CHARACTERIZATION OF *Mycobacterium tuberculosis* TCA CYCLE ENZYMES ISOCITRATE DEHYDROGENASE AND ACONITASE

Thesis submitted to

Department of Biochemistry
School of Life Sciences
University of Hyderabad
Hyderabad

For the degree of

Doctor of Philosophy

SHARMISTHA BANERJEE

Centre for DNA Fingerprinting and Diagnostics
Hyderabad
2004

Registration Number: 2LKBPH03

CERTIFICATE

This is to certify that this thesis entitled "Characterization of *Mycobacterium tuberculosis* TCA cycle enzymes isocitrate dehydrogenase and aconitase" comprises the work done by *Sharmistha Banerjee* under my guidance at Centre for DNA Fingerprinting and Diagnostics. The work is original and has not been submitted in part or full for any other degree or diploma of any other University.

Dean, School of Life Sciences
University of Hyderabad

Dean
School of Life Sciences
University of Hyderabad
Hyderabad-500 134. (India)

Dr. Seyed E Hasnain
Supervisor

Dr Seyed E Hasnain
Director
Centre for DNA Fingerprinting & Diagnostics
DBT Ministry of Science & Technology, Govt. of India
Hyderabad 500 076

Head, Department of Biochemistry
University of Hyderabad

HEAD
DEPARTMENT OF BIOCHEMISTRY
SCHOOL OF LIFE SCIENCES,
UNIVERSITY OF HYDERABAD,
HYDERABAD-500 046. (A.P.)

CDFD सी डी एफ डी
CENTRE FOR DNA FINGERPRINTING AND DIAGNOSTICS
(An Autonomous Centre of the Department of Biotechnology, Ministry of Science & Technology, Govt. of India)
डीएनए फिंगरप्रिंटिंग एवं निदान केन्द्र
(जैव प्रौधोगिकी विभाग, विज्ञान एवं तकनीकी मंत्रालय, भारत सरकार की स्वायत्त संस्था)
Nacharam नाचारम, Hyderabad हैदराबाद - 500 076, India भारत

DECLARATION

I hereby declare that this thesis entitled "Characterization of *Mycobacterium tuberculosis* TCA cycle enzymes isocitrate dehydrogenase and aconitase" has been carried out by me under the supervision of *Dr. Seyed E. Hasnain* at Centre for DNA Fingerprinting and Diagnostics, Hyderabad. The work is original and has not been submitted in part or full for any other degree or diploma of any other University earlier.

Sharmistha Banerjee
Sharmistha Banerjee
Candidate

Tel. दूर भाष : (040) 27151344 -46-47-56, 27150727
: 27155607-08-12, 27171491, 27171503
Fax फैक्स : (91-040) 27155479 & 27155610
Website : http://www.cdfd.org.in

To ...

Prasenjit

CONTENTS

ACKNOWLEDGEMENTS

I am greatly indebted to my research supervisor *Dr. Seyed E. Hasnain*, for his constant guidance, support and encouragement throughout the course of this study and for showing confidence in my abilities, especially during times when my self-confidence was low and for imbibing in me the desire to aim high.

My heartfelt gratitude to my seniors and lab colleagues, who have been extremely supportive, for making my graduate experience both intellectually rewarding and personally enriching. I have cherished every moment working at LMCB, the cordial and fun-filled atmosphere of which has made me evolve into a better person.

Thanks also goes to all my friends at CDFD & NIN who had helped me in one way or other in completing this work.

A very special thanks to my best friend and colleague, M. Senthil Kumar, without whose help, constant encouragement and wonderful coffee, I could not have pulled through my graduation.

I am thankful to my family, for their love, consistent support and patience through out these years. My heartfelt thanks to my husband, *Prasenjit*, for believing in my dreams, for supporting in everything I did and for his unshakeable confidence in my abilities. He instilled in me the desire to aim for excellence.

I wish to thank our collaborators, Dr. VM Katoch and Dr. KJR Murthy for their kind help.

Financial assistance from the Council of Scientific and Industrial Research (CSIR) is gratefully acknowledged.

Sharmistha Banerjee

ABBREVIATIONS

°C	Degree centigrade
Acn	Aconitase
AFLP	Amplified fragment length polymorphism
AIDS	Acquired immunodeficiency syndrome
APS	Ammonium persulphate
ATP	Adenosine-5'-triphosphate
ATT	Antitubercular therapy
BCG	Bacillus of Calmette Guerin
BLAST	Basic Local Alignment Search Tool
bp	Base pair
BSA	Bovine Serum Albumin
Ci	Curies
cm	Centimeter (10^{-2} meter)
dATP	2'-deoxyadenosine-5'-triphosphate
dCTP	2'-deoxycytidine-5'-triphosphate
DEPC	Diethyl pyrocarbonate
dGTP	2'-deoxyguanosine-5'-triphosphate
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphate
DOTS	Directly observed therapy short course
DR	Direct repeat
DTT	1,4-Dithiothreitol
dTTP	2'-deoxythymidine-5'-triphosphate
E. coli	*Escherechia coli*
EDTA	Ethylenediaminetetraacetic acid (disodium salt)
ELISA	Enzyme linked immunosorbent assay
EtBr	Ethidium Bromide
FAFLP	Fluorescent amplified fragment length polymorphism
Fe–S	Iron–Sulphur
FPLC	Fast protein liquid chromatography
gm	Gram

HIV	Human immunodeficiency virus
Hsp	Heat shock protein
Icd	Isocitrate dehydrogenase
INH	Isoniazid
IPTG	Isopropyl-β-D-thiogalactopyranoside
IRE	Iron responsive element
IRP	Iron responsive protein
IS	Insertion sequence
kb	Kilobase pair
KCl	Potassium Chloride
kDa	Kilodaltons
Ki	Inhibitor constant
Km	Michaelis-Menten constant
KV	Kilovolts
L	Litre
LB	Luria-Bertani
LJ	Lowenstein Jensen
M	Molar
M. bovis	*Mycobacterium bovis*
M.tb	*Mycobacterium tuberculosis*
MCS	Multiple cloning site
MDR TB	Multidrug resistant tuberculosis
mg	Milli gram (10^{-3} gram)
Mg^{++}	Magnesium ion
$MgCl_2$	Magnesium Chloride
$MgSO_4$	Magnesium sulphate
MIC	Minimum inhibitory concentration
min	Minutes
ml	Millilitre
ml	Millilitres (10^{-3} litres)
mM	Millimolar
mmol	Millimoles (10^{-3} moles)
Mn^{++}	Manganese ion
MPTR	Major polymorphic tandem repeats
NaCl	Sodium Chloride
NADP	Nicotinamide adenine dinucleotide phosphate
NADPH	Nicotinamide adenine dinucleotide phosphate (reduced)
NBDCl	7-chloro-4-nitrobenzo-2-oxa-1,3-diazole
ng	Nano gram (10^{-9} gram)
nt	Nucleotide

NTM	Non-Tuberculous Mycobacterium
OD	Optical density
OFL	Ofloxacin
OPD	Out patients department
ORF	Open reading frame
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
PE	Proline-Glutamine
PEG	Polyethylene glycol
PGRS	Polymorphic glycine rich regions
pmoles	Picomoles (10^{-12} moles)
PPD	Purified Protein Derivative
PPE	Proline-Proline-Glutamine
RBS	Ribosome Binding site
RE	Restriction endonuclease
RFLP	Random fragment length polymorphism
RFM	Rifampicin
RNA	Riboxynucleic acid
RNTCP	Revised National Tuberculosis Control Programme
rpm	Revolutions per minute
RT-PCR	Reverse transcriptase polymerase chain reaction
SCOTS	Selective capture of transcribed sequences
SDS	Sodium dodecyl sulphate
Sec	Seconds
SEM	Standard Error of Mean
TAE	TrisCl Acetic acid EDTA buffer
Taq	*Thermus aquaticus*
TB	Tuberculosis
TBE	TrisCl Boric acid EDTA buffer
TCA	Tri-carboxylic acid
TE	TrisCl EDTA
TEMED	N,N,N',N'-tetramethylethylenediamine
TGE	TrisCl Glucose EDTA buffer
TrisCl	Tris-(hydroxymethyl) aminomethane hydrochloride
U	Units
UTR	Untranslated region
UV	Ultra Violet
Ve	Elution volume
Vmax	Velocity (maximum)
Vo	Void volume

WHO	World Health Organization
Zn^{++}	Zinc ion
mg	Micro gram (10^{-6} grams)
mj	Micro joules (10^{-6} joules)
mM	Micromolar (10^{-6} molar)

Chapter 1

Introduction

CHAPTER 1
INTRODUCTION

- Brief History
- *Mycobacterium tuberculosis* in post genomic era
- Genomic clues to adaptive responses and virulence in *Mycobacterium tuberculosis*
  - Physiological changes in *Mycobacterium tuberculosis* in response to host infection
    * Adaptation to altered carbon source
    * Adaptation for maintaining metal ion homeostasis
    * Biosynthesis of essential amino acid and purines for *in vivo* survival
    * Genes important in anaerobic respiration
    * Coping with oxidative stress
    * Transcriptional regulators and regulatory machinery
  - Genomic clues to the immune response to tuberculosis
    * Humoral response to *M. tb* antigens
    * T-cell mediated immunity to *M. tb* antigens
  - Energy Metabolism and Energy cycle Enzymes in *Mycobacterium tuberculosis*
    * *Mycobacterium tuberculosis* Isocitrate Dehydrogenases, *Mycobacterium tuberculosis* Aconitase

Fathomed as the 'ducks' of the bacterial world, *Mycobacteria*, are unicellular, aerobic, Gram-positive bacteria, taxonomically related to the antibiotic-producing streptomycetes. They are exceptional among bacteria in having an extremely thick, hydrophobic cell wall that protects them from desiccation. Many mycobacteria are innocuous and ecologically useful as they can fix nitrogen and degrade organic matter in soil. Better known paradoxically, are the few pathogenic mycobacteria which cause tuberculosis (*M. tuberculosis, M. bovis M. microti and M. africanum,*) and leprosy (*M. leprae*).

Mycobacterium tuberculosis (*M. tb*) is a facultative intracellular pathogen usually infecting mononuclear phagocytes (e.g. macrophages). An obligate aerobe, it grows most successfully in tissues with high oxygen content, such as the lungs. They are slow growers, with a generation time of 12 to 18 hours. *M. tb* spreads through aerosol, but not till 16th century was it known to be contagious. Earlier efforts to trace the evolutionary history of tuberculosis pointed their origin from *Mycobacterium bovis*. *M. bovis* was causing TB in the animal kingdom long before invading humans. However, after the domestication of cattle between 8000-4000 BC, there is archaeological evidence of human infection by *M. bovis* probably thought to be through milk consumption (Bates and Stead, 1993). *M. tb* is likely a human-specialized form of *M. bovis* that developed among milk-drinking population, a hypothesis that is debatable in the present post genomics era. It later spread to the rest of the world due to extensive migration. By 1000 BC, *M. tb* and pulmonary TB had spread throughout the world. At the end of 19th century, Dr. Robert Koch isolated the dreaded organism that was the etiological agent of tuberculosis (Koch, 1882; Barnes 2000). Since then, there had been desultory experiments with a variety of medical applications in the management of pulmonary tuberculosis. The first success, marking the antibiotic era, came in the year 1944, with introduction of streptomycin for tuberculosis treatment. In 1952, the wonder drug isoniazid was introduced which was found to be far more effective than streptomycin. Years 1954, 1962 and 1963 saw the introductions of pyrazinamide, ethambutol and rifampicin,

respectively. The joy of triumph over TB by the combinatorial therapy (Heym and Cole, 1996) did not last very long with reports of emerging multi-drug resistant strains in the early 1970s (Cohn *et al*, 1997; Chowgule *et al*, 1998; Das *et al*, 1995). Adding to the misery is the synergism of *M. tb* with Human Immunodeficiency Virus (HIV) (Iseman and Sabarbaro, 1992; Pitchenik *et al*, 1990; Edlin *et al*, 1992; Coronado *et al*, 1993; Ahmed *et al*, 2003; Siddiqi *et al*, 1998; Raman *et al*, 2000, Ahmed *et al*, 1998). Today, despite being a completely curable disease, the morbidity and mortality statistics of tuberculosis is so extravagant that in the world someone dies of TB every 15 seconds (WHO report 2003). Even today, TB remains a global emergency.

The statistical records maintained around the world highlight a significant character of this pathogen, that is, its ability to escape the host immune system and remain undetected in lungs for decades. For instance, in India, every day, more than 20,000 people become infected with the TB bacillus but only 5000 develop the disease (RNTCP report, 2004). In the remaining, *Mycobacterium tuberculosis* is either cleared or remains undetected by the host immune system. Every year 18 lakh (or 1.8 million) people develop TB, of which nearly 8 lakh (0.8 million) are infectious (sputum-positive) and the remaining 10 lakh remain asymptomatic (RNTCP report, 2004). These asymptomatic cases remain undetected and untreated. Untreated pulmonary TB patients spread infection to others in the community -each infectious patient can infect 10-15 persons in a year unless effectively treated. Hence, the major threat to mankind from *Mycobacterium tuberculosis* is its persistence and latency in human host.

If persistence and latency of *M. tb* is a major problem for the human host, it is also a major challenge for the pathogen itself that encounters diverse environmental pressures and an over abundance of hostile conditions inside the human macrophages. As in other model of host-microorganism interactions such as *Rhizobium meliloti* (symbiosis), *Vibrio cholerae* or *Pseudomonas aeruginosa*

(pathogenesis) adaptive response is likely to play a crucial role in the virulence, persistence and latency of *Mycobacterium tuberculosis*.

Efforts to fight tuberculosis have intensified recently, and an important tool for researchers is the genome sequence of *Mycobacterium tuberculosis* H37Rv (Cole *et al*, 1998; Cole *et al*, 1994). Since its isolation in 1905, the H37Rv strain of *M. tb* has found extensive, worldwide application in biomedical research because it has retained full virulence in animal models of tuberculosis, unlike some clinical isolates; it is also susceptible to drugs and pliable to genetic manipulations. *Mycobacterium tuberculosis* research got tremendous boost from the whole genome sequencing of H37Rv (Cole *et al*, 1998) marking the beginning of post genomics era in tuberculosis research and disease management. Researchers around the world are now using this information to look for genes that alter the pathogen's metabolism to adapt inside human lung macrophages, their natural niche (Chakhaiyar *et al*, 2004; Hasnain, 2003; Ahmed and Hasnain, 2004). The facts disclosed by the genome sequence, including the presence of 13 different sigma factors, clearly suggest that the bacterium is well equipped for adaptation to environmental changes at the transcriptional level. The post genomics efforts, assisted by powerful bioinformatic tools, are attempting to understand *Mycobacterium tuberculosis* in terms of its adaptive physiology and evolution.

The present work is an effort to understand two important adaptations by *M. tb* inside host macrophages with respect to TCA cycle enzymes isocitrate dehydrogenase and aconitase. These are (I) adaptation to controlled energy flux during altered carbon source and (II) adaptation for maintaining metal ion homeostasis. Before going into the details of the work done, it is important to present an overview of genomic clues to altered physiology and immunology of *Mycobacterium tuberculosis* inside host to understand its adaptive responses, virulence and immunogenicity.

Genomic clues to adaptive responses and virulence in *Mycobacterium tuberculosis*

Mycobacterium tuberculosis has co-evolved with the human host and has established itself as one of the most efficient pathogen in the history of mankind. In doing so, it has mastered the art to adjust its physiology according to the need for intracellular survival and host immune evasion. Therefore, there exists a distinct difference in physiology of *Mycobacterium tuberculosis* during intracellular and extracellular growth and understanding these differences can lead to clear perception of the survival strategy applied by the pathogen and hence rational validation of antitubercle targets. *M. tb* faces a very hostile condition inside host macrophages where it confronts altered carbon source, absence or altogether low levels of essential nutrients like amino acids, purines and pyrimidines or altered levels of essential metal cofactors, like magnesium, zinc or iron, deficiency or excess of which can be lethal for the pathogen (Boon *et al*, 2002; Boon *et al*, 2001).

One of the most remarkable advantages of the completion of the *M. tb* genome and its annotation is the promotion of comparative genomics through development of DNA arrays (Behr *et al*, 1999; Brosch *et al*, 2002; Brosch *et al*, 2000; Cole *et al*, 2001; Mariani *et al*, 2000; Wei *et al*, 2000; Ahmed *et al*, 2003; Majeed *et al*, 2004). Microarrays have made it possible to get a global picture of the expression profile of all the *M. tb* genes under varied conditions in considerable less time (Schoolnik 2002, Mariani *et al*, 2000; Stewart *et al*, 2002; Wernisch *et al*, 2003). The post genomics efforts have seen researchers creating mutants for genes encoding enzymes in the biosynthetic, biodegradative or acquisition pathways for elucidating their role in pathogenesis. Energy flux being an important criteria, genes encoding respiratory enzymes and enzymes that protect against oxidative stress have been mutated to understand their importance during normal aerobic respiration, as well those created by infected hosts. Also, several other techniques, such as, subtractive hybridization, signature-tagged mutagenesis, complementation and selective capture of transcribed sequences

(SCOTS) have been applied, aided by genome information, to elucidate the details of host-pathogen interactions, antigenic variations and immune response to *Mycobacterium tuberculosis*. For pragmatically understanding the pathogenesis of *Mycobacterium tuberculosis*, one has to understand its genetics, physiology and immunology. Presented below is a synopsis of some of the important genes involved in altered physiology and immunology of the pathogen.

Physiological changes in *Mycobacterium tuberculosis* in response to host infection

Adaptation to altered carbon source: *M. tb* shifts from a metabolism that preferentially uses carbohydrates when growing *in vitro* to one that utilizes fatty acids when growing in the infected host (Segal *et al*, 1956; Segal *et al*, 1957). Inside macrophages, *M. tb* is more lipolytic than lipogenic owing to abundance of fatty acids as carbon source. These earlier observations are consistent with the genome information that found more than 250 genes involved in fatty acid metabolism as compared to only 50 in *Escherichia coli*. Some of these genes have been proved essential in such altered conditions experimentally.

Isocitrate lyase (Rv0467) is an enzyme that converts isocitrate to succinate in the glyoxylate shunt. This ultimately links to the Kreb's cycle *via* malate synthase and allows *M. tb* to grow on acetate or fatty acids as sole carbon sources. Initial observations showed that isocitrate lyase activity increases as the cells reach stationary phase (Wayne *et al*, 1982) in an *in vitro* model and also its mRNA level increases when *M. tb* infects human macrophages (Dubnau *et al*, 2002; Graham *et al*, 1999). Allelic replacement of *icl* in *M. tb* Erdman strain showed that mutant cannot grow on C2 carbon sources and hence cannot survive inside activated macrophages (McKinney *et al*, 2000; Shi *et al*, 2003).

Comparative proteomics of H37Rv and H37Ra showed that *fadD33* (Rv1345), a possible polyketide synthase, involved in lipid degradation, is expressed at much higher levels in virulent H37Rv than in the avirulent H37Ra strain. There are 36 genes annotated as putative orthologs of *E. coli fadD* that catalyze the first step in fatty acid β-oxidation by adding a Coenzyme A moiety to free fatty acids. Rv1345 gene was inactivated and the mutant showed tissue specific attenuation in mice (Rindi *et al*, 2002; Dubnau *et al*, 2002).

Pantothenate operon (PanC/PanD) participates in a crucial step that synthesizes coenzme A (CoA) that is central to fatty acid biosynthesis as well as degradation, intermediary metabolism, and other cellular processes. The mutants were attenuated for virulence (Sambandamurthy *et al*, 2002). When injected into mice, the mutant was also protective against an aerosol challenge with virulent *M. tb*, showing protection similar to that conferred by *M. bovis* BCG.

Phospholipases cluster *plcABC* (*plcA*, *plcB*, *plcC*, *plcD*) is characteristic of *M. tb* and is absent in *M. bovis* and its BCG derivatives (Raynaud *et al*, 2002). *plcA*, *plcB*, and *plcC*, are closely linked to each other in *M. tb* genome, but *plcD* is not. *plcD* is polymorphic in different strains of *M. tb*. The triple and quadruple *plc* mutants of *M. tb* grew normally in THP-1 macrophage cell lines but were found to be attenuated in mice (Fuangthong *et al*, 2002).

Adaptation for maintaining metal ion homeostasis: Metals form important cofactors in majority of biochemical pathways. Some are highly crucial, like, replication requires magnesium and iron is an important part of all iron-sulphur cluster and heme proteins that participate in a variety of life supporting functions. However, paradoxically, there always exists a love-hate relation between the concentration of these metal ions and both excess and deficit of these can be detrimental. Thus, genes responsible for altered physiology to maintain homeostasis inside macrophage may be very important to the pathogen and mutants that affect the uptake of these metals cause the attenuation of virulence. For example, MgtC

(Rv1811) is a transporter involved in magnesium uptake (Moncrief *et al*, 1998) and is essential for *M. tb* to grow in low-magnesium media and in macrophages, indicating that this environment is limiting for this divalent cation (Groisman, 2001; Buchmeier *et al*, 2000).

