped nuclei will precess in the plane of xy before returning to z-axis **[C]** NMR spectrometer main units.

The signal in the NMR spectrum is typically known as a resonance, the spectrum shows the intensity of signals against the radio frequency, where the relative frequency of a resonance to a standard compound that has frequency of 0 ppm is referred to as its chemical shift. This arises from the fact that each nucleus is partially shielded due to the electrons around it, from the external magnetic field; which means that each nucleus will experience a different magnetic field and thus resonate at different frequencies. The relative frequencies are extremely small, and thus chemical shifts are measured by part per million unit. Many studies in recent years have tried to use chemical shifts as restraint to calculate protein structures (Williamson, 2012).

The external magnetic field is not the only force that governs the chemical shift of nuclei, another magnetic field can be generated from neighboring nuclei because they have a magnetic moment. Although this field is weak in comparison to the external magnetic field, the distance is very close between nuclei and hence the effect can be sufficiently strong to be detected, especially in the high field. This affects the chemical shifts, and is known as dipolar coupling, which is the main source of nuclear relaxation. This type of dipole-dipole interaction through space also causes the Nuclear Overhauser Effect (NOE). The information obtained from residual dipolar coupling is valuable for protein structure calculations (Williamson, 2012).

Spins can also interact by *J*-coupling, also known as scalar coupling, where the effect occurs not through space but instead through bond interaction. The *J*-coupling effect is mutual due to transformation through bonding electrons, which can extend up to three bonds. This is used to assign spectra, and can be quite useful to provide information about protein backbone dihedral angles and side-chain direction (Williamson, 2012).

5.3 WhiB1 isotopic labelling

WhiB1 samples for NMR experiments need an expression medium enriched with ^{15}N or ^{15}N and ^{13}C isotopic sources. NMR experiments require stable, NMR active, nuclei with a non-null nuclear spin such as ^{1}H, ^{13}C, ^{15}N, ^{19}F… etc. Natural isotopes prominent in proteins (^{14}N and ^{12}C) are invisible in NMR, where ^{14}N has broad signals and ^{12}C has no spin; hence enrichment with NMR active isotopes is essential. Nuclei of ^{1}H, ^{13}C and^{15}N act as magnetic dipoles when an intense magnetic field is applied and the transition between the energy levels created by the applied magnetic field can be studied by applying pulses of electromagnetic radiation; and the properties of these nuclei and how their nuclear dipoles influence each other can be investigated. Information such as frequencies of nuclei (chemical shifts), which are dependent on their chemical environment, scalar coupling through bonds, and the dipole interactions (dipolar coupling) in space are important for studying protein characteristics by NMR. The nuclear resonance of ^{1}H, ^{13}C and ^{15}N isotopes are in different frequency ranges allowing researchers to design experiments in which specific active nuclei can be studied in an environment that has several NMR active nuclei (Mossakowska and Smith, 1997; Rattle, 1995; Reid et al., 1997). Thus, the first step was to look for an appropriate minimal medium suitable for *M. smegmatis* growth and overproduction of labelled WhiB1. No suitable reference was found for this purpose; however, after several optimizations of mycobacteria 7H9 medium a recipe for medium preparation enriched with ^{13}C glycerol, ^{13}C glucose and $(^{15}NH_4)_2SO_4$ or $^{15}NH_4Cl$, was achieved (Section 2.3.5). Subsequent protein overproduction and purification was as described in Section 3.4.1, with the exception of increasing the incubation time of cultures to 40 h (instead of 24 h) before induction with 0.2% acetamide and a further 24 h of incubation (instead of 12 h) at 37ºC with shaking at 250 rpm. Removing imidazole for protein preparation was done using 0.5 ml Zeba™ spin desalting columns (Thermo Fisher) pre-equilibrated with 25 mM NaH_2PO_4, pH 7.0 containing 0.25 M NaCl. Where necessary replacing H_2O with D_2O was done by protein freeze-drying method.

To investigate the efficiency of isotopic labelling and suitability of the protein for NMR, 1D ^{1}H-NMR and 2D ^{15}N-HSQC spectra were obtained and studied. The

1D ^{1}H-NMR (Section 3.4.5) confirmed that holo-WhiB1^{Ms} was well-folded. Moreover, 1D ^{1}H-NMR also showed in the downfield region six well resolved peaks at 16.6, 15.0, 13.0, 12.4, 12.0 and 10.0 ppm likely to be cysteine (Hβ) iron-sulfur cluster ligand resonances as observed for *Desulfovibrio africanus* ferredoxin I and *Clostridium acidic uric* ferredoxin (Bertini et al., 1994; Davy et al., 1995) (Figure 5.3A).

In the 2D ^{15}N-HSQC spectrum, each ^{1}H-^{15}N group, which is covalently bonded, shows a single peak. Each peak is composed of a chemical shift for the ^{1}H and another one for the ^{15}N nucleus. A well-structured protein shows one peak for each amide proton in the protein backbone (except for proline which lacks the amide proton when it is incorporated into a protein). Thus, the ^{15}N-HSQC spectrum provides a fingerprint of the protein, where the spectrum should include one signal for each residue (side-chains of tryptophan residues show a peak while pairs of peaks for asparagine and glutamine side-chains can be seen as well). Peaks in ^{15}N-HSQC spectrum for a folded protein are typically well-dispersed and can be easily distinguished; unlike an unfolded protein spectrum, where peaks overlap in the middle area of the spectrum (Bodenhausen and Ruben, 1980; Kwan et al., 2011; Pascal, 2008).

The ^{15}N-HSQC spectrum of ^{15}N-labelled WhiB1^{Ms} (200 µM) exhibited well-dispersed peaks, with the number of amide signals close to the number of WhiB1 residues (90 peaks for 99 residues not counting Proline) (Figure 5.3Ba). This showed that WhiB1^{Ms} was well-folded and appropriate for NMR structure determination. Leaving the protein sample in the NMR machine for four days at room temperature showed no change in the 2D spectrum (Figure 5.3Bb), suggesting the protein was stable.

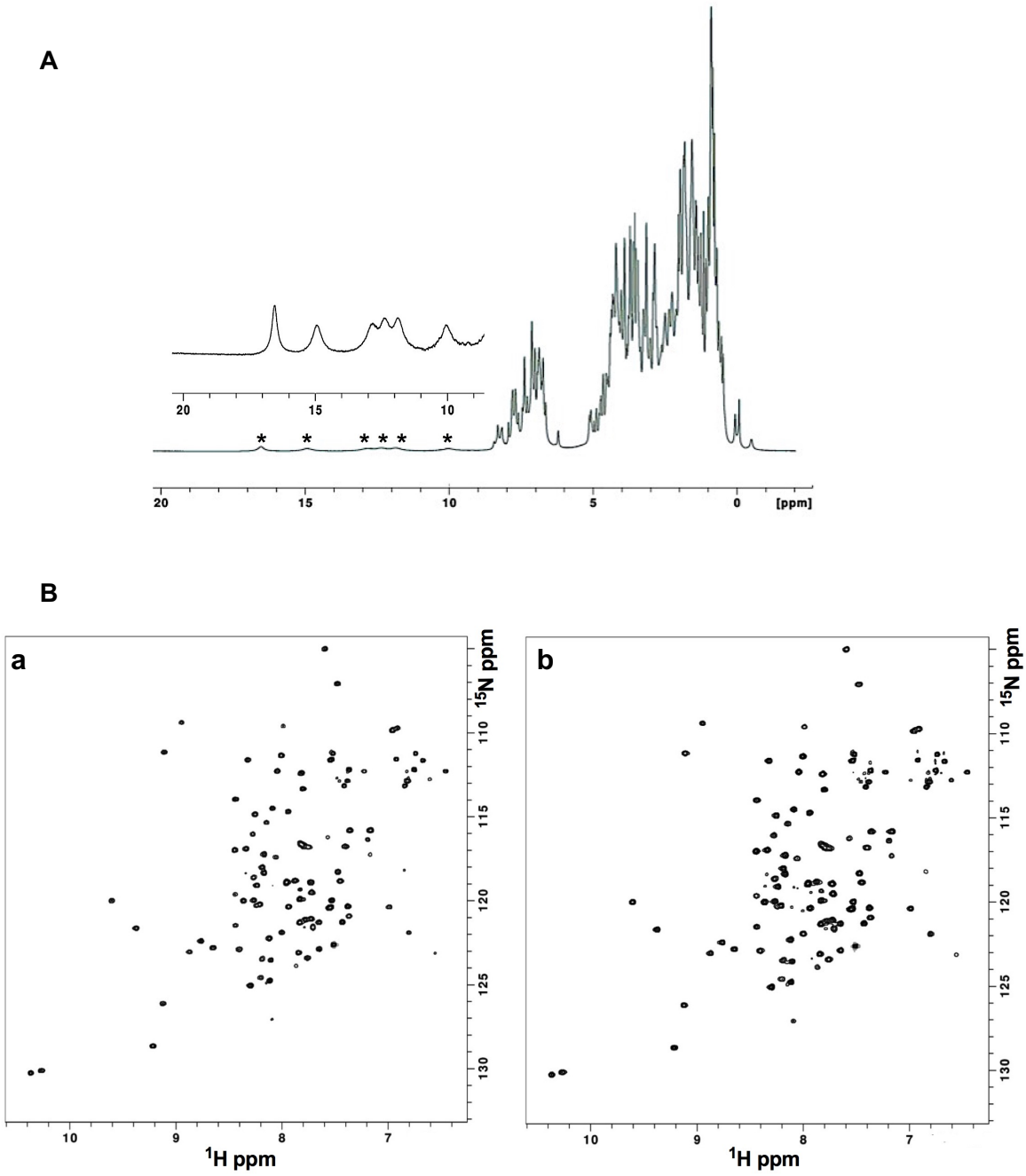


Figure 5.3: [A] 1D ^{1}H-NMR spectrum of unlabelled WhiB1^{Ms}. [B] The ^{15}N-HSQC spectrum of ^{15}N labelled WhiB1^{Ms}. WhiB1^{Ms} (0.2 mM) was isolated in 25 mM NaH_2PO_4 (pH 7.0) containing 0.25 M NaCl. The 1D spectrum was recorded in the presence of 100% D_2O at 800 MHz, while the 2D spectrum was in 10% D_2O. Peaks likely to represent cysteine iron-sulfur cluster ligand resonances are indicated (*) in **[A]**; the downfield region of the spectrum is shown in the inset. The residue NH peaks in the 2D spectra are shown at the beginning of spectrum recording **(a)**, and after four days in panel **(b)**.

5.4 WhiB1 backbone assignment

The assignment of backbone resonance to determine the frequencies of H_N, N_H, CO, Cα, Cβ of WhiB1^{Ms} amino acid residues was carried out in collaboration with Mrs. Andrea Hounslow (University of Sheffield, department of Molecular Biology and Biotechnology, NMR facility). The ^{15}N-HSQC spectrum is the initial step in backbone assignment, where in six triple resonance (^{13}C, ^{15}N, ^{1}H) 3D experiments, every NH amide peak of every residue was correlated to its CO, Cα, Cβ of its own and the preceding residues. This was done through using variable sets of different pulse sequences. Exciting the ^{1}H nuclei is the first step of these experiments because it is highly sensitive compared to ^{15}N or ^{13}C in NMR. The magnetization is transfered from ^{1}H to ^{15}N and ^{13}C and then returns to ^{1}H for acquisition of the signal; where, every transfer allows the magnetization labelling of each type of nuclei frequencies. By using different *J*-couplings between ^{15}N-^{13}Cα, ^{13}Cα-^{13}CO and ^{15}N-^{13}CO different experiments can be designed to measure the nuclear frequencies that are essential for protein structure determination. The experiments used for WhiB1 back bone assignment are as follows. The name of each experiment depends on magnetization transfer direction (the frequency information for the atoms in brackets is not obtained) (Figure 5.4) (Clubb et al., 1992; Grzesiek and Bax, 1992).

1) HNCO correlates NH_i to CO_{i-1}
2) HN(CA)CO correlates NH_i to CO_i and CO_{i-1}
3) HNCA correlates NH_i to $C\alpha_{i-1}$ and $C\alpha_i$
4) HN(CO)CA correlates NH_i to $C\alpha_{i-1}$
5) HNCACB correlates NH_i to $C\alpha_i$, $C\alpha_{i-1}$, $C\beta_i$ and $C\beta_{i-1}$
6) CBCA(CO)NH correlates $C\beta_i$ to $C\alpha_i$ and $C\alpha_i$ to NH_{i+1}

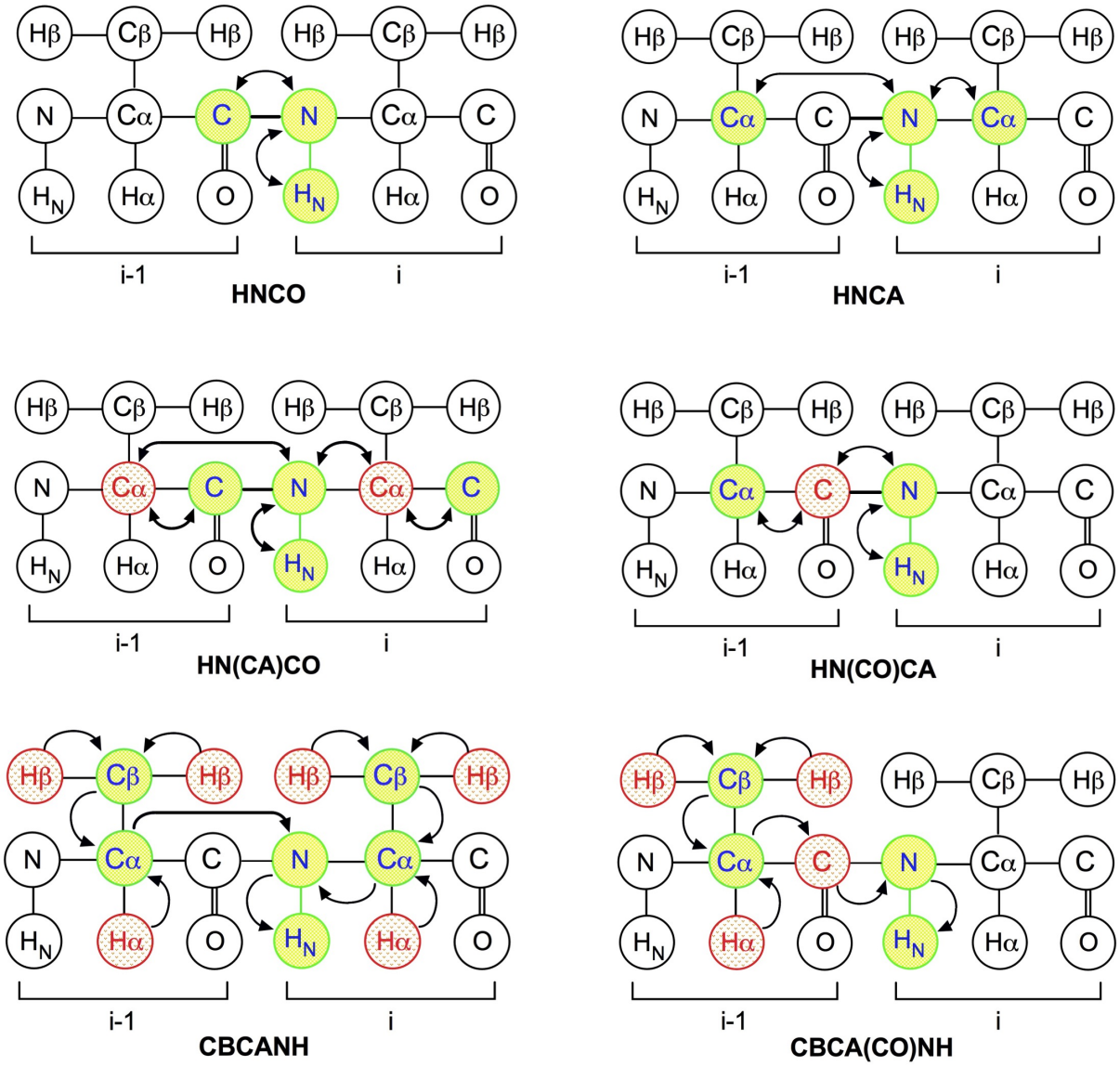


Figure 5.4: NMR experiments used for WhiB1 backbone assignment. Each experiment name is indicated beneath each scheme. While, each scheme shows the direction pathway of magnetization, the obtained frequencies of nuclei (light green), and the nuclei through which the magnetization flows without obtaining frequencies information, are indicated in red and in between brackets. The arrows refer to the direction of magnetization transfer.

A set of amide pair resonances from the ^{15}N-HSQC with their correlated resonance in the backbone experiment are called a spin system. Spin systems are numbered at the beginning of the experiments in groups depending on available data until finishing the assignment where each residue and the preceding residue in the WhiB1 sequence are linked to its identified frequency data. This was achieved by matching the chemical shifts of CO, Cα, Cβ using the *astool* package. This program uses two types of information for matching purposes, the amino acid sequence of the protein and the chemical shifts of every spin system (Figure 5.5). The residue types are usually identified using a table of assigned amino acid chemical shifts in a random coil (Wishart et al., 1995). The frequencies of WhiB1 ^{13}C for pCα, Cα, pCβ, Cβ, pCO, and CO were obtained from 83 identified spin systems, and *astool* placed these spin systems after several iterations in order of the amino acid sequence of WhiB1 depending on the carbon chemical shifts of paired (own and preceding) spin systems and by comparing the frequencies of amino acids in the WhiB1 sequence to the values of the chemical shifts of amino acids in a random coil. The results of these experiments produced a table containing the chemical shift values of H_N, N, Cα, Cβ, and CO (Appendix 6). By using these chemical shifts in the TALOS-N program, predictions for the phi, psi and chi dihedral angles were obtained (Section 5.7). These are essential for NMR structural calculations (Section 5.9). The His-tag region was not assigned and V8, V42, and T43 were missed because they are close to the cysteine residues and are thus affected by the iron-sulfur cluster paramagnetic properties. In addition, G61 and G62 could not be assigned suggesting that these are also close to the cluster. Cysteine residues that participate in anchoring the cluster were also not assigned here, while the remaining amino acids were unambiguously assigned (Figure 5.6).

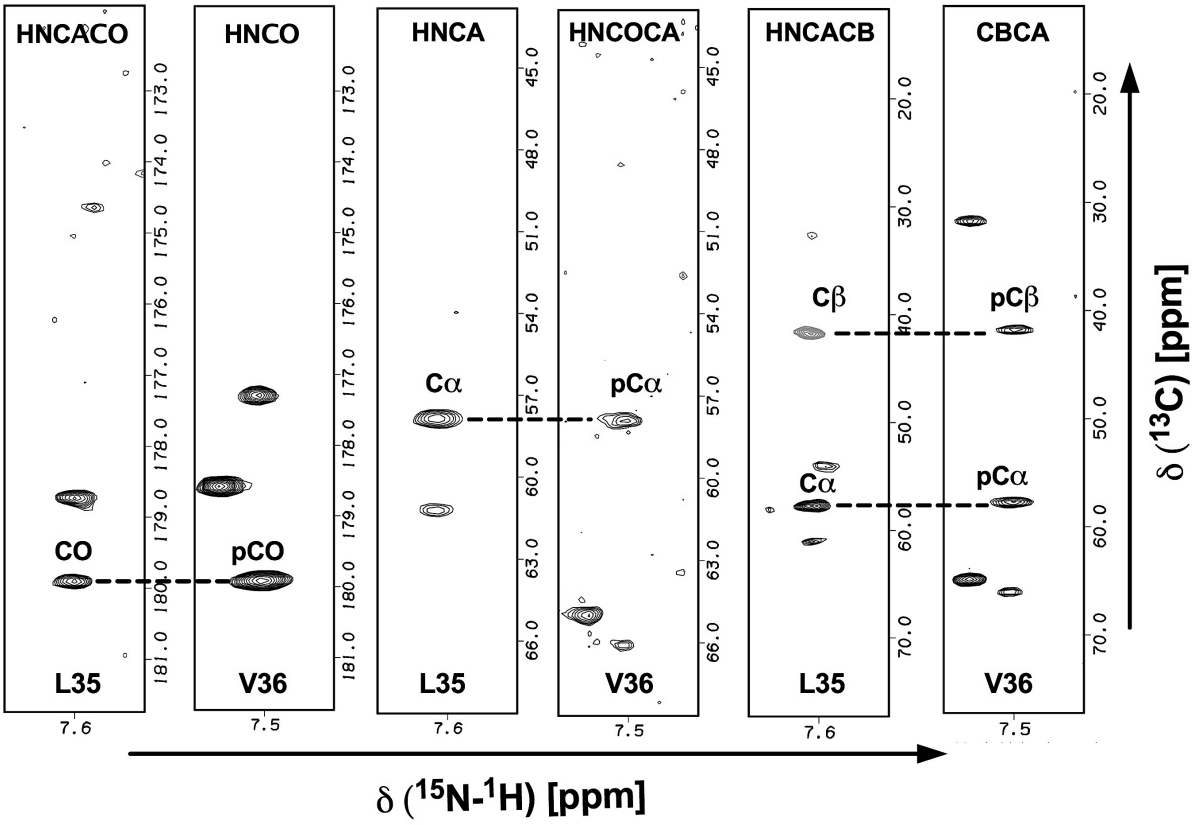


Figure 5.5: An example of backbone assignment achieved by astool package. The WhiB1 backbone assignment was done using data from six different triple resonance experiments; where, ^{15}N-^{1}H pairs from ^{15}N-HSQC were correlated dependent on the type of experiment with Cα, Cβ, CO and preceding nuclei. In this example, the sequential assignment of L35-V36 is demonstrated, where the preceding frequencies data of V56 match the intraresidue frequencies of L35.