Iron is essential for most life forms and a delicate balance between free and bound iron is maintained inside cell. Since most of the time iron is in the form of the insoluble ferric salts in the environment, iron uptake systems are required to solubilize and transport the iron into the cell. In pathogenic bacteria, siderophores are the iron acquisition systems that chelate and solubilize iron from host iron-containing protein and the solubilized iron is internalized by high affinity transporters. The host, as a defense mechanism, sequesters iron to prevent bacterial growth (Litwin *et al*, 1993; Lounis *et al*, 2001). The *mbt* operon of *M. tb* comprises of ten ORFs, *mbtA* through *mbtJ*, and is central to synthesis of mycobactin and carboxymycobactin (De Voss *et al*, 2000; Quadri *et al*, 1998), the major siderophores in *M. tb* (Ratledge *et al*, 1996). This regulon is repressed by IdeR in high-iron conditions (Gold *et al*, 2001; Rodriguez *et al*, 2002). IdeR is an iron dependent regulator that binds to specific DNA regions depending upon the availability of iron or related divalent cations (Pohl *et al*, 1999; Schmitt *et al*, 1995; Wong *et al*, 1999). IdeR is an essential gene in *M. tb* and is the major mycobacterial regulator of iron uptake and storage genes, repressing the former and activating the latter (Dussurget *et al*, 1996; Gold *et al*, 2001; Rodriguez *et al*, 1999; Rodriguez *et al*, 2002).

Biosynthesis of essential amino acid and purines for *in vivo* survival: Some of the amino acids have been found to be essential for the survival inside host macrophages and inactivation of enzymes in the biosynthesis of these amino acids have resulted in attenuated growth inside macrophages. Interestingly, some of these mutants could protect mice against *M. tb* infection just like BCG. Some of these genes include anthranilate phosphoribosyl transferase (Rv2192c, *trpD*) (Parish *et al*, 1999; Smith *et al*, 2001), pyrroline-5-carboxylate reductase (Rv0500, *proC*) involved in proline

biosynthesis, isopropylmalate isomerase (Rv2987c, *leuD*), an enzyme that functions in the biosynthesis of leucine (Hondalus *et al*, 2000; McAdam *et al*, 1995; Bange *et al*, 1996), and 1-phosphoribosylaminoimidazole- succinocarboxamide synthase (Rv0780, *purC*) (Jackson *et al*, 1999).

Genes important in anaerobic respiration: A strict aerobe, yet *M. tb* encounters microareophilic environment in lung granulomas during later stages of infection. Genome sequence and subsequent experiments on *M. tb* have shown the presence of *narG* (Rv1161) which is a subunit of nitrate reductase, an enzyme involved in anaerobic respiration (Wayne, 1994). Anaerobic nitrate reductase activity increases when *M. tb* becomes microaerophilic. A *narG* mutant of *M. bovis* BCG showed a significant virulence in infected mice, but growth remain unaffected either under aerobic or anaerobic conditions *in vitro* (Weber *et al*, 2000). These results have not been established in *M. tb* yet, it is proposed that anaerobic or microaerophilic growth is a significant property of *M. tb* physiology during infection.

Coping with oxidative stress: Peroxides are the by-products of aerobic respiration. These, when accumulate, give rise to highly toxic reactive oxygen intermediates (ROIs). Therefore, every aerobic organism is equipped with a battery of peroxidases and catalases. These enzymes are also important for the response to various external oxidative stress. Since phagocytic cells produce ROIs to kill invading bacteria, it is expected that these enzymes are important for *M. tb* virulence. KatG is the only enzyme with catalase activity in *M. tb* that degrades H_2O_2 and organic peroxides. In addition to its catalase activity, it converts isoniazid from its prodrug form to activated form. Mutation in *katG* has been correlated with resistance to isoniazid (Marcinkeviciene *et al*, 1995; Siddiqi *et al*, 2002; Siddiqi *et al*, 2001; Siddiqi *et al*, 1998). In *M. smegmatis* (Zahrt *et al*, 2001) and *M. tb* (Pym *et al*, 2001) *katG* is negatively regulated by the FurA protein, whose gene is directly upstream of *katG*. The interdependence of Fur A and KatG is still not clear and neither any

experimental evidence is available to show that FurA is a virulent factor. AhpC is an alkyl hydroperoxide reductase, a member of family of enzymes that detoxify organic hydroxyperoxides. Sherman *et al.* demonstrated that all KatG mutant *M. tb* strains overexpressed AhpC protein at significantly higher levels than INH-sensitive strains. It was hypothesized that AhpC can compensate for the lack of KatG activity in *M. tb*. However, levels of AhpC could not be correlated with the virulence of *katG* mutants (Heym *et al*, 1997). SodA is the iron-factored while SodC is the Cu, Zn-factored superoxide dismutase of *M. tb* that degrades superoxides produced due to aerobic respiration within the pathogen or superoxides produced by the host macrophages upon phagocytosis. Sod A inhibits induction of the cellular immune response by inhibiting the redox signaling by macrophages. Hence, it is an essential enzyme for intracellular survival.

Transcriptional regulators and regulatory machinery

The intricacy of the environmental conditions encountered by *M. tb* necessitates comprehensive regulatory machinery. But contrary to other bacteria, *M. tb* has only 11 complete pairs of sensor histidine kinases and response regulators, and a few other regulatory genes. Thirteen putative sigma factors govern gene expression at the level of transcription initiation and more than 100 regulatory proteins are predicted. The relative scarcity in signal transduction pathways is probably balanced by the presence of a phosphorelay system comprising of eukaryotic-like serine/ threonine protein kinases (STPKs). This also indicates a complexity of the regulatory machinery involving cross talk between proteins of regulatory pathways. In the last few years, much effort has been put in understanding these regulatory factors that act as 'switches' and allow the pathogen to selectively transcribe genes required for the new environment (Das Gupta *et al*, 1993; Dellagostin *et al*, 1995; DeMaio *et al*, 1996; Primm *et al*, 2000). Some of them are discussed below.

Sigma factors: The promoter specificity of RNA polymerases is because of different sigma factors that recognize specific promoters

and allow the transcription of genes suitable for the new 'lifestyle' of the pathogen. *M. tb* genome has also shown presence of several such sigma factors and response regulators. Sigma A (Rv2703, *sigA*) is the essential principal mycobacterial sigma factor and is presumably necessary for most mycobacterial housekeeping gene transcription (Gomez *et al*, 1998; Predich *et al*, 1995). Sigma E (Rv1221, *sigE*) dominantly controls expression of the genes that are induced as bacterial response to various environmental stresses like high temperature or detergent stress (Manganelli *et al*, 1999). *M. tb* encounters these stresses as the outcome of host inflammatory response. Hybridization-based methods showed that *sigE* mRNA levels increased during *M. tb* growth in human macrophages (Graham *et al*, 1999; Jensen-Cain *et al*, 2001). DNA arrays of the *sigE* mutant and the wild-type parent were compared. It showed that during normal growth, (Brosch *et al*, 2002) genes are under regulatory function of sigma E, while 23 additional genes are upregulated by Sig E after sodium dodecyl sulfate (SDS) stress.

Two-component systems: Elaborate signal transduction pathways occur in *M. tb*, apart from various regulatory sigma factors, to adapt to hostile conditions inside the host. These are the "two-component systems" where sensor proteins (usually histidine kinases) receive and respond to the environmental signal and in turn, activate related effector proteins called the response regulators, which are usually transcriptional regulatory factors. Each two-component system responds to a specific stimulus. Some of these are PhoP (Rv0757, *phoP*) (Groisman, 2001; Perez *et al*, 2001), PrrA (Rv0903c, *prrA*) (Graham *et al*, 1999; Ewann *et al*, 2002) and mycobacterial persistence regulator (Rv0981, *mprA*) (Zahrt *et al*, 2001).

Amongst other transcriptional regulators are HspR (Rv0353, *hspR*) a repressor of heat shock genes like *hsp70* (Gomez *et al*, 2000; Lee *et al*, 1995; Stewart *et al*, 2001) and WhiB (Rv3416) speculated to be participating in persistence or latency of *M. tb* (Gomez *et al*, 2000; Soliveri *et al*, 2000).

Genomic clues to the immune response to tuberculosis

Immune responses to tuberculosis have been largely studied in the animal models that are capable of developing and expressing resistance to infection with virulent mycobacteria when vaccinated with *M. bovis* BCG. The immunity against *M. tb* parted by BCG vaccination is measured in animal models reduction in the size and severity of tissue damage in the primary lung lesion (Smith *et al*, 1989). The 'host' factors play a very important role in the kind of immune response to *M. tb* infection and actually may decide if the infected person will be an asymptomatic carrier or will have full blown TB (van Crevel *et al*, 2002).

Humoral response to *M. tb* antigens

Antibody response to *M. tb* has not yet been explored substantially as humoral response is thought not to contribute to resistance in tuberculosis (Kaufmann, 1989). Animal model (guinea pig) infected with either BCG alone or aerosolized with *M. tb* H37Rv alone were studied extensively for immune response to *M. tb*. Antibodies against both recombinant *M. bovis* Hsp65 and *M. tb* Hsp 70 antigen were detected in dot blot immune assay within 1 to 2 weeks after primary infection (Bartowr *et al*, 1990). These results indicated that B-cell epitopes are indeed recognized in infected animal sera. The regulatory or modulatory importance of these circulating secreted or released immune complexes became evident by experiments in animal models (Bartowr *et al*, 1990). It was postulated that these act through lymphocytes expressing crystallizable fragment Fc receptor for IgG. The removal of FcγR+ lymphocytes from PPD-induced cultures *in vitro* enhanced the proliferation of non-Fc receptor bearing T-lymphocytes (Mcmurray *et al*, 1992). It was experimentally shown (Mcmurray *et al*, 1992) that removal of immune complexes from such serum by polyethylene glycol precipitation relieved the suppression mediated by FcγR+ lymphocytes. Thus, humoral response may modulate cellular reactivity and, by inference, also modulate antituberculous resistance. This also rationalizes why

attenuated live vaccines are more effective in generating an antibody response than killed vaccines.

Culture filtrate proteins:

A large number of proteins, approximately 200 (Sonnenberg *et al*, 1997), have been observed to be present in the culture medium in which *M. tb* is grown. These are called culture filtrate proteins or CFPs, for many of which the mechanism of secretion is not defined. This also includes some of the cell wall associated proteins and also proteins that have no leader sequence required for protein secretion. Recently, CFPs have been given much importance as many of them could be recognized by the sera of TB patients, indicating that these are released inside the host as well and possibly elicit B-cell response. Many of the proteins found in the culture filtrate are involved in general metabolism of *M. tb*, such as KatG (catalase-peroxidase) (Braunstein *et al*, 2000), GlnA (glutamine synthase) (Tullius *et al*, 2001), SodA (superoxide dismutase) (Tullius *et al*, 2001) etc. These proteins are released during early stages of growth, indicating that their localization in medium is physiological and is not related to lysis of the cell. Amongst the proteins that are used as autolysis marker are malate dehydrogenase and isocitrate dehydrogenase of TCA cycle. (Anderson, 1991). Amongst the thoroughly studied enzymes are KatG (catalase-peroxidase) and SodA (superoxide dismutase). Presence of these enzymes that degrade ROIs in the culture filtrate may be of significance in *M. tb* survival *via* more efficient detoxification of injurious molecules produced by the host in the infected phagosome (Braunstein *et al*, 2000). However, experiments indicate only highly stable and overexpressed proteins such as GlnA and SodA are found in early culture filtrates (Tullius *et al*, 2001) and the probable explanation could be bacterial leakage or lysis. Therefore, it has been claimed that live attenuated *M. tb* vaccines are better than those made from heat-killed cells because during growth in the host, *M. tb* releases several proteins that stimulate host immune mechanisms (Andersen, 1991). Some of the other proteins in the culture filtrate include HspX (Rv2031c, *hspX*) (Wayne, 1994). Esat6/CF-10 (Rv3875, Rv3874)

(Skjot *et al*, 2000; Renshaw *et al*, 2002), 19-kD protein (Rv3763, *lpqH*). (Noss *et al*, 2001; Thoma-Uszynski *et al*, 2001) and several others (Choudhary *et al*, 2003; Choudhary *et al*, 2004; Chakhaiyar *et al*, 2004).

The Esat6 and CF-10 fall in the RD-1 region, a region that is essentially absent in all BCG strains (Lewis *et al*, 2003). Absence of RD1 in BCG as compared to other members of *M. tb* complex could account for the attenuation and hence these ORFs could be a probable answer to the virulence of pathogenic mycobacteria (Mahairas *et al*, 1996; Pym *et al*, 2002).

The PE and PPE multigene families

The most remarkable discovery in the *M. tb* genome is the presence of a hitherto new and unique family of proteins called the PE and PPE multigene families. Almost 10% of the genome is devoted to these two unknown families of proteins that exhibit distinct motifs of Pro-Glu (PE) and Pro-Pro-Glu (PPE) at their highly conserved N-terminals. PE family consists of 99 members with a highly conserved ~110 amino acids at N-terminal while the C-terminal is variable and differ in the copy number of repeats. The PE proteins are further classified into sub-classes, the largest of which has 61 members characterized by the presence of highly repetitive poly glycine rich regions (PGRS). This sub-class of proteins, called PE-PGRS, is glycine rich (up to 50%) due to multiple tandem repetitions of Gly-Gly-Ala or Gly-Gly-Asn motifs. PPE family, on the other hand, has 68 members with ~180 amino-acid residues conserved at N-terminal. Atleast three groups have been marked in this class, of which the multiple tandem repeat (MPTR) class is characterized by tandem copies of the motif Asn-X-Gly-X-Gly-Asn-X-Gly, second group has conserved motif at 350 position and the third group houses the unrelated PPE proteins. The C-terminal of these proteins is highly variable and for this reason it was suggested that these proteins are involved in conferring antigenic variations amongst *M. tb* strains (Banu *et al*, 2002; Cole *et al*, 1998; Choudhary *et al*, 2003; Choudhary *et al*, 2004).

T-cell mediated immunity to *M. tb* antigens

M. tb is controlled by the cooperative actions of CD4, CD8 T cells and macrophages mediated by cytokines. The uptake of mycobacteria by macrophages is by conventional phagocytosis (Armstrong *et al*, 1975; Schlesinger, 1993). Binding of complement components from the surrounding serum may be the key mediators of this event (Schlesinger *et al*, 1990; Schlesinger *et al*, 1991). In addition, mannose receptors on the macrophages may also play a role (Schlesinger *et al*, 1993; Stokes *et al*, 1993; Chatterjee *et al*, 1992) and are directed preferentially to the lipoarabinomannan (LAM) structure on the mycobacterial cell wall (Schlesinger *et al*, 1993; Stokes *et al*, 1993; Crowle *et al*, 1991).

After mycobacterial antigens are induced into an endosomic pathway in the macrophage, peptides of the antigen are subsequently presented on the macrophage cell surface in the cleft of class II MHC encoded molecules. These complexes are then recognized by T-cells bearing complementary receptors that trigger acquired resistance. The cascade of cytokines that is produced during the infectious process is the key immunomodulator that drive these events. Some of these cytokines are IL-1, IL-6, IL-8, factors with antimicrobial activity like TNF, IL-10, IL-12, T-cell growth factor beta (TGFβ), macrophage colony stimulating factor (M-CSF), granulocyte-macrophage colony stimulating factor (GM-CSF), granulocyte colony stimulating factor (G-CSF), platelet-derived growth factor (PDGF) and others (Cooper *et al*, 1993; Flynn *et al*, 1993; Barnes *et al*, 1992).

An overall model for T-cell generation during the course of tuberculosis infection (mouse-model) is summarized in Figure 1.1. In an early protective phase, CD4 T-cells recognize secreted or exported proteins of the bacillus presented in association with class II MHC-encoded molecules and begin to secrete IFN-γ and other important cytokines resulting in macrophages activation and monocyte recruitment. γδ T-cells also accumulate and may augment protective

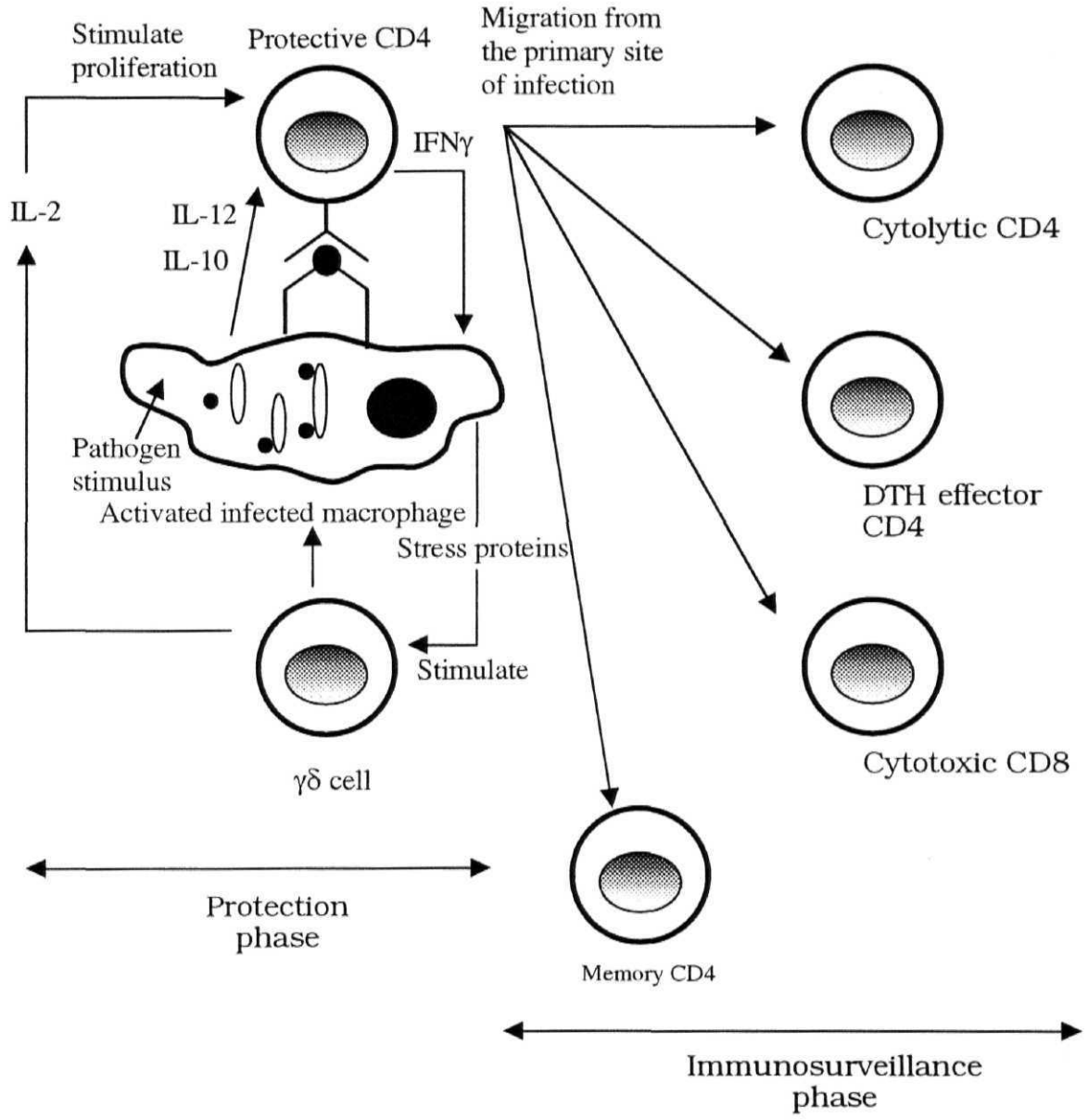
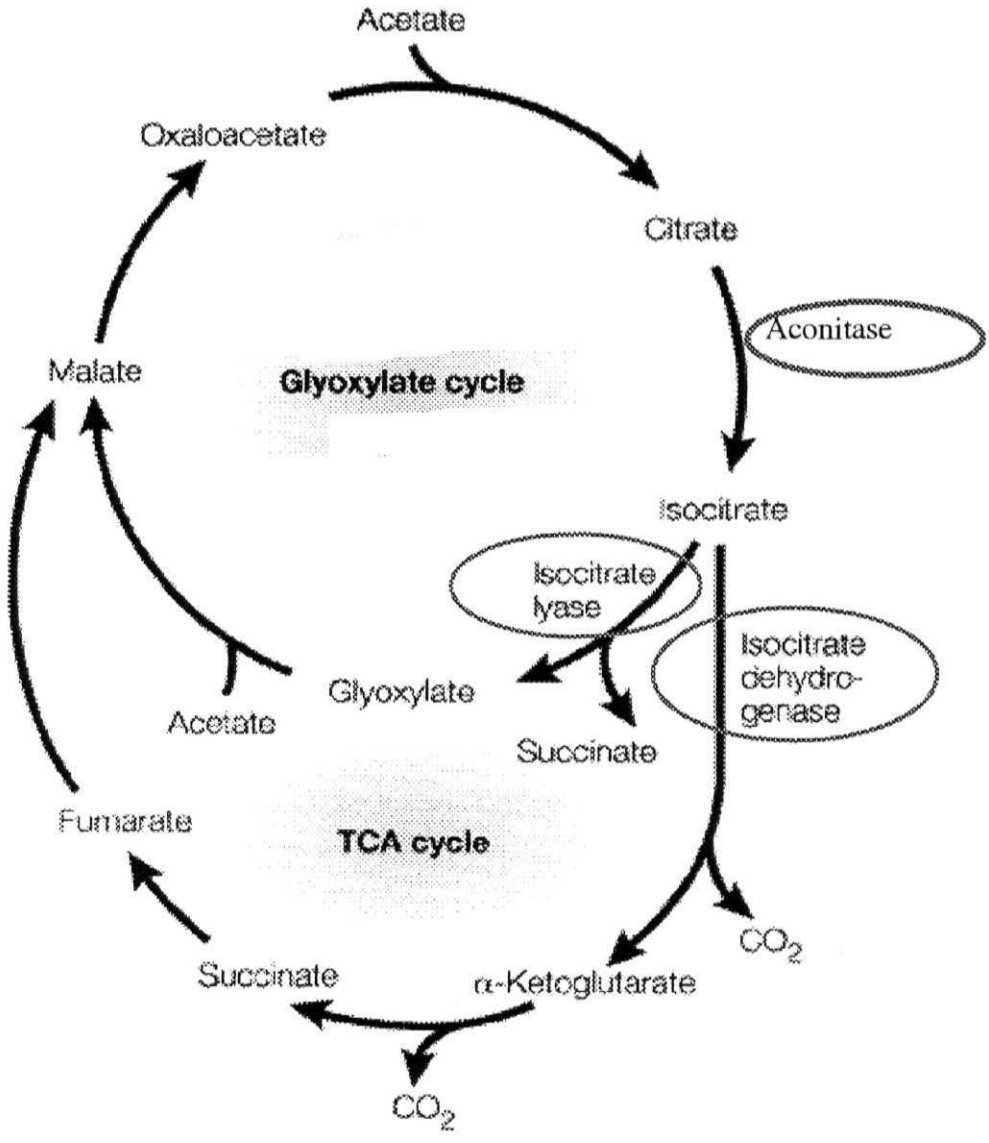


Figure 1.1: T-cell generation during the course of tuberculosis infection (mouse-model) [After Rom WN and Garay SM (Eds). Tuberculosis. 1995. 1st Edition, Little, Brown and Company (Inc), USA.]

T-cell proliferation by secreting IL-2 (Pancholi *et al*, 1993). What actually triggers γδ cells is unclear, but it is postulated that it may be induced by stress proteins expressed by the infected macrophages. As the infection progresses, some sensitized T-cells leave the site of infection and differentiate into longer-lived population. These cells participate in immune surveillance and try to contain the spread of the infection. The DTH effector T-cells recirculate and on recognition of the bacilli antigens, rapidly mediate the initiation of granulamatous response at the site. Other cells have cytolytic functions, presumably to release bacillus from crippled or heavily infected macrophages (CD4 cells) or from cells which can present only class I MHC molecules (CD 8 cells) (Xu *et al*, 1994; Cox *et al*, 1989; Orme, 1988; Hussein *et al*, 1987).

Energy Metabolism and Energy cycle Enzymes in *Mycobacterium tuberculosis*

The central metabolic pathways in bacteria, especially in *E.coli*, have been extensively studied to understand the physiology of the organisms under altered carbon sources (Holms, 2001). In prokaryotes where an enormous diversity exists in terms of environmental parameters, energy production and expenditure along with substrate utilization are closely monitored and regulated. Tricarboxylic acid cycle is the universal pathway for energy production. The enzymes of the universal, tri-carboxylic acid energy cycle have been found to be operative in mycobacterial species (Katoch *et al*, 1987; Prabhakaran, 1986; Sharma *et al*, 1985; Kannan *et al*, 1985). In most bacteria including *Mycobacterium tuberculosis*, apart from the TCA cycle, there exists an alternate pathway to metabolize two carbon sources called the glyoxylate shunt pathway. Degradation of fatty acids leads to formation of acetyl coenzyme A (acetyl-CoA) (Lakshmi *et al*, 1978). To grow on the acetyl- CoA produced when fatty acids are present as the sole carbon source, as is the case in *M. tb* infected macrophages, a mechanism must be available to replenish the dicarboxylic acids drained from the tricarboxylic acid cycle for cellular biosynthesis (Kornberg 1966). This is accomplished by the glyoxylate shunt (Figure 1.2). The two



Nature Reviews Molecular Cell Biology 2: 569-586 (2001)

Figure 1.2: Energy metabolism: Tricarboxylic acid (TCA) cycle and Glyoxylate shunt pathways in bacteria.

unique enzymes of the glyoxylate shunt, isocitrate lyase and malate synthase, are induced when acetate or fatty acids serve as the sole carbon source (Kornberg, 1966; McKinney *et al*, 2000; Shi *et al*, 2003).

When *M. tb* infects macrophages, the carbon source is altered from six-carbon to two-carbon, owing to abundance of fatty acids in human macrophages. McKinney *et al* (2000) showed that persistence of *M. tb* in mice is facilitated by isocitrate lyase, the first enzyme in glyoxylate shunt mechanism. They could also show that *icl* gene disruption attenuated bacterial persistence and virulence in immune-competent mice without affecting bacterial growth during the acute phase of infection. This was apparent at the level of the infected macrophages where activation of infected macrophages increased expression of ICL, and the *icl* mutant was markedly attenuated for survival in activated but not resting macrophages. These data suggested that the metabolism of *M. tb in vivo* required shift from TCA cycle to Glyoxylate shunt pathway.

***Mycobacterium tuberculosis* Isocitrate Dehydrogenases**

One of the key regulatory enzymes of the TCA cycle is isocitrate dehydrogenase that allosterically regulates the conversion of oxidative decarboxylation of D-isocitrate to α-ketoglutarate and CO_2 in the presence of a cofactor (Stryer, 1995). This rate-limiting step is the first NADPH-yielding reaction of the TCA cycle (Stryer, 1995). To sustain aerobic growth and persistence within the macrophages, the pathogen has to swap from TCA cycle to Glyoxylate shunt. Glyoxylate shunt pathway enzyme isocitrate lyase and TCA cycle enzyme isocitrate dehydrogenase share a common substrate, isocitric acid. However, isocitrate lyase has a relatively high Km value (and hence lower affinity) for the substrate isocitric acid and has to compete with isocitrate dehydrogenase for the same. The flux of isocitrate at this point is largely regulated by altered biochemistry of isocitrate dehydrogenase. It is known in other bacteria that phosphorylation of isocitrate dehydrogenase results in four fold decrease in Vmax of the

enzyme resulting in accumulation of isocitric acid which then becomes available to isocitrate lyase and the energy production shifts from TCA cycle to glyoxylate shunt, resulting in persistence of the pathogen. Thus, knowledge about the biochemical characters of *M. tb* isocitrate dehydrogenases becomes crucial for understanding the survival strategy of the bacillus inside the host.

M. tb genome has two isoforms of isocitrate dehyrogenase, Rv0066c (ICD-2) is annotated as $NADP^+$ dependent monomeric isocitrate dehydrogenase while Rv3339 (ICD-1) is a probable isocitrate dehydrogenase based on homology (Cole *et al*, 1998). The primary structure of the two isoforms shows only ~19% identity but the signature residues of the ICD family are conserved. *M. tb* isocitrate dehydrogenase has been evaluated in earlier reports for its immunogenic properties (Ohman and Ridell, 1996; Ohman and Ridell, 1995) and as a marker of autolysis during the late logarithmic growth phase (Anderson *et al*, 1991). However, little is known about the enzymological behaviour of *M. tb* isocitrate dehydrogenases.

***Mycobacterium tuberculosis* Aconitase**

Another important enzyme that is common between the TCA cycle and Glyoxylate shunt is Aconitase. *M. tb* aconitase (Acn) is an enzyme involved in the energy cycle pathway that belongs to a family of monomeric, Fe-S cluster containing proteins. Such proteins are known to act as iron regulatory and sensor proteins. The iron sulphur cluster of these proteins not only participate in electron transport during reversible isomerization of citrate and isocitrate in citric acid cycle (Beinert *et al*, 1996) but also serve as iron and oxygen sensors of the cell (Beinert *et al*, 1997; Kang *et al*, 2003). The binary activities are exerted through the assembly and dissasembly of iron and sulfur cluster. Protein with the intact [4Fe-4S] cluster, is the holoprotein, that functions as aconitase, whereas the apoprotein, [3Fe-4S] is an RNA-binding translational regulator (Rouault and Klausner 1997; Haile *et al*, 1992). The stability and functionality of aconitases as a translation regulator is affected not only by iron levels, but also by oxidative stress, which induces these Iron

Regulatory Proteins (IRPs) to bind to Iron Responsive Elements (IREs) and maintain iron homeostasis (Rouault *et al*, 1996; Fillebeen *et al*, 2002). IRPs maintain iron homeostasis by post transcriptional binding to conserved RNA stem-loop structures, Iron Responsive Elements (IREs), either at 5' or 3' end of untranslated regions of mRNA. Depending on whether the IRE is present on 3' or 5' end, binding of IRPs to IREs either protects the mRNA from degradation or inhibits their translation (Alén *et al*, 1999). For instance, ferritin, an iron storage protein, under low iron condition needs to be down regulated. Hence binding of IRP to its 5' end results in inhibition of translation by preventing ribosome to bind (Koeller *et al*, 1989). On the other hand, mammalian TfR, coding for transferrin receptor, is required to be upregulated under low iron conditions for uptake of iron. This is accomplished by binding to five IREs present at 3' end, which stabilizes the mRNA and prevents its degradation (Koeller *et al*, 1989). Thus, IRPs ensure coordinated regulation of iron utilization and uptake, maintaining threshold levels of iron in a cell. IRE like sequences are present in the UTRs of atleast two enzymes of citric acid cycle, aconitase and succinate dehydrogenase in mammals, implying that IRPs play an important role in mediating iron regulation of mitochondrial energy production (Gray *et al*, 1996). IRPs have also been shown to be activated by both hydrogen peroxide and iron mediated oxidative stress (Nunez *et al*, 2003). The reactivity of H_2O_2 with iron (Fenton reaction) intimately connects oxidative stress and cellular iron metabolism (Pantopoulos *et al*, 1995). Thus, recruitment of IRPs constitutes a highly effective strategy employed by the pathogen for cell survival.

Aconitases and IRPs are related with respect to the conserved amino acid residues across the family. This became evident when active site residues identified in pig heart mitochondria aconitase crystal structure were found to be conserved across mammalian IRPs (Frishman and Hentze, 1996). Based on primary structure similarity, all bacterial aconitases, including the α - proteobacterial aconitases are categorized mainly either into the aconitase group similar to eukaryotic IRP or cytosolic aconitase (Acn A / IRP group), or the

aconitase group found only in bacteria (AcnB) (Walden, 2002; Baughn *et al*, 2002). Several bacteria such as, *E. coli*, have two isoforms of aconitase, AcnA and AcnB, with different physiological properties and expression profiles (Jordan *et al*, 1999; Tang *et al*, 2002; Varghese *et al*, 2003), while prokaryotes like *Bacillus* or *Xanthomonas*, have only one aconitase in their genome. *Mycobacterium tuberculosis* also carries a single copy of aconitase gene (*acn*) coded by Rv1475c. Sequence comparisons of *Mycobacterial* aconitase (Swissprot ID: O53166) with *Escherichia coli* aconitase AcnA (Swissprot ID: P25516) and AcnB (Swissprot ID: P36683) showed that mycobacterial aconitase has closer identity to AcnA of *E. coli* (60% identity) rather than AcnB (20% identity). Earlier reports showed that AcnA of *E. coli* is induced in stationary phase or during oxidative stress (Jordan *et al*, 1999, Tang *et al*, 2002,). It is to be noted that AcnA / IRP group is less sensitive to oxygen mediated inactivation (Walden, 2002; Varghese *et al*, 2003) and therefore, it is logical to argue that aerobic respiration in *M. tb* would use the more stable AcnA / IRP aconitases for energy metabolism, a feature important for survival under oxidative stress. The fact that *M. tb* aconitase expression was downregulated 4.74 folds in starvation model (Betts *et al*, 2002) and upregulated at high iron stress (Wong *et al*, 1999) associates it to both energy as well as iron metabolism.

The work presented in this thesis constitutes complete biochemical characterization of the two isoforms of *M. tb* ICDs, Rv3339 (ICD-1) and Rv0066c (ICD-2). The studies provide experimental proof confirming that Rv3339 codes for ICD. Unusual behaviour of *M. tb* ICDs became evident when they were found to elicit strong B-cell response in TB-infected patient sera. As discussed earlier proteins that are released from *Mycobacterium tuberculosis* during late logarithmic growth phase are often considered as candidate components of immunogenic or autolysis markers. Isocitrate dehydrogenase (ICD) was found to be one such protein along with enzymes like superoxide dismutase etc. The immunological properties of the M. tb ICDs were evaluated through detection of anti-*M. tb* ICD antibody in sera of different categories of TB patient

sera through enzyme linked immunosorbent assays. The sensitivity and specificity of ICDs were compared to distinguish TB and non-TB patients *vis-à-vis* the conventional control – *M. tb* Hsp 60 and purified protein derivative (PPD). The result points to the advantage of the *M tb* ICDs in discriminating BCG vaccinated healthy control from TB patients. In the present work, the possibilities of *M. tb* aconitase being a bifunctional protein have also been explored, showing enzyme activity when reconstituted by iron and RNA binding activity when the enzyme is deprived of iron. The present work further elaborates upon the functional oligomeric state and basic biochemical properties of *M. tb* aconitase as a TCA cycle enzyme and also defines its ability to bind to selected IRE like squences in *M. tb* genome. The identity of the likely amino acid residues associated with the functional specificity of *M. tb* aconitase as a TCA cycle enzyme is highlighted. These results also point out that the two properties of this protein are mutually independent.

Chapter 2

Biochemical Characterization of *Mycobacterium tuberculosis* Isocitrate Dehydrogenases

INTRODUCTION

Controlled energy flux and metabolism are critical for pathogenic bacteria that encounter a plethora of hostile conditions inside host. One of the key regulatory enzymes in the tri-carboxylic acid energy cycle is the isocitrate dehydrodenase that allosterically regulates the conversion of oxidative decarboxylation of D-isocitrate to α-ketoglutarate and CO_2 in presence of a cofactor (Stryer, 1995). This rate-limiting step is the first NADPH yielding reaction of the TCA cycle (Stryer, 1995). ICD lies at the branch-point between citric acid cycle and glyoxylate shunt pathway (Nimmo *et al*, 1987; Holms, 1987). In *E.coli*, complex patterns of control of isocitrate dehydrogenase have been shown where glyoxylate bypass and citric acid cycle operate concurrently (Holms, 1987; Holms, 1996). Isocitrate dehydrogenase belongs to a family of enzymes that exhibits diversity with regard to amino acid composition, cofactor specificity, metal ion requirement and oligomeric state. ICD from different organisms has been phylogenetically affiliated to three subfamilies (Steen *et al*, 2001). Majority of the bacterial ICDs fall into subfamily I that includes archaeal and bacterial NADP dependent ICDs (Steen *et al*, 2001).

Mycobacterium tuberculosis, an intracellular pathogen, is exposed to different environmental constraints during its course of infection that necessitates a tight regulation of substrate utilization for energy flux. The glyoxylate shunt pathway in *M. tb* has been observed under hostile conditions inside macrophages where C_2 substrates, such as fatty acids, are the only carbon source (McKinney *et al*, 2000). To sustain aerobic growth and persistence within the macrophages the glyoxylate shunt pathway enzyme isocitrate lyase, which has relatively high Km value for isocitrate, has to compete with isocitrate dehydrogenase of citric acid cycle for the available substrate isocitric acid. The flux of isocitrate at this point is largely regulated by phosphorylation of ICD that results in four fold decrease in Vmax of the enzyme resulting in accumulation of isocitrate (Singh *et al*, 2002; Kornberg *et al*, 1966). Therefore, information on the biochemistry of

isocitrate dehydrogenase becomes crucial for understanding the survival strategy of the bacillus inside the host.

Two isoforms of isocitrate dehydrogenase, has been noted in *M. tb* genome. Rv0066c or *icd 2* (ICD-2) was annotated as $NADP^+$ dependent monomeric isocitrate dehydrogenase while Rv3339 or *icd-1* (ICD-1) is a probable isocitrate dehydrogenase based on homology (Cole *et al*, 1998). The percentage identity at amino acid level of two isoforms is only about 19% but the signature residues of the ICD family are conserved. In the present work the possible reasons for the presence of two isoforms of isocitrate dehydrogenase in *M. tb* genome have been explored. Further questions pertaining to their expression *in vivo*, differences in the biochemical properties, functional oligomeric state and differential expression of the two ICDs under different stages of growth have been addressed.

MATERIALS AND METHODS

RNA extraction and Reverse Transcriptase PCR: All the glassware were treated and the solutions were prepared with DEPC- water to inhibit RNAse activity. H37Rv was grown in Middlebrooks 7H9 media and cells were periodically harvested on 3^{rd} (early log), 7^{th}, 10^{th} (log), 12^{th} (late log) and 18^{th} (stationary phase) day by centrifuging at 2300Xg at 4°C for 10 minutes. RNA from H37Rv was extracted using the Qiaquick total RNA extraction kit (Qiagen, USA). The RNA was stored in alcohol at -70°C till further use. Internal primers for *M. tb icd-1* (FP- gcg gcg ggg ctg acc tac gag; RP- gta atc cgc ggg tcc agg caa aga) and *M. tb icd-2* (FP- gta ccg ccc gca cga gaa cga; RP- tgc gcg gcc agc tct tgt g) were used to confirm *in vivo* transcription of these genes using the Access RT-PCR kit (Promega Inc., USA). A control pair of primers (FP- gcc gcc gac tcg ccg ccc cag ac and RP- tca ccg ccg ccg acc aca cta aac) was designed to amplify a 245 bp region of the house keeping gene Rv1437 coding for phosphoglycerate kinase. A 50µl RT-PCR reaction contained ~200 ng of RNA, 100 ng each of both the primers, 0.2mM dNTPs, 1 X of the reaction buffer, 1.5mM $MgSO_4$, 0.05 U of AMV reverse transcriptase and *Tfl* polymerase. First strand

synthesis was carried out at 48°C for 45 minutes followed by an incubation of 94°C for 2 minutes to denature the reverse transcriptase. Second strand synthesis was carried out for 40 cycles at 94°C – 30 seconds, 60°C - 30 seconds and 68°C - 1 minutes. A final hold at 68°C for 7 minutes was included. The resulting RT-PCR product of 310 bp (*M. tb icd*-1) and ~ 530 bp (*M. tb icd*-2) and ~245 bp (Rv1437) were resolved by electrophoresis in a 2% agarose gel and the products' band intensities were compared using 'Quantity One (4.1.1)' gel documentation software (BioRad, USA).

Cloning, expression and purification of *M.tb* ICD-1 and *M.tb* ICD-2: The ORFs, corresponding to *M.tb* ICD-1 (Rv3339c, 1.230 kb) and *M.tb* ICD-2 (Rv0066c, 2.238 kb) were PCR amplified from the genomic DNA of H37Rv. *Bam*HI and *Hin*dIII restriction sites were incorporated in the 5' end of forward and reverse primers respectively for both *M.tb* ICD-1 and *M.tb* ICD-2. The primers and parameters for thermal cycle amplification have been tabulated below

Primers	Sequence	PCR Parameters	Amplicon Size
*M.tb icd-1*FP: *M.tb icd-1*RP:	ggatccATGTCCAACGCACCCAAGATA aagcttCTAATTGGCCAGCTCCTTTTC	94°C for 2' (35 Cycles) 94°C for 30" 50°C for 1' 72°C for 3' 72°C for 7'	~1.2 Kb
M.tb icd-2 FP: *M.tb icd-2* RP:	AGCTTggatccATGAGCGCCGAACAGCC CATGGaagcttTCAGCCTTGGACAGCCT	94°C for 2' (10 Cycles) 94°C for 30" 50°C for 30" 72°C for 3.30' (25 cycles) 94°C for 30" 58°C for 30" 72°C for 3.30' 72°C for 7'	~2.23Kb

Each 50µl PCR reaction contained 250 ng of template DNA (genomic DNA of H37Rv), 200 ng of each primer with 1U of *Taq* DNA

polymerase (Amersham, USA), 200 μ moles of all 4 dNTPs, 1.5mM Magnesium chloride and 1X PCR buffer supplied with the enzyme. The template DNA was amplified in GeneAmp PCR system 9700 (PE Applied Biosystems, USA). The PCR products were gel extracted using Qiagen Gel extraction kit (Qiagen Inc. USA) and stored at -20°C.