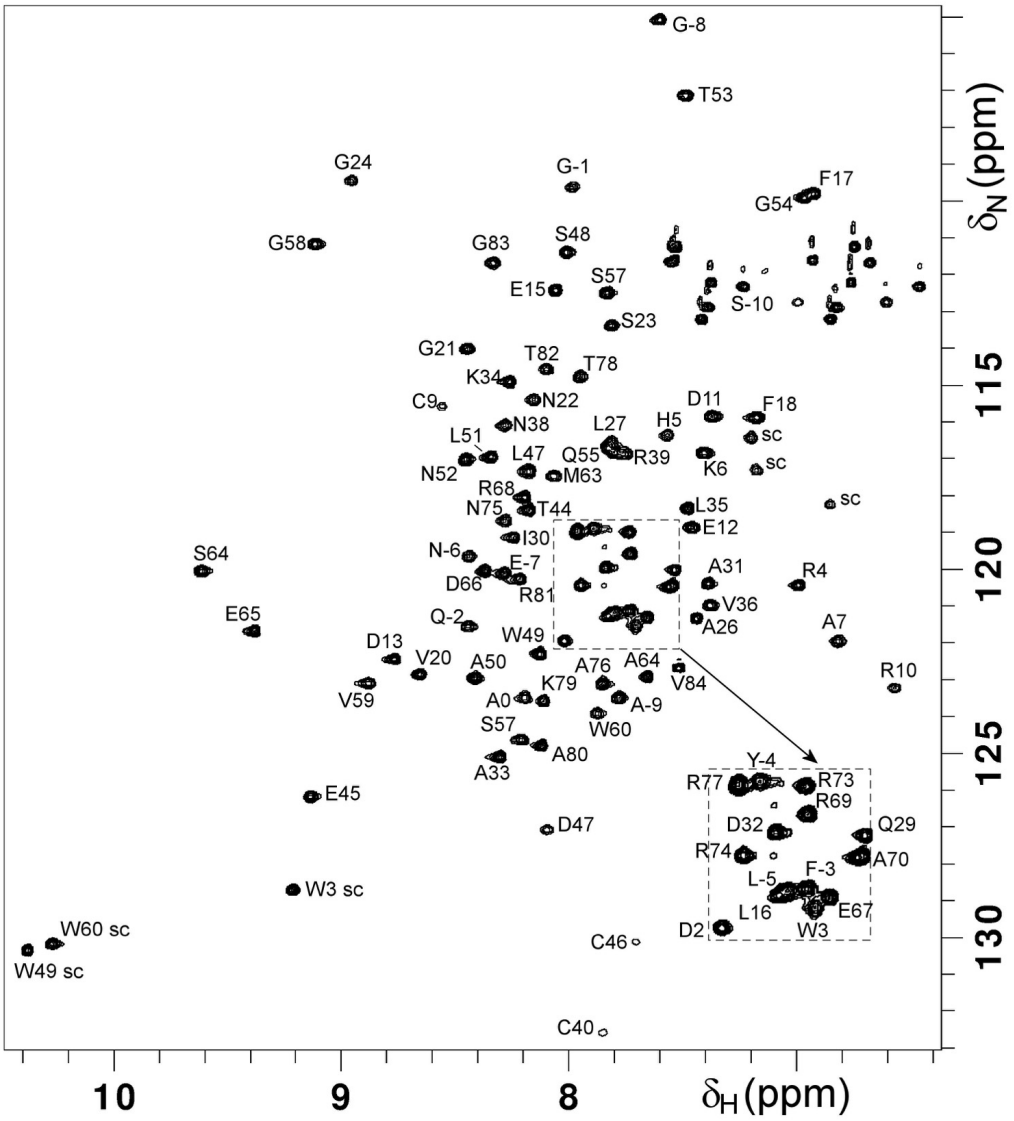


Figure 5.6: ^{1}H-^{15}N HSQC 2D spectrum of WhiB1. The spectrum shows amide resonances of assigned backbone residues labelled by single letter code and position in the WhiB1 sequence; sc represents the resonance of side-chain amide. Residues labelled from -1 to -20 represent the His_6-TEV-tag.

5.5 Side-chain assignment of WhiB1 residues

The 3D CCH- and HCCH-TOCSY experiments were carried out in collaboration with Mrs. Andrea Hounslow to obtain the chemical shifts of WhiB1 side-chain nuclei of aliphatic residues such as Cγ, Cδ, Cε, … etc, and their attached protons. Chemical shifts obtained here and from the backbone assignment are essential to obtain the dihedral angles and for NMR calculations (Section 5.9). These TOCSY experiments are designed to correlate the side-chain aliphatic protons with the resonance of ^{13}C nuclei via *J* coupling. The magnetization pathway and the nuclei involved are shown in Figure 5.7A. The sequence of pulses are as described in Clore et al., (1990) and Fesik et al., (1990). Felix2007 software was used for analysis of the spectra. The analysis was carried out, where the available chemical shifts of Cα and Cβ from the backbone assignment, and Hα and Hβ from ^{15}N TOCSY-HSQC (magnetization is mixed isotropically in the ^{15}N TOCSY-HSQC between proton nuclei of each residue, and then is transferred to NH) (Figure 5.7B) (Marion et al., 1989) and HAHB(CO)NH (which correlates $H\beta_i$ to $H\alpha_i$ and $H\alpha_i$ to NH_{i+1}) (Figure 5.7C) (Grzesiek and Bax, 1993) for identification of the rest of the nuclei in the side-chain of each WhiB1 residue. The assignment of side-chains was achieved for about 50% of WhiB1 residues (Figure 5.8) (Appendix 6). Many residues had strongly overlapping or missing signals, limiting how many residues could be assigned.

Side-chain assignment of WhiB1 nuclei in aromatic residues was carried out here as well. The chemical shifts of nuclei in aromatic rings were obtained using the (Hβ)Cβ(CγCδ)Hδ and (Hβ)Cβ(CγCδCε)Hε experiments, which correlates Cβ to Hδ and Hε respectively (Figure 5.9A, B). The pulse sequences are as described in Yamazaki et al. (1993). Moreover, the aromatic $^{1}H/^{13}C$ HSQC experiment, which correlates protons with carbons in the aromatic rings (Figure 5.9C), and the aromatic ^{13}C-NOESY-HSQC experiment, which correlates Hβ with the aromatic protons that are close in space, in addition to the correlation between the $^{1}H/^{13}C$ of rings via *J* coupling (Figure 5.9D) were also carried out to get the frequencies of nuclei in aromatic residues. The assignment of side-chains was achieved for about 50% of WhiB1 aromatic

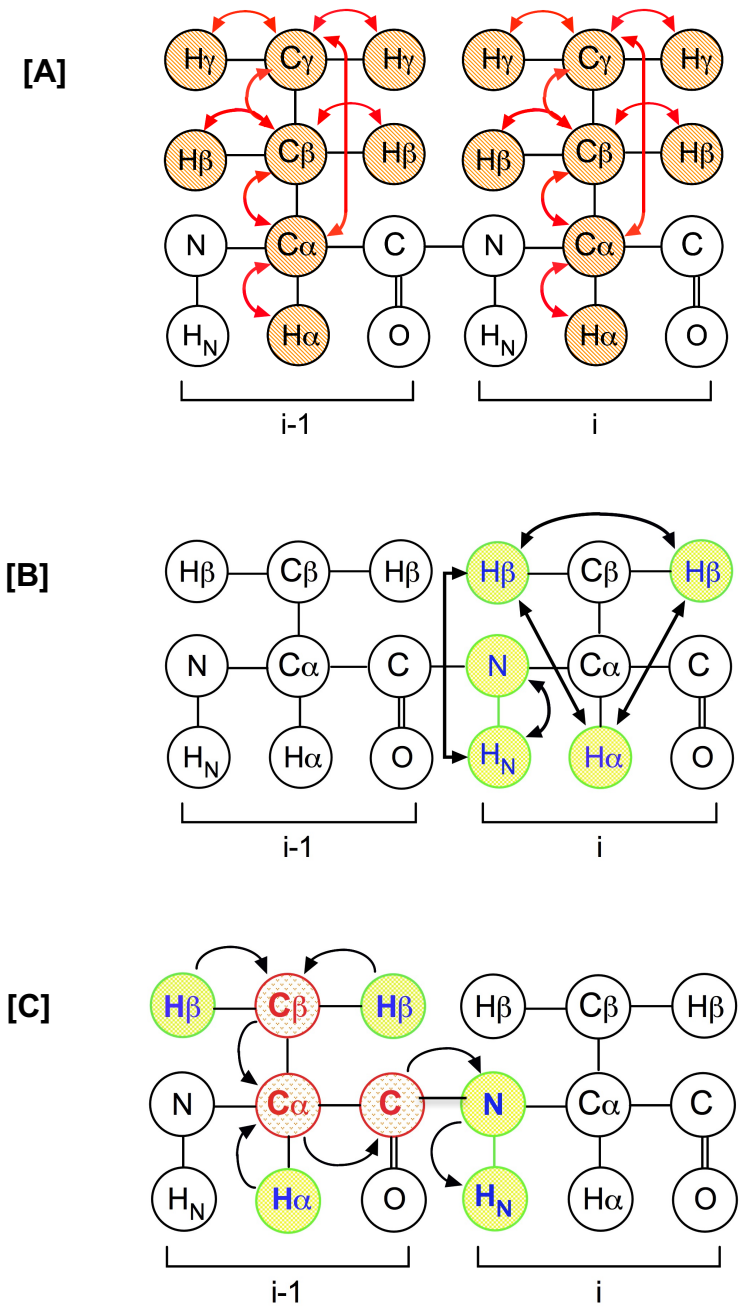


Figure 5.7: Schematic diagram of NMR experiments: [A] CCH-TOCSY and HCCH-TOCSY, [B] ^{15}N TOCSY-HSQC which were used for side-chain assignment of WhiB1 residues. [C] HAHB(CO)NH used to assign Hα and Hβ. The schemes show the direction pathway of magnetization and the obtained frequencies of nuclei, orange in **[A]**, and yellow in **[B]** and **[C]**. the nuclei through which the magnetization flows without obtaining frequencies information are indicated in red and in between brackets. The arrows refer to the direction of magnetization transfer.

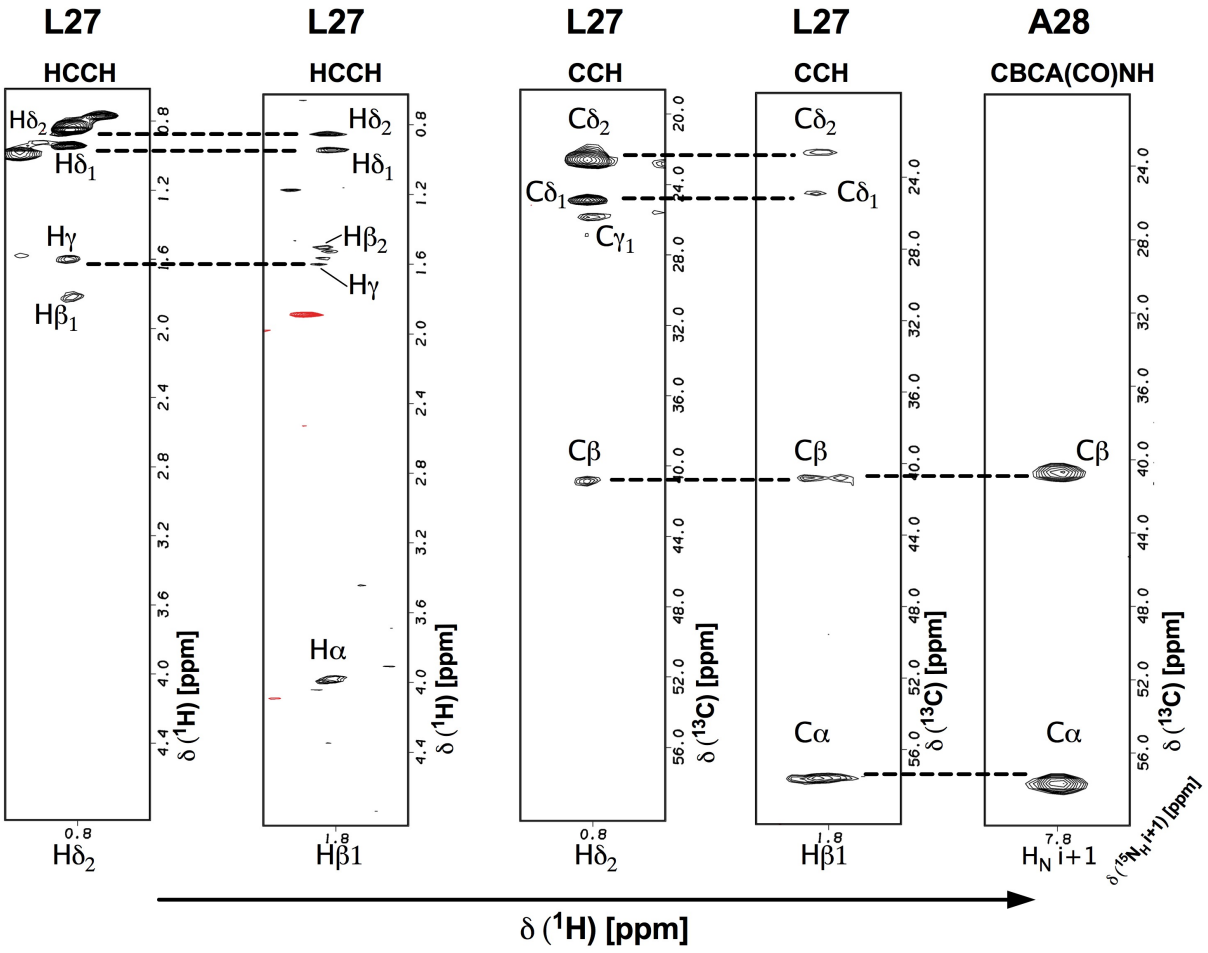


Figure 5.8: An example of the assignment of Leu27 side-chain using HCCH and CCH TOCSY experiments achieved by Felix2007 package. The available data of Cα, Cβ and Hα, Cβ from backbone assignment were used as the initial step of the side-chain assignment. The experiment name, the identified nuclei in this example and the detected matches are indicated in each slice.

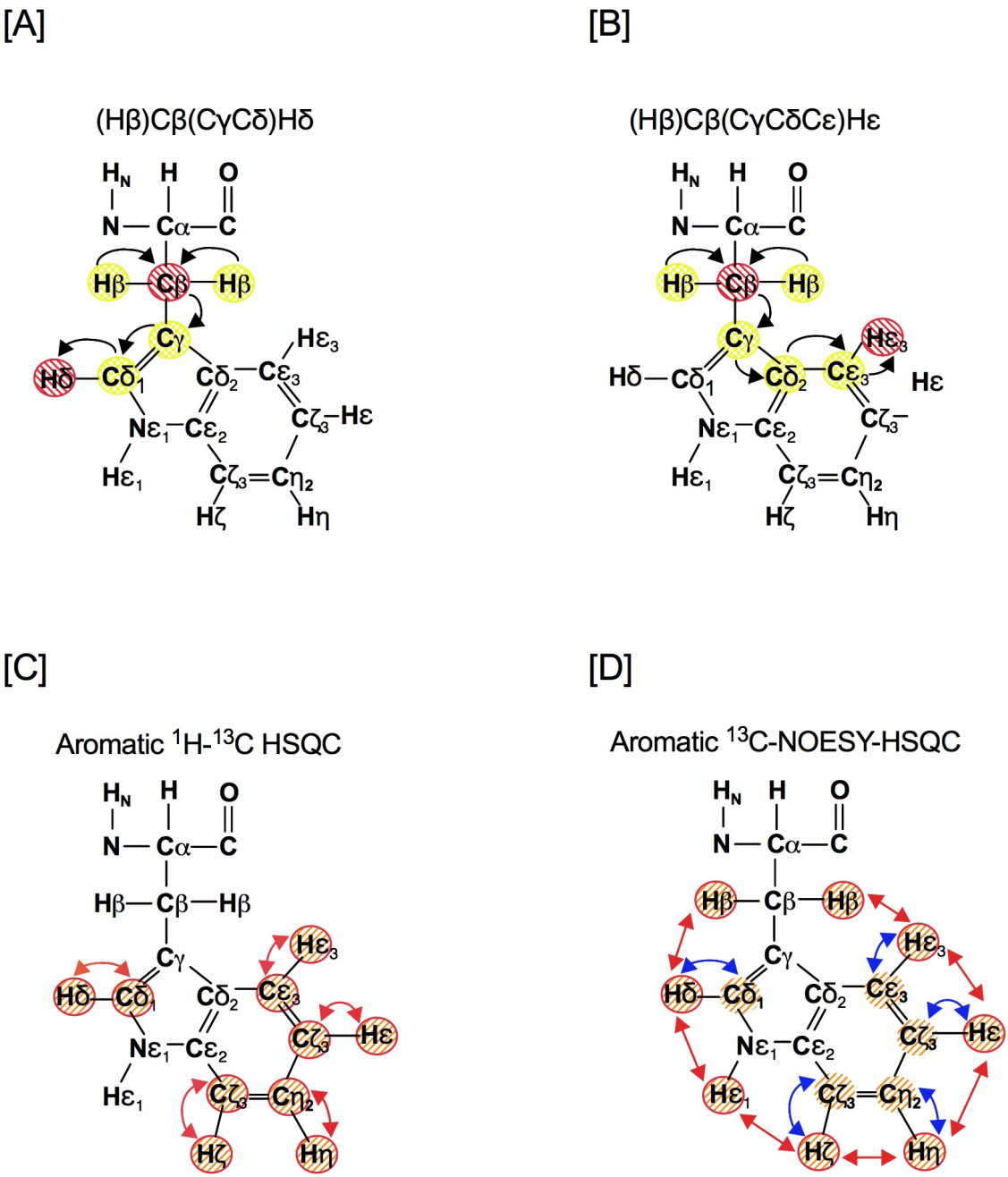


Figure 5.9: NMR experiments used for aromatic side-chain assignment of WhiB1 residues. The schemes show the pathway of magnetization, the obtained frequencies of nuclei are highlighted in red, nuclei through which the magnetization flows in panels **[A]** and **[B]** are labelled in yellow. Blue pathway in scheme **[D]** shows the second step of correlation between the $^{1}H/^{13}C$ of rings via *J* coupling.

residues (Appendix 6). Additional assignment using fast HSQC was made to see the very rapidly relaxing amide signals of residues near the iron-sulfur cluster. This was achieved with a recycle delay of 0.3 s, acquisition time of 17 ms, and 1/4J INEPT transfer delays of 833 μs. NOESY spectra used simultaneous acquisition of ^{13}C and ^{15}N with a mixing time of 100 ms. The *R1* relaxation rates of amide protons were measured using standard experiments, plus an experiment optimized for rapid relaxation (Bertini et al., 1996). This helped to see the two missing cysteine residues, as well as increasing the peak intensity of C40 and C46 (Figure 5.10). Measuring the *T1* relaxation times of amide protons for WhiB1^{Ms} residues showed efficient relaxation of residues close to the four cysteine residues (Figure 5.11). These observations are typical of ferredoxin, which in its $[4Fe\text{-}4S]^{2+}$ form exists in a low energy paramagnetic excited state, such that there is significant unpaired electron density on the iron atoms at room temperature, with consequent paramagnetic broadening of nearby nuclei (Davy et al., 1998).

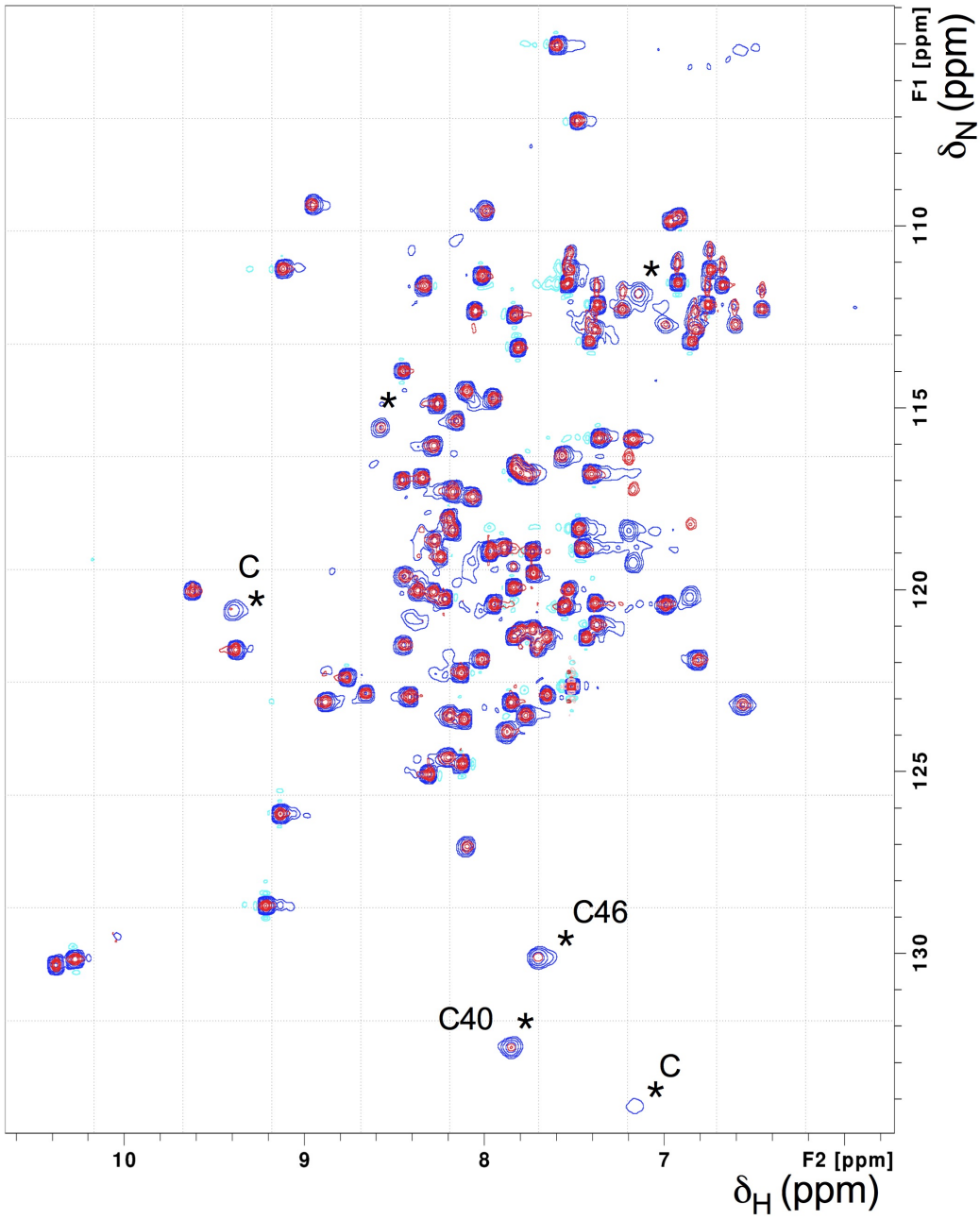


Figure 5.10: Fast ^{15}N-^{1}H HSQC 2D spectra of WhiB1. WhiB1^{Ms} (0.2 mM) was isolated in 25 mM NaH_2PO_4 (pH 7.0) containing 0.25 M NaCl and used. The spectrum shows the appearance of some missing signals, due to fast relaxation of nuclei near the cluster centre, after shortening the time of experiment with a recycle delay of 0.3 s, acquisition time of 17 ms, and INEPT transfer delays of 833 μs. The spectrum under standard parameters is shown in red. New peaks and those which increase in intensity are indicated (*). Peaks indicated with (*) and (C) are assumed to be the missing Cys HN of C9 and C37.

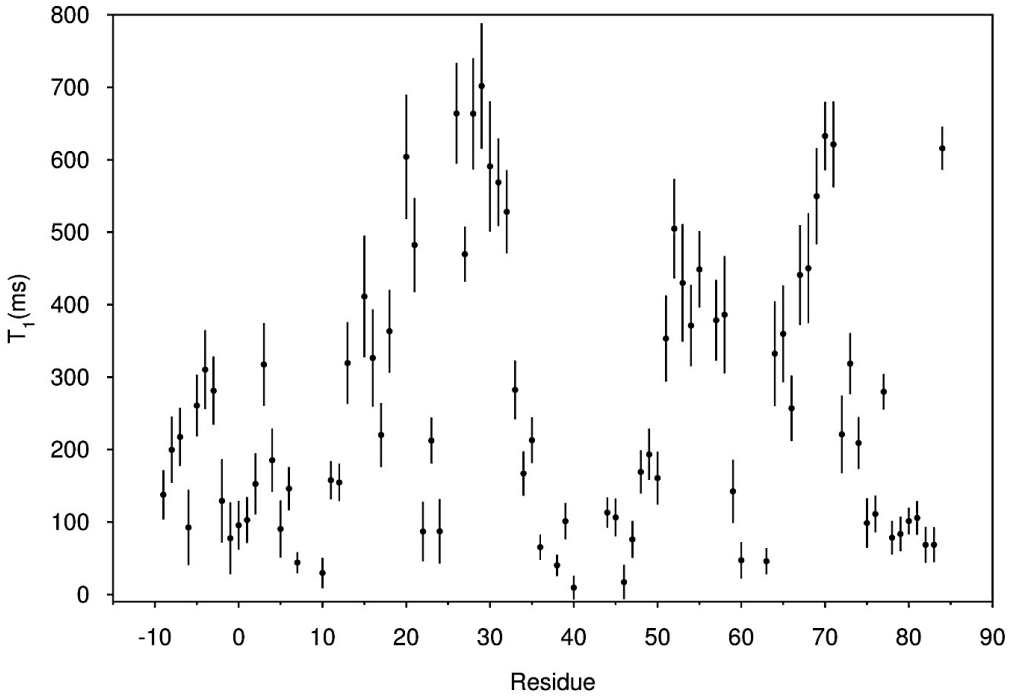


Figure 5.11: T_1 relaxation times of amide protons in WiB1^{Ms}. Nuclei with *T1* values of <50 ms are A7, R10, N38, C40, C46, W60, M63. The first 10 residues in this graph represent the TEV tag of WhiB1^{Ms}.