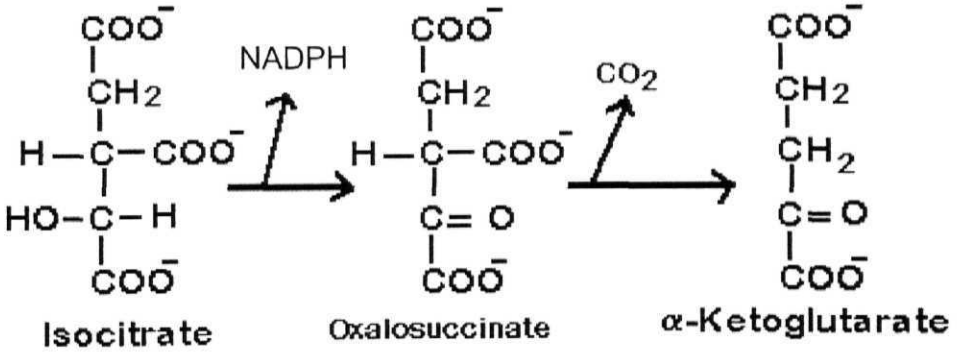
The amplicons carrying the full length *M.tb* ICD-1 and *M.tb* ICD-2 were cloned at the *Bam*HI and *Hin*dIII sites of the expression vector pRSET-A (Invitrogen) with a six histidine sequence tag at N-terminal. For the ligation reaction, 200ng of vector (pRSET A), 400ng of insert (*icd-1* or *icd-2*), 1 unit of T4 DNA ligase and 1X ligation buffer (30 mM Tris-HCl pH 7.8, 10mM $MgCl_2$, 10mM DTT, 2mM ATP, 5% PEG 8000) were mixed together and volume made to 20 μl with distilled water. The tube was incubated at 16°C for 16 hours or overnight. The ligation mix was used to transform DH5α strain of *E. coli*. The clones were confirmed by double digestion with *Bam*HI and *Hin*dIII first and later by sequencing using the T7 promoter primer, on an ABI prism 377 DNA sequencer (PE Biosystems, USA). The generated constructs 'setAicd1' and 'setAicd2' were further transformed into the BL21 (DE3) strain of *E.coli*.

The genes were overexpressed in the pRSET-A/ *E.coli* BL-21 (DE3) expression system Through induction by 0.1mM Isopropyl-beta-D-thiogalactopyranoside (IPTG). The setAicd1 BL-21 transformants were grown in ampicillin (100μg /ml) containing LB media supplemented with glycylglycine as described by Ghosh *et al*, 2004. The cells were grown to an OD_{600} 0.4 at 37°C, cooled to 27°C and induced by 0.1mM IPTG and allowed to grow overnight at 27°C for overexpression of the protein. The setAicd2 transformed BL-21 cells were grown in Terrific Broth (TB) containing ampicillin (100μg /ml) to an OD_{600} of 0.4 to 0.5 at 37°C, cooled to 27°C, induced with 0.1mM IPTG and grown overnight at 27°C.

The overexpressed his-tagged recombinant protein was purified by Ni^{2+}-nitrilotriacetate affinity chromatography from the soluble fraction. The cells were lysed by sonication, followed by centrifuging at 16000Xg for 30 minutes at 4°C. The clear lysate was loaded onto Ni^{2+}-NTA column, which was then washed with 50mM NaH_2PO_4, 300mM NaCl, 20mM imidazole, pH 8. The protein was eluted in the same buffer supplemented with 200mM imidazole. Purity of each protein was checked on 10% SDS-PAGE by Commasie Blue staining. The purified recombinant proteins were dialyzed against 20mM TrisCl, pH7.5 with 100mM NaCl and 3% glycerol, concentrated using Centricon tubes with 10kDa cut-off filter (Millipore, USA) and quantified using Bradford's Reagent (Bradford, 1976).

Dehydrogenase kinetics / Biochemical assays for isocitrate dehydrogenase

Dehydrogenase activity was measured spectrophotometrically by monitoring the time dependent reduction of NADP+ to NADPH at 25°C in Unicam UV/Vis spectrometer at 340nm, the absorbance maximum of NADPH.



The standard assay solution per 400µL contained 20mM triethanolamine chloride buffer pH 7.5, 2mM $NADP^+$, 0.03mM DL-isocitrate, 10mM $MgCl_2$ / 10mM $ZnCl_2$, 100mM NaCl and 50 -100 pico moles of the enzyme. Environmental parameters for the enzymes were measured by altering the pH of the buffer (range 4 - 10), temperature (20 - 65°C), concentration of substrate (0.01mM - 0.18mM), cofactor (0.1 - 2mM), metal ion (1 - 12.5mM), salt (100mM – 500mM) and metal ion requirement (Mg^{++}, Zn^{++}, Mn^{++}, 10mM each).

The pH dependence of the enzyme was measured using the following buffers: 30mM Na-acetate buffer (pH 4.0 to pH 5.5), 20mM phosphate buffer (pH 5.7 to pH 7), 30mM imidazole buffer (pH 6 to pH 7) and 20mM Tris buffer (pH 7.5 to pH 10). The cofactor specificity was checked with both $NADP^+$ and NAD^+.

The Km was determined by altering the concentration of either the substrate or the coenzyme. The substrate concentration gradient varied from 0.01mM to 0.75mM, while $NADP^+$ concentration was taken from 0.1mM to 2mM. The values were plotted as V vs S for calculating Km and Vmax for this first order reaction. The results were counter checked by double inverse Lineweaver-Burk plot. Competitive inhibition was performed with reduced NADP (NADPH) versus $NADP^+$ to estimate inhibitor constant, Ki. Standard Km analysis was performed followed by repeating the assay with NADPH. Two concentrations of the inhibitor were tested, 0.002mM and 0.005mM. The uninhibited run provided the value of Km for the reaction and the inhibited run provided the apparent Km (Km *app*) for the reaction. Ki for the competitive inhibition was calculated by the formula: Ki = (Km) (I) /Km *app* – Km.

Size exclusion chromatography

Size exclusion chromatography was performed at room temperature using FPLC equipped with Superdex-200 HR 10/30 column (Amersham Pharmacia Biotech). Calibration of the column was performed using protein molecular-mass standards for gel-filtration (Sigma, USA), namely thyroglobulin (669kDa), ferritin (443kDa), β-amylase (200kDa), alcohol dehydrogenase (150kDa), albumin (66kDa) and carbonic anhydrase (29kDa). The column was equilibrated with standard proteins in 50mMTris/HCl, pH7.5 with desired concentration of NaCl and elution volume (V_e) for each protein was determined. The void volume (V_o) was determined by running Dextran Blue on the column. The calibration curve was plotted as Ve/Vo versus log of molecular mass. No difference in the elution volume or the calibration curve was observed when different concentrations of NaCl (No NaCl; 150mM; 300mM; 1.5M) were used

to calibrate the column. The column was equilibrated with three bed volumes of the elution buffer prior to each run. Protein elution was monitored at A_{280}. A 2.4mg/ml (for ICD-1) and 1.06mg/ml (for ICD-2) concentration of recombinant proteins were used for all gel filtration experiments.

Sequence alignment and phylogenetic analysis

The amino acid sequence of *M. tb* ICD-1 and *M. tb* ICD-2 were compared against the NCBI protein database. The sequences with the BLAST score upto e-153 or 65% identity were selected for construction of the phylogenetic tree. MegAlign, the multiple sequence alignment package of DNA*star was used to determine the conserved residues across the aligned sequences. The extremely variable regions at both N- and C-terminal were ignored to further check the relatedness of isocitrate dehydrogenases across species. Only the conserved sequences were aligned and used to construct a phylogenetic tree using CLUSTAL program. The sequence alignment file with PHYLIP to construct the unrooted tree (Altschul *et al* 1990; Retief *et al* 2000).

The amino acid sequences showing more than 65% identity with *M. tb* ICDs were used for CLUSTAL alignment and tree construction. Protein sequence accession numbers for ICDs included in phylogenetic analysis of *M. tb* ICD-1 are: 23326780 [*Bifidobacterium longum*]; 3747089 [*Glycine max*]; 1750380 [*Eucalyptus globulus*]; 2623962 [*Apium graveolens*]; 15982950 [*Prunus persica*]; 5764653 [*Citrus limon*]; 5007084 [*Oryza sativa*];13928690 [*Rattus norvegicus*]; 4105615 [*Microtus mexicanus*]; 284570 [mitochondrial – pig]; 12003362 [*Mus musculus*]; 479431[mitochondrial – bovine]; 4504575 [*Homo sapiens*] ; 2564042 [*Candida tropicalis*]; 2117471 [*Sphingomonas yanoikuyae*]; 16126761 [*Caulobacter crescentus* CB15]; 15889170 [*Agrobacterium tumefaciens*]; 13470817 [*Mesorhizobium loti*]; 23502076 [*Brucella suis* 1330]; 8133104 [*Sinorhizobium meliloti*].

Protein sequence accession numbers for ICDs included in phylogenetic analysis of *M. tb* ICD-2 are: 7649645 [*Streptomyces coelicolor* A3(2)]; 4927874 [*Streptomyces coelicolor* A3(2)]; 15791892 [*Campylobacter jejuni*]; 15676815 [*Neisseria meningitidis* MC58]; 15794063 [*Neisseria meningitidis* Z2491]; 15597820 [*Pseudomonas aeruginosa*]; 15828442 [*Mycobacterium leprae*]; 1170476 [*Colwellia maris*]; 8118494 [*Vibrio cholerae*]; 15641154 [*Vibrio cholerae*]; 1708408 [*Azotobacter vinelandii*]; 15839289 [*Xylella fastidiosa* 9a5c]; 20086389 [*Corynebacterium efficiens*]; 19551894 [*Corynebacterium glutamicum*].

RESULTS

The ORFs encoding *M. tb* hypothetical ICD-1 (Rv3339) and the annotated ICD2 (Rv0066c) are expressed *in vivo*

The *M. tb icd-1*, annotated as the probable isocitrate dehydrogenase based on the sequence homology with other members of isocitrate dehydrogenase family, has so far been considered a hypothetical protein of the bacillus. The *M. tb icd-1* was indeed found to be expressed at mRNA level as evident from RT-PCR analysis (Figure 2.1). Since *M.tb icd-2* is the annotated ORF for which experimental evidence exists; expression at the level of mRNA was not tested. However anti-ICD1 and anti-ICD2 antibodies were detected in the sera of TB-infected patients. ELISA based assay revealed the presence of antibody titers against both the purified proteins ICD-1 and ICD-2 in the infected sera samples. A total of 125 patient sera were tested with 14 healthy controls. Figure 2.2 is a sample representation of ten patients each for ICD-1 (sample 1 –10) and ICD-2 (sample 11 –20). A high immunogenic response against these recombinant proteins was evident when compared to those with BCG-vaccinated healthy sera (Figure 2.2, 21 – 25: healthy control reactions for ICD-1 and 26 – 30: healthy control reactions for ICD-2). These results demonstrated that the hypothetical ORF Rv3339 is indeed transcribed and both the ORFs encoding ICDs are expressed at the protein level *in vivo*.

RT-PCR analysis and serological evidence in favour of expression of the proteins ICD-1 and ICD-2 *in vivo*

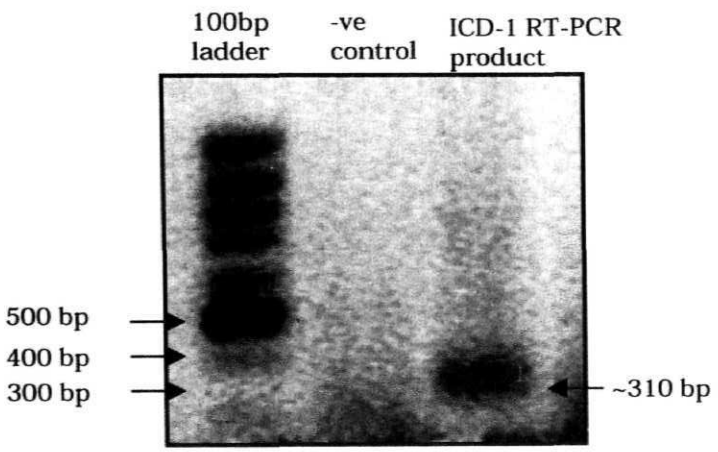


Figure 2.1: RT-PCR of ICD-1 (Rv3339) using RNA extracted from the *Mycobacterium tuberculosis* (H37Rv).

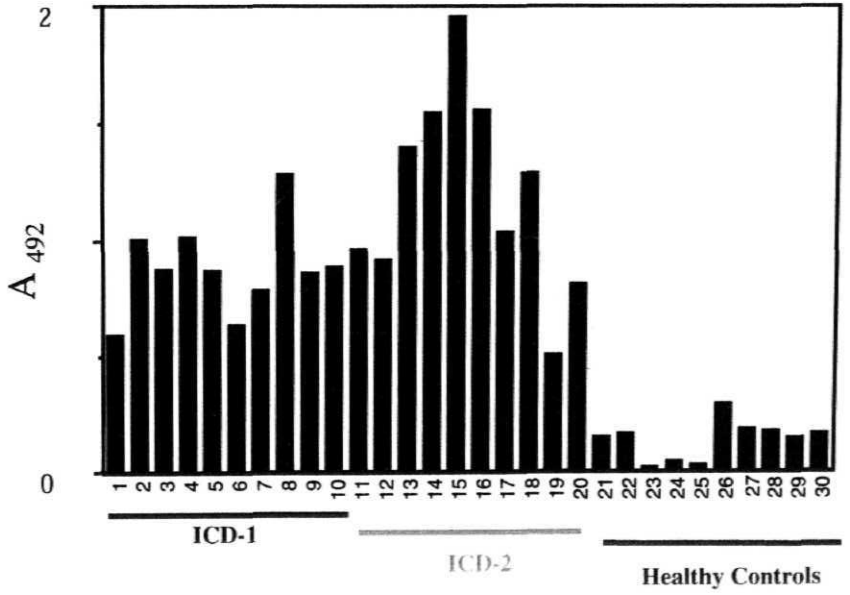


Figure 2.2: Antibody response to *M. tb* ICD-1 and ICD-2 as determined by ELISA.

PCR amplification and cloning of *M. tb icd-1* and *icd-2* in expression vector pRSET A

M. tb icd-1 and *icd-2* were cloned in expression vector pRSET A (Figure 2.3A and 2.3B). The amplified PCR products corresponding to sizes ~1.2 kb (*M. tb icd-1*) and ~2.23 kb (*M. tb icd-2*) was excised from 1.2% agarose gel and extracted (Figure 2.4A). The PCR product and the vector pRSET A were digested with *Bam*HI and *Hind*III and ligated. The ligation mix was used to transform competent DH5α cells and the colonies were screened for the positive clone by double digestion with *Bam*HI and *Hind*III to check for the respective insert fall out (Figure 2.4B).

Expression, purification and quantification of *M. tb* ICD-1 and ICD-2

The genes were overexpressed in the pRSET-A/ *E. coli* BL-21 (DE3) expression system through induction by 0.1mM Isopropyl-beta-D-thiogalactopyranoside. The over-expressed N-terminal His-tagged *M. tb* ICD-1 and *M. tb* ICD-2 were purified on a Nickel affinity column to 95% and 90% homogeneity respectively (Figure 2.5A and 2.5B). The purification was carried out under native conditions for both the proteins from soluble fractions with yields of 3.25mg per 500 ml start culture for *M. tb* ICD-1 and 20.42mg from 1000 ml start culture for *M. tb* ICD-2. The molecular mass of the recombinant *M. tb* ICD-1 and *M. tb* ICD-2 proteins were determined by SDS-PAGE analyses and were found to be ~ 49 kDa and ~ 86 kDa, respectively.

Biochemical characterization reveals differences between ICD-1 and ICD-2

Metal ions as co-factors

The enzymes were tested for metal ion requirement with respect to three divalent metal ions, namely, Mg^{++}, Zn^{++} and Mn^{++}. It was apparent that *M. tb* ICD-1 accepts both Mg^{++} and Zn^{++} as a divalent metal ion components for enzyme activity (Figure 2.6A) but shows no activity in presence of Mn^{++}. This is unlike *M. tb* ICD-2, which accepts only Mg^{++} as metal ion and shows no activity in presence of

Cloning of *M. tb icd-1* and *icd-2* in expression vector pRSET A

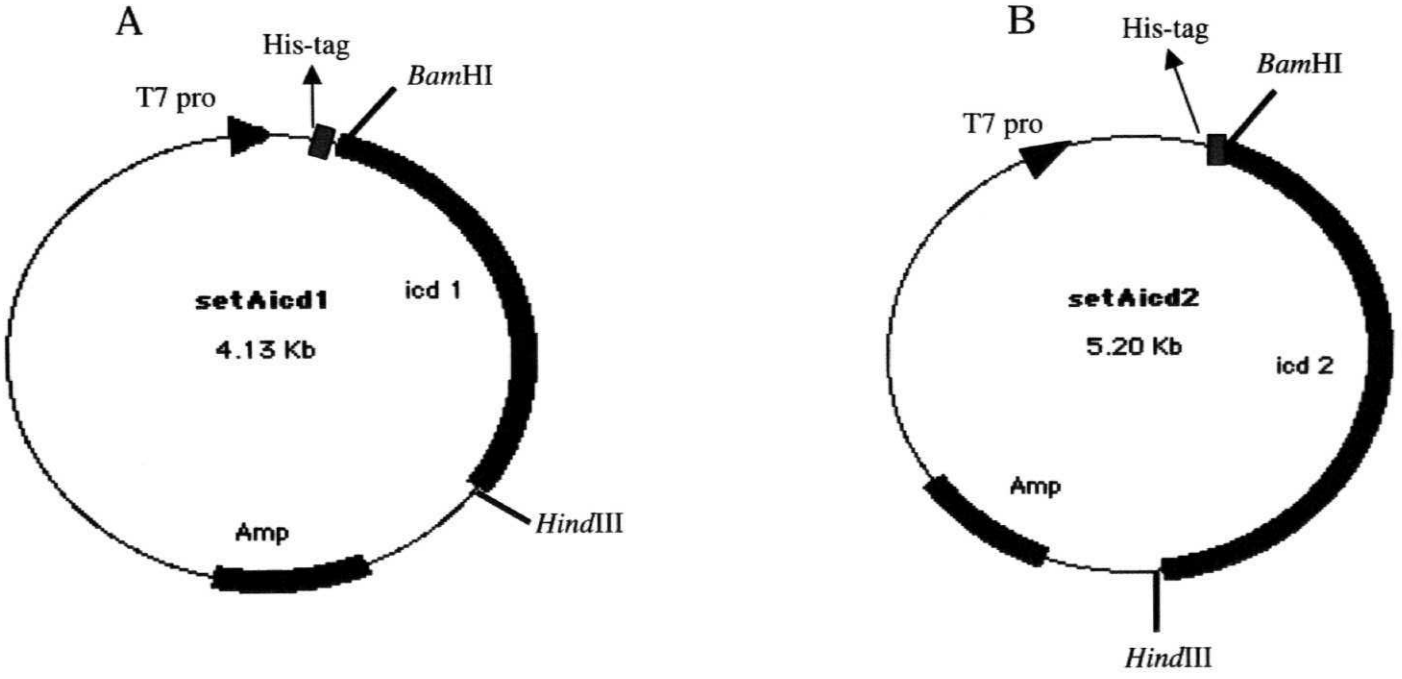


Figure 2.3

PCR amplification and Cloning of *M.tb* ICD-1 and ICD-2 in expression vector pRSET A

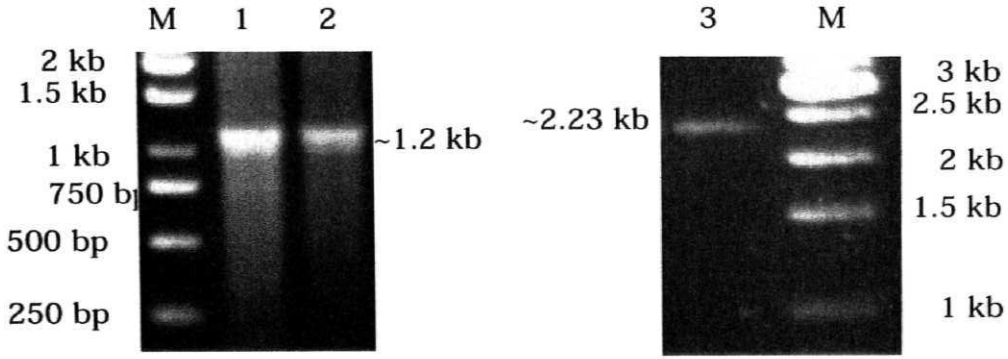


Figure 2.4A: PCR amplification of *icd-1* (Rv3339c) and *icd-2* (Rv0066) from H37Rv genome. Lane M: 1 Kb DNA ladder, Lane 1 and 2: *icd-1* PCR product, Lane 3: *icd-2* PCR product