5.6 Hydrogen bond restraint

Hydrogen bond restraints have a significant impact on protein structure and are usually imposed in well-characterized regions of secondary structure. Hydrogen bonds in α-helices or 3_{10}-helices are located between backbone oxygen of the carbonyl group of the i residue and the amide proton of i+4 or i+3 residues, respectively (Hubbard and Haider, 2001). Amide protons involved in hydrogen bond formation can be experimentally distinguished by a H_2O/D_2O exchange; in the case of WhiB1, exchange was achieved by freeze-drying the protein and then dissolving in 100% D_2O. The presence of D_2O allows a constant exchange of amide protons (^1H) with the inactive NMR protons of the solvent (^2H). However, protons buried in the core of proteins, and protons involved in protein secondary structure, will exchange more slowly than amide protons located in loops or on the protein surface (Zhang et al., 1995). The ^1H-^{15}N correlation experiment (^{15}N-HSQC) was used to monitor amide proton exchange. After nearly one hour, 18 residues were resistant to exchange, and after 4 days only 4 residues resisted the exchange (Figure 5.12). The residues which showed slow H/D exchange indicated that they are involved in forming H-bonds and are needed for protein secondary structure. These residues were used as restrains during structure calculations (Section 5.9). Hydrogen bonds which were involved in the secondary structure of WhiB1 were mainly diagnosed by TALOS-N (Section 5.7), and the H/D exchange experiment confirmed some of the TALOS-N prediction data.

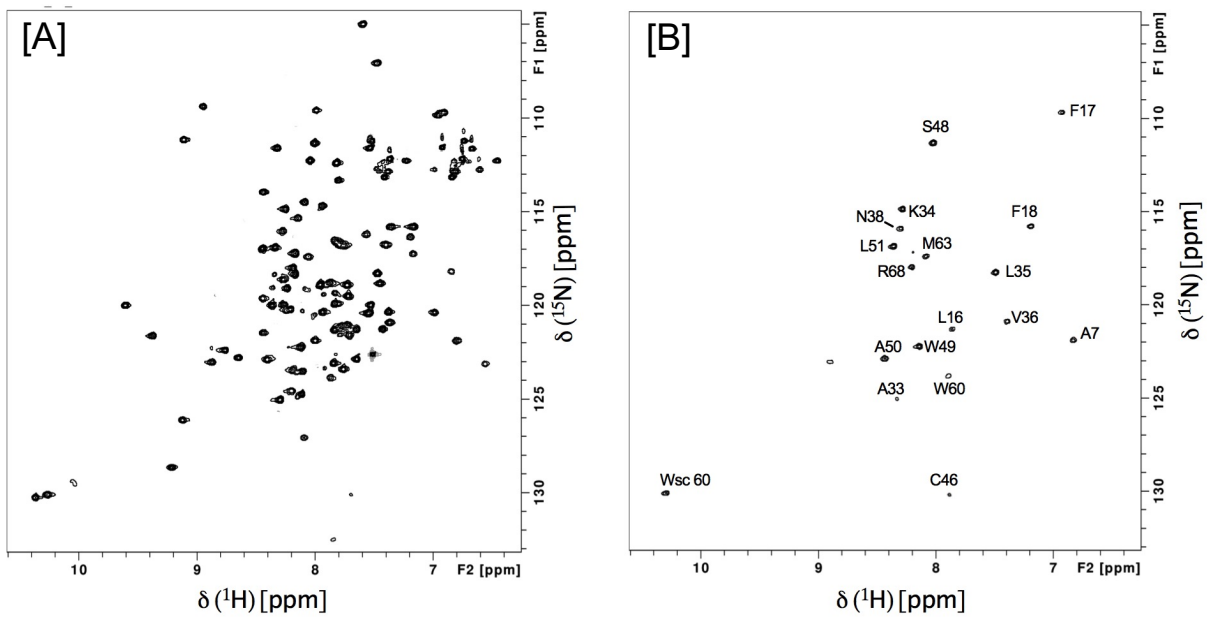


Figure 5.12: 2D ^{1}H-^{15}N HSQC spectra of ^{15}N-labelled WhiB1. WhiB1^{Ms} (0.25 mM) was isolated in 25 mM NaH_2PO_4 (pH 7.0) containing 0.25 M NaCl. Spectra were recorded in the presence of 10% D_2O **[A]**, and in 100% D_2O with incubation time for 1 h before spectrum reading **[B]**.

5.7 Dihedral angles restraints from TALOS-N

TALOS-N (Torsion Angle Likelihood Obtained from Shifts and sequence similarity) (Shen and Bax, 2013) is a database system used to predict torsion angles of protein backbone and side-chains based on NMR chemical shift data and protein sequence. This software package uses the protein sequence along with chemical shifts information to provide predicted values of φ, ψ angles and side-chain angles χ1 as well as a value of uncertainty for each prediction. This software is also able to predict protein secondary structure. The database of TALOS-N contains the torsion angles for proteins with structures solved by X-ray crystallography, that also have nearly complete NMR resonance assignments of 580 proteins. The software database contains experimental secondary chemical shifts from the 580 proteins and chemical shifts information for 9523 proteins processed by another software called SPARTA+. TALOS-N analyzes the input data, which contains both protein sequence and nuclei chemical shifts. This analysis searches the database for heptapeptide [*i-3,…,i,…,i*+3] of similar adjacent residues and similar chemical shifts to those of the input. Then 25 heptapeptide are selected by the software with the best matches to the target protein sequence for each heptapeptide. If these 25 heptapeptide display similar torsion backbone angles then their average can be estimated. TALOS-N is a powerful prediction system which can guarantee a high rate of prediction (~90% with an error less than 3.5%), while the error in angle is about ±12°. In addition, this software can predict about 50% of side-chain χ1 angles with a consistency of 89% to reference structures. Moreover, when a reliable prediction cannot be produced by TALOS-N, the software can produce a “generous” prediction with a prediction rate of 77% of backbone angles (Shen and Bax, 2013).

The chemical shifts of WhiB1 residues from the NMR experiment were uploaded on the server of TALOS-N (https://spin.niddk.nih.gov/bax/nmrserver/talosn/). The software predicted four α-helices and a very short β-strand. The α-helices were detected at positions of P14 to F17, G24 to N38, T44 to N52 and E65 to N75, where the average probability of confidence was more than 90%. Residues 79 and 80 were predicted to form β-strand (Figure 5.13C; 5.14).

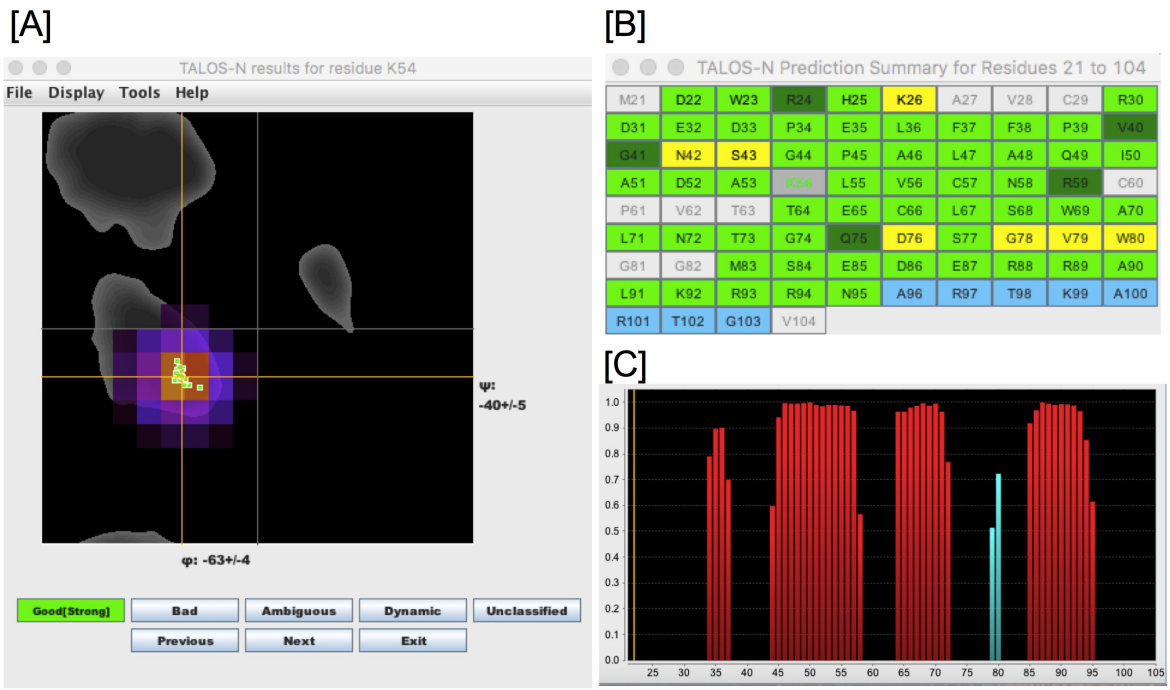


Figure 5.13: TALOS-N user interface representing the WhiB1 K54 residue results. **[A]** the plot panel displays the φ/ψ angles plot of 25 closest matches of the program database as green symbols overlaid on a Ramachandran map showing in gray. **[B]** The user interface shows the WhiB1 sequence, and the residues are marked in different colour depending on the classification of each resides (φ/ψ) prediction data; light grey means no prediction; light and dark green means that the consistent predictions are “Strong’’ and ‘‘Generous’’ respectively; yellow colour refer to the ambiguous prediction, and dynamic residues are indicated in blue. **[C]** The graphic user interface shows the predicted secondary structure of WhiB1, where, red bars represent helix and aqua coloured bars are β-sheet, the bars height reflecting the probability prediction, and the base of each bar represent the type of residues in WhiB1 sequence. The first 20 residues represent the His_6-TEV tag, and WhiB1 sequence start at M21.

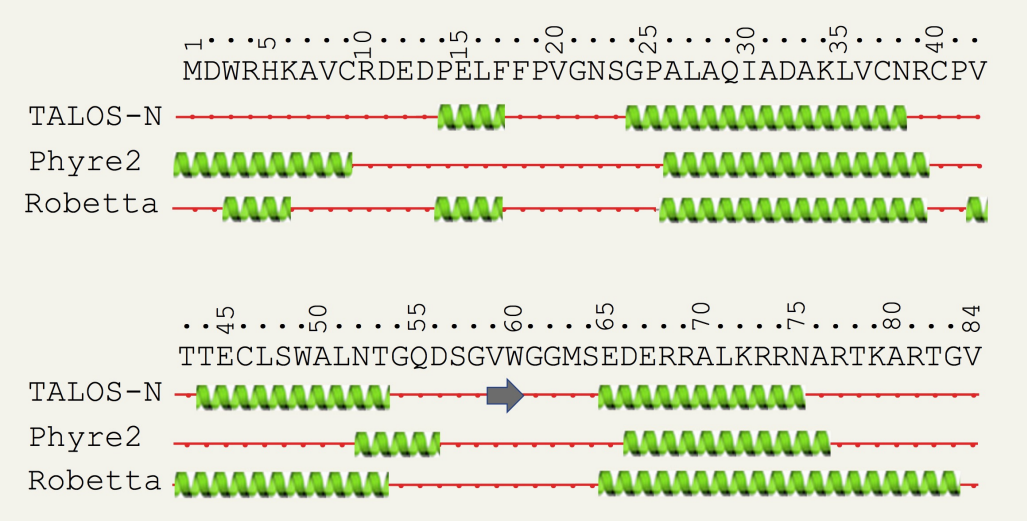


Figure 5.14: Secondary structure distribution among WhiB1 residues predicted by using TALOS-N, Phyre2, and Robetta servers.

Secondary structure prediction based on sequence homology, which was achieved by searching in protein data bank, by Phyre-2 (Kelly et al., 2015) and Robetta (Kim et al., 2004) servers showed a high similarity to TALOS-N prediction which was done mainly based on NMR data (Figure 5.14). The values of WhiB1 backbone torsion angles obtained by TALOS-N are shown in Appendix 7.

5.8 NOE restraints

Protein structure determination using NMR depends on an intense network of inter-proton distance restraints arising from nuclear Overhauser effects (NOEs). The essential role of NOEs in defining protein secondary and tertiary structures comes from the importance of NMR data obtained by NOESY experiments, which connect hydrogen atoms of different residues close together in space (up to 5 Å) but could be separate in the primary structure. The NOE relies on a distinct type of relaxation, which is known as dipole-dipole cross-relaxation. In an NMR spectrum when one signal is irradiated, the intensity of physically nearby signals could be changed. Feeding the original signal (a spin system) with energy causes a signal perturbation and as a result the system uses a dipole-dipole cross-relaxation processes in an attempt to return to equilibrium; this changes the intensity of other signals. The magnitude of gained signal by protons relies on the distance between the irradiated and observed proton(s), where the effect is very small when the distance is more than 5 Å. This process is quite complicated due to the spin diffusion phenomenon and because of the presence of hundreds to thousands of protons in a protein. Manual NOESY spectral assignments are the most difficult and time consuming processes in protein structure determination (Reid, 1997; Williamson, 2012). Measuring the inter-proton distance is usually achieved using spectra of multidimensional NOESY, where the processes of cross-relaxation led to NOE cross peaks in spectra which provide information about the correlation between protons close in space. By using the chemical shifts of protons which were previously obtained by backbone and side-chain assignments, the NOE cross peaks can be matched with known chemical shifts to characterize protons close in space (within 5 Å), which are very useful for protein structure determination. The ^{13}C-NOESY-HSQC and ^{15}N-NOESY-

HSQC spectra which correlate the H-C group of side-chain or H_N from the amide group respectively with protons close in space is the core source of protein structural information, and here the data collection and assignment for WhiB1 was carried by Mrs. Andrea Hounslow. The total unambiguous distance restraints value was 218; where, the NOE cross-peak intensities were divided into three groups with upper bounds of 3.0, 4.0 and 5.0 Å.

5.9 Structure calculations

Structure calculations were carried out by Professor Michael Williamson (University of Sheffield, Molecular Biology and Biotechnology, NMR facility) using the standard simulated annealing protocol of CNS (Brunger et al., 1998) with standard torsion angle refinement. The backbone chemical shift assignments were used to generate φ, ψ and $\chi 1$ dihedral angle restraints using TALOS-N (Shen and Bax, 2013). Further restraints were obtained from NOESY spectra. In addition, nuclei that could not be observed in NMR spectra (N of V8, C9, C37, V42, T43, G61 and G62; CA of C9, C46, and CB of C9, E45, C46) were restrained to be within 6.0 Å of the centre of the iron-sulfur cluster; and nuclei with T1 values of < 50 ms (N of A7, R10, N38, C40, C46, M63) were also restrained to be within 6.5 Å of the centre of the cluster. A small number of nuclei (CB of R10, E12, K6; CG of V8, R10, E12, K6, T43; CD of R10 and K6) that were not noticeably broadened were restrained to be more than 6.5 Å from the centre of the cluster to improve convergence of the calculation. Hydrogen bonds were added when TALOS-N, Phyre2, and Robetta (Kelly et al., 2015) predicted the location of α-helices. The cluster was represented by a single (nominally) zinc atom at the centre of the cluster, restrained to have typical [4Fe-4S] geometry, i.e. with Zn-Cys S distances of 3.9 Å and S-Zn-S angles of 110 Å. The structure (Figure 5.15) contains a core of three α-helices held together by the [4Fe-4S] cluster, with a fourth *C*-terminal helix and a short additional helical segment. Analysis of the chemical shifts suggests that residues preceding M1 or following A76 are dynamically averaged.

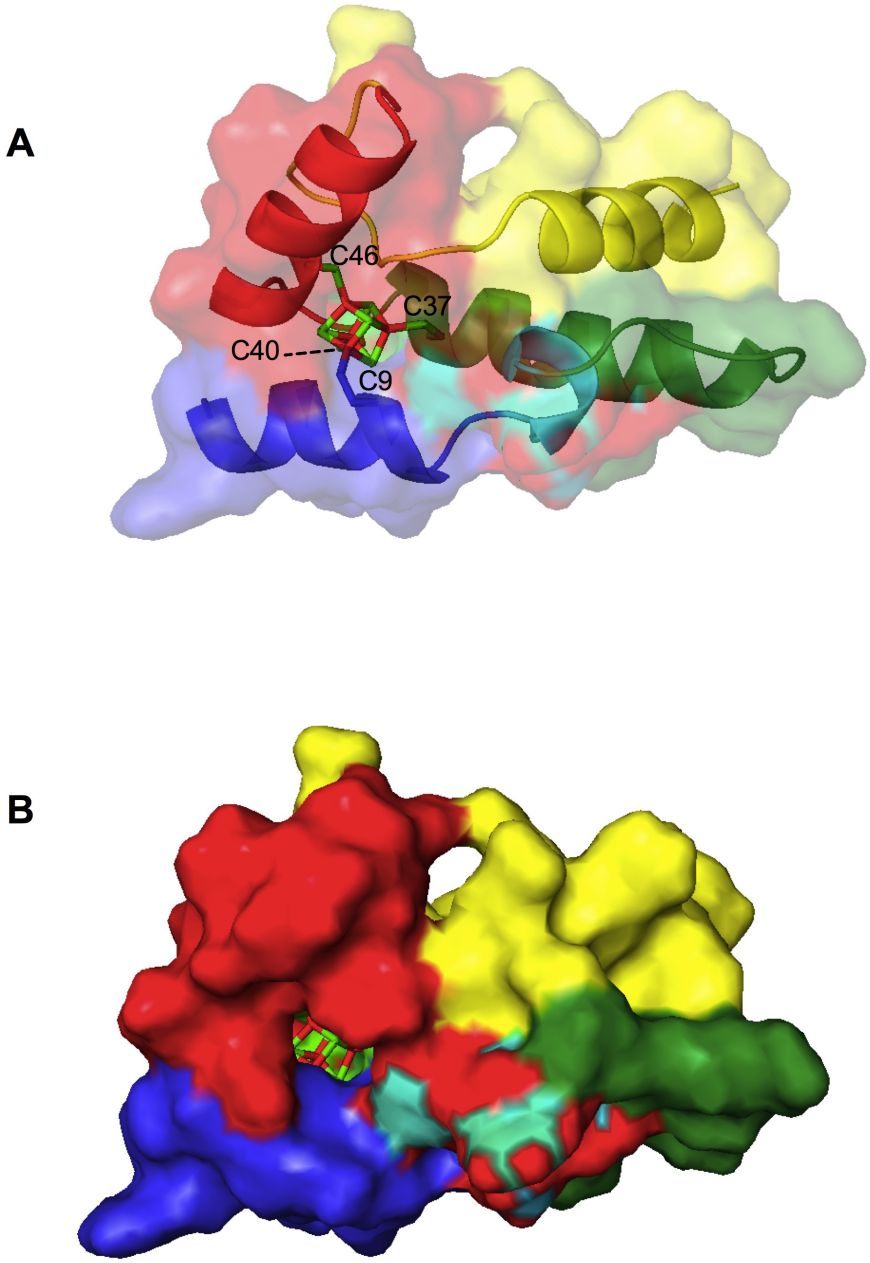


Figure 5.15: 3D Structural model of WhiB1 calculated using NMR restraints. The chain is coloured from blue at the *N*-terminus to yellow at the *C*-terminus. The structure shown starts at residue 1, and finishes at residue 76, excluding the first 20 residues of the tag. **[A]** Ribbon/cartoon presentation, **[B]** Surface presentation.

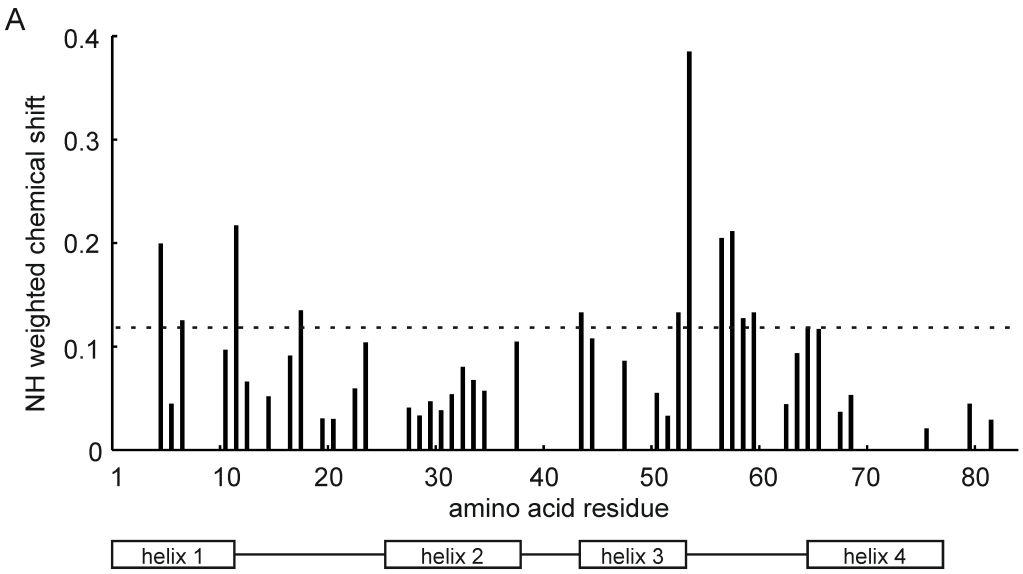
5.10 The *C*-terminal region of WhiB1, which is implicated in DNA-binding, does not interact with SigA

Studying the binding sites of SigA in WhiB1 was carried out by overproducing an ^{15}N labelled complex of WhiB1:$SigA_{C82}$, which was achieved by growing *E. coli* JRG6857 carrying both pGS2566 and pGS2560 (Section 4.5.2), in a minimal medium supplemented with $^{15}NH_4Cl$ (Section 2.3.2). This was followed by a 2D ^{15}N-HSQC experiment.

For defining the location of SigA binding based on chemical shift change, the most suitable technique was by measuring the shift differences between an assigned WhiB1 signal and an unassigned signal in the complex (using ^{15}N-HSQC spectrum). This experiment was used here because the WhiB1-His_6-$SigA_{C82}$ complex was expressed and purified as a complex, meaning that chemical shift assignments for WhiB1 could not be made by following titration shifts.

Changes in the chemical shifts observed in two-dimensional (^1H and ^{15}N) NMR spectra provide a means to define regions of interaction between proteins. Comparison of NH weighted chemical shift changes between His_6-TEV-WhiB1 and the WhiB1-His_6-$SigA_{C82}$ complex indicated that the structure of the *C*-terminal region of WhiB1, which is implicated in DNA-binding, was similar in both (Figure 5.16). This suggests that SigA does not interact with the DNA-binding region of WhiB1. The largest weighted NH chemical shift changes for the WhiB1 amino acids that could be assigned were for His5, Glu12, Phe18 and Gly54, suggesting that the *N*-terminal region and the conserved predicted β-turn (^{54}GQDSGVW^{60}) of WhiB1 participate in interactions with $SigA_{C82}$.

The next step was to target some possible residues in WhiB1 and mutate them in order to define the key residue(s) essential for binding with SigA. According to the results of this section it was assumed that the positively-charged residues are crucial for DNA-binding (Smith *et al.* 2010), while some negative



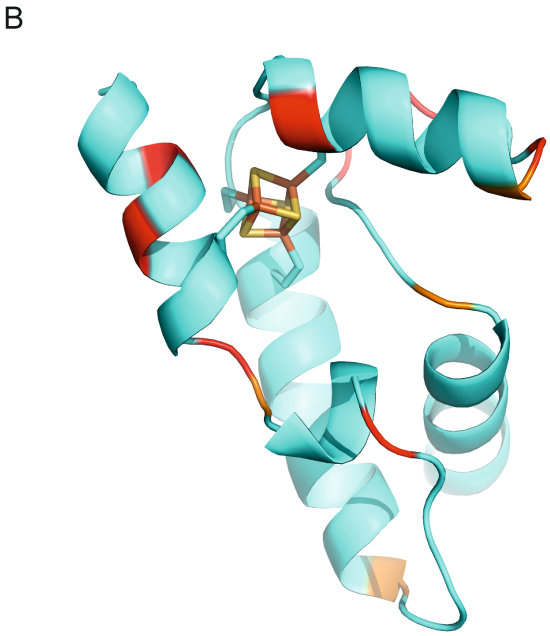


Figure 5.16: Weighted NH chemical shift differences for His-TEV-WhiB1 compared to the WhiB1:His-SigA_{C82} complex suggest that the DNA-binding region of WhiB1 is not involved in interaction with SigA_{C82}. (A) The chart shows the weighted ($\Delta\delta = \sqrt{(\delta H)^2 + (\delta N/5)^2}$) chemical shifts for the WhiB1:His-SigA_{C82} complex compared to the His-TEV-WhiB1 protein. The *C*-terminal region of WhiB1 shows little change upon complex formation with SigA_{C82}. **(B)** Shifts are mapped onto the structure of WhiB1. The largest shift changes (red) are in the β-turn sequence and *N*-terminal region. The *C*-terminal DNA-binding helix (bottom right) shows very little shift change.

residues in the *N*-terminal region and near the predicted β-turn might be needed for interaction with SigA. WhiB1 has clusters of negatively-charged residues such as (^{2}D,^{11}DEDPE^{15}) and (^{65}EDE^{67}). In the literature there are some examples (but not in the case of WhiB1) that show that SigA R515 is essential for interaction and some others that show that the interaction happens mainly via negative residues in SigA partner proteins (Burian et al., 2013; Feng et al., 2016; Gautam et al., 2014; Steyn et al., 2002). Therefore, targeting the WhiB1 negatively-charged residues showing the biggest NH weighted chemical shift change was investigated by mutation to positive residues. Site-directed mutagenesis was carried out as described previously (Section 2.5.12). The pGS2560, plasmid which encoded a non-tagged WhiB1, was used as a template to create the following plasmids; pGS2570, pGS2571, pGS2574, pGS2576, pGS2577, which encode WhiB1 E12K, D13R, E45K, E65K, and D66R respectively. Expanding the investigation by targeting the rest of WhiB1 negatively-charged residues using pGS2569, pGS2572, pGS2573, pGS2575, pGS2578, which encode WhiB1 D2R, E15K, D32R, D56R, and E67K respectively was also attempted. These plasmids, which each carry a single mutation, were individually co-transformed with pGS2566, encoding His_6-ThCS-$SigA_{C82}$, into the electro-competent *E. coli* JRG5302 strain. Overproduction and purification of the corresponding mutant WhiB1:$SigA_{C82}$ complexes were as descried in Section 4.2.1. Results showed that none of the mutations disrupted the complex. This suggests that other amino acids, or more than one residue might be needed for binding with SigA. Interestingly, all these mutations did not impair the ability of WhiB1 to anchor an iron-sulfur cluster, as all the complexes of mutant WhiB1:$SigA_{C82}$ contained a [4Fe-4S] cluster. The effect of WhiB1 D11R was not studied due to the presence of several unintended mutations.

5.11 WhiB1 structure description and discussion

The structure of WhiB1 based on NMR experiments is the first structure of a Wbl protein. Moreover, mapping the binding interface between WhiB1 and the *C*-terminal domain of SigA ($SigA_{C82}$) has been studied. Stabilizing the [4Fe-4S] cluster of WhiB1 due to the presence of the TEV-tag or the partner protein (SigA) was the breakthrough that enabled structural studies by NMR for the first time.

The initial result of 1D ^{1}H NMR of TEV-tagged WhiB1 (Figure 3.19) showed that the protein was well-structured, unlike the apo-form of WhiB1, suggesting that the cluster can hold the domains of the protein together and change the protein structure. The target of this chapter therefore was to understand how the cluster organizes the protein, which was achieved by building an NMR structure.

The NMR structure was built using the WhiB1 nuclei chemical shifts obtained from the assignments of different NMR spectra. However, some chemical shifts of the residues that anchor the cluster and some others close to the cluster could not be assigned due to the paramagnetic effect of the cluster (Section 5.9). Similar obstacles were reported in the case of *Desulfovibrio africanus* Ferredoxin I; where, the quality of the data was affected due to the paramagnetic nature of the cluster (Davy et al., 1998). Fast nuclear relaxation rates due to paramagnetism causes broadening of the NMR peaks. This happens due to the coupling of nuclei with cluster unpaired electrons, which hamper the detection of NOEs and *J*-coupling constraints (nuclear connectivities) that are important for structure determination (Bertini et al., 1999). Many other examples showed similar paramagnetic effects of FeS clusters such as *Chromatium vinosum* [4Fe-4S]-HiPIP (Banci et al., 1995; Bertini et al., 1995a), *Clostridium pasteurianum* 2[4Fe-4S] ferredoxin (Bertini et al., 1995b), and *Bacillus schlegelii* [4Fe-4S] and [3Fe-4S] ferredoxin (Aono et al., 1998).

The TALOS-N server was used here to obtain the torsion angles of WhiB1 using the chemical shifts of the protein residues, torsion angles were used later for structure calculations along with NOEs. Based on the chemical shifts of

WhiB1 and the protein sequence, TALOS-N also predicted the secondary structure of WhiB1 (Section 5.7, Figure 5.13), where, hydrogen bonds of α-helices were defined at the beginning. Phyre2 and Robetta servers were used to predict secondary structure using the protein sequence only by searching in protein data bank for sequence homology. The WhiB1 predicted secondary structure by Phyre2 and Robetta increased the confidence of predicted secondary structure by TALOS-N due to the high similarity (Figure 5.14). However, no similar protein to WhiB1 has a solved three-dimensional structure in the protein database, and predicted secondary structures by Phyre2 and Robetta were made based on a combination of different domains in different proteins. Nevertheless, after calculating the structure and using NOE restraints the secondary structure of WhiB1 was like the predicted models, which improves confidence in the NMR structure (Figure 5.17).

The WhiB1 structure is a novel structure, where three-dimension structural comparison using the DALI server (http://ekhidna.biocenter.helsinki.fi/dali_server/start), which searches the protein data bank for similar proteins based on structure, produced hits to numerous helical bundles, but nothing of high similarity (Holm and Rosenström, 2010).

WhiB1 has four α-helices with a fifth short helical segment, but there were no beta strands. The first α-helix (Helix1) located at the *N*-terminal region of WhiB1 is composed from (^{2}DWRHKAV<u>C</u>RD$_{NH}$^{11}). The thiol group of Cys9 side-chain represents the first ligand that anchors the WhiB1 [4Fe-4S] cluster. Helix1 is followed by a short loop of (^{11}$_{CO}$DEDP$_{NH}$^{14}), where the main NH group of Glu12 in this turn forms a hydrogen bond with carbonyl oxygen groups of Cys9; in contrast the main CO and COδ_2 groups of Asp13 accept hydrogen bonds from main NH groups of Leu16 and Glu15 respectively. As a result, and due to the presence of these H-bonds, it is unlikely that this region between Helix1 and the short helical segment is a random loop, and is more likely to form a turn. Moreover, the interaction of SigA$_{C82}$ with WhiB1 showed a big change in the weighted NH chemical shift of WhiB1 Glu12 and Asp13 in addition to His 5 and Ala 7 (in the helix) compared to free WhiB1 (Figure 5.16), suggesting that these sites could be involved in the interaction with SigA. However, WhiB1 E12K, or D13R separately still interact with SigA and the

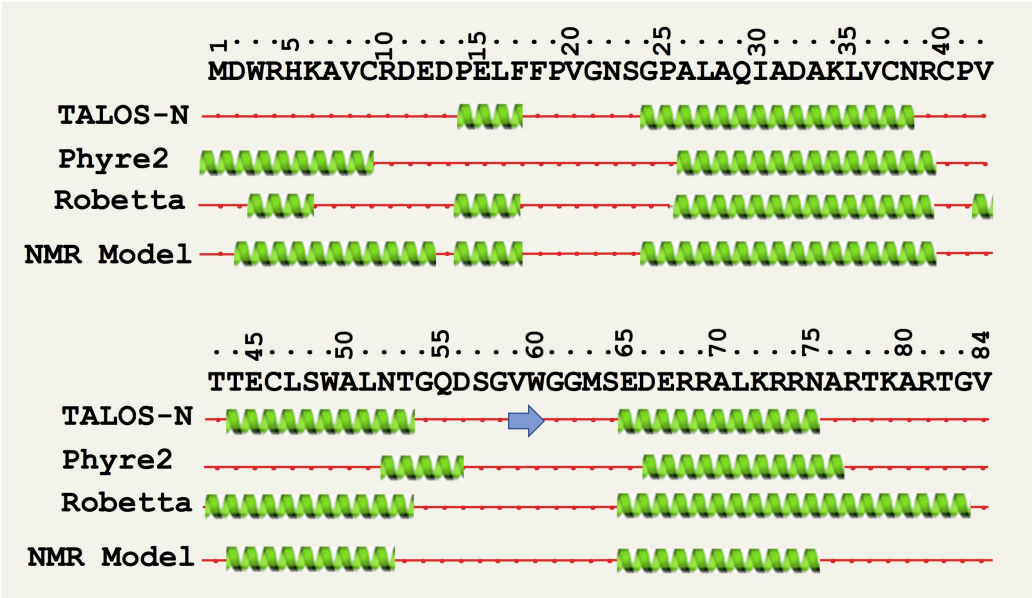


Figure 5.17: Distribution of secondary structure over WhiB1 sequence generated by prediction server; TALOS-N, Phyre2, and Robetta along with WhiB1 structural model built by NMR method.

mutations did not impair the interaction, suggesting that these residues are not essential individually for disrupting the interaction. An α-helical segment composed from ($^{14}_{CO}$PELF$_{NH}^{17}$) is located after the short turn, however this segment is also followed by a loop of ($^{17}_{CO}$FFPVGNSG$_{NH}^{24}$). The latter loop has no hydrogen bonds between residues except between CO group of Gly21 and the side-chain (Hδ_2) group of Asn22. This loop is followed by the second WhiB1 α-helix (Helix2) ($^{24}_{CO}$GPALAQIADAKLV<u>C</u>NR$_{NH}^{39}$) leading to a helix-loop-helix structure. Where the sulfhydryl group of Cys37 forms the second ligand that binds the [4Fe-4S] cluster of WhiB1. The third cysteine (Cys 40) ligand of the cluster is located at the loop region of ($^{39}_{CO}$R<u>C</u>PVTT$_{NH}^{44}$). While, the final Cys47 ligand is present in the next α-helix (Helix3) ($^{44}_{CO}$TE<u>C</u>LSWALNT$_{NH}^{53}$). Clearly, the cluster holds the three α-helices together and folds the protein correctly. Once the protein loses the cluster it also changes its tertiary structure. Moreover, a surface representation of WhiB1 shows that a sulfur atom of the [4Fe-4S] cluster is exposed at the bottom of a channel whose mouth is formed by Arg10, Phe17, Glu45 and Trp49 (~9 Å diameter) (Figure 5.18). This structural feature is perhaps consistent with cluster disassembly initiating with sulfur ejection (Figure 3.21) and provides a route for NO to access the cluster.

Alam et al. (2009) reported that all *M. tuberculosis* Wbl (except WhiB5) are protein disulfide reductases, this result was assumed based on reduction of an insulin disulfide, fast migration of oxidized proteins on SDS-PAGE, and measuring the protein weight by mass spectrometry after alkylating the protein by iodoacetamide which suggested that all cysteine residues were in an oxidized form. They assumed that apo-WhiB1 (and the others) tend to form two disulfide bonds after losing the cluster, one between Cys37 and Cys40, and a second between Cys9 and Cys46. The same group also published that WhiB1 specifically interacted with α(1,4)-glucan branching enzyme (GlgB) to reduce its disulfide bonds (Garg et al., 2009, 2007a). If the disulfide formation arrangement is true in apo-WhiB1 as reported by Alam and Garg, and according to WhiB1 structure reported here, the Helix1 and Helix3 would be attached by one S-S bond while the cysteine side-chain (SH group) of Cys37 and Cys40 located in the second α-helix will form the second bond; which

could lead to loss of the link between the three helices leading to changing the protein structure after losing the cluster. However, Smith et al, (2010), reported that the oxidized apo-WhiB1 is less likely to be a general protein disulfide reductase, as the reaction rate of insulin reduction in the presence of WhiB1 was slow compared to the rate of reaction in the presence of thioredoxin in a much lower concentration, and similar to the reaction rate when no WhiB1 is present. In addition, the reaction of apo-WhiB1 with iodoacetamide suggested that mixed states of reduced and oxidized cysteine were present in WhiB1 after losing the cluster (Section 3.3.5, Figure 3.9), indicating that the oxidation state of apo-WhiB1 cysteines is still under the debate.

The structure of WhiB1 also showed another α-helix (Helix4) ($^{65}_{CO}$EDERRALKRRN^{75}) located at the *C*-terminal region, which is not involved in supporting the cluster. WhiB1 also has a loop region ($^{53}_{CO}$TGQDS<u>GVWGG</u>MSE$_{NH}^{65}$) located between Helix3 and Helix4 forming another helix-loop-helix. The underlined residues in the loop region were suggested for long time to form a β-turn, these residues are well conserved among several *M. tuberculosis* Wbl proteins and some other members of actinomycetes (Smith et al., 2012; Soliveri et al., 2000). The holo-WhiB1 structure at this loop region does not show any hydrogen bonds indicating the absence of this predicted β-turn. Smith et al. (2012) showed that apo- and nitrosylated-WhiB1 were able to bind DNA but not the holo-form, and Gly58, Gly61, and Gly62 were each essential for DNA-binding but not for the cluster incorporation. Moreover, each of the Lys72, Arg73, and Arg74, which are located at the *C*-terminal α-helix, when they were mutated to Glu impaired maintenance DNA-binding by WhiB1 but did not affect cluster reconstitution. In contrast, any single replacement of the four cysteine residues, that are essential for cluster holding, with Ala did not affect the apo-WhiB1 binding to DNA. Here, Arg73 and Arg74 in the DNA-binding domain are pointing inwards to the body of the holo-protein and are therefore not accessible to DNA. Accordingly, a mechanism for switching between holo- and DNA-binding forms has been suggested: when WhiB1 holds the cluster Helix3 will be restrained with the other two helices of the *N*-terminal domain, but once the protein loses the cluster Helix3 will be free which might produce a beta turn formation in the

next loop region. This change in folding might allow the *C*-terminal domain to bind DNA. Interestingly, the complex of WhiB1:$SigA_{C82}$ caused changes in the weighted NH chemical shift of 2D NMR spectrum at the predicted region of β-turn, suggested that this loop region could play a dual function, where for holo-WhiB1 it is involved in the interaction with SigA, and in the apo-form could aid DNA-binding.

Although apo-WhiB1 was found to be dominated by random coils as judged by CD and 1D ^{1}H NMR spectra (Figure 3.14, 3.19), it is possible that the apo-protein can be folded once it binds with its target DNA. This was seen clearly in the small acid-soluble spore proteins (SASP) such as *B. megaterium* SASP-A*, B. subtilis* $SspC^{Tyr}$, *B. cereus* BceI, and *C. bifermentans* CbiC, which all have been reported to lack secondary structure properties when there is no DNA but upon binding with DNA they become significantly rich with α-helices (Hayes et al., 2000).

Several mutations of negatively-charged residues of WhiB1 did not impair the interaction with SigA, which later on after building the WhiB1 NMR structure were found that some of them are either buried or not entirely on the surface of WhiB1 such as Glu12, Asp13 and Glu67 (Figure 5.19). It is possibly that more than one mutation is needed to disrupt the interaction or perhaps different types of residues are required. Moreover, due to the possible overlap of WhiB1 surfaces interacting with SigA and the TEV-tag, perhaps the comparison of chemical shifts changes between TEV-tagged-$WhiB1^{Ms}$ and the complex is not an ideal experiment to investigate the sites of interaction. Perhaps using a random mutagenesis protocol with two-hybrid BACTH system might provide a direct answer to identify the binding sites. However, the interaction of WhiB1 with SigA and the TEV-tag on the cluster stability need further structural-functional investigations.

Work here has produced a WhiB1 structural model for the first time among all the Wbl proteins. Getting a structure of these proteins and understanding their function could provide tremendous information about the mechanisms of dormancy and reactivation of *M. tuberculosis* during the infection, which could lead to new therapeutics. Work to further understand the WhiB1 structure after

nitrosylation, interaction with SigA, or in a complex with DNA may provide very important insights to understand how WhiB1 interacts with SigA and regulates genes in response to NO.

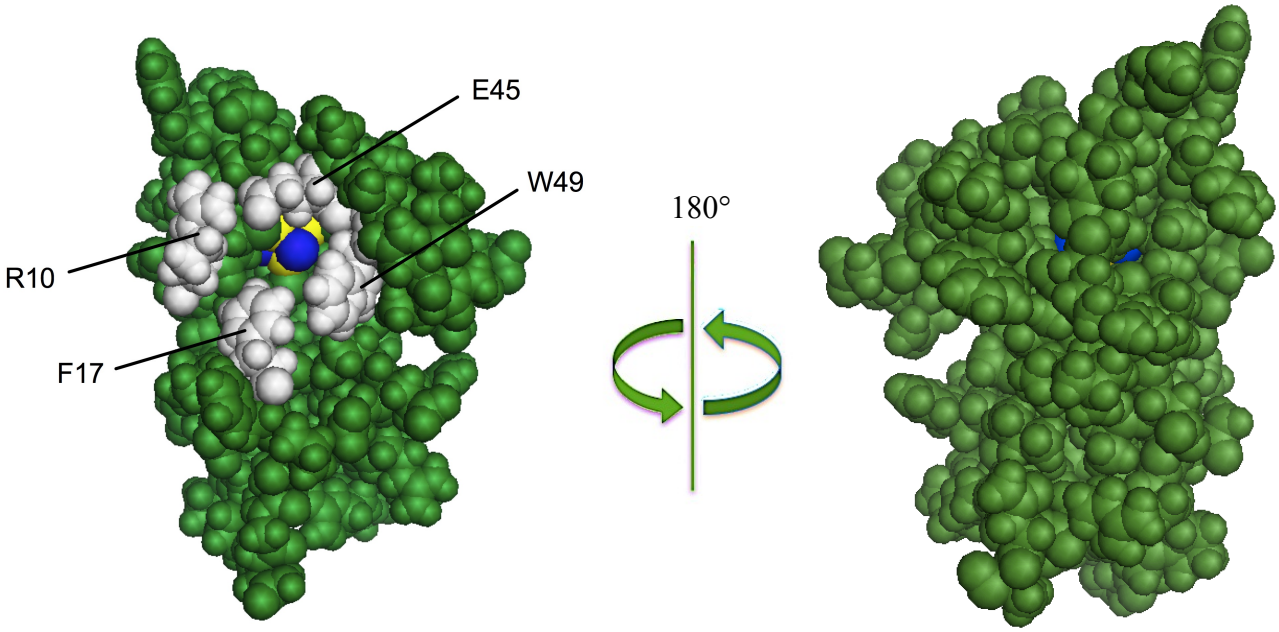


Figure 5.18: Surface representation of WhiB1 showing solvent access to the iron-sulfur cluster. The WhiB1 surface is shown in green, iron as yellow spheres, sulfur as blue spheres and residues of Arg10, Phe17, Glu45 and Trp49 that form the mouth of the channel leading to the exposed sulfur atom of the cluster are highlighted in white and single letter amino acid codes.

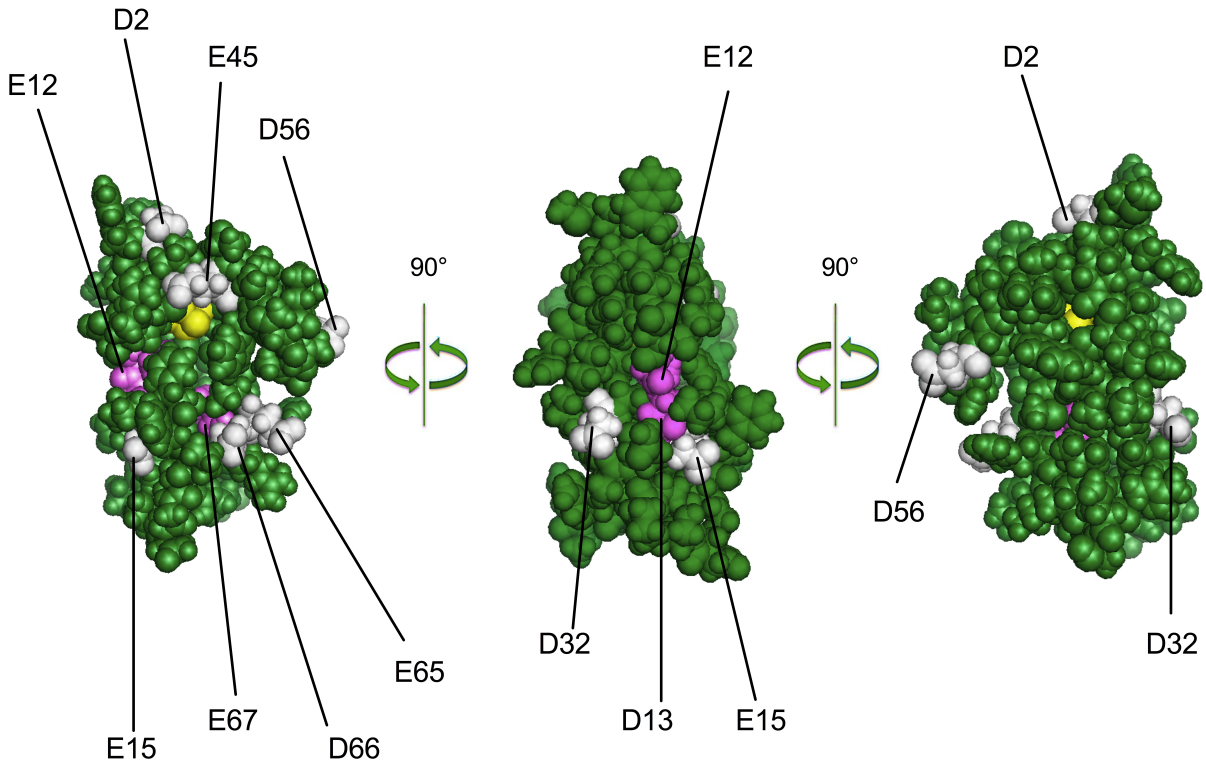


Figure 5.19: Surface representation of WhiB1 showing the positions of negatively charged residues. The WhiB1 surface is shown in green, iron-sulfur cluster as yellow spheres, and the highlighted spheres in white and violet and single letter amino acid codes represents the negative charged residues where each Glu and Asp were mutated to Lys and Arg respectively. Glu 12, Asp13, and Glu 67 which are highlited in violet, are buried and not available on surface.