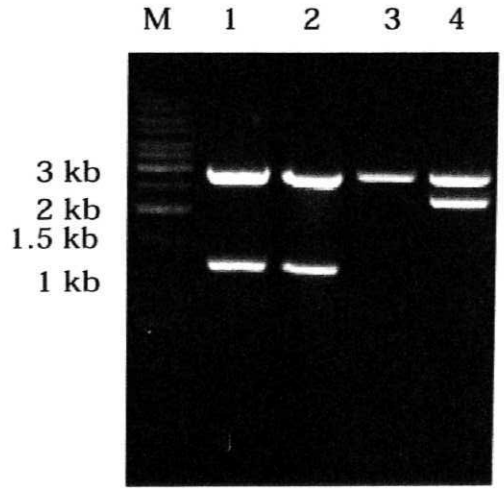
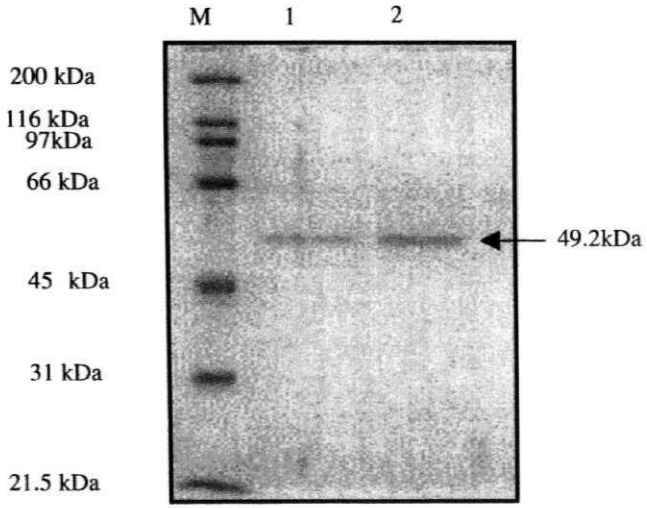
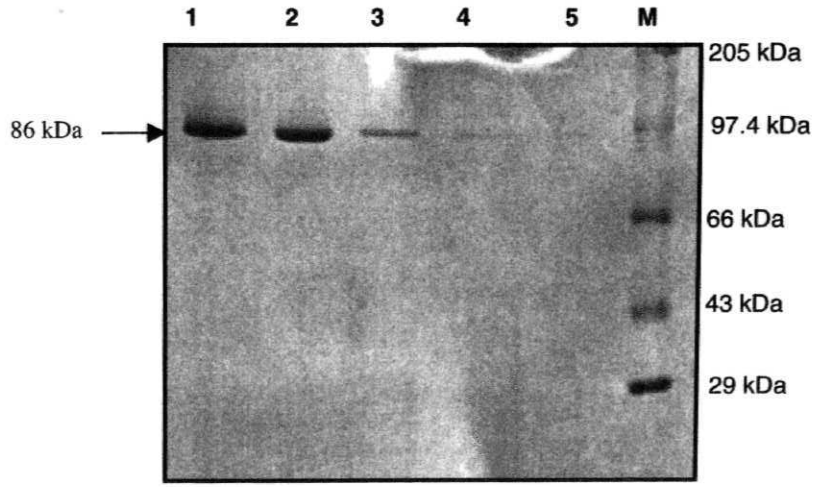


Figure 2.4B: Identification of the positive clone of *icd-1* (Rv3339c) and *icd-2* (Rv0066) in pRSET-A vector by restriction digestion with *Bam*HI and *Hind*III. Lane M: 1 Kb DNA ladder, Lane 1 and 2: *icd-1* cloned in pRSET A, Lane 3: vector pRset A, Lane 4: *icd-2* cloned in pRSET A

Purification of *M. tb* ICD-1 and ICD-2 under native conditions by Ni-NTA column



(A) *M. tb* ICD-1



(B) *M. tb* ICD-2

Figure 2.5: Affinity purification of *M. tb* ICD-1 and *M. tb* ICD-2. Histidine-tagged recombinant protein was purified by nickel column chromatography under native condition and stained with Coomassie Blue following electrophoresis on 10% SDS polyacrylamide gels.
(A) Lane M: protein molecular size markers (200 kDa, 116 kDa, 97 kDa, 66 kDa, 45 kDa, 31 kDa and 21.5 kDa); lanes 1 and 2: purified recombinant *M. tb* ICD-1.
(B) Lane M: protein molecular size markers (205 kDa, 97.4 kDa, 66 kDa, 43 kDa and 29 kDa); lanes1-5: purified recombinant *M. tb* ICD-2 fractions

either Zn^{++} or Mn^{++} (Figure 2.6B). The saturation kinetics indicates a complete saturation at 10mM of metal ion for both Mg^{++} and Zn^{++}. Km [isocitrate] in presence of either Mg^{++} or Zn^{++} was therefore calculated taking 10mM of metal ions in the reaction mix so that metal ion is not limiting for the enzyme kinetics. Km [isocitrate], thus calculated, for *M. tb* ICD-1 were10μM ± 5 in presence of Mg^{++} and 22μM ± 7 in presence of Zn^{++}. For *M. tb* ICD-2, the value is 20μM ± 1 in presence of Mg^{++}. This suggested that *M. tb* ICD-1 has higher affinity for Mg-substrate complex than Zn-substrate and also more readily binds to metal-substrate complex than *M. tb* ICD-2. The observation that these enzymes showed no activity upon pre-incubation with isocitrate and $NADP^+$ in absence of metal ions suggests that they are unable to utilize free isocitrate as a substrate

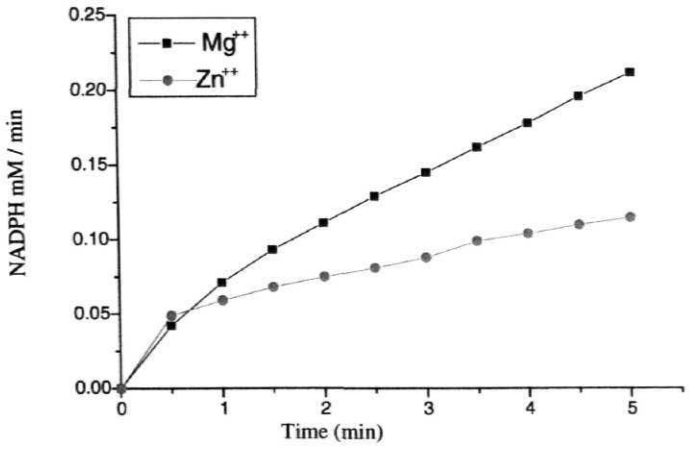
Optimum temperature and thermostability

Activity profiles as a function of temperature of *M. tb* ICD-1 and *M. tb* ICD-2 were studied to interpret optimum temperature and thermostability of the enzymes. The temperature varied over a range of 20°C to 75°C with incubation time of 30 minutes for each reaction. The optimum temperature for the activity of *M. tb* ICD-1 ranged from 37°C to 45°C (Figure 2.7A). The enzyme was partially active till 60°C. The thermal inactivation remained irreversible after incubation at 65°C for half an hour. For *M. tb* ICD-2, activity remained the same across a temperature range of 25°C to ~ 45°C. The enzyme could be renatured upon slow cooling till 55°C where partial activity was restored. Complete denaturation of the protein was observed at about 60°C (Figure 2.7B). These results demonstrate the fine differences between the two ICDs in terms of thermostability.

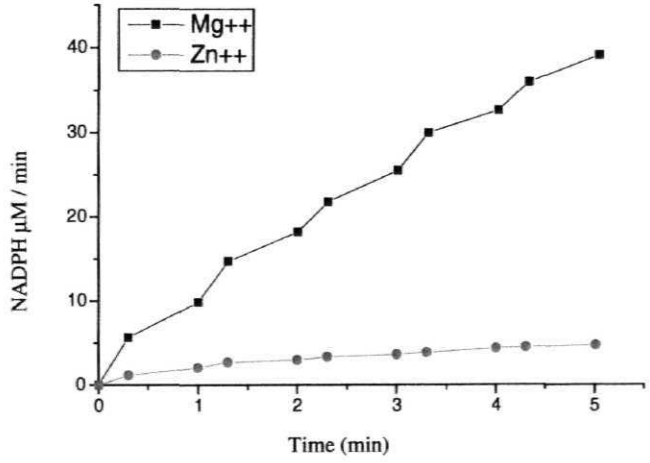
pH tolerance: *M. tb* ICD-1 was active across a wider range of pH than *M. tb* ICD-2

While the optimum pH was found to be 7.5 for *M. tb* ICD-2, *M. tb* ICD-1 was found to be active even at alkaline pH (upto 9) (Figure 2.8). As compared to the physiological pH 7, the tolerance to a low pH range of 5.3 to 5.7 of the reaction buffer by *M. tb* ICD-1 was also

M. tb ICD-1accepts Zn^{+2} ion as metal ion component for catalytic reactions while *M. tb* ICD-2 does not



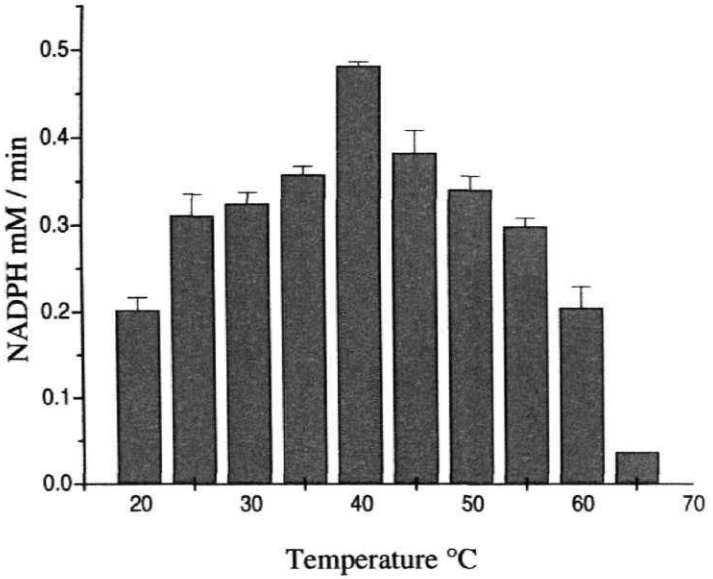
(A) *M. tb* ICD-1



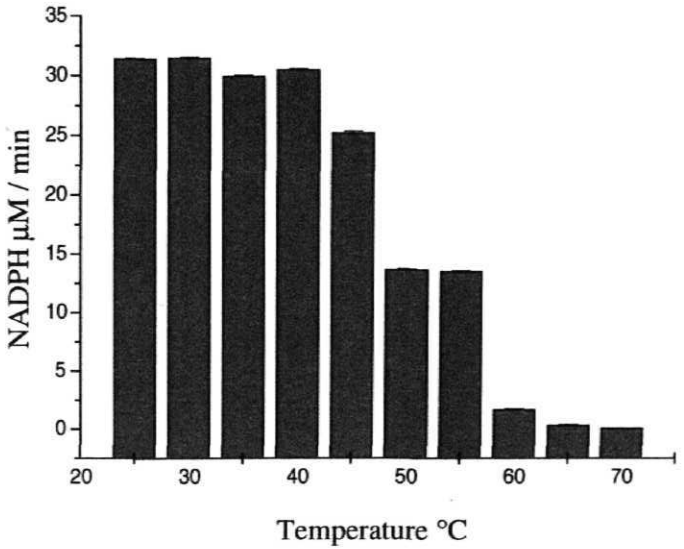
(B) *M. tb* ICD-2

Figure 2.6: Comparative rate curves for the enzyme activities in presence of Mg^{++} and Zn^{++} as cofactors. A) Rate curves of *M. tb* ICD-1 and B) Rate curves of *M. tb* ICD-2 The details of the reactions are described in the method section

Effect of temperature on the activity of *M. tb* ICD-1 and *M. tb* ICD-2



(A) *M.tb* ICD-1

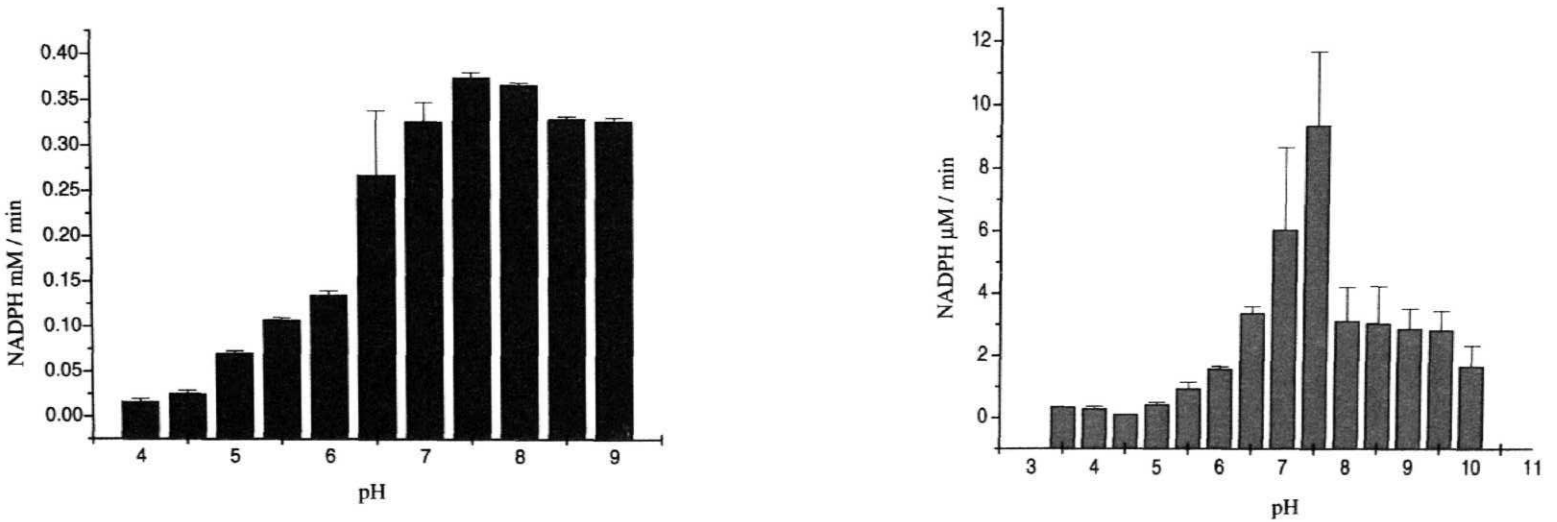


(B) *M.tb* ICD-2

Figure 2.7: ICD-1 and ICD-2 exhibit differential activity as a function of temperature. The enzyme activity was assayed at different temperatures, A) *M. tb* ICD-1, 20°C to 65°C and B) *M. tb* ICD-2, 25°C to 70°C.

The optimum temperature for *M. tb* ICD-1:	37°C to 40°C
Thermal inactivation:	65°C for half an hour
The optimum temperature for *M. tb* ICD-2:	25°C to 40°C.
Thermal inactivation:	60°C- 65°C for half an hour

M. tb ICD-1 was active across a wider range of pH than *M. tb* ICD-2



(A) *M. tb* ICD-1

(B) *M. tb* ICD-2

Figure 2.8: The two ICDs exhibit different pH sensitivity. Activity of *M. tb* ICD-1 (A) and *M. tb* ICD-2 (B) was tested as a function of pH (4 to 10). The buffers used for the experiment were: 30mM Na-acetate buffer (pH 4.0 to pH 5.5), 20mM phosphate buffer (pH 5.7 to pH 7), 30mM imidazole buffer (pH 6 to pH 7) and 20mM Tris buffer (pH 7.5 to pH 10).
M. tb ICD-1: 33 –35 % of the enzymatic activity was retained at a low pH 5.3 to 5.7
M. tb ICD-2: only 10.22% activity at a pH of 5.5 which gradually decreased to only 3.4% at pH 4.

noted where 33 –35 % of the enzymatic activity was retained (Figure 2.8A). In contrast, *M. tb* ICD-2 could retain only 10.22% of activity at a pH of 5.5 which gradually decreased to only 3.4% at pH 4 (Figure 2.8B). Unlike *M. tb* ICD-1, the ICD-2 was marginally active under alkaline pH.

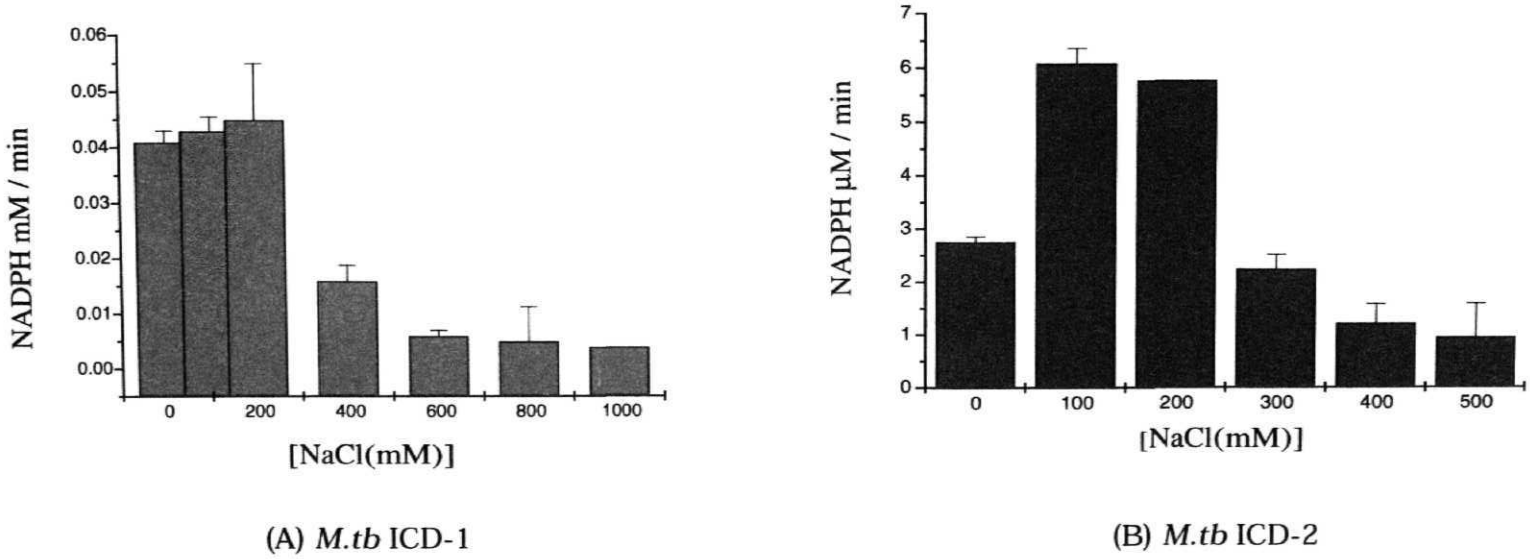
Salt requirements for the stability and activity of the proteins

The effect of NaCl on the stability as well as activity of the enzymes was assessed. The absence of NaCl in the reaction buffer did not affect the activity of *M. tb* ICD-1 drastically but *M. tb* ICD-2 showed higher activity in presence of upto 200mM of NaCl. Presence of salt in the reaction buffer probably maintains the integrity of the enzymes, especially for *M. tb* ICD-2, Higher concentrations of salt (above 200mM) proved to be detrimental for the activities of both the enzymes (Figure 2.9A and 2.9B). The decrease in the activity at higher salt concentration can either be due to the disruption of ionic bonds that is essential for the tertiary structure of the proteins or ionic interactions are involved in catalysis. The reversibility of the enzyme activities after removal of excess of salt has not been tested as the enzymes had a tendency to precipitate upon long exposure to room temperature or even 4°C during dialysis.

Coenzyme specificity

The coenzyme specificity of *M. tb* ICD-1 and *M. tb* ICD-2 was confirmed by checking the activity with both $NADP^+$ and NAD^+ (Figure 2.10A and 2.10B). The activity curves clearly indicates that *M. tb* ICDs are $NADP^+$ - dependent members of the isocitrate dehydrogenase family and shows no activity whatsoever in presence of NAD^+. Km [NADP] values in presence of Mg^{++} for *M. tb* ICD-1 and *M. tb* ICD-2 are 125μM ± 5 and 19.6μM ± 6, respectively. These values verify the stronger affinity of *M. tb* ICD-2 for the coenzyme $NADP^+$ as compared to *M. tb* ICD-1.

The decrease in the activity at higher salt concentration indicates Involvement of ionic interactions during catalysis

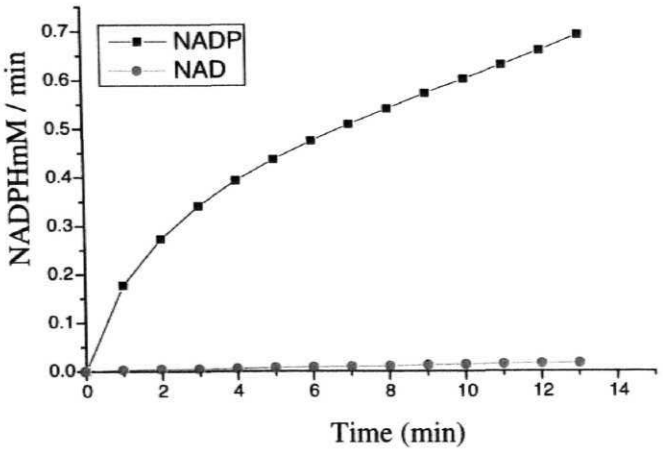


(A) *M.tb* ICD-1

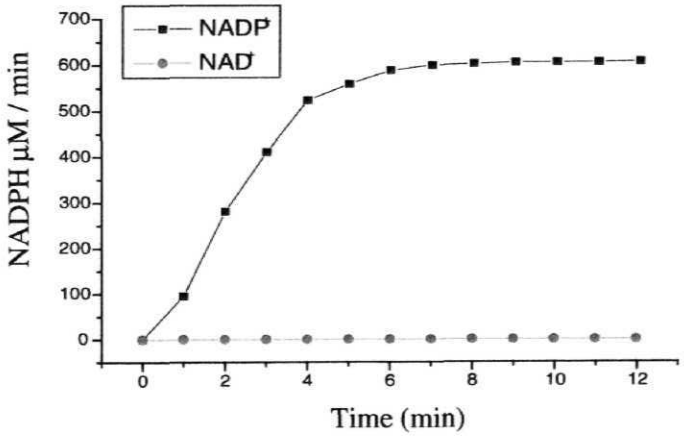
(B) *M.tb* ICD-2

Figure 2.9: Both the ICDs are most active in presence of 200mM NaCl. Activity of the enzymes ICD-1 (A) and ICD-2 (B) was measured by reduction of NADP at different concentrations of NaCl. The enzymes were most active in presence of 200mM NaCl above which the activity rapidly decreases.