Chapter six

General discussion

6.0 General discussion

The work presented in this study has defined for the first time that the [4Fe-4S] cluster is indeed the cofactor present in *M. tuberculosis* WhiB1; it is essential for interaction of WhiB1 with the *C*-terminal domain of major sigma factor SigA, the dissociable unit of RNA polymerase; the iron-sulfur cluster can react with NO, the major "weapon" of host macrophages, and this reaction leads to release of SigA from the complex; the structural analysis of WhiB1 indicates that the cluster is crucial for stabilizing the structure of the protein, and the WhiB1 domains that hold the cluster are required for interaction with SigA.

This work and that of others raises several questions about the function of WhiB1 in the bacilli during switching between dormant and replicative states. Perhaps, the most important question which previously was unknown is: why WhiB1 is essential for *M. tuberculosis*? Finding the purpose behind the interaction of WhiB1 with SigA could be useful to answer this question. As previously discussed in Chapter four (Section 4.6), it is unlikely that WhiB1 acts as an anti-SigA factor, because SigA is required for transcribing housekeeping genes. Thus, it is more likely that WhiB1 redirects transcription by targeting a sub-set of gene(s) under special circumstances (Figure 6.1). Moreover, it is possible that this process is governed by another as yet unidentified regulatory factor(s) that act as an anti-WhiB1 factor. Switching WhiB1 release and binding to RNA polymerase or anti-WhiB1 factor(s) is important for regulating essential processes during *M. tuberculosis* dormancy and reactivation (Figure 6.1). The theory of the presence of another WhiB1 partner protein is supported by the fact that both TEV tag and SigA stabilize the protein; however, there is no sequence similarity between the tag and the SigA protein. Moreover, several examples in actinomycetes and other families of bacteria showed the presence of anti- anti-sigma factors (Beaucher et al., 2002; Lee et al., 2004). An investigation of the presence of another WhiB1 partner protein and designing *in vitro* transcription assays in the presence of WhiB1 might provide new insight into the mechanisms of transcription regulation by WhiB1. The other possible scenario is the regulation of essential genes occurs in response to specific stimuli such as NO. As reported previously by Smith et al. (2010) and confirmed here, WhiB1 is not an O_2 sensor but it reacts very quickly with eight

molecules of NO forming an octa-nitrosylated form. This reaction switches WhiB1 from a non-DNA binding form to DNA-binding active form (Smith et al., 2010). Here it was also found that NO triggers WhiB1 release from a complex with SigA. Thus, it is likely that switching between binding with SigA and DNA in response to NO plays a regulatory role in *M. tuberculosis* genetic network that may be essential for the pathogen under stress conditions (Figure 6.1).

Although, the sensitivity of WhiB1 to NO and O_2 are well studied here and by others; the presence of another external signal(s) associated with infection could regulate or react with the WhiB1 cluster is possible. Signals that are more abundant in macrophage or granuloma under anaerobic conditions are ROI like hydrogen peroxide and superoxide. Indeed, cyclic AMP is another signal that is important in regulating WhiB1, which is present in macrophages at high level during the infection (Bai and Schaak, 2009). The expression of *whiB1* is regulated by cAMP receptor protein (CRP), which binds directly upstream of *whiB1* (Agarwal *et al.,* 2006). The complex mechanism of this regulation depends on the available level of this signal. In the presence of high level of cAMP, two binding sites are occupied by CRP downregulating expression of *whiB1,* but at the low level of cAMP only one site is occupied by CRP, and *whiB1* transcription is upregulated (Stapleton *et al.,* 2010). Thus, it is more likely that in the presence of low level of NO (high concertation can kill *M. tuberculosis*) and high level of cAMP, the pathogen enters a dormant state; where nitrosylated-WhiB1 downregulates *whiB1* gene. This downregulation could lead to changes in the transcription state of other regulators directly or in directly that are important for dormancy such as PhoPR and DosS/T/R systems which might explain why WhiB1 is essential.

In conclusion, the essential nature of WhiB1 and the ability to bind SigA, in addition to its response and sensitivity to NO which leads to dissociation of the WhiB1:SigA complex and switching the protein to a DNA-binding form provide a new information about the significance of this regulator. Thus, WhiB1 is an important NO sensor protein that possibly plays a vital role in *M. tuberculosis* during the adaptation processes at the harsh environments within the host. Therefore, WhiB1 is a potential factor that can be targeted for more research, which could lead to potential new treatments against this pathogen.

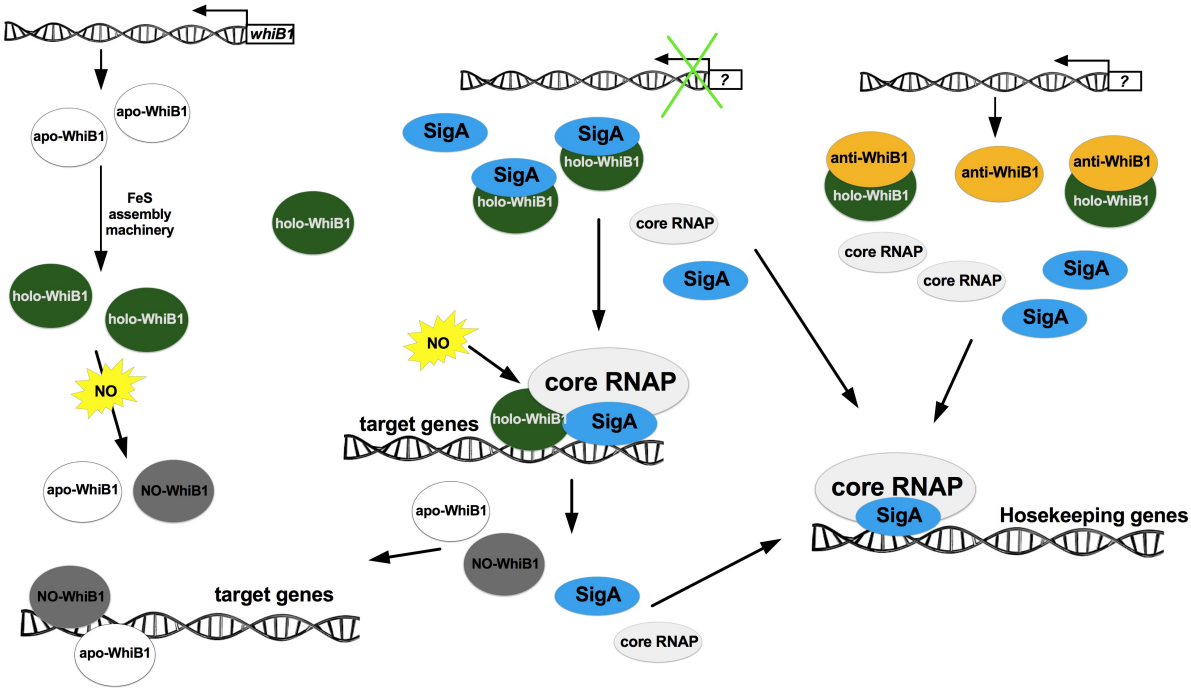


Figure 6.1: Model depicting the possible pathway of WhiB1 forms in regulating transcription during the developmental processes in *M. tuberculosis*. When the transcription of a gene(s) that encodes anti-WhiB1 protein(s) is switched on and activate protein(s) production, a complex of anti-WhiB1:holo-WhiB1 is formed. This complex might be required for another virulence or gene-regulatory functions. As a result, a complex of SigA:holo-WhiB1 will not be formed; but when the anti-WhiB1 is downregulated, the SigA:holo-WhiB1 complex is formed and redirects transcription of target sub-set of gene(s). SigA level is possibly higher than holo-WhiB1 during infection (see Section 4.6), thus housekeeping genes will continuously be transcribed by free holo-RNAP whether WhiB1 is bound with anti-WhiB1 or some SigA. In the presence of NO the transcription governed by holo-WhiB1:holo-RNAP will be inhibited due to dissociation of WhiB1:SigA complex. Nitric oxide reaction generates apo- or nitrosylated-WhiB1, switching the protein to DNA-binding forms which can target another sub-set of gene(s). Apo-WhiB1 however can also be transformed to holo-WhiB1 via iron-sulfur assembly machinery system.

References

References

Agarwal, N., Raghunand, T.R., and Bishai, W.R. (2006). Regulation of the expression of *whiB1* in *Mycobacterium tuberculosis*: Role of cAMP receptor protein. *Microbiology, 152*: 2749–2756.

Aguiló, N., Marinova, D., Martín, C., and Pardo, J. (2013). ESX-1-induced apoptosis during mycobacterial infection: to be or not to be, that is the question. *Front Cell Infect Microbiol*, *3*: 88.

Alam, M.S., Garg, S.K., and Agrawal, P. (2007). Molecular function of WhiB4/Rv3681c of *Mycobacterium tuberculosis* H37Rv: A [4Fe-4S] cluster co-ordinating protein disulphide reductase. *Mol Microbiol*, *63***(5):** 1414–1431.

Alam, M.S., Garg, S.K., and Agrawal, P. (2009). Studies on structural and functional divergence among seven WhiB proteins of *Mycobacterium tuberculosis* H37Rv. *FEBS J, 276***(1):** 76–93.

Aono, S., Bentrop, D., Bertini, I., Donaire, A., Luchinat, C., Niikura, Y., and Rosato, A. (1998). Solution structure of the oxidized Fe_7S_8 Ferredoxin from the thermophilic bacterium *Bacillus schlegelii* by ^1H NMR spectroscopy. *Biochemistry*, *37***(27):** 9812–9826

Armstrong, J., and Hart, P. (1975). Phagosome-lysosome interactions in cultured macrophages infected with virulent tubercle bacilli. Reversal of the usual nonfusion pattern and observations on bacterial survival. *J Exp Med,* 1; *142***(1):** 1-16

Bai, G., Schaak, D., McDonough, K. (2009). cAMP levels within *Mycobacterium tuberculosis* and *Mycobacterium bovis* BCG increase upon infection of macrophages. *FEMS Immunol Med Microbiol, 55***(1):** 68–73.

Banci, L., Bertini, I., Dikiy, A., Kastrau, D.H., Luchinat, C., and Sompornpisut, P. (1995). The three-dimensional solution structure of the reduced high-potential iron-sulfur protein from *Chromatium vinosum* through NMR. *Biochemistry*, 34**(1):** 206-219.

Beaucher, J., Rodrigue, S., Jacques, P.-E., Smith, I., Brzezinski, R., and Gaudreau, L. (2002). Novel *Mycobacterium tuberculosis* anti- σ factor antagonists control σ^F activity by distinct mechanisms. *Mol Microbiol*, 45***(6):*** 1527-1540.

Bertini, I., Capozzi, J.F., Luchinat, C., Piccioli, M., and Vilatvt, A.J. (1994). The Fe_4S_4 centers in Ferredoxins studied through proton and carbon hyperfine coupling. Sequence-specific assignments of cysteines in Ferredoxins from *Clostridium acidi urici* and *Clostridium pasteurianum*. *J Am Chem Soc, 116***(2):** 651-660.

Bertini, I., Dikiy, A., Kastrau, D.H.W., Luchinat, C., and Sompornpisut, P. (1995a). Three-dimensional solution structure of the oxidized high potential iron-sulfur protein from *Chromatium vinosum* through NMR. Comparative

analysis with the solution structure of the reduced species. *Biochemistry, 34***(31):** 9851-9858.

Bertini, I., Donaire, A., Feinberg, B.A., Luchinat, C., Picciolp, M., and Yuan, H. (1995b). Solution structure of the oxidized 2[4Fe-4S] Ferredoxin from *Clostridium Pasteurianum*. *Eur J Biochem, 232***(1):** 192-205.

Bertini, I., Couture, M.M.J., Donaire, A., Eltis, L.D., Felli, I.C., Luchinat, C., Piccioli, M., and Rosato, A. (1996). The solution structure refinement of the paramagnetic reduced high-potential iron-sulfur protein I from *Ectothiorhodospira halophila* by using stable isotope labeling and nuclear relaxation. *Eur J Biochem*, *241*: 440-452.

Bertini, I., Rosato, A., and Turano, P. (1999). Solution structure of paramagnetic metalloproteins. *Pure Appl Chem, 71***(9):** 1717–1725.

Bhat, S.A., Singh, N., Trivedi, A., Kansal, P., Gupta, P., and Kumar, A. (2012). The mechanism of redox sensing in *Mycobacterium tuberculosis*. *Free Radic Biol Med*, *53***(8):** 1625-1641.

Bodenhausen, G., and Ruben, D.J. (1980). Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy. *Chem Phys Lett*, *69***(1):** 185–189.

Bradford, M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem, 72*: 248-254.

Brunger, A.T., Adams, P.D., Clore, G.M., Delano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J. S., Kuszewski, F., 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 Cryst, 54*: 905-921.

Burian, J., Ramón-García, S., Sweet, G., Gómez-Velasco, A., Av-Gay, Y., and Thompson, C.J. (2012). The mycobacterial transcriptional regulator *whiB7* gene links redox homeostasis and intrinsic antibiotic resistance. *J Biol Chem, 287***(1):** 299-310.

Burian, J., Yim, G., Hsing, M., Axerio-Cilies, P., Cherkasov, A., Spiegelman, G.B., and Thompson, C.J. (2013). The mycobacterial antibiotic resistance determinant WhiB7 acts as a transcriptional activator by binding the primary sigma factor SigA (RpoV). *Nucleic Acids Res*, *41***(22):** 10062-10076.

Cai, M., Huang, Y., Sakaguchi, K., Clore, G.M., Gronenborn, A.M., and Craigie, R. (1998). An efficient and cost-effective isotope labeling protocol for proteins expressed in *Escherichia coli*. *J Biomol NMR*, *11*: 97-102.

Cao, G., Howard, S.T., Zhang, P., Wang, X., Chen, X.L., Samten, B., and Pang X. (2015). EspR, a regulator of the ESX-1 secretion system in

Mycobacterium tuberculosis, is directly regulated by the two-component systems MprAB and PhoPR. *Microbiology*, *161***(3):** 477-489.

Casonato, S., Cervantes Sánchez, A., Haruki, H., Rengifo González, M., Provvedi, R., Dainese, E., Jaouen, T., Gola, S., Bini, E., Vicente, M., Johnsson, K., Ghisotti, D., Palù, G., Hernández-Pando, R., and Manganelli, R. (2012). WhiB5, a transcriptional regulator that contributes to *Mycobacterium tuberculosis* virulence and reactivation. *Infect Immun*, *80***(9):** 3132–3144.

Chandra, G., Chater, K.F., and Bornemann, S. (2011). Unexpected and widespread connections between bacterial glycogen and trehalose metabolism. *Microbiology, 157*: 1565–1572.

Chawla, M., Parikh, P., Saxena, A., Munshi, M., Mehta, M., Mai, D., Srivastava, A.K., Narasimhulu, K. V, Redding, K.E., Vashi, N., Kumar, D., Steyn, A.J., Singh, A. (2012). *Mycobacterium tuberculosis* WhiB4 regulates oxidative stress response to modulate survival and dissemination in vivo. *Mo Microbiol*, *85***(6):** 1148–1165.

Chen, P., Ruiz, R.E., Li, Q., Silver, R.F., and Bishai, W.R. (2000). Construction and characterization of a *Mycobacterium tuberculosis* mutant lacking the alternate sigma factor gene, sigF. *Infect Immun*, *68***(10):** 5575–5580.

Cheng, H., and Markley, J.L. (1995). NMR spectroscopic studies of paramagnetic proteins: Iron-sulur proteins. *Annu Rev Biophys Biomol Strucure, 24*: 209–237.

Clore, G.M., Bax, A., Driscoll, P.C., Wingfield, P.T., and Gronenborn, A.M. (1990). Assignment of the side-chain ^{1}H and ^{13}C resonances of Interleukin-I β using double- and triple-resonance heteronuclear three-dimensional NMR spectroscopy. *Biochemistry*, 29**(35):** 8172–8184.

Clubb, R.T., Thanabal, V., and Wagner, G. (1992). A constant-time three-dimensional triple-resonance pulse scheme to correlate intraresidue $^{1}H^{N}$, ^{15}N, and $^{13}C'$ chemical shifts in ^{15}N-^{13}C-labelled proteins. *J Magn Reson, 97*: 213–217.

Cole, S. T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S. V., Eiglmeier, K., Gas, S., Barry III, C. E., 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.

Collins, D.M., Kawakami, R.P., de Lisle, G.W., Pascopella, L., Bloom, B.R., and Jacobs, W.R. (1995). Mutation of the principal sigma factor causes loss of virulence in a strain of the *Mycobacterium tuberculosis* complex. *Proc Natl Acad Sci USA, 92***(17):** 8036–8040.

Corpet, F. (1988). Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res*, *16***(22):** 10881–10890.

Crack, J.C., Hengst, C.D. den, Jakimowicz, P., Subramanian, O., Johnson, M.K., Buttner, M.J., Thomson, A.J., and Le Brun, N.E. (2009). Characterization of [4Fe-4S]-containing and cluster-free forms of *Streptomyces* WhiD. *Biochemistry 48***(51):** 12252-12264.

Crack, J.C., Smith, L.J., Stapleton, M.R., Peck, J., Watmough, N.J., Buttner, M.J., Buxton, R.S., Green, J., Oganesyan, V.S., Thomson, A.J., Le Brun, N.E. (2010). Mechanistic insight into the nitrosylation of the [4Fe-4S] cluster of WhiB-like proteins. *J Am Chem Soc, 133***(4):** 1112–1121.

Crack, J.C., Svistunenko, D.A., Munnoch, J., Thomson, A.J., Hutchings, M.I., and Le Brun, N.E. (2016). Differentiated, promoter-specific response of [4Fe-4S] NsrR DNA binding to reaction with nitric oxide. *J Biol Chem*, *291***(16):**8663–8672.

van Crevel, R., Ottenhoff, T.H.M., and van der Meer, J.W.M. (2002). Innate Immunity to *Mycobacterium tuberculosis*. *Clin Microbiol Rev*, *15***(2):** 294–309.

Cruz-Ramos, H., Crack, J., Wu, G., Hughes, M.N., Scott, C., Thomson, A.J., Green, J., and Poole, R.K. (2002). NO sensing by FNR: Regulation of the *Escherichia coli* NO-detoxifying flavohaemoglobin, Hmp. *EMBO J*, *21***(3):** 3235–3244.

Davis, N., and Chater, K. (1992). The *Streptomyces coelicolor whiB* gene encodes a small transcription factor-like protein dispensable for growth but essential for sporulation. *Mol Gen Genet MGG,* **232(3):**351–358.

Davy, L., Osborne, M.J., Breton, J., Moore, G.R., Thomson, A.J., Bertinib, I., and Luchinat, C. (1995). Determination of the $[Fe_4S_4]Cys_4$ cluster geometry of *Desulfovibrio africanus* ferredoxin I by ^1H NMR spectroscopy. *FEBS Lett, 363*: 199–204.

Davy, S.L., Osborne, M.J., and Moore, G.R. (1998). Determination of the structure of oxidised *Desulfovibrio africanus* ferredoxin I by H-1 NMR spectroscopy and comparison of its solution structure. *J Mol Bio,* 277**(3):** 683-706.

DeLano, W. (2002). The PyMOL Molecular Graphics System. Version 1.5.0.4 Schrödinger, LLC.

den Hengst, C.D., and Buttner, M.J. (2008). Redox control in actinobacteria. *Biochim Biophys Acta*, *1780***(11):** 1201–1216.

Depamphilis, B. V, Averill, B.A., Que, L., and Holm, R.H. (1974). Synthetic analogs of the active sites of iron-sulfur proteins. VI. Spectral and redox characteristics of the tetranuclear clusters $[Fe_4S_4(SR)_4]^{2-}$. *J Am Chem Soc, 96**(13)*:** 4159–4167.

Ding, H., and Demple, B. (2000). Direct nitric oxide signal transduction via nitrosylation of iron-sulfur centers in the SoxR transcription activator. *Proc Natl Acad Sci U S A*, *97***(10):** 5146-5150.

Dombroski, A.J., Walter, W.A., and Gross Carol A (1993). Amino-terminal amino acids modulate it-factor DNA-binding activity. *Genes Dev, 7***(12A):** 2446-2455.

Dove, S.L., Darst, S. a, and Hochschild, A. (2003). Region 4 of sigma as a target for transcription regulation. *Mol Microbiol*, *48***(4):** 863–874.

Ducati, R.G., Ruffino-netto, A., and Basso, L.A. (2006). The resumption of consumption – A review on tuberculosis. *Mem Inst Oswaldo Cruz,101***(7):** 697-714.

Feng, L., Chen, Z., Wang, Z., Hu, Y., and Chen, S. (2016). Genome-wide characterization of monomeric transcriptional regulators in *Mycobacterium tuberculosis*. *Microbiology, 162***(5):** 889–897.

Fennelly, K.P., Jones-López, E.C., Ayakaka, I., Kim, S., Menyha, H., Kirenga, B., Muchwa, C., Joloba, M., Dryden-Peterson, S., Reilly, N., Okwera, A., Elliott, A.M., Smith, P.G., Mugerwa, R.D., Eisenach, K.D., Ellner, J.J. (2012). Variability of infectious aerosols produced during coughing by patients with pulmonary tuberculosis. *Am J Respir Crit Care Med, 186***(5):** 450-457

Fesik, S.W., Eaton, H.L., Olejniczak, E.T., Zuiderweg, E.R.P., McIntosh, L.P., and Dahlquist, F.W. (1990). 2D and 3D NMR spectroscopy employing ^{13}C-^{13}C magnetization transfer by isotropic mixing. Spin system identification in large proteins. *J Am Chem Soc, 112***(2)**: 886–888.

Firmani, M.A., and Riley, L.W. (2002). Reactive nitrogen intermediates have a bacteriostatic effect on *Mycobacterium tuberculosis in vitro*. *J Clin Microbiol, 40***(9):** 3162–3166.

Flynn, J.L., Scanga, C.A., Tanaka, K.E., Chan, J. (1998). Effects of aminoguanidine on latent murine tuberculosis. *J Immunol, 160***(4):** 1796–1803.

Flynn, J.L., and Chan, J. (2001). Tuberculosis : Latency and reactivation. *Infect Immun*, *69***(7):** 4195–4201.

Fortune, S.M., Jaeger, a, Sarracino, D. a, Chase, M.R., Sassetti, C.M., Sherman, D.R., Bloom, B.R., and Rubin, E.J. (2005). Mutually dependent secretion of proteins required for mycobacterial virulence. *Proc Natl Acad Sci USA*, *102***(30):**10676-10681.

Frigui, W., Bottai, D., Majlessi, L., Monot, M., Josselin, E., Brodin, P., Garnier, T., Gicquel, B., Martin, C., Leclerc, C., Cole, S.T., Brosch, R. (2008). Control of *M. tuberculosis* ESAT-6 secretion and specific T cell recognition by PhoP. *PLoS Pathog*, *4***(2):** e33.

Gao, L.-Y., Guo, S., McLaughlin, B., Morisaki, H., Engel, J.N., and Brown, E.J. (2004). A mycobacterial virulence gene cluster extending RD1 is required for cytolysis, bacterial spreading and ESAT-6 secretion. *Mol Microbiol*, *53***(6):** 1677–1693.

Garg, S., Alam, M.S., Bajpai, R., Kishan, K.R., and Agrawal, P. (2009). Redox biology of *Mycobacterium tuberculosis* H37Rv: protein-protein interaction between GlgB and WhiB1 involves exchange of thiol-disulfide. *BMC Biochem*. *10*:1.

Garg, S.K., Suhail Alam, M., Soni, V., Radha Kishan, K. V, and Agrawal, P. (2007a). Characterization of *Mycobacterium tuberculosis* WhiB1/Rv3219 as a protein disulfide reductase. *Protein Expr Purif, 52***(2):** 422–432.

Garg, S.K., Alam, M.S., Kishan, K.V.R., and Agrawal, P. (2007b). Expression and characterization of a-(1,4)-glucan branching enzyme Rv1326c of *Mycobacterium tuberculosis* H37Rv. *Protein Expr Purif*, *51***(2):**198–208.

Garton, N.J., Waddell, S.J., Sherratt, A.L., Lee, S.M., Smith, R.J., Senner, C., Hinds, J., Rajakumar, K., Adegbola, R. a, Besra, G.S., Butcher, P.D., and Barer, M.R.(2008). Cytological and transcript analyses reveal fat and lazy persister-like bacilli in tuberculous sputum. *PLoS Med, 5***(4):** e75.

Gaskell, A.A., Crack, J.C., Kelemen, G.H., Hutchings, M.I., and Le Brun, N.E. (2007). RsmA is an anti-sigma factor that modulates its activity through a [2Fe-2S] cluster cofactor. *J Biol Chem, 282**(43)*:** 31812–31820.

Gautam, U.S., Sikri, K., Vashist, A., Singh, V., and Tyagi, J.S. (2014). Essentiality of DevR/DosR interaction with SigA for the dormancy survival program in *Mycobacterium tuberculosis*. *J Bacterio*, *196***(4):** 790–799.

Gebhard, S., Hümpel, A., McLellan, A.D., and Cook, G.M. (2008). The alternative sigma factor SigF of *Mycobacterium smegmatis* is required for survival of heat shock, acidic pH and oxidative stress. *Microbiology, 154*: 2786–2795.

Geiman, D., Kaushal, D., and Ko, C., Tyagi. S., Manabe, Y.C., Schroeder, B.G., Fleischmann, R.D., Morrison, N.E., Converse, P.J., Chen, P., Bishai, W.R. (2004). Attenuation of late-stage disease in mice infected by the *Mycobacterium tuberculosis* mutant lacking the SigF alternate sigma factor and identification of SigF-dependent genes by microarray analysis. *Infect Immun*, *72***(3):**1733-17345.

Geiman, D.E., Raghunand, T.R., Agarwal, N., and Bishai, W.R. (2006). Differential gene expression in response to exposure to antimycobacterial

agents and other stress conditions among seven *Mycobacterium tuberculosis whiB*-like genes. Antimicrob Agents Chemother, *50***(8):** 2836–2841.

Gomez, J.E., and Bishai, W.R. (2000). *whmD* is an essential mycobacterial gene required for proper septation and cell division. *Proc Natl Acad Sci U S A, 97***(15):** 8554–8559.

Gomez, M., Doukhan, L., Nair, G., and Smith, I. (1998). *sigA* is an essential gene in *Mycobacterium smegmatis*. *Mol Microbiol, 29***(2):** 617–628.

Gonzalo Asensio, J., Maia, C., Ferrer, N.L., Barilone, N., Laval, F., Soto, C.Y., Winter, N., Daffé, M., Gicquel, B., Martín, C., Jackson, M. (2006). The virulence-associated two-component PhoP-PhoR system controls the biosynthesis of polyketide-derived lipids in *Mycobacterium tuberculosis*. *J Biol Chem*,*281***(3):** 1313–1316.

Gruber, T.M., and Gross, C.A. (2003). Multiple sigma subunits and the partitioning of bacterial transcription space. *Annu Rev Microbiol***,** *57*: 441–466.

Grzesiek, S., and Bax, A. (1992). Improved 3D triple-resonance NMR techniques applied to a 31 kDa protein. *J Magn Reson*, *96***(2):** 432–440.

Grzesiek, S., and Bax, A. (1993). Amino-acid type determination in the sequential assignment procedures of uniformly $^{13}C/^{15}N$-enriched proteins. *J Biomol NMR, 3***(2):**185–204.

Guinn, K.M., Hickey, M.J., Mathur, S.K., Zakel, K.L., Grotzke, J.E., Lewinsohn, D.M., Smith, S., and Sherman, D.R. (2004). Individual RD1-region genes are required for export of ESAT-6/CFP-10 and for virulence of *Mycobacterium tuberculosis*. *Mol Microbiol*, *51***(2):** 359–370.

Han, J.C., and Han, G.Y. (1994). A procedure for quantitative determination of tris(2-carboxyethyl) phosphine, an odorless reducing agent more stable and effective than dithiothreitol. *Anal Biochem*, *220***(1):** 5–10.

Hayes, C.S., Peng, Z.Y., and Setlow, P. (2000). Equilibrium and kinetic binding interactions between DNA and a group of novel, nonspecific DNA-binding proteins from spores of *Bacillus* and *Clostridium* species. *J Biol Chem, 275***(45):** 35040–35050.

He, H., Hovey, R., Kane, J., Singh, V., and Zahrt, T.C. (2006). MprAB Is a stress-responsive two-component system that directly regulates expression of sigma factors SigB and SigE in *Mycobacterium tuberculosis*. *J Bacteriol*, *188***(6):** 2134-2143.

Hinckley, G.T., and Frey, P.A. (2006). Cofactor-dependence in reduction potentials for $[4Fe\text{-}4S]^{2+/1+}$ in lysine 2,3-Aminomutase. *Biochemistry*, *45***(10)**:3219-25.

Holm, L., and Rosenström, P. (2010). Dali server: Conservation mapping in 3D. *Nucleic Acids Res*, *38*: 545–549.

Homolka, S., Niemann, S., Russell, D.G., and Rohde, K.H. (2010). Functional genetic diversity among *Mycobacterium tuberculosis* complex clinical isolates: delineation of conserved core and lineage-specific transcriptomes during intracellular survival. *PLoS Pathog. 6***(7):** e1000988.

Houben, E.N.G., Korotkov, K. V, and Bitter, W. (2014). Take five - Type VII secretion systems of Mycobacteria. *Biochim Biophys Acta*, *1843***(8):** 1707–1716.

Hubbard, R.E., and Haider, M. k (2001). Hydrogen bonds in proteins: role and strength. *eLS 1*: 1–6.

Hümpel, A., Gebhard, S., Cook, G.M., and Berney, M. (2010). The SigF regulon in *Mycobacterium smegmatis* reveals roles in adaptation to stationary phase, heat, and oxidative stress. *J Bacteriol*, *192***(10):** 2491–2502.

Ibrahim, S.A., Crack, J.C., Rolfe, M.D., Acuña, J.M.B., Thomson, A.J., Le Brun, N.E., Schobert, M., Stapleton, M.R., and Green, J. (2015). Three *Pseudomonas putida* FNR family proteins with different sensitivities to O_2. *J Biol Chem*, *290***(27):** 16812–16823.

Ishihama, A. (2000). Functional modulation of Escherichia coli RNA polymerase. *Annu Rev Microbiol, 54*: 499–518.

Jackett, P., Aber, V., and Lowrie, D. (1978). Virulence of *Mycobacterium tuberculosis* and susceptibility to peroxidative killing systems. *J Gen Microbiol, 107:* 273–278.

Jakimowicz, P., Cheesman, M.R., Bishai, W.R., Chater, K.F., Thomson, A.J., and Buttner, M.J. (2005). Evidence that the *Streptomyces* developmental protein WhiD, a member of the WhiB family, binds a [4Fe-4S] cluster. *J Bio Chem, 280***(9):** 8309–8315.

Johnson, K. a, Verhagen, M.F., Brereton, P.S., Adams, M.W., and Amster, I.J. (2000). Probing the stoichiometry and oxidation states of metal centers in iron-sulfur proteins using electrospray FTICR mass spectrometry. *Anal Chem*, *72***(7):** 1410–1418.

Karimova, G., Pidoux, J., Ullmann, A., and Ladant, D. (1998). A bacterial two-hybrid system based on a reconstituted signal transduction pathway. *Proc Natl Acad Sci USA*, *95***(10):**5752-5756.

Karls, R.K., Guarner, J., McMurray, D.N., Birkness, K.A., and Quinn, F.D. (2006). Examination of *Mycobacterium tuberculosis* sigma factor mutants using low-dose aerosol infection of guinea pigs suggests a role for SigC in pathogenesis. *Microbiology, 152*:1591–1600.

Kelly, L.A., Mezulis, S., Yates, C., Wass, M., and Sternberg, M. (2015). The Phyre2 web portal for protein modelling, prediction, and analysis. *Nat Protoc, 10***(6):** 845–858.

Kelly, S.M., Jess, T.J., and Price, N.C. (2005). How to study proteins by circular dichroism. *Biochim Biophys Acta*, *1751***(2):** 119–139.

Khoroshilova, N., Popescu, C., Munck, E., Beinert, H., and Kiley, P.J. (1997). Iron-sulfur cluster disassembly in the FNR protein of *Escherichia coli* by O2: [4Fe-4S] to [2Fe-2S] conversion with loss of biological activity. *Proc Natl Acad Sci*, *94***(12):** 6087–6092.

Kim, D.E., Chivian, D., and Baker, D. (2004). Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Res*, *32(Web Server issue):* 526–531.

Kindler, V., Sappino, A., Grau, G., Piguet, P., and Vassalli, P. (1989). The inducing role of tumor necrosis factor in the development of bactericidal granulomas during BCG infection. *Cell*, *56***(5):** 731–740.

Konar, M., Alam, M.S., Arora, C., and Agrawal, P. (2012). WhiB2/Rv3260c, a cell division-associated protein of *Mycobacterium tuberculosis* H37Rv, has properties of a chaperone. *FEBS J*, *279***(15):** 2781–2792.

Kumar, A., Toledo, J.C., Patel, R.P., Lancaster, J.R., and Steyn, A.J.C. (2007). *Mycobacterium tuberculosis* DosS is a redox sensor and DosT is a hypoxia sensor. *Proc Nat Acad Sci U S A, 104***(28):** 11568–11573.

Kumar, A., Farhana, A., Guidry, L., Saini, V., Hondalus, M., and Steyn, A.J.C. (2011). Redox homeostasis in mycobacteria: the key to tuberculosis control? *Expert Rev Mol Med, 13*, e39.

Kumar, S., Badireddy, S., Pal, K., Sharma, S., Arora, C., Garg, S.K., Alam, M.S., Agrawal, P., Anand, G.S., and Swaminathan, K. (2012). Interaction of *Mycobacterium tuberculosis* Rsha and Sigh is mediated by salt bridges. *PLoS One, 7***(8):** e43676

Kurthkoti, K., and Varshney, U. (2012). Distinct mechanisms of DNA repair in mycobacteria and their implications in attenuation of the pathogen growth. *Mech Ageing Dev*. *133***(4):** 138–146.

Kwan, A.H., Mobli, M., Gooley, P.R., King, G.F., and MacKay, J.P. (2011). Macromolecular NMR spectroscopy for the non-spectroscopist. *FEBS J, 278***(5):** 687–703.

Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, *227*: 680–685.

Larsson, C., Luna, B., Ammerman, N.C., Maiga, M., Agarwal, N., and Bishai, W.R. (2012). Gene expression of *Mycobacterium tuberculosis* putative transcription factors WhiB1-7 in redox environments. *PLoS One, 7***(7):** e37516.

Lee, E., Cho, Y., Kim, H., Ahn, B., and Roe, J. (2004). Regulation of σ^B by an anti- and an anti-anti-Sigma factor in *Streptomyces coelicolor* in response to osmotic stress. *J Bacteriol, 186***(24):** 8490–8498.

Leistikow, R.L., Morton, R. a, Bartek, I.L., Frimpong, I., Wagner, K., and Voskuil, M.I. (2010). The *Mycobacterium tuberculosis* DosR regulon assists in metabolic homeostasis and enables rapid recovery from nonrespiring dormancy. *J Bacteriol*. *192***(6):** 1662–1670.

Lewis, K., Liao, R., Guinn, K.M., Hickey, M.J., Smith, S., Behr, M.A., and Sherman, D.R. (2003). Deletion of RD1 from *Mycobacterium tuberculosis* mimics bacille Calmette-Guérin attenuation. *J Infect Dis, 187***(1)** 117–123.

Lonetto, M., Gribskov, M., and Gross, C.A. (1992). The sigma 70 family: sequence conservation and evolutionary relationships. *J Bacteriol*, *174***(12):** 3843–3849.

MacMicking, J.D., North, R.J., LaCourse, R., Mudgett, J.S., Shah, S.K., and Nathan, C.F. (1997). Identification of nitric oxide synthase as a protective locus against tuberculosis. *Proc Natl Acad Sci U S A*, *94***(10):** 5243–5248.

Manganelli, R., Dubnau, E., Tyagi, S., Kramer, F.R., and Smith, I. (1999). Differential expression of 10 sigma factor genes in *Mycobacterium tuberculosis*. Mol Microbiol, *31***(2):** 715–724.

Manganelli, R., Provvedi, R., Rodrigue, S., Beaucher, J., Gaudreau, and L., Smith, I. (2004). σ factors and global gene regulation in *Mycobacterium tuberculosis*. *J Bacteriol, 186***(4):** 895–902.

Marion, D. (2013). An introduction to biological NMR spectroscopy. *Mol Cell Proteomics, 12***(11):** 3006–3025.

Marion, D., Driscoll, P.C., Kay, L.E., Wingfield, P.T., Bax, A., Gronenborn, A.M., and Clore, G.M. (1989). Overcoming the overlap problem in the assignment of ^{1}H NMR spectra of larger proteins by use of three-dimensional heteronuclear ^{1}H-^{15}N Hartmann-Hahn-multiple quantum coherence and nuclear Overhauser-multiple quantum coherence spectroscopy: application to interleukin 1 beta. *Biochemistry, 28***(15):** 6150–6156.

Miller, J.H. (1972). Assay of β-galactosidase. In: Experiments in molecular genetics, Cold Spring Harbor, N.Y, pp. 352-355.

Morris, R.P., Nguyen, L., Gatfield, J., Visconti, K., Nguyen, K., Schnappinger, D., Ehrt, S., Liu, Y., Heifets, L., Pieters, J., Schoolnik, G., and Thompson, C.J. (2005). Ancestral antibiotic resistance in *Mycobacterium tuberculosis*. *Proc Natl Acad Sci U S A,102***(34):** 12200–12205.

Mossakowska, D.E., and Smith, R.A.G. (1997). Production and characterization of recombinant proteins for NMR structural studies. In: Protein NMR techniques, D.G. Reid, Totowa, NJ: Humana Press, pp. 325–335.

Niederweis, M., Wolschendorf, F., Mitra, A., and Neyrolles, O. (2015). Mycobacteria, metals, and the macrophage. *Immunol Rev, 264***(1):** 249–263.

O'Garra, A., Redford, P.S., McNab, F.W., Bloom, C.I., Wilkinson, R.J., and Berry, M.P.R. (2013). The immune response in tuberculosis. *Annu Rev Immunol*, 31:475-527.

Page, R., Peti, W., Wilson, I.A., Stevens, R.C., and Wüthrich, K. (2005). NMR screening and crystal quality of bacterially expressed prokaryotic and eukaryotic proteins in a structural genomics pipeline. *Proc Natl Acad Sci U S A, 102***(6):** 1901–1905.

Paget, M.S. (2015). Bacterial sigma factors and anti-sigma factors: Structure, function and distribution. *Biomolecules 5***(3):** 1245–1265.

Panaghie, G., Aiyar, S.E., Bobb, K.L., Hayward, R.S., and de Haseth, P.L. (2000). Aromatic amino acids in region 2.3 of Escherichia coli sigma 70 participate collectively in the formation of an RNA polymerase-promoter open complex. *J Mol Biol, 299***(5):** 1217–1230.

Pang, X., Samten, B., Cao, G., Wang, X., Tvinnereim, A.R., Chen, X.L., and Howard, S.T. (2013). MprAB regulates the espA operon in Mycobacterium tuberculosis and modulates ESX-1 function and host cytokine response. *J Bacteriol, 195***(1):** 66–75.

Pascal, S.M. (2008). NMR Primer: An HSQC-based Approach with Vector Animations. Chichester: IM Publications.

Peters, W., and Ernst, J.D. (2003). Mechanisms of cell recruitment in the immune response to *Mycobacterium tuberculosis*. *Microbes Infect, 5***(2):** 151–158.

Piccioli, M., and Turano, P. (2015). Transient iron coordination sites in proteins: Exploiting the dual nature of paramagnetic NMR. *Coord Chem Rev*, *284*: 313–328.

Dalton, D.K., Pitts-Meek, S., Keshav, S., Figari, I.S., Bradley, A., and Stewart, T.A. (1993). Multiple defects of immune cell function in mice with disrupted interferon-gamma genes. *Science*, *259***(5102):** 1739–1742.

Provvedi, R., Kocíncová, D., Donà, V., Euphrasie, D., Daffé, M., Etienne, G., Manganelli, R., and Reyrat, J.M. (2008). SigF controls carotenoid pigment production and affects transformation efficiency and hydrogen peroxide sensitivity in *Mycobacterium smegmatis. J Bacteriol, 190***(23):** 7859–7863.

Py, B., and Barras, F. (2010). Building Fe-S proteins: bacterial strategies. *Nat Rev Microbiol, 8***(6):** 436–446.

Pym, A.S., Brodin, P., Brosch, R., Huerre, M., and Cole, S.T. (2002). Loss of RD1 contributed to the attenuation of the live tuberculosis vaccines *Mycobacterium bovis* BCG and *Mycobacterium microti*. *Mol Microbiol, 46***(3):** 709–717.

Raghunand, T.R., and Bishai, W.R. (2006a). *Mycobacterium smegmatis whmD* and its homologue *Mycobacterium tuberculosis whiB2* are functionally equivalent. *Microbiology, 152*: 2735–2747.

Raghunand, T.R., and Bishai, W.R. (2006b). Mapping essential domains of *Mycobacterium smegmatis* WhmD: insights into WhiB structure and function. *J Bacteriol, 188***(19):** 6966–6976.

Raja, A. (2004). Immunology of tuberculosis. *Indian J Med Res*, *120**(4)*:** 213–232.

Rattle, H.W.E. (1995). An NMR Primer for Life Scientists. Partnership Press, pp.1-30.

Reid, D.G., MacLachlan, L.K., Edwards, A.J., Hubbard, J.A., and Sweeney, P.J. (1997). Introduction to the NMR of Proteins. In Protein NMR Techniques, D.G. Reid, Totowa, NJ: Humana Press, pp. 1–28.

Renshaw, P.S., Lightbody, K.L., Veverka, V., Muskett, F.W., Kelly, G., Frenkiel, T. a, Gordon, S. V, Hewinson, R.G., Burke, B., Norman, J., Williamson, R.A., Carr, M.D. (2005). Structure and function of the complex formed by the tuberculosis virulence factors CFP-10 and ESAT-6. *EMBO J, 24***(14):** 2491–2498.

Rickman, L., Scott, C., Hunt, D.M., Hutchinson, T., Menéndez, M.C., Whalan, R., Hinds, J., Colston, M.J., Green, J., Buxton, R.S. (2005). A member of the cAMP receptor protein family of transcription regulators in *Mycobacterium tuberculosis* is required for virulence in mice and controls transcription of the rpfA gene coding for a resuscitation promoting factor. *Mol Microbiol*, *56***(5):**1274-1286

Ross, W., Gosink, K., Salomon, J., Igarashi, K., Zou, C., Ishihama, A., Severinov, K., and Gourse, R. (1993). A third recognition element in bacterial promoters: DNA binding by the alpha subunit of RNA polymerase. *Science*, *262***(5138):**1407–1413.

Russell, D.G., Cardona, P., Kim, M., Allain, S., and Altare, F. (2010). Foamy macrophages and the progression of the human TB granuloma. *Nature immunology*,*10***(9):** 943-948.

Rybniker, J., Nowag, A., van Gumpel, E., Nissen, N., Robinson, N., Plum, G., and Hartmann, P. (2010). Insights into the function of the WhiB-like protein of mycobacteriophage TM4 – a transcriptional inhibitor of WhiB2. *Mol Microbiol, 77***(3):** 642–657.

Sachdeva, P., Misra, R., Tyagi, A.K., and Singh, Y. (2010). The sigma factors of *Mycobacterium tuberculosis*: Regulation of the regulators. *FEBS J*, *277***(3):** 605–626.

Samten, B., Wang, X., and Barnes, P. (2011). Immune regulatory activities of early antigenic target of 6-kD protein of *Mycobacterium tuberculosis* and

implications for tuberculosis vaccine design. *Tuberculosis (Edinb), 91S1*: S114–S118

Sanderson, A., Mitchell, J.E., Minchin, S.D., and Busby, S.J.W. (2003). Substitutions in the *Escherichia coli* RNA polymerase σ^{70} factor that affect recognition of extended -10 elements at promoters. *FEBS Lett, 544*: 199–205.

Serrano, P.N., Wang, H., Crack, J.C., Prior, C., Hutchings, M.I., Thomson, A.J., Kamali, S., Yoda, Y., Zhao, J., Hu, M.Y., Alp, E.E., Oganesyan, V.S., Le Brun, N.E., and Cramer, S.P. (2016). Nitrosylation of nitric-oxide-sensing regulatory proteins containing [4Fe-4S] clusters gives rise to multiple iron-nitrosyl complexes. *Angew Chemie Int Ed Engl, 55***(47):** 14575–14579.

Shen, Y., and Bax, A. (2013). Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J Biomol NMR, 56***(3):** 227–241.

Shiloh, M., Manzanillo, P., and Cox, J. (2008). *Mycobacterium tuberculosis* senses host-derived carbon monoxide during macrophage infection. *Cell Host Microbe*, *3***(5):** 323–330.

Singh, A., Guidry, L., Narasimhulu, K. V, Mai, D., Trombley, J., Redding, K.E., Giles, G.I., Lancaster, J.R., and Steyn, A.J.C. (2007). *Mycobacterium tuberculosis* WhiB3 responds to O_2 and nitric oxide via its [4Fe-4S] cluster and is essential for nutrient starvation survival. *Proc Natl Acad Sci U S A, 104***(28):** 11562–11567.