M. tb ICDs are NADP+ - dependent members of the isocitrate dehydrogenase family



(A) *M.tb* ICD-1
Km (NADP): 125µM ±5



(B) *M.tb* ICD-2
Km (NADP): 19.6µM ±6

Figure 2.10: The *M. tb* ICDs are NADP+ dependent. Coenzyme specificity of *M. tb* ICD-1(A) and *M. tb* ICD-2. (B) was assayed in presence of NADP+ and NAD+ as coenzymes. The activity curves clearly indicates that *M. tb* ICD-1 as well as *M. tb* ICD-2 are NADP+ dependent enzymes.

The basic enzyme kinetics parameters Km, Vmax and Kcat for DL-isocitrate and $NADP^+$ were determined for both ICD-1 and ICD-2 and have been tabulated below as Table 1.

Table 1: Kinetic parameters, Km and Vmax, for *M. tb* ICD-1 and *M. tb* ICD-2 with respect to Mg^{++} and Zn^{++} as metal ion supplements

Kinetic parameters	With $MgCl_2$	With $ZnCl_2$
***M. tb* ICD-1**		
Km[isocitrate]	10μM ±5	22μM ±7
Vmax [isocitrate]	380μM NADPH/min	190μMNADPH/min
Kcat [isocitrate]	3.8μM NADmPH/min /pico mole enzyme	1.9μMNADPH/min /pico mole enzyme
Km [$NADP^+$]	125μM ± 5	-
Vmax [$NADP^+$]	400μM NADPH/min	-
Kcat [$NADP^+$]	4μM NADPH/min /pico mole enzyme	-
***M. tb* ICD-2**		
Km [isocitrate]	20μM ± 1	No activity
Vmax [isocitrate]	371.3μM NADPH/min	
Kcat [isocitrate]	37.13μM NADmPH/min /pico mole enzyme	
Km [$NADP^+$]	19.6μM ± 6	
Vmax [$NADP^+$]	374μM NADPH/min	
Kcat [$NADP^+$]	37.4μM NADmPH/min/ pico mole enzyme	
Ki (NADPH)	0.46 X 10^{-5} M.	

Feedback inhibition by NADPH

Competitive inhibition was observed with NADPH versus $NADP^+$ for *M. tb* ICD-2 (Figure 2.11). The mean inhibitor constant, Ki, was calculated to be 0.46×10^{-5} M. Since *M. tb* ICD-1 showed low affinity for $NADP^+$, it could neither be efficiently nor significantly inhibited by NADPH. Glyoxylate and oxaloacetic acid, the known inhibitors of isocitrate dehydrogenase, when used in the competition experiments could inhibit both ICD-1 and ICD-2 mediated enzyme reactions.

The recombinant *M. tb* ICDs display different oligomeric state

Size exclusion chromatography was performed to check the oligomeric assembly of the recombinant *M. tb* ICD-1 and *M. tb* ICD-2 (Figure 2.12). The experiment was performed in presence of 20mM TrisCl pH 7.5 and 100mM NaCl. The chromatogram for *M. tb* ICD-1 showed two distinct peaks indicating presence of two oligomeric species in the solution (Figure 2.12A). The first peak corresponds to a mass of ~200 kDa while the major peak showed migration of the molecule as a mass of ~100 kDa. The results showed the existence of the recombinant protein in two existing forms, the minor tetrameric and major dimeric states.

In order to evaluate the functional significance of oligomerization in catalysis of *M. tb* ICD-1, each fraction was collected separately and checked for the activity. While the dimeric fraction showed complete activity, the tetrameric fraction displayed a rather insignificant activity (Figure 2.12A). A few dimeric species that might have originated due to disintegration of the tetramer may account for the feeble activity of the collected tetramer fraction. The monomer form was previously checked for the activity and was found to be inactive. Thus, the oligomeric functional form of the *M. tb* ICD-1 is a dimer.

Despite the fact that *M. tb* ICD-2 is annotated as a monomeric isocitrate dehydrogenase, repeated results from size exclusion chromatography showed a single peak corresponding to ~180 kDa indicating a dimeric state of the enzyme. The dimeric fraction collected showed ICD activity. Even though it could not be

Competitive inhibition of *M. tb* ICD-2 by NADPH indicates the regulation of this enzyme by feedback mechanism

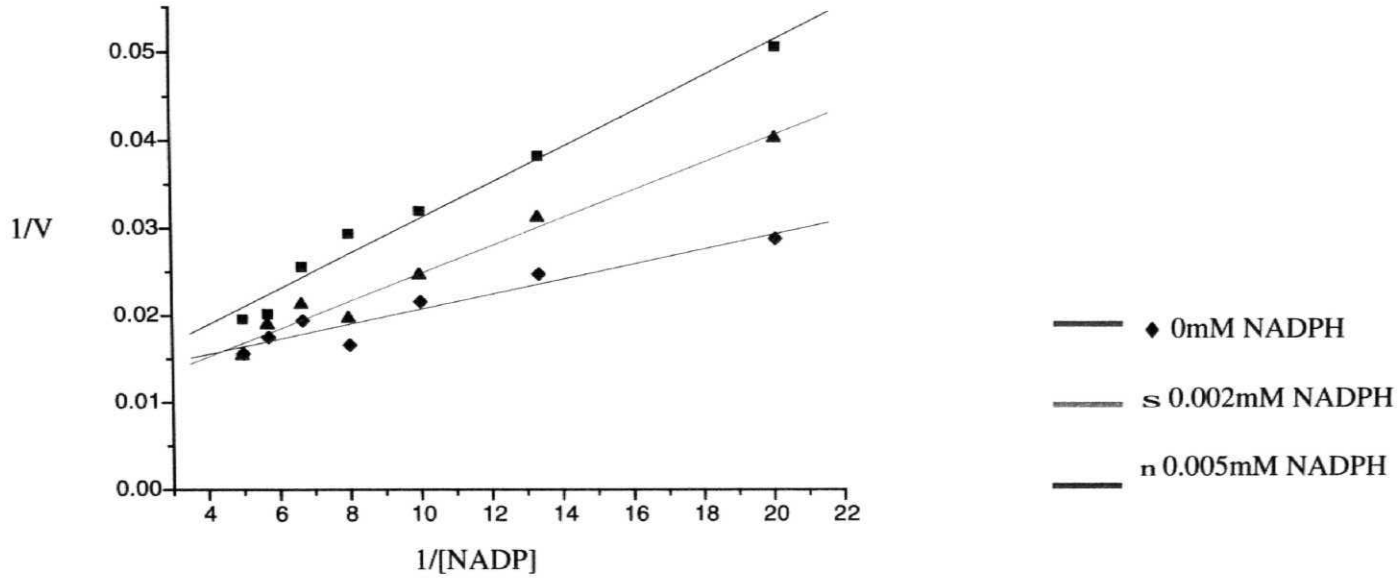
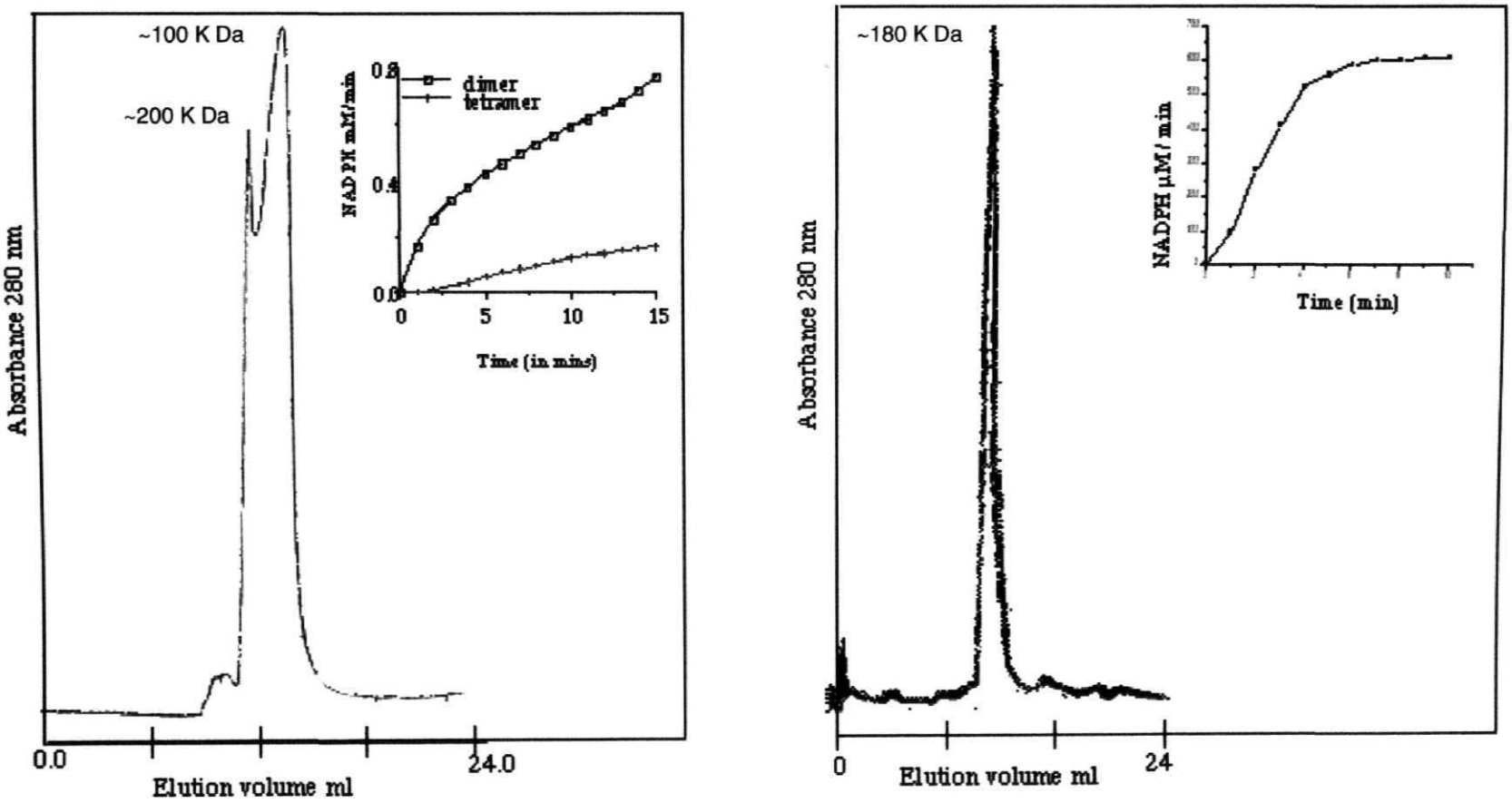


Figure 2.11: Lineweaver Burk plot for competitive inhibition of *M. tb* ICD-2. **Ki(NADPH): 0.46 X 10^{-5} M.**

M.tb ICD-1is a homodimer, while *M.tb* ICD-2, annotated as a monomer exists as a dimer in reaction mixture



(A) *M. tb* ICD-1 (B) *M. tb* ICD-2

Figure 2.12: The ICDs exist in different oligomeric state as evident from gel filtration analysis. The elution profiles of recombinant *M. tb* ICD-1 (A) and *M. tb* ICD-1 (B) on a Superdex-200 HR 10/30 column

determined how critical the dimeric state is for the enzyme activity of the protein, size exclusion chromatography evidence demonstrate the existence of *M. tb* ICD-2 in a dimeric state (Figure 2.12B).

Sequence alignment and phylogenetic analysis of *M. tb* ICD-1 and *M. tb* ICD-2 reveal different phylogenetic lineage

The ICD-1 and ICD-2 were aligned at the protein sequence level, with the corresponding sequences from a range of organism from both prokaryotes and eukaryotes, after a BLAST search. Based on the sequence alignment, phylogenetic analyses were carried out. Results (Figure 2.13) reveal a closer relationship of the functionally conserved residues of *M. tb* ICD-1 with eukaryotic $NADP^+$ dependent isocitrate dehydrogenases rather than those of prokaryotes in general (Figure 2.13). It can also be seen that the major cluster that included *M. tb* ICD-1 also included $NADP^+$ dependent mitochondrial isocitrate dehydrogenase of *Homo sapiens*. The closest prokaryotic neighbour of *M. tb* ICD-1 was *Bifidobacterium longum*. The NADP-dependent isocitrate dehydrogenases of the following prokaryotes were found to cluster with *M. tb* ICD-1: *Spingobium yanoikuyae, Caulobacter crescentes, Agrobacterium tumefacians, Brucella sp., Sinnorhizobium meliloti* etc (Figure 2.13).

CLUSTAL alignment of the amino acid sequences of these closely related ICDs with *M. tb* ICD-1, represented by the discontinuous alignment file shown as Figure 2.14A, 2.14B, 2,14C and 2.14D, revealed a highly conserved region of 378 amino acids starting from 6^{th} to 384^{th} position in *M. tb* ICD-1. This region accounted for almost 80% of the homology between these NADP-dependent ICDs from a wide range of organisms. The N-terminal, however, was highly variable. Putative active site amino acid residues in the modeled protein based on structural homology with that of *E.coli* were found to be highly conserved across these organisms. The conserved amino acids include Mg^{+2}-isocitrate binding residues Arg-103, 135, Asp-278, 282, Asn-99, $NADP^+$-interacting Asp-378, Ser-329, Thr-330, Gly-292, 313, Leu-291, Lys-377 and Ala-311. In *M. tb* ICD-1, Ile-381 that is also predicted to be a part of the $NADP^+$-interacting site is not

M. tb ICD-1 is closer to eukaryotic NADP+ dependent isocitrate dehydrogenases

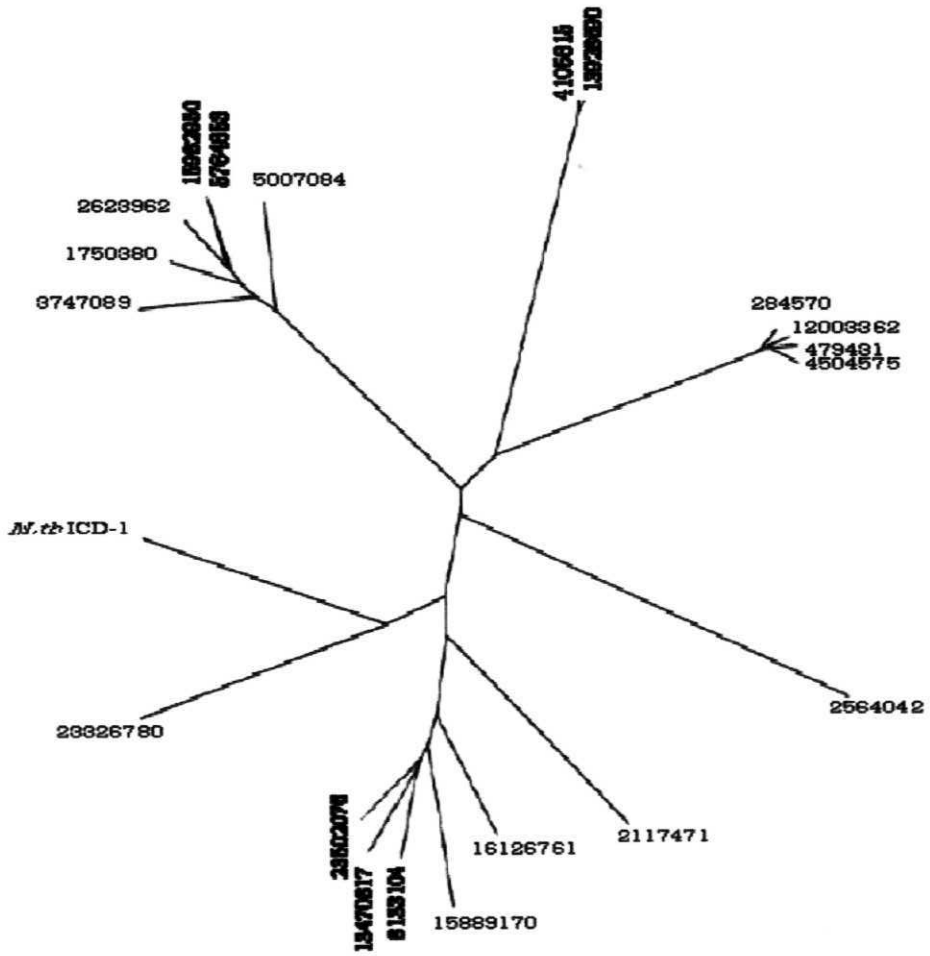


Figure 2.13: Unrooted phylogenetic tree of *M. tb* ICD-1

Catalytically important residues across isocitrate dehydrogenase family are conserved in M. tb ICD-1

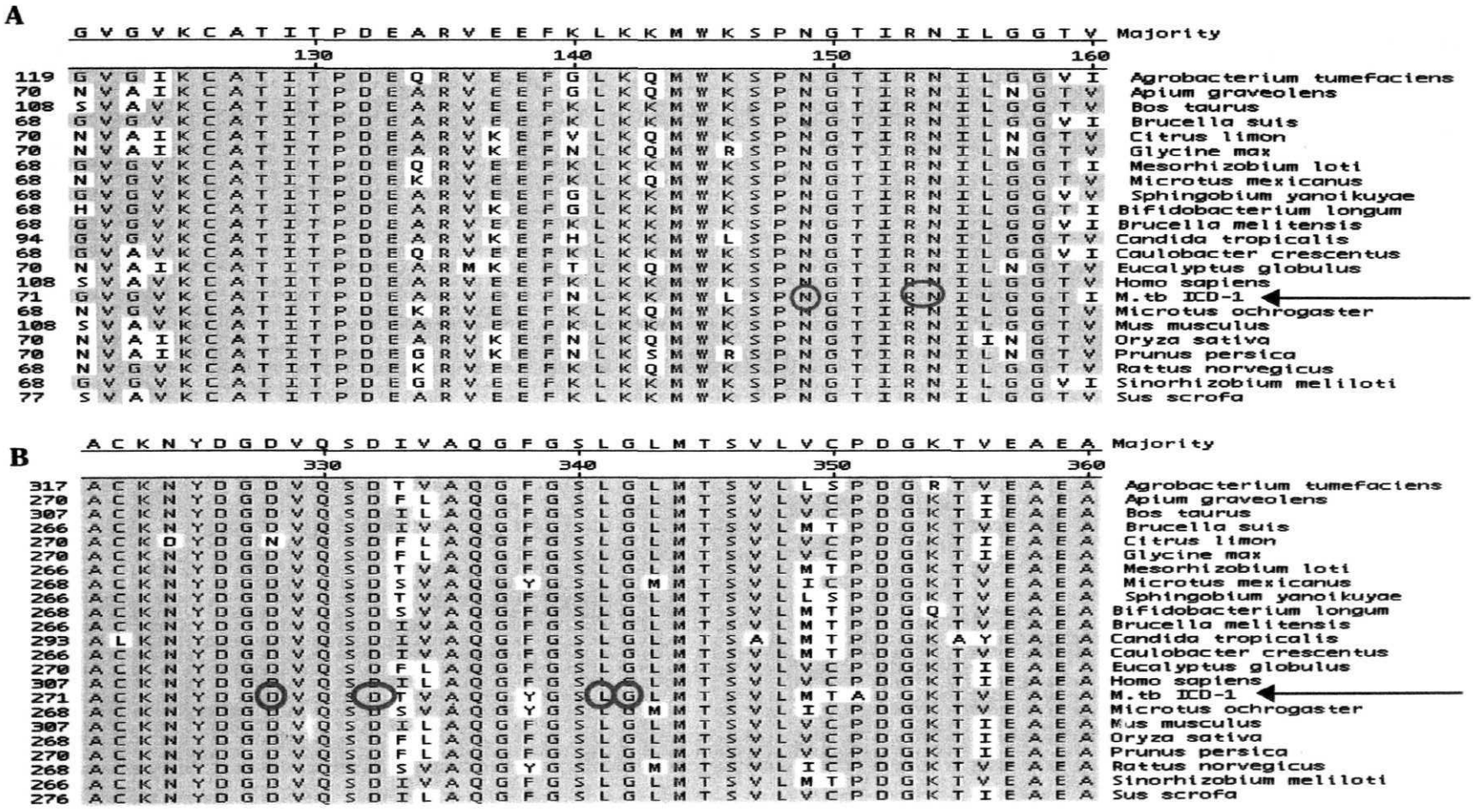


Figure 2.14: Discontinuous sample alignment report of *M. tb* ICD-1 to highlight a few catalytically important conserved amino acid residues: indicating Mg^{+2}-isocitrate binding residues Arg-103, Asp-278, 282, Asn-99. $NADP^{+}$-interacting Gly-292 and Leu-291. **Cont.**