Singh, A., Crossman, D.K., Mai, D., Guidry, L., Voskuil, M.I., Renfrow, M.B., and Steyn, A.J.C. (2009). *Mycobacterium tuberculosis* WhiB3 maintains redox homeostasis by regulating virulence lipid anabolism to modulate macrophage response. *PLoS Pathog, 5***(8):** e1000545.

Smith, I. (2003). *Mycobacterium tuberculosis* pathogenesis and molecular determinants of virulence. *Clin Microbiol Rev*, *16***(3):**463-496.

Smith, L.J., Stapleton, M.R., Fullstone, G.J.M., Crack, J.C., Thomson, A.J., Le Brun, N.E., Hunt, D.M., Harvey, E., Adinolfi, S., Buxton, R.S., Green J. (2010). *Mycobacterium tuberculosis* WhiB1 is an essential DNA-binding protein with a nitric oxide-sensitive iron-sulfur cluster. *Biochem J*, *432***(3):** 417–427.

Smith, L.J., Stapleton, M.R., Buxton, R.S., and Green, J. (2012). Structure-function relationships of the *Mycobacterium tuberculosis* transcription factor WhiB1. *PLoS One, 7***(7):** e40407.

Solans, L., Aguiló, N., Samper, S., Pawlik, A., Frigui, W., Martín, C., Brosch, R., and Gonzalo-Asensio, J. (2014). A specific polymorphism in *Mycobacterium tuberculosis* H37Rv causes differential ESAT-6 expression and identifies WhiB6 as a novel ESX-1 component. *Infect Immun, 82***(8):** 3446–3456.

Soliveri, J.A., Gomez, J., Bishai, W.R., and Chater, K.F. (2000). Multiple paralogous genes related to the *Streptomyces coelicolor* developmental regulatory gene whiB are present in *Streptomyces* and other actinomycetes. *Microbiology*, *146*: 333–343.

Sritharan, M. (2016). Iron homeostasis in *Mycobacterium tuberculosis*: Mechanistic insights into siderophore-mediated iron uptake. *J Bacteriol, 198***(18):** 2399–2409.

Stapleton, M., Haq, I., Hunt, D.M., Arnvig, K.B., Artymiuk, P.J., Buxton, R.S., and Green, J. (2010). *Mycobacterium tuberculosis* cAMP receptor protein (Rv3676) differs from the *Escherichia coli* paradigm in its cAMP binding and DNA binding properties and transcription activation properties. *J Biol Chem, 285***(10):** 7016–7027.

Steyn, A.J.C., Collins, D.M., Hondalus, M.K., Jacobs, W.R., Kawakami, R.P., and Bloom, B.R. (2002). *Mycobacterium tuberculosis* WhiB3 interacts with RpoV to affect host survival but is dispensable for in vivo growth. *Proc Natl Acad Sci U S A*, *99***(5):** 3147–3152.

Sweeney, W. V, and Rabinowitz, J.C. (1980). Proteins containing 4Fe-4S: an overview. *Annu Rev Biochem, 49*: 139–161.

Tiso, M., and Schechter, A.N. (2015). Nitrate reduction to nitrite, nitric oxide and ammonia by gut bacteria under physiological conditions. *10***(3):** e0119712.

Troxler, H., Kuster, T., Rhyner, J.A., Gehrig, P., and Heizmann, C.W. (1999). Electrospray ionization mass spectrometry: Analysis of the Ca^{2+}-binding properties of human recombinant α-Parvalbumin and nine mutant proteins. *Anal Biochem*, *268***(1):** 64–71.

Tsai, M.C., Chakravarty, S., Zhu, G., Xu, J., Tanaka, K., Koch, C., Tufariello, J., Flynn, J., and Chan, J. (2006). Characterization of the tuberculous granuloma in murine and human lungs: cellular composition and relative tissue oxygen tension. *Cell Microbiol, 8***(2):**218–232.

Tucker, N.P., Hicks, M.G., Clarke, T.A., Crack, J.C., Chandra, G., Le Brun, N.E., Dixon, R., and Hutchings, M.I. (2008). The transcriptional repressor protein NsrR senses nitric oxide directly via a [2Fe-2S] cluster. *PLoS One, 3***(11):** e3623

Vassylyev, D.G., Sekine, S., Laptenko, O., Lee, J., Vassylyeva, M.N., Borukhov, S., and Yokoyama, S. (2002). Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 A resolution. *Nature, 417***(6890):** 712–719.

Veenstra, D., Johnson, L., Tomlinson, A.J., Craig, A., and Naylor, S. (1998). Zinc-induced conformational changes in the DNA-binding domain of the vitamin D receptor determined by electrospray ionization mass spectrometry. *J Am Soc Mass Spectrom, 9**(1):*** 8-14.

Vergne, I., Chua, J., Singh, S.B., and Deretic, V. (2004). Cell biology of *Mycobacterium tuberculosis* phagosome. *Annu Rev Cell Dev Biol, 20*: 367–394.

Volpe, E., Cappelli, G., Grassi, M., Martino, A., Serafino, A., Colizzi, V., Sanarico, N., and Mariani, F. (2006). Gene expression profiling of human macrophages at late time of infection with *Mycobacterium tuberculosis*. Immunology *118***(4):** 449–460.

Walters, S.B., Dubnau, E., Kolesnikova, I., Laval, F., Daffe, M., and Smith, I. (2006). The *Mycobacterium tuberculosis* PhoPR two-component system regulates genes essential for virulence and complex lipid biosynthesis. *Mol Microbiol, 60***(2):** 312–330.

Wang, X., Wang, H., and Xie, J. (2011). Genes and regulatory networks involved in persistence of *Mycobacterium tuberculosis*. *Sci China Life Sci, 54***(4):** 300–310.

Wayne, L., and Hayes, L. (1996). An in vitro model for sequential study of shiftdown of *Mycobacterium tuberculosis* through two stages of nonreplicating persistence. *Infect Immun, 64***(6):** 2062-2069.

Wayne, L.G., and Sohaskey, C.D. (2001). Nonreplicating persistence of mycobacterium tuberculosis. *Annu Rev Microbiol, 55*: 139–163.

WHO (2016). Global tuberculosis report 2016. WHO Press. Geneva Switz. 2016. WHO/HTM/TB/2016.13 214.

Williams, E.P., Lee, J.H., Bishai, W.R., Colantuoni, C., and Karakousis, P.C. (2007). *Mycobacterium tuberculosis* SigF regulates genes encoding cell wall-associated proteins and directly regulates the transcriptional regulatory gene *phoY1*. *J Bacteriol*, *189***(11):** 4234–4242.

Williamson, M. (2012). How proteins work. Taylor and Francis Group, pp.391-404

Wishart, D.S., Bigam, C.G., Holm, A., Hodges, R.S., and Sykes, B.D. (1995). ^{1}H, ^{13}C and ^{15}N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects. *J Biomol NMR, 5***(1):** 67–81.

Wösten, MM.. (1998). Eubacterial sigma-factors. *FEMS Microbiol Rev, 22***(3):**127–150.

Wu, S., Howard, S.T., Lakey, D.L., Kipnis, A., Samten, B., Safi, H., Gruppo, V., Wizel, B., Shams, H., Basaraba, R.J., Orme, I.M., Barnes, P.F. (2004). The principal sigma factor sigA mediates enhanced growth of *Mycobacterium tuberculosis* in vivo. *Mol Microbiol, 51***(6):** 1551–1562.

Yamazaki, T., Forman-kay, J.D., and J, L.E.K. (1993). Two-dimensional NMR experiments for correlating $^{13}C\beta$ and $^{1}H\delta/\varepsilon$ chemical shifts of aromatic residues

in ^{13}C-Labeled proteins via scalar couplings. *J Am Chem Soc, 115***(23):** 11054–11055.

Yang, C.S., Yuk, J.M., and Jo, E.K. (2009). The role of nitric oxide in mycobacterial infections. *Immune Netw, 9***(2):** 46–52.

Yu, Z., Zhang, N., Wu, X., Gao, Z., Zhu, Q., and Kong, F. (1993). The Production and photodissociation of iron-sulphur clusters. *J Chem Phys, 99*: 1765.

Yukl, E., Elbaz, M., Nakano, M., and Moënne-Loccoz, P. (2008). Transcription factor NsrR from *Bacillus subtilis* senses nitric oxide with a 4Fe−4S cluster. *Biochemistry, 47***(49):** 13084–13092.

Zhang, B., Crack, J.C., Subramanian, S., Green, J., Thomson, A. J., Le Brun, N.E., and Johnson, M.K. (2012). Reversible cycling between cysteine persulfide-ligated [2Fe-2S] and cysteine-ligated [4Fe-4S] clusters in the FNR regulatory protein. *Proc Natl Acad Sci, 109***(39):** 15734–15739.

Zhang, Y.Z., Paterson, Y., and Roder, H. (1995). Rapid amide proton exchange rates in peptides and proteins measured by solvent quenching and two-dimensional NMR. *Protein Sci, 4***(4):** 804–814.

Zheng, L., White, R.H., Cash, V.L., Jack, R.F., and Dean, D.R. (1993). Cysteine desulfurase activity indicates a role for NIFS in metallocluster biosynthesis. *Proc Nat Acad Sci USA, 90***(7):** 2754–2758.

Appendix

Appendix 1

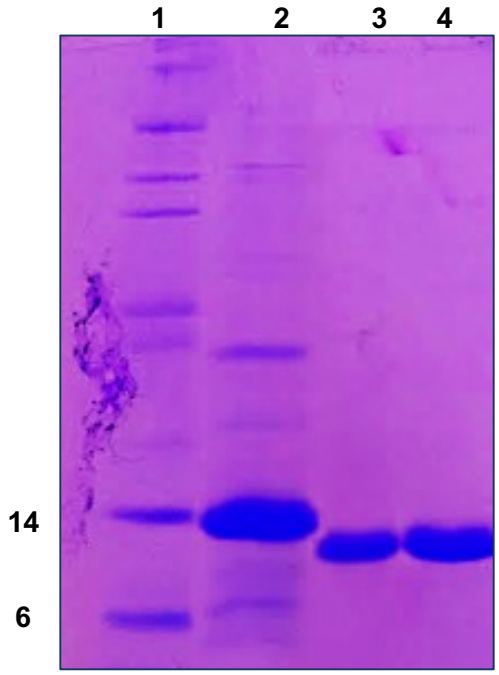


Figure S.1: Coomassie blue stained SDS-PAGE gel of WhiB1^{Ec} treated with Thrombin protease. Lane 1, protein ladder (sizes in kDa are indicated); lane 2, apo-WhiB1^{Ec} treated with 0% protease; lanes 3 and 4, apo-WhiB1^{Ec} digested with Thrombin protease; where 10 units of protease was used for each 100 nmole of WhiB1^{Ec} with overnight incubation at 4°C. Digestion was carried out in 25 mM NaH_2PO4 (pH 7.5) with 0.1 M NaCl.

Appendix 2

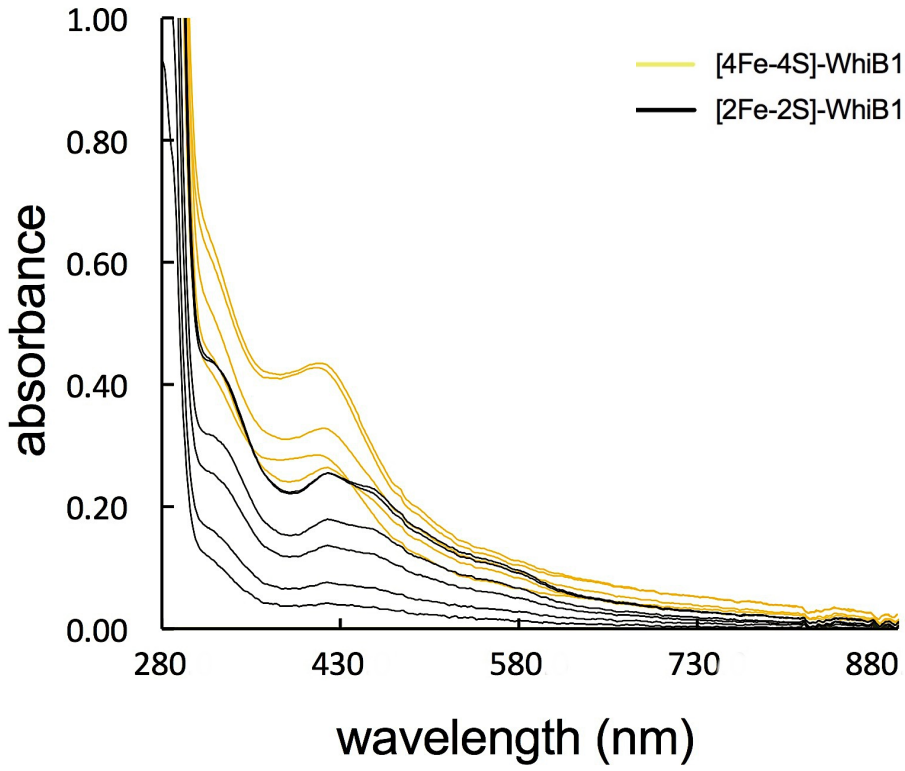


Figure S.2: UV-visible spectra of WhiB1^{Ec} eluted fractions from nickel-charged column. WhiB1^{Ec} overproduced in growth medium supplemented with 0.5 mg/l ferric chloride shows only two forms of protein; [4Fe-4S]-WhiB1 (yellow traces) and [2Fe-2S]-WhiB1 (black traces) after protein purification. Each trace represents an eluted fraction of the protein, each in a different concentration.

Appendix 3

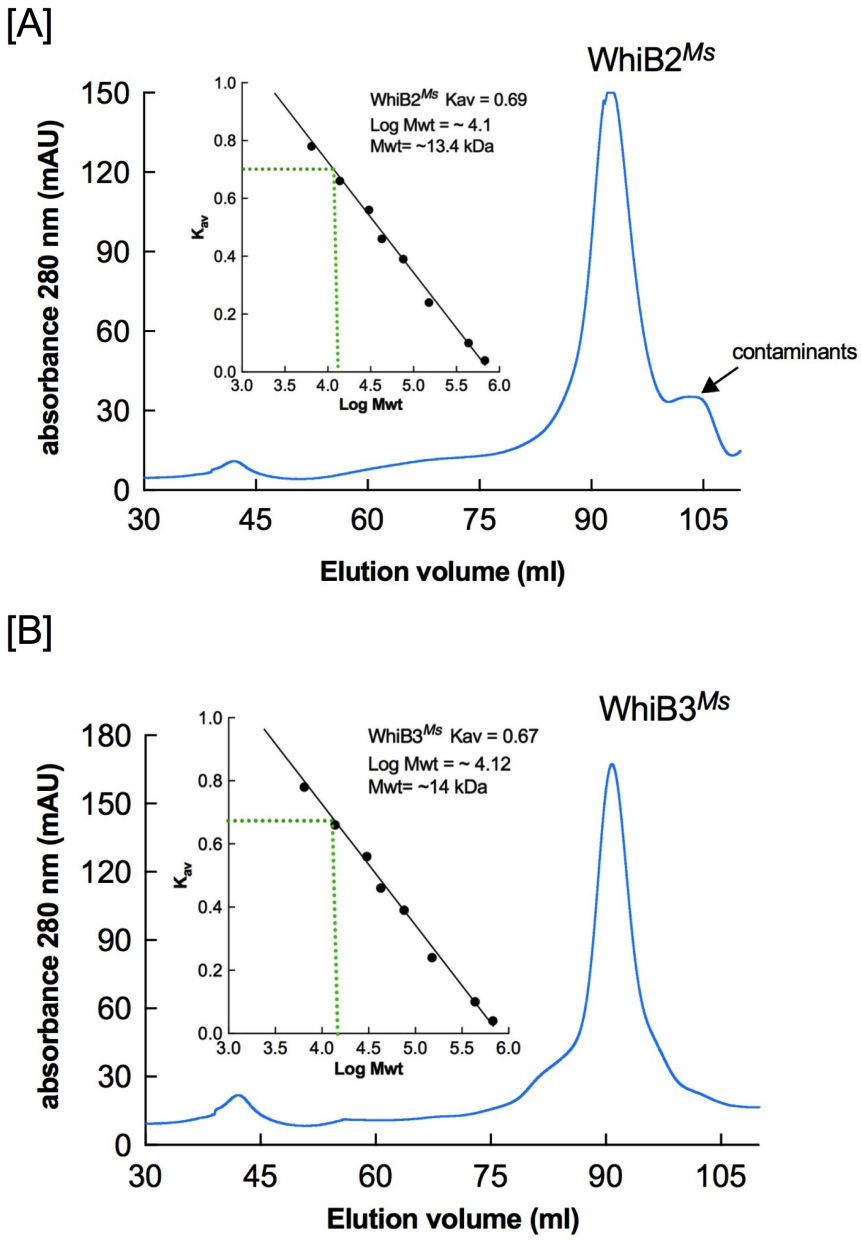


Figure S.3: Size-exclusion chromatography profile of holo-WhiB2^{Ms} (A) and holo-WhiB3^{Ms} (B). HiLoad 16/600 Superdex column pre-equilibrated with anaerobic buffer of 50 mM Tris-HCl (pH 8.0) containing 0.5 M NaCl and 10 mM DTT was used for analysis the protein oligomerization state. Holo-WhiB2 and holo-WhiB3 were eluted as monomers at estimated molecular weights of ~13 kDa and ~14 kDa respectively.

Appendix 4

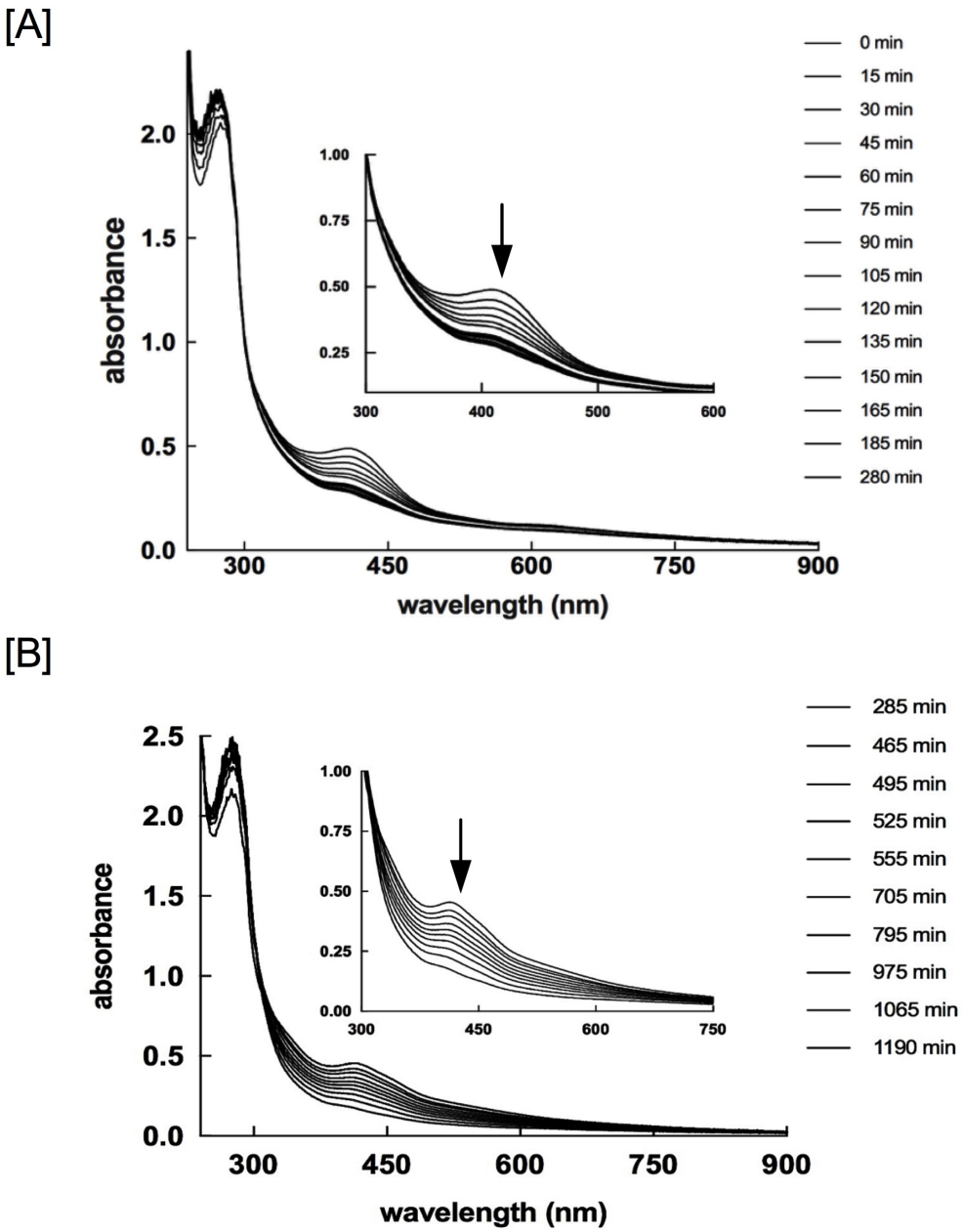


Figure S.4: Reaction of the WhiB2^{Ms} [4Fe-4S] with O_2. Oxygen-saturated buffer (50 mM Tris pH 7.4, 0.5 M NaCl, and 5% glycerol) with final concentration of 110 µM of O_2 was introduced to reconstituted [4Fe-4S] WhiB2^{Ms} (30 µM cluster) and the UV-visible spectra were obtained at the indicated times (0 to 280 min in figure **[A],** and 285 to 1190 min in figure **[B]**) at 25ºC. Arrows indicate the direction of change within period of incubation.