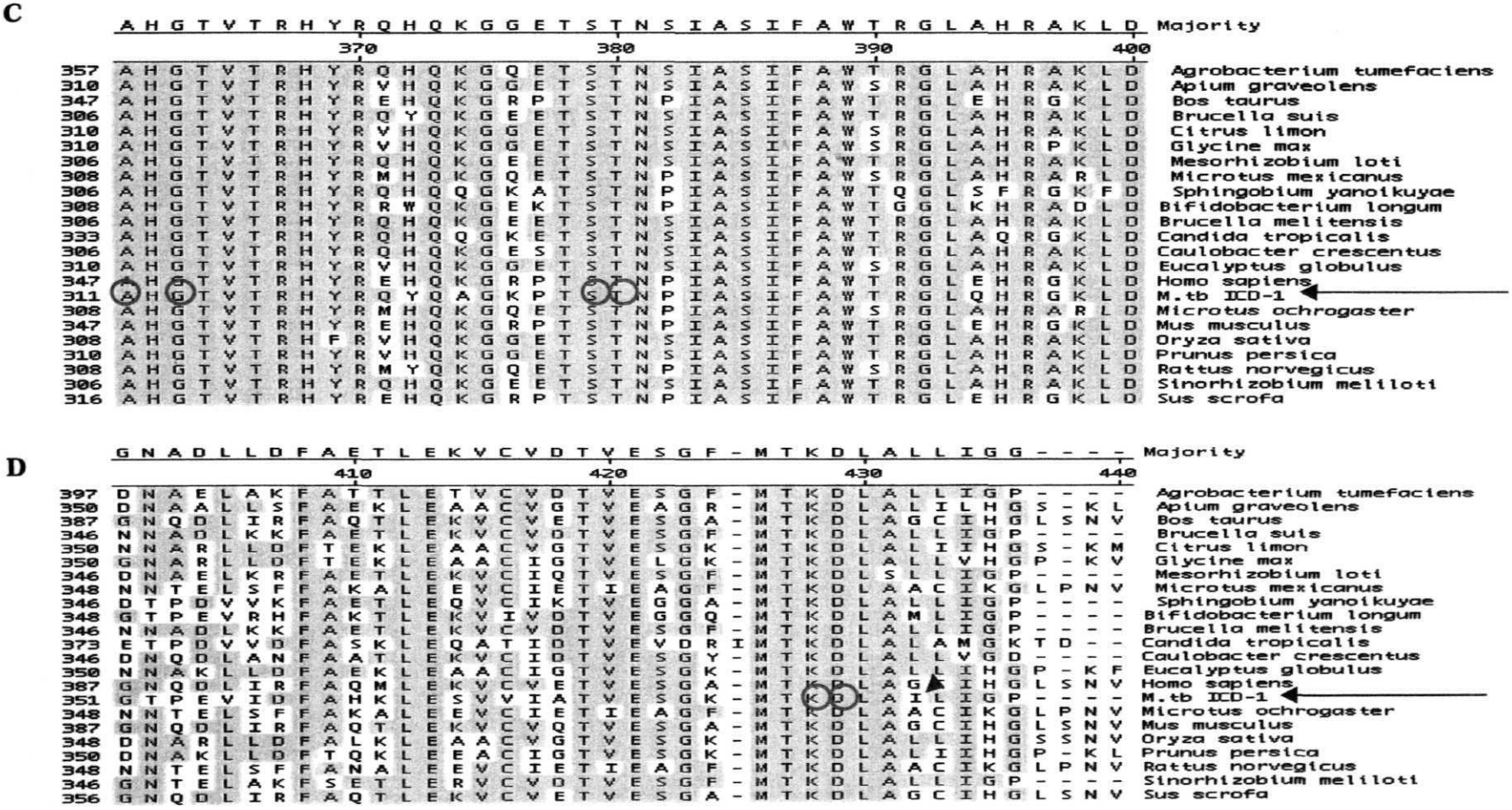


Figure 2.14: Discontinuous sample alignment report of *M. tb* ICD-1 to highlight a few catalytically important conserved amino acid residues. $NADP^+$-interacting Asp-378, Ser-329, Thr-330, Gly- 313, Lys377 and Ala-311. In *M. tb* ICD-1, Ile-381 (marked by an arrow head) that is also predicted to be a part of the $NADP^+$-interacting site is not conserved and is a replacement for the leucine seen as a consensus amino acid in other organisms.

conserved and is a replacement for the leucine seen as a consensus amino acid (Figure 2.14).

The unrooted phylogenetic tree of *M. tb* ICD-2 however, is a total contrast of *M. tb* ICD-1, where it clusters with other $NADP^+$ dependent ICDs of gram-ve bacteria (Figure 2.15). *M. tb* ICD-2 has closest homology with *Mycobacterium leprae* showing 85.4% similarity at the level of primary protein structure. *M. tb* ICD-2 is absent in *E. coli*, but is significantly similar to that of human pathogens like *Pseudomonas aeruginosa* (65%), *Vibrio cholerae* (60%), *Neisseria menengitidis* (59%) etc. CLUSTAL alignment of the amino acid sequences of these closely related ICDs is represented by the discontinuous alignment file shown in Figure 2.16A and 2.16B.

Differential Expression of *M. tb* ICD-1 and *M. tb* ICD-2 at different stages of growth

ICD1 and ICD2 have been reported to be non-essential genes by Himar1-based transposon mutagenesis in H37Rv and CDC1551 strains (Sassetti *et al.*, 2003 and Lamichhane *et al.*, 2003). Since both ICD1 and ICD2 are enzymatically active TCA cycle enzymes *in vitro*, it could be hypothesized that one can compensate for the activity of the other. It was shown that the two isoforms have differences in oligomeric state, coenzyme affinity, pH tolerance and phylogenetic affiliation. Under such observations, a differential expression of the two enzymes under different stages of growth could not be ruled out. Accordingly, the differential expression of *M. tb* ICD-1 and *M. tb* ICD-2 by RT-PCR based assays at different stages (early log, log, late log and stationary phase) of growth in H37Rv was checked. Total RNA was extracted from H37Rv cultures and RT-PCR was performed using internal primers. Phosphoglycerate kinase (Rv1437), a constitutively expressed gene was used as internal control for quantification of RT-PCR products which were run in 2% agarose gel. The experiment pointed to a temporal expression of the two ICDs. ICD-1 is expressed only till the 7^{th} day and reappears alongwith ICD-2 after 10^{th} day. Expression of ICD-1 was less than ICD-2 on 12^{th} and 18^{th} day quantitatively (Figure 2.17). The tolerance

M. tb ICD-2 is closer to prokaryotic NADP+ dependent isocitrate dehydrogenases

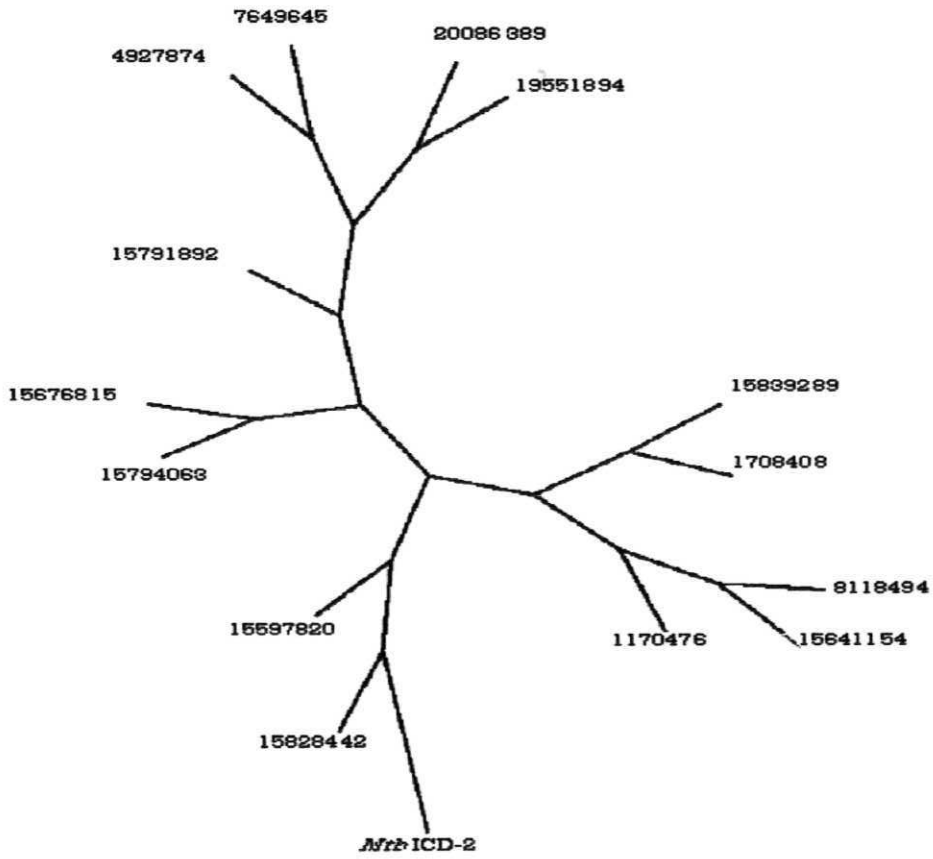


Figure 2.15: Unrooted phylogenetic tree of *M. tb* ICD-2

Multiple sequence alignment file for *M. tb* ICD-2

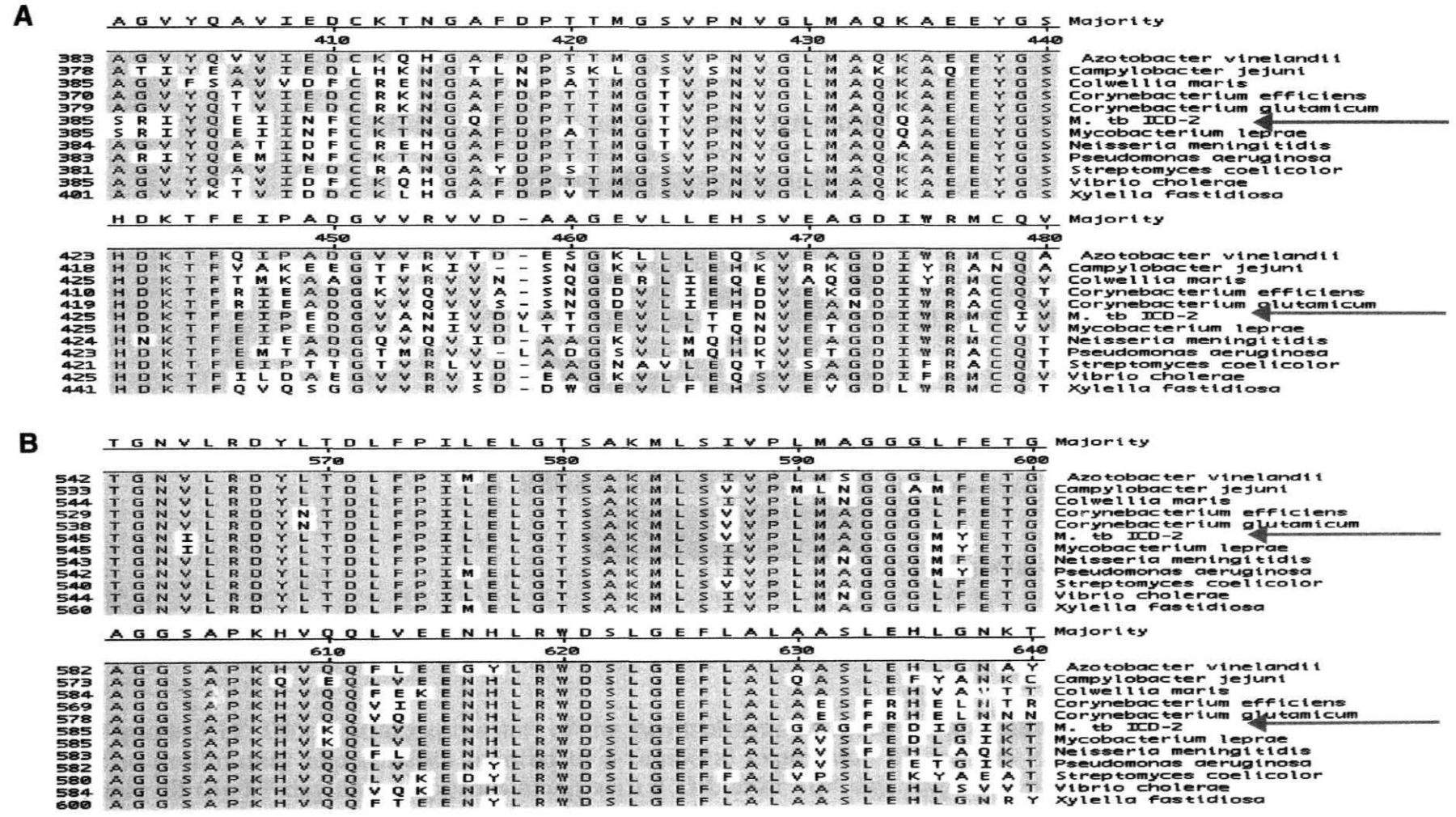


Figure 2.16: CLUSTAL alignment of the amino acid sequences of ICDs closely related to *M. tb* ICD-2 is represented by the discontinuous alignment file shown here. Shaded box represent residues similar to the consensus or majority.

Differential expression of *M. tb* ICD1 and I CD2

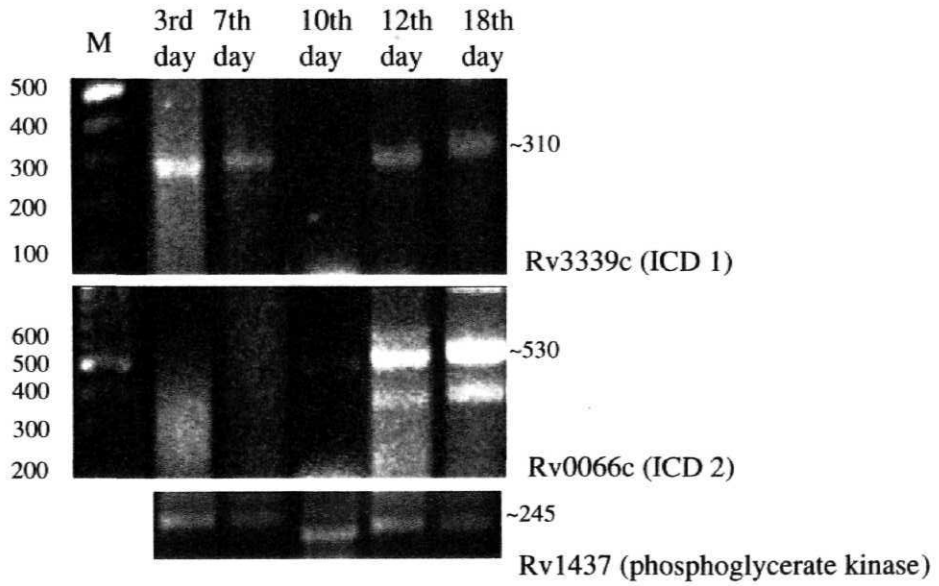


Figure 2.17: Differential expression of the two isoforms of *M. tb* ICDs at different stages of growth. The preliminary experiments pointed to a temporal expression of the two ICDs. ICD1 is expressed only till the 7th day and reappears along with ICD 2 after 10th day. Expression of ICD 1 was less than ICD 2 on 12th and 18th day quantitatively

to a broader range of pH by ICD-1 may be a logical explanation to its reappearance with ICD-2 expression, especially during late log and stationary phases when pH inside the cell is much more variable than log phase.

DISCUSSION

Mycobacterium tuberculosis during latent infection remains isolated in granulomas with limited access to nutrient and oxygen (Saunders *et al*, 1999; Saunders *et al*, 2000). Adaptation of the pathogen for long term survival under such hostile conditions requires coordinated alterations in metabolic network and readjusted energy flux. Evaluation of a nutrient starvation model showed that there is a decrease of TCA cycle enzymes at the protein level (Betts *et al*, 2002). Alterations in the aerobic respiratory apparatus of bacteria during stasis or growth arrest have been reported as an essential part of their defense against oxygen radicals and controlled utilization of endogenous resources (Nystrom *et al*, 1998; Nystrom *et al*, 1996). In *E.coli*, where glyoxylate bypass and citric acid cycle operate concurrently (Holms, 1987), activity of isocitrate dehydrogenase has been reported to be closely monitored. In *M. tb*, however, glyoxylate bypass is observed inside macrophages where C_2 substrate is the main carbon source (McKinney *et al*, 2000). Therefore isocitrate dehydrogenase in *M. tb* lies at a crucial branchpoint where the pathogen decides to either opt for glyoxylate shunt or citric acid cycle. The switchover largely depends upon the alteration in the biochemical parameters of the enzyme isocitrate dehydrogenase during survival inside macrophages. It has been reported that phosphorylation of ICD brings down the Vmax of the enzyme, thereby resulting in accumulation of isocitrate inside cell (Singh *et al*, 2002; Kornberg, 1966). Higher titers of isocitrate lead to its accessibility to isocitrate lyase and energy cycle diverts from TCA cycle to glyoxylate shunt pathway. The occurrence of two isoforms of ICD in *M. tb* genome with the possibility of each having characteristic biochemical properties is interesting under such circumstances.

The biochemical profiles of *M. tb* ICD-1 and *M. tb* ICD-2, in terms of kinetic parameters, substrate specificity, metal ion requirements and oligomeric state represent an attempt to understand the functioning of this very important TCA cycle enzyme. *M. tb* ICD-1 and ICD-2 follow a first order reaction and exhibit a typical "saturation kinetics", that is, at higher concentrations of substrate a point is reached wherein the substrate binding site of every molecule of the enzyme has a molecule of substrate bound to it, hence no further increase in the activity of the enzyme is registered. Km [isocitrate] value of *M. tb* ICD-1 clearly indicates a high affinity of this enzyme for isocitric acid as compared to *M. tb* ICD-2. Km [isocitrate] of some of the known $NADP^+$ dependent ICD has been presented in Table II (page: **61**) for reference and comparison (Kanao *et al*, 2002; Muro-Pastor *et al*, 1992; Reeves *et al*, 1972; Vasquez *et al*, 1981; Chen *et al*, 2000; Balamir *et al*, 1983; Jennings *et al*, 1997; Soundar *et al*, 2000; Dhariwal *et al*, 1987). Several other probable substrates for *M. tb* ICD-1 with close structural similarity were tested. Two such substrates were aspartate and glutamate. Both the compounds have close similarity with isocitric acid. However, the poor activity of the enzyme with these substrates confirmed that the enzyme is specific to only isocitrate as a substrate.

Km [$NADP^+$] of *M. tb* ICD-1 showed poor affinity for the co-factor $NADP^+$ as compared to *M. tb* ICD-2 and other known NADP-dependent isocitrate dehydrogenases (Table II). The poor affinity of *M. tb* ICD-1 to $NADP^+$ warranted an investigation on whether dual co-factor specificity occurs in *M. tb* ICD-1 as reported in some archaeal bacteria (Steen *et al*, 2001). Therefore the enzymatic activity of both the enzymes in presence of $NADP^+$ and NAD^+ was compared (Figure 2.10A and 2.10B). It can be clearly seen that *M. tb* ICD-1 as well as *M. tb* ICD-2 accepts $NADP^+$ as a co-factor and do not acquire NAD^+ as a proton acceptor thereby pointing to the fact that *M. tb* ICD-1 and *M. tb* ICD-2 are $NADP^+$ dependent isocitrate dehydrogenases.

The range of tolerance for the enzyme to different pH conditions (Figure 2.8A and 2.8B) was also studied. It was noted that even

though optimum pH for enzyme activity was 7.5, *M. tb* ICD-1 could tolerate pH as low as 5.5. The enzyme is not completely inactivated at lower pH, a property that holds significance in terms of immediate survival of the pathogen before activation of proton pumps when engulfed by phagocytic macrophages. Even though, *M. tb* is known to prevent acidification of phagosomes and prevent lysosomal fusion (Clemens *et al*, 1995; McDonough *et al*, 1993; Sturgill-Koszycki *et al*, 1994) it initially encounters a low pH (pH 5 – 5.5) of the phagosomes. It therefore appears that the broad pH range of *M. tb* ICD-1 could be an interesting feature of this enzyme. This is in contrast to *M. tb* ICD-2 that has a narrow range of pH tolerance. Competitive inhibition of *M. tb* ICD-2 by NADPH indicates the regulation of this enzyme by feedback mechanism (Figure 2.11). Owing to poor affinity for NADP by ICD-1, ICD-1 did not show such kind of inhibition.

The homodimeric state of *M. tb* ICD-1 is the functionally active species, even though residual activity was noticed in tetrameric fraction which could be a reflection of the presence of a few dimeric species as a consequence of disintegration of the tetrameric forms. Despite sequence homology with monomeric ICDs, the ICD-2 generated an exclusive peak corresponding to a dimer as seen in the chromatogram (Figure 2.12B). The fraction corresponding to the peak was collected and checked for activity as well as for monomeric status on SDS-PAGE. The protein was found to be catalytically active and the size corresponded to monomeric 83 kDa.

Classical nomenclature of isocitrate dehydrogenase family categorizes the enzyme from different organisms into three major sub-families (Steen et al, 2001). Accordingly, M. tb ICDs come under subfamily I along with other archaeal and bacterial ICDs with preference to $NADP^+$. The earlier work reported by Steen and co-workers do not describe the affiliation of M. tb ICD-2. The results presented here on the phylogenetic analysis of M. tb ICD-1 revealed a closer relationship with eukaryotic $NADP^+$ dependent ICDs (Figure 2.13) with more than 65% identity with that of Glycine max, Sus scrofa, Bos and Homo sapiens. M. tb ICD-1, indeed, is correctly

Table II: Comparison of $NADP^+$-dependent isocitrate dehydrogenases from different organisms.