Appendix 5

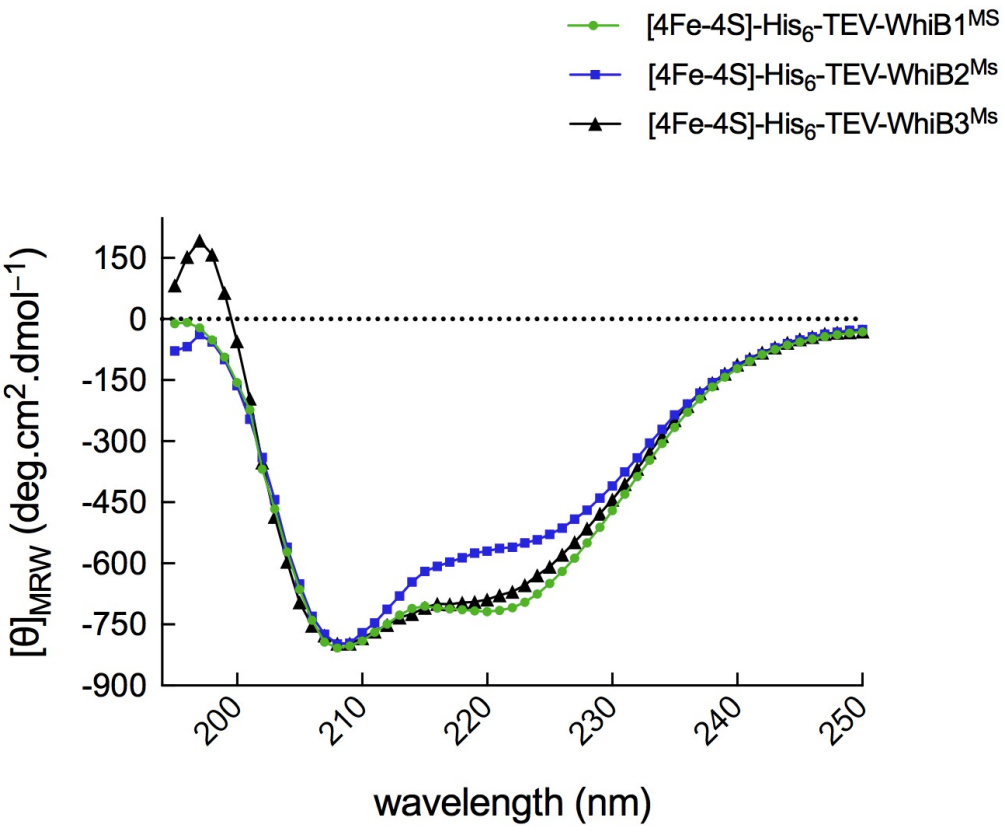


Figure S.5: The far UV CD-spectra of [4Fe-4S] WhiB1, 2, and 3^{Ms}. The spectra were recorded using 30 μM of each sample in 20 mM NaH_2PO_4, pH 7.4 containing 0.1 M NaCl at 25°C. The trace of WhiB2^{Ms} (■) shows a feature at 208 nm, and a shoulder at 215-225, while both traces of WhiB3^{Ms} (▲) and WhiB1^{Ms} (●) show two features at 208 and 222 nm. WhiB3 ^{Ms} shows another feature at ~198 nm.

Appendix 6

Table S.1: Nuclei chemical shifts of WhiB1 residues obtained from backbone and side-chain assignments

Residue	HN	N	Cα	Cβ	C	Others
-10 A	7.88	123.53	52.67	19.1	179.37	(Hβ,1.55)
-9 G	7.71	105.11	46.39		174.29	($H\alpha_1$, 3.89), ($H\alpha_2$, 3.78)
-8 E	8.38	120.08	56.53	31.55	176.46	(Hα, 4.24), ($H\beta_1$, 2.02), ($H\beta_2$, 1.91), (Cγ, 32.33), ($H\gamma_1$, 2.12), ($H\gamma_2$, 2.19)
-7 N	8.56	119.74	53.57	38.12	175.02	($H\beta_1$, 2.71), ($H\beta_2$, 2.71)
-6 L	7.9	121.24	55.43	42.19	176.71	(Hα, 4.10), ($H\beta_1$, 1.14), ($H\beta_2$, 1.30), (Cγ, 26.7), (Hγ, 1.28), ($C\delta_1$, 23.40), ($H\delta_1$, 0.60), ($C\delta_2$, 24.70), ($H\delta_2$, 0.69)
-5 Y	7.99	118.93	57.79	38.2	175.57	(Hα, 4.43), ($H\beta_1$, 2.90), ($H\beta_2$, 2.75), (Cδ, 133.0), ($H\delta_1$, 6.95), (Cε, 118.1), (Hε, 6.74)
-4 F	7.85	121.19	57.64	39.54	175.58	(Hα, 4.55), ($H\beta_1$, 3.07), ($H\beta_2$, 3.07), (Cδ, 131.3), (Hδ, 7.24)
-3 Q	8.56	121.58	55.9	28.87	176.2	(Hα, 4.56), ($H\beta_1$, 2.14), ($H\beta_2$, 1.94), ($H\gamma_1$, 2.28), ($H\gamma_2$, 2.28)
-2 G	8.11	109.72	45.31		174.25	($H\alpha_1$, 3.84), ($H\alpha_2$, 4.07)
-1 A	8.31	123.59	52.95	18.96	177.76	(Hα, 4.23), (Hβ, 1.31)
1 M	8.29	117.38	54.93	32	176.38	(Hα, 4.36), ($H\beta_1$, 2.31), ($H\beta_2$, 2.31)
2 D	8.12	121.99	55.59	41.96	-	(Hα, 4.42), ($H\beta_1$, 2.72), ($H\beta_2$, 2.90)
3 W	7.83	121.69	58.14	27.83	177.45	(Hα, 4.35), ($H\beta_1$, 3.18), ($H\beta_2$, 3.13), ($C\delta_1$, 127.95), ($H\delta_1$, 7.45), ($N\varepsilon_1$, 128.78), ($H\varepsilon_1$, 9.36), ($C\zeta_2$, 114.5), ($H\zeta_2$, 6.82), ($C\eta_2$, 126.0), ($H\eta_2$, 7.38)
4 R	7.11	120.49	59.52	28.59	178.64	(Hα, 5.01), ($H\beta_1$, 3.10), ($H\beta_2$, 3.19)
5 H	7.69	116.33	58.28	29.84	176.14	(Hα, 4.25), ($H\beta_1$, 3.08), ($H\beta_2$, 3.08), ($C\delta_2$, 118.5), ($H\delta_2$, 6.35), ($C\varepsilon_1$, 137.1), ($H\varepsilon_1$, 7.91)
6 K	7.52	116.9	55.33	33.48	175.7	(Hα, 4.28), ($H\beta_1$, 1.84)
7 A	6.92	122	55.98	-	175.73	(Hα, 4.32)
8 V	-	-	-	-	-	-
9 C	-	-	-	-	174.89	-
10 R	6.68	123.24	58.65	30.12	176.79	(Hα, 4.21), ($H\beta_1$, 1.84), ($H\beta_2$, 1.68)
11 D	7.48	115.92	54.55	40.93	175.83	(Hα, 4.87), ($H\beta_1$, 2.98), ($H\beta_2$, 2.58)
12 E	7.57	118.94	53.98	30.72	174.65	(Hα, 4.47), ($H\beta_1$, 1.96), ($H\beta_2$, 1.96), (Cγ, 36.83), ($H\gamma_1$, 2.32), ($H\gamma_1$, 2.19)
13 D	8.88	122.52	51.53	41.3	175.49	(Hα, 4.47), ($H\beta_1$, 2.86), ($H\beta_2$, 2.63)
14 P			64.66	32.641	178.759	-
15 E	8.16	112.38	58.23	28.55	179.43	(Hα, 3.95), ($H\beta_1$, 1.95), ($H\beta_2$, 2.06), ($H\gamma_1$, 2.35), ($H\gamma_2$, 2.46)
16 L	7.95	121.42	57.73	41.29	176.84	(Hα, 3.78), ($H\beta_1$, 1.47), ($H\beta_2$, 1.57), (Cγ, 26.4), ($C\delta_1$, 25.65), ($H\delta_1$, 0.54), ($C\delta_2$, 22.75), ($H\delta_2$, 0.76)
17 F	7.03	109.84	58.47	37.66	172.69	(Hα, 3.41), ($H\beta_1$, 1.91), ($H\beta_2$, 2.59), (Cδ, 133.4), (Hδ, 7.21)
18 F	7.29	115.93	56.46	39.57	-	(Hα, 4.73), ($H\beta_1$, 3.17), (Cδ, 131.7), (Hδ, 7.13)
19 P		-	63.006	31.484	176.957	(Hα, 4.49), ($H\beta_1$, 2.26), ($H\beta_2$, 2.26), (Cγ, 27.47),

						($H\gamma_1$, 2.0), ($C\delta_1$, 50.78), ($H\delta_1$, 3.95)
20 V	8.77	122.91	62.13	32.58	175.94	($H\alpha$, 4.23), ($H\beta$, 2.04), ($C\gamma_1$, 21.057), ($H\gamma_1$, 0.953), ($C\gamma_2$, 20.845), ($H\gamma_2$, 1.04)
21 G	8.56	114.08	44.47		173.02	($H\alpha_1$, 4.28), ($H\alpha_2$, 3.79)
22 N	8.27	115.46	52.03	38.917	174.53	($H\alpha$, 4.78)
23 S	7.92	113.45	57.44	65	174.06	($H\alpha_1$, 4.60), ($H\beta$, 3.99), ($H\beta_2$,4.06)
24 G	9.07	109.51	46.96		175.17	($H\alpha_1$, 4.29), ($H\alpha_2$, 3.82)
25 P		-	64.879	31.84	178.59	-
26 A	7.55	121.38	54.96	18.27	179.49	($H\alpha_1$, 4.13), ($H\beta_1$, 1.31)
27 L	7.91	116.88	57.81	40.75	180.53	($H\alpha$, 3.99), ($H\beta_1$, 1.82), ($H\beta_2$, 1.50), ($C\gamma$, 26.0), ($C\delta_1$, 25.0), ($H\delta_1$, 0.94), ($C\delta_2$, 22.6), ($H\delta_2$, 0.85)
28 A	7.77	122.98	54.53	17.61	179.13	($H\alpha$, 4.16), ($H\beta_1$, 1.47)
29 Q	7.65	120.1	59.06	-	178.13	-
30 I	8.36	119.2	66.12	38.19	177.29	($H\alpha$, 3.34), ($H\beta$, 1.91), ($C\gamma_1$, 27.0), ($H\gamma_{11}$, 0.75), ($H\gamma_{12}$, 0.75), ($C\gamma_2$, 17.79), ($H\gamma_2$, 0.851), ($C\delta_1$, 13.89), ($H\delta_1$, 0.935
31 A	7.5	120.44	55.37	17.48	180.42	(3.98), ($H\beta$, 1.495)
32 D	7.95	120.02	57.34	39.83	179.04	($H\alpha$, 4.38), ($H\beta_1$, 2.78), ($H\beta_2$, 2.51)
33 A	8.41	125.14	55.25	17.28	179.43	($H\alpha$, 3.48), ($H\beta$, 0.93)
34 K	8.37	114.98	61.19	32.71	178.73	($H\alpha$, 3.63)
35 L	7.59	118.4	57.86	41.76	179.91	($H\alpha$, 4.09), ($H\beta_1$, 1.56), ($H\beta_2$, 1.98), ($H\gamma$, 1.94), ($C\delta_1$, 25.2), ($H\delta_1$, 0.97), ($C\delta_2$, 23.1), ($H\delta_2$, 0.91)
36 V	7.49	121.03	66.34	30.99	178.08	-
37 C	-	-	-	-	177.51	-
38 N	8.4	116.14	55.18	-	176.65	-
39 R	7.87	116.93	55.04	30.15	175.93	-
40 C	7.97	132.64	57.28	-	-	-
41 P	-	-	-	-	-	-
42 V	-	-	-	-	-	-
43 T	-	-	63.0	68.468	177.483	($H\alpha$, 3.59), ($H\beta$, 4.25), ($C\gamma_2$, 22.0), ($H\gamma_2$, 1.32)
44 T	8.29	118.45	66.75	68.17	176.76	($H\alpha$, 3.83), ($H\beta$, 3.96), ($C\gamma_2$, 21.56), ($H\gamma_2$, 1.13
45 E	9.24	126.24	60.92	-	179.59	-
46 C	7.82	130.22	66.2	-	175.09	-
47 L	8.29	117.3	57.94	41.57	178.51	($H\alpha$, 3.88), ($C\gamma$, 26.77), ($H\gamma$, 1.92), ($C\delta_1$, 22.8), ($H\delta_1$, 0.97), ($C\delta_2$, 25.9), ($H\delta_2$, 1.0
48 S	8.12	111.46	61.83	62.64	176.26	($H\alpha$, 3.98), ($H\beta_1$ 3.89), ($H\beta_2$, 3.81)
49 W	8.24	122.36	62.21	29.07	179.61	($H\alpha$, 4.19), ($H\beta_1$, 3.81), ($H\beta_2$, 3.52), ($C\delta_1$, 127.2), ($H\delta_1$, 7.47), ($N\varepsilon_1$, 130.34), ($H\varepsilon_1$, 10.49), ($C\zeta_2$, 115.3), ($H\zeta_2$, 7.55), ($C\eta_2$, 124.7), ($H\eta_2$, 7.22)
50 A	8.52	123.01	55.08	17.18	180.23	($H\alpha$, 3.09), ($H\beta$, 1.07)
51 L	8.46	117.02	57.29	41.81	180.24	($H\alpha$, 3.99), ($H\beta_1$, 1.83), ($C\gamma$, 26.6), ($H\gamma$, 1.80), ($C\delta_1$, 26.0), ($H\delta_1$, 0.80), ($C\delta_2$, 22.8), ($H\delta_2$, 0.84)
52 N	8.57	117.1	55.51	38.65	177.23	($H\alpha$, 4.41), ($H\beta_1$, 2.78), ($H\beta_2$, 2.78
53 T	7.6	107.19	61.46	68.99	175.1	($H\alpha$, 3.84), ($H\beta$, 4.18), ($C\gamma_2$, 19.08), ($H\gamma_2$, 0.14
54 G	7.08	109.97	47.1		175.37	($H\alpha_1$, 3.72), ($H\alpha_2$, 3.72
55 Q	7.94	116.72	53.09	-	175.35	($H\alpha$, 4.28)

56 D	8.21	127.18	55.24	41.46	175.36	(Hα, 4.59)
57 S	7.94	112.52	56.98	65.79	175.05	(Hα, 5.0), ($H\beta_1$, 4.04), ($H\beta_2$, 4.04)
58 G	9.23	111.28	45.31		173.89	($H\alpha_1$, 4.02), ($H\alpha_2$, 3.66)
59 V	8.99	123.16	63.04	32.26	173.58	(Hα, 3.62), (Hβ, 2.27), ($C\gamma_1$, 19.29), ($H\gamma_1$, 0.972), ($C\gamma_2$, 20.63), ($H\gamma_2$, 1.025)
60 W	7.99	123.97	61.06	33.17	177.55	(Hα, 4.36), (Cβ, 33.10), ($C\delta_1$, 127.7), ($H\delta_1$, 6.97), ($N\varepsilon_1$, 130.24), ($H\varepsilon_1$, 10.39), ($C\zeta_2$, 114.3), ($H\zeta_2$, 7.91), ($H\eta_2$, 124.3), ($H\eta_2$, 7.46)
61 G	-	-	-	-	-	-
62 G	-	-	48.17	-	173.17	-
63 M	8.18	117.5	54.7	37.53	176.01	(Hα, 5.15), ($H\beta_1$, 2.06), ($H\beta_2$, 2.06), (Cγ, 33.57), ($H\gamma_1$, 2.56), ($H\gamma_2$, 2.66
64 S	9.73	120.12	57.22	65.43	174.86	(Hα, 4.44), (Hβ1, 2.81), (Hβ2, 2.81
65 E	9.49	121.75	59.76	28.52	179.49	(Hα, 4.01)
66 D	8.48	120.13	57.57	-	179.21	-
67 E	7.77	121.4	59.04	-	180.26	-
68 R	8.3	118.11	60.62	29.25	178.13	(Hα, 4.087), ($H\beta_1$, 2.68), ($H\beta_2$, 2.68)
69 R	7.84	119.61	60.05	29.96	178.92	(Hα, 3.98), ($H\beta_1$, 1.99), ($H\beta_2$, 1.89)
70 A	7.66	120.48	54.54	17.88	179.76	(Hα, 4.13), (Hβ, 1.50)
71 L	7.67	120.53	57.58	42.13	178.54	(Hα, 4.10), ($H\beta_1$, 1.82), ($H\beta_2$, 1.74)
72 K	8.08	119.13	59.03	32.24	178.82	(Hα, 4.03), ($H\beta_1$, 1.88), ($H\beta_2$, 1.88
73 R	7.85	119.02	58.14	30.21	178.01	(Hα, 4.18), ($H\beta_1$, 1.89), ($H\beta_2$, 1.78), (Cγ, 26.71), ($H\gamma_1$, 1.62), ($H\gamma_2$, 1.71), (Cδ, 43.28), ($H\delta_1$, 3.2)
74 R	8.06	120.47	58.14	30.09	177.89	(Hα, 4.15), ($H\beta_1$, 1.93), ($H\beta_2$, 1.93)
75 N	8.39	118.75	53.85	38.28	176.01	(Hα, 4.60), ($H\beta_1$, 2.83), ($H\beta_2$, 2.68)
76 A	7.96	123.2	53.34	18.73	178.45	(Hα, 4.27), (Hβ, 1.47)
77 R	8.08	119.03	56.77	30.41	177.15	(Hα, 4.31), (Hβ1, 1.81), (Hβ2, 1.90), (Cγ, 27.19), (Hγ1, 1.75), (Hγ2, 1.64), (Cδ, 43.3), (Hδ1, 3.34
78 T	8.06	114.81	62.48	69.53	174.82	(Hα, 4.27), (Hβ, 4.25), ($C\gamma_2$, 21.55), ($H\gamma_2$, 1.23)
79 K	8.22	123.63	56.53	32.68	176.46	(Hα, 4.26), ($H\beta_1$, 1.81), ($H\beta_2$, 1.83), (Cγ, 24.67), ($H\gamma_1$, 1.42), (Cδ, 28.80), ($H\delta_1$, 1.66), ($H\delta_2$, 1.75), (Cε, 41.92), ($H\varepsilon_1$, 2.92), ($H\varepsilon_2$, 2.97
80 A	8.23	124.87	52.57	18.99	177.81	(Hα, 4.3), (Hβ, 1.40)
81 R	8.33	120.33	56.19	30.59	176.57	(Hα, 4.39), ($H\beta_1$, 1.78), ($H\beta_2$, 1.88), (Cγ, 27.11), ($H\gamma_1$, 1.62), ($H\gamma_2$, 1.67), (Cδ, 43.22), ($H\delta_1$, 2.84), ($H\delta_2$, 3.19)
82 T	8.21	114.6	61.8	69.8	175.06	(Hα, 4.39), (Hβ, 4.26), ($C\gamma_2$, 21.43), ($H\gamma_2$, 1.20
83 G	8.44	111.73	45.35		173.26	($H\alpha_1$, 3.99), ($H\alpha_1$, 3.99)
84 V	7.63	122.73	63.53	32.81	181.19	(Hα, 4.08), (Hβ, 2.11), (Cγ1, 19.8), (Hγ1, 0.86), (Cγ2, 21.5), (Hγ2, 0.90

Appendix 7

Table S.2: WhiB1 torsion backbone angles predicted by TALOS-N. The table shows the degree of PHI (φ) and PSI (ψ) dihedral angles predicted based on chemical shifts obtained by NMR backbone assignment; the estimated standard deviation of errors in the angles (Δφ, Δψ); the "COUNT" represents database heptapeptide number used to predict the torsion angles; and finally, the "Class", where, "Strong/Generous" means the high consensus due to majority database matches and the prediction is more likely to be very good; "Warn", means no database matches consensus; "Dyn" the prediction is more likely to be dynamic conformation.

Residue	φ	ψ	Δφ	Δψ	Count	Class
2D	-67.163	145.072	6.761	9.981	25	Strong
3W	-83.634	140.048	15.297	16.871	25	Strong
4R	-68.541	137.576	13.423	13.733	10	Generous
5H	-62.365	-28.682	5.62	8.54	25	Strong
6K	-82.985	-7.111	12.726	13.197	9	Warn
10R	-61.532	-31.32	6.485	7.755	25	Strong
11D	-98.33	6.665	7.85	13.816	25	Strong
12E	-95.208	154.732	16.446	14.963	25	Strong
13D	-64.667	134.794	8.351	14.279	25	Strong
14P	-56.543	-34.015	4.111	7.396	25	Strong
15E	-63.258	-24.741	4.891	6.815	25	Strong
16L	-67.512	-27.501	6.957	9.51	25	Strong
17F	-92.41	-1.816	7.901	10.569	25	Strong
18F	-77.373	140.37	13.526	12.387	25	Strong
19P	-66.881	144.297	8.304	8.929	25	Strong
20V	-77.012	134.771	12.13	10.094	10	Generous
21G	-110.714	169.081	35.11	19.994	10	Generous
22N	-108.007	0.139	12.513	14.996	8	Warn
23S	-109.714	164.529	41.712	7.46	25	Warn
24G	-56.344	-50.805	5.615	5.659	25	Strong
25P	-59.749	-39.449	6.534	5.517	25	Strong
26A	-67.72	-39.711	3.955	4.977	25	Strong
27L	-68.355	-35.106	3.289	6.238	25	Strong
28A	-66.813	-38.221	3.42	3.658	25	Strong
29Q	-65.274	-40.118	3.578	4.913	25	Strong
30I	-64.675	-44.498	3.533	2.999	25	Strong
31A	-64.6	-38.567	2.311	4.205	25	Strong
32D	-65.499	-41.474	3.579	4.823	25	Strong
33A	-65.785	-40.199	2.93	3.837	25	Strong
34K	-63.576	-40.157	4.099	5.162	25	Strong
35L	-64.305	-37.991	5.563	4.237	25	Strong
36V	-66.705	-38.537	4.495	5.248	25	Strong
37C	-67.315	-29.506	5.221	5.475	25	Strong

38N	-67.227	-25.727	7.085	6.675	25	Strong
39R	-93.746	0.842	9.492	9.322	10	Generous
44T	-66.168	-36.064	4.243	4.125	25	Strong
45E	-66.752	-40.311	3.827	4.435	25	Strong
46C	-66.176	-40.159	3.259	4.314	25	Strong
47L	-64.344	-42.154	5.303	4.949	25	Strong
48S	-64.939	-38.714	2.77	4.003	25	Strong
49W	-64.271	-44.252	4.116	2.743	25	Strong
50A	-64.302	-40.738	3.193	4.757	25	Strong
51L	-66.053	-36.572	4.12	4.306	25	Strong
52N	-67.521	-26.851	4.204	5.798	25	Strong
53T	-97.101	5.272	8.514	4.848	25	Strong
54G	69.888	29.019	6.676	6.262	25	Strong
55Q	-99.391	143.282	11.194	10.773	10	Generous
56D	-72.115	145.941	8.249	10.579	7	Warn
57S	-135.314	156.645	17.632	11.164	25	Strong
58G	149.122	-169.074	49.307	20.959	6	Warn
59V	-86.516	138.839	20.181	11.259	6	Warn
60W	-92.792	133.544	22.887	13.051	5	Warn
63M	-129.527	151.251	9.424	13.23	25	Strong
64S	-74.286	161.593	7.684	6.1	25	Strong
65E	-59.657	-37.605	5.963	5.241	25	Strong
66D	-67.226	-38.18	3.517	5.282	25	Strong
67E	-67.863	-40.856	2.899	3.745	25	Strong
68R	-64.254	-40.243	2.715	3.835	25	Strong
69R	-64.41	-43.251	2.276	2.766	25	Strong
70A	-65.372	-37.491	4.622	4.683	25	Strong
71L	-68.327	-42.602	4.387	5.421	25	Strong
72K	-67.397	-37.221	3.878	6.6	25	Strong
73R	-68.491	-37.679	4.529	5.54	25	Strong
74R	-68.135	-35.729	4.653	7.484	25	Strong
75N	-68.281	-28.907	4.939	8.153	25	Strong
76A	-68.565	-21.528	6.94	9.761	25	Dyn
77R	-84.819	-10.76	10.795	10.593	25	Dyn
78T	-75.841	141.769	18.821	12.483	10	Dyn
79K	-86.407	137.633	26.621	17.783	5	Dyn
80A	-104.034	146.234	32.917	25.19	4	Dyn
81R	-78.204	-15.997	19.139	22.194	10	Dyn
82T	-79.812	-9.826	20.33	44.536	8	Dyn
83G	-80.242	-17.757	17.311	24.304	4	Dyn