	Km(isocitrate)	Km($NADP^+$)
Blastocladiella emersonii	20μM	10μM
Chlorobium limicola	45μM ± 13	27μM ±10
Bacillus substilis	5.9μM ± 0.9	14.5μM ± 2.2
Escherichia coli	4.9μM ± 0.2	19.6μM ± 3.6
Synechocystis sp PCc 6803	59μM	12μM
Pyrococcus furiosus	-	4400μM
Aeropyrum pernix	-	30μM
Thermotoga moritima	-	55.2μM
Aeropyrum fulgidus	-	30μM
Mycobacterium phlei (ATCC-354)	74μM	53μM
***M. tb* ICD-1**	10μM ±5	125μM ± 5
***M. tb* ICD-2**	20μM ± 1	19.6μM ± 6
Beef liver $NADP^+$-IDH	1.7μM	7.3μM
Rat liver (cytosolic)	9.7μM ± 2.9	11.5μM ± 0.2
Porcine heart $NADP^+$-IDH	-	5μM ± 0.19

placed in subfamily II that includes eukaryotic $NADP^+$ dependent ICDs and a single bacterial ICD (Sphingomonas yanoikuyae) (Steen et al, 2001). With $NADP^+$ dependent isocitrate dehydrogenase of Sphingomonas yanoikuyae, M. tb ICD-1 has more than 65% identity at primary structure level. Phylogenetic analysis of *M. tb* ICD-2 showed that the classical nomenclature applies to ICD-2 and it can be placed in subfamily I, the closest being *M. leprae* (Figure 2.15). The closest bacterial relative of *M. tb* ICD-1 as inferred by our study is $NADP^+$ dependent isocitrate dehydrogenase of *Bifidobacterium longum* (Figure 2.13). *Bifidobacterium* sp. are gram positive, anaerobic, natural components of human intestinal microbiota (Xu *et al*, 2003). This might be argued as a case of horizontal transfer or lateral transfer of gene amongst unrelated organisms across the boundaries of phylogenetic domains. Horizontal transfer of genes is a common occurrence in nature and accounts for almost 10-50% of genes in bacteria (Schnarrenberger *et al*, 2002; Lorencz *et al*, 1994).

ICD1 and ICD2 have been reported to be non-essential genes by Himar1-based transposon mutagenesis in H37Rv and CDC1551 strains (Sassetti *et al.*, 2003 and Lamichhane *et al.*, 2003). The present result show that both ICD1 and ICD2 are enzymatically active TCA cycle enzymes *in vitro*, therefore, can compensate for the activity of the other when one is mutated or knocked out. It was also shown that the two isoforms have differences in oligomeric state, coenzyme affinity, pH tolerance and phylogenetic affiliation. Under such observations, a differential expression of the two enzymes under different stages of growth could not be ruled out. The co-expression of ICD-1 and ICD-2 during late log phase and stationary phase was observed (Figure 2.17). The tolerance to a broader range of pH by ICD-1 may be a rational explanation to its reappearance with ICD-2 expression when pH inside the cell is much more variable than log phase. However, ICD-2 seems to be the primary isocitrate dehydrogenase expressed by *M. tb* during log phase (Figure 2.17).
The data presented represent the first attempt to characterize this important member of the TCA cycle of *Mycobacterium tuberculosis*.

Conclusively, our studies reveal that both ICD-1 and ICD-2 are $NADP^+$ dependent and the former is a homodimer having closer homology with eukaryotic ICDs while *M. tb* ICD-2, unlike the expected monomeric state, exists as a dimer in the reaction mixture. The two isoforms differ in their affinity for coenzyme NADP as represented by their Km[$NADP^+$] values (Table I) and also with respect to pH tolerance and phylogenetic affiliations. *M. tb* ICD-2 is a more efficient enzyme as inferred by comparing Vmax[$NADP^+$]/Km[$NADP^+$] ratios for the two enzymes. The two isoforms also show differential expression during different stages of growth. Additionally, these results reflect an attempt to trace the evolution of ICDs and may help to understand the adaptive role of isocitrate dehydrogenase in intracellular persistence of this pathogen.

Chapter 3

Immunological Characterization of *Mycobacterium tuberculosis* Isocitrate Dehydrogenases

INTRODUCTION

Mycobacterium tuberculosis, the causative organism for TB, has evolved as one of the most efficient pathogen in the history of mankind and despite combinatorial drug regime, remains one of the major cause of mortality and morbidity factor. Roughly 2 – 3 millions deaths every year are attributed to TB worldwide (Dye *et al*, 1999; Bloom *et al*, 1992). It can practically remain undetected in lungs for decades, efficiently escaping detection by the host immune system. Only in 10% of the infected people, the number being higher in immuno-compromised patients, TB erupts as a full-blown disease (Helmuth, 2000). Every year 18 lakh (or 1.8 million) people develop TB in India, of which nearly 8 lakh (0.8 million) are infectious (sputum-positive) and the remaining 10 lakh remain asymptomatic (RNTCP report, 2004). Delay in diagnosis of these asymptomatic patients and hence their treatment impedes the downstream management and control of the disease. These patients act as potential reservoir of *M. tb* and spread infection in the community, infecting 10-15 persons in a year unless effectively treated. With the increasing emergence of multi drug resistant strains and co-infection with HIV the problem is getting further compounded (Dye *et al*, 2002; Siddiqi *et al*, 2002; Ahmed *et al*, 2003). Early diagnosis, therefore, is a matter of utmost concern not just for TB disease management but also for epidemiological investigations (Ahmed *et al*, 2004). Current diagnostic tools for tuberculosis often lack sensitivity and can be time consuming. TB diagnosis in developing countries largely banks upon tuberculin skin test and staining and culture methods. The epidemiological relevance of tuberculin test with purified protein derivative (PPD) is questionable in areas where BCG vaccination is compulsory because PPD is not sensitive enough to distinguish between vaccinated and infected individual (Roche *et al*, 1994). Microscopic determination of the bacilli in the sputum samples is a direct way of examining pulmonary tuberculosis (Dye *et al*, 2002). This however requires high titers of bacilli (5000 – 10000 / ml) in sputum – a condition seen only in full blown tuberculosis patients.

Culture techniques can detect very low titers but are time consuming taking approximately 3 – 6 weeks (Laidlaw, 1989).

Since humoral response is thought not to contribute to resistance in tuberculosis (Kaufmann, 1989), B cell activation and reactivity to *M. tb* antigens have not yet been explored substantially. However, the importance of the major extracellular proteins of the pathogen as candidate components of a subunit vaccine has been reported earlier (Horwitz *et al*, 1995). Current discovery of the RD1 locus in the *M. tb* genome, encoding mainly the proteins actively secreted by mycobacteria into the culture medium, such as CFP-10 and ESAT-6, have further encouraged immunological tests as an adjunct to conventional diagnosis (Mustafa, 2002; Louise *et al*, 2001; Trajkovic *et al*, 2004; Mori *et al* 2004).

Proteins that are released from *Mycobacterium tuberculosis* during late logarithmic growth phase, such as, superoxide dismutase and isocitrate dehydrogenase are employed as autolysis markers (Anderson *et al*, 1991). The use of isocitrate dehydrogenase as a

potential antigen for serodiagnosis along with malate dehydrogenase has been suggested (Ohman *et al*, 1996; Florio *et al*, 2002). The *Mycobacterium tuberculosis* genome carries two isoforms of isocitrate dehydrogenase, *M.tb* ICD-1 and *M.tb* ICD-2. The relevance of ICDs as potent immunogenic markers for TB infection has been evaluated through detection of anti-*M.tb* ICD antibody in sera of different well characterized categories of TB patient using enzyme linked immunosorbent assays. This chapter describes the results of experiment on the sensitivity and specificity of ICDs to distinguish TB patients from those vaccinated with BCG, and also from those patients infected with non-tuberculous mycobacteria or other pathogens *vis-à-vis* the conventional antigen – HSP 60 (Perschinka *et al*, 2003) and purified protein derivative (PPD).

MATERIALS AND METHODS

Cloning, expression and purification of *M.tb* ICD-1 and *M.tb* ICD-2: The ORFs, corresponding to *M.tb* ICD-1 (Rv3339c, 1.230 kb) and *M.tb* ICD-2 (Rv0066c, 2.238 kb) were PCR amplified from the genomic DNA of H37Rv. *Bam*HI and *Hin*dIII restriction sites were incorporated in the 5' end of forward and reverse primers respectively for both *M.tb* ICD-1 and *M.tb* ICD-2. The primers and parameters for thermal cycle amplification have been provided in materials and methods section of chapter 2.

Each amplicon, *M.tb* ICD-1 and *M.tb* ICD-2, were cloned at the *Bam*HI and *Hin*dIII sites of the expression vector pRSET-A (Invitrogen, USA) with six-histidine sequence tag at N-terminal. The generated constructs 'setAicd1' and 'setAicd2' were further transformed into the BL21 (DE3) strain of *E.coli*. The clones were confirmed by sequencing using the T7 promoter primer, on an ABI prism 377 DNA sequencer (PE Biosystems, USA).

The genes were over-expressed in the pRSET-A/ *E.coli* BL-21 (DE3) expression system. The over-expressed his-tagged recombinant protein was purified by Ni^{2+}-nitrilotriacetate affinity chromatography. The cells transformed with the constructs were grown in Terrific Broth (TB) containing ampicillin (100μg /ml) to an OD_{600} of 0.4 to 0.5 at 37°C, cooled to 27°C, induced with 0.1mM isopropyl β-D-thiogalactoside and grown overnight at 27°C. The cells were lysed by sonication, followed by centrifuging at 13000 rpm for 30 minutes at 4°C. The clear lysate was loaded onto Ni^{2+}-NTA column, which was then washed with 50mM NaH_2PO_4, 300mM NaCl, 20mM imidazole, pH 8. The protein was eluted in the same buffer supplemented with 200mM imidazole. The proteins were 90-95% pure as seen on 10% SDS-PAGE followed by Commasie Blue staining. The purified recombinant proteins were dialyzed against 20mM TrisCl, pH7.5 with 100mM NaCl and 3% glycerol and quantified using Bradford Reagent (Bradford, 1976).

Human Sera: The study population (n= 215) comprised of the *M.tb* infected human sample population reporting to Mahavir Hospital and Research Centre, Hyderabad and Central JALMA Institute for Leprosy, Agra. These were categorized into three groups, namely group 1 (n=42 patients), group 2 (n=32 patients) and group 3 (n=35 patients). In addition to the above 44 clinically healthy donors, 30 NTM cases and 32 non - TB patients who were proven culture negative for acid fast bacteria were also included as controls in this study. Group 1 comprised of patients with fresh infection with no history of TB treatment. Group 2 comprised of patients with relapsed cases, i.e. those who were treated earlier for tuberculosis but the symptoms re-emerged after the completion of the treatment. Group 3 included patients with extra-pulmonary tuberculosis. Group1 and group 2 patients were diagnosed by sputum examination (acid-fast bacillus smear positive and negative) while the extra-pulmonary cases were confirmed by tissue biopsy. Clinically healthy donors were *M. bovis* BCG vaccinated and had no symptoms of TB at the time of sera collection. Randomly picked individuals from the population of healthy controls were subjected to PCR test for TB and were found to be PCR negative. Mycobacteria other than *M. leprae* that are not included in the *M. tuberculosis* complex are referred to as nontuberculous mycobacteria or NTM (Saiman, 2004). However, the group referred to as NTMs in this study included sera collected from patients infected with non-tuberculous mycobacterial species (n= 14), such as, *M. avium*, *M. xenopi*, and *M. fortuitum* as well as sera from patients with *M. leprae* infection (n= 16). The non- TB patient category included infected individuals who were tested negative for acid-fast bacteria by staining and culture based techniques. These patients were also negative for HIV and HBV. These randomly picked patients were suffering from either pneumonia, lower respiratory infections, septicemia, urinary tract infections, gastrointestinal infections, cirrhosis or fever of unknown origin. The study population had no sex or age bias. This study was approved by the Institutional Ethics Committee.

Immunosorbent assays: Enzyme linked immunosorbent assays (ELISA) were performed to check the B cell immune response in human to the *M.tb* ICD-1 and ICD-2 proteins and control antigen HSP 60 and PPD. HSP 60 used was *M. tb* HSP 65/GroEL. In brief, the 96 well microtitre plates (Corning, Costar, USA) were coated with ~500ng of either control antigens or recombinant *M.tb* ICD-1 and *M.tb* ICD-2. The plates were incubated overnight at 4°C, washed thrice with phosphate buffer saline (PBS) and blocked with 100 µl of blocking buffer (2% BSA in PBS) for 2 hour at 37°C. The plates were then washed thrice with wash buffer PBST (0.05% Tween 20 in 1 X PBS). The *M. tuberculosis* infected human sera belonging to different clinical groups were diluted 200 times in blocking buffer (1% BSA in PBS). 50 µl of sera were added to antigen coated wells followed by incubation for 1hr at 37°C. The plates were thoroughly washed with PBST and further incubated with anti-human IgG-horseradish peroxidase (HRP) (Sigma, USA) at 37°C for 1hr. HRP activity was detected using a chromogenic substance *o*-phenylenediamine tetrahydrochloride (Sigma, USA) in citrate-phosphate buffer (pH 5.4) and H_2O_2 (Qualigens, India) as 1 µl/ml. The reactions were terminated using 1N H_2SO_4, and the absorbance values were measured at 492 nm in an ELISA reader (BioRad, USA). Each ELISA was repeated at least twice with some randomly picked sera samples tested thrice for confirmation, with and without replicates for each sample within individual ELISA.

Data analysis: t-test was performed to compare the means of two variable groups, healthy and infected groups, using the online scientific calculator of GraphPad (http://www.graphpad.com/quickcalcs/ttest1.cfm) to calculate means, standard error of means (SEM) and p values.

RESULTS

Expression and purification of *M.tb* ICD-1 and *M.tb* ICD-2

The over-expressed N-terminal His-tagged *M.tb* ICD-1 was purified to 95% homogeneity on a Nickel affinity column (Figure 2.5A, chapter 2). The molecular size of the recombinant ICD-1 was determined to be 49.2 kDa. The purification was carried out under native conditions from soluble fraction with an yield of 3.25 mg protein per 500 ml of start culture. Similarly *M.tb* ICD-2, a 83 kDa protein, was purified to 90-95% homogeneity (Figure 2.5B, chapter 2) with an yield of about 20.4 mg per 1000 ml of start culture.

M.tb ICD-1 and *M.tb* ICD-2 show high reactivity to patient sera as opposed to BCG-vaccinated healthy controls

Humoral immune responses directed against the *M. tb* ICD-1 and *M. tb* ICD-2 were compared between patients with tuberculosis (including all groups) and BCG-vaccinated healthy controls (Figure 3.1A and 3.1B). The recombinant proteins were used to screen the infected and the healthy sera by ELISA using anti-human IgG-HRP as conjugates. A similar exercise was performed for non-TB patient sera that were essentially culture negative for tubercle bacillus, and also patients with atypical TB (NTMs) the results for which have been discussed later. The patient sera were also tested against HSP 60 and the purified protein derivative (PPD) (Figure 3.1C and 3.1D). The immunoreactivity of ICD-1, ICD-2, HSP 60 and PPD were statistically analyzed and compared with respect to both infected and healthy sera (Table III). The mean antibody titers as depicted in terms of ELISA reactivity for *M. tb* ICD-1 from TB patients and healthy controls were 0.49466 ± 0.02464 and 0.27631 ± 0.01574 respectively, with p value < 0.0001. The similar values for *M. tb* ICD-2 from TB patients as against healthy controls were 0.23239 ± 0.01762 and 0.08776 ± 0.00595 respectively, with p value < 0.0001. These data demonstrate that sera of all the infected patients mounted a statistically significant ($p < 0.0001$) antibody response against recombinant proteins *M. tb* ICD-1 and *M. tb* ICD-2 as

M.tb ICD-1 and ICD-2 show high B-cell reactivity and can differentiate BCG vaccinated healthy controls from TB patients

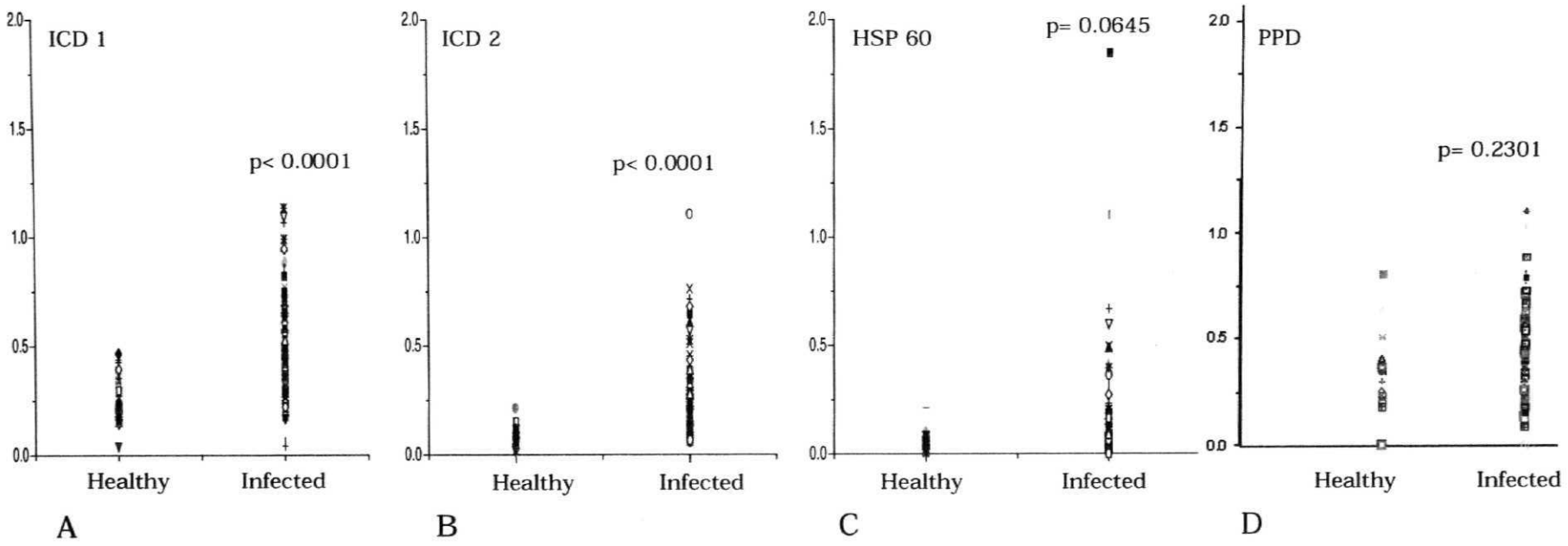


Figure 3.1: The humoral immune responses directed against the recombinant proteins, *M. tb* ICD-1 (A), *M. tb* ICD-2 (B), HSP 60 (C) and antigen PPD (D) were compared between patients with tuberculosis (including all groups) and BCG vaccinated healthy controls.

compared to that of the healthy controls. HSP 60 could also discriminate between healthy and infected sera, however the p value 0.0645 suggests that it is not as statistically significant as ICDs. PPD, on the other hand, reacted against both healthy as well as TB infected sera with a mean antibody titers of 0.36850 ± 0.05244 and 0.45051 ± 0.01752, respectively. It is interesting to note that as compared to ICD-1 and ICD-2 ($p<0.0001$) the difference in the reactivity of PPD to infected and healthy sera was negligible and statistically insignificant ($p = 0.2301$).

A correlation between reactivity against *M.tb* ICD-1 and *M.tb* ICD-2 in patient sera with the state of disease, fresh or relapse, was attempted by comparing the antibody responses to *M.tb* ICD-1 and *M.tb* ICD-2 between various clinical categories (Figure 3.2A and 3.2B respectively). *M.tb* ICD-1 failed to discriminate between fresh, relapsed and extra-pulmonary TB cases as no significant differences in immunoreactivity in different patient groups were observed (Figure 3.2A). Yet as compared to BCG-vaccinated healthy controls, each category yielded p values less than 0.0001 indicating that *M.tb* ICD-1 can differentiate substantially between BCG-vaccinated healthy population and any category of *M.tb* infected patients, pulmonary or non-pulmonary. *M.tb* ICD-2, on the other hand, could also discriminate relapsed cases from both fresh infections ($p< 0.0001$) and extrapulmonary infections ($p= 0.0003$). Like *M.tb* ICD-1, *M.tb* ICD-2 could also distinguish substantially between BCG-vaccinated healthy population and any category of *M.tb* infected patients. Surprisingly, HSP 60, even though, could discriminate Group1 and Group 2 from healthy controls ($p= 0.0011$ and 0.0036 respectively), failed to distinguish the extrapulmonary infections from BCG vaccinated healthy controls ($p= 0.2177$). These results (Figure 3.2) demonstrate that (i) recombinant *M.tb* ICD-1 and ICD-2 proteins could differentiate sera from TB infected patients vis-à-vis healthy BCG vaccinated controls, (ii) the extrapulmonary infections that could not be distinguished from healthy controls by HSP 60, could

M.tb ICD-1 and ICD-2 show high B-cell reactivity to sera from TB infected patients from different groups as opposed to BCG vaccinated healthy controls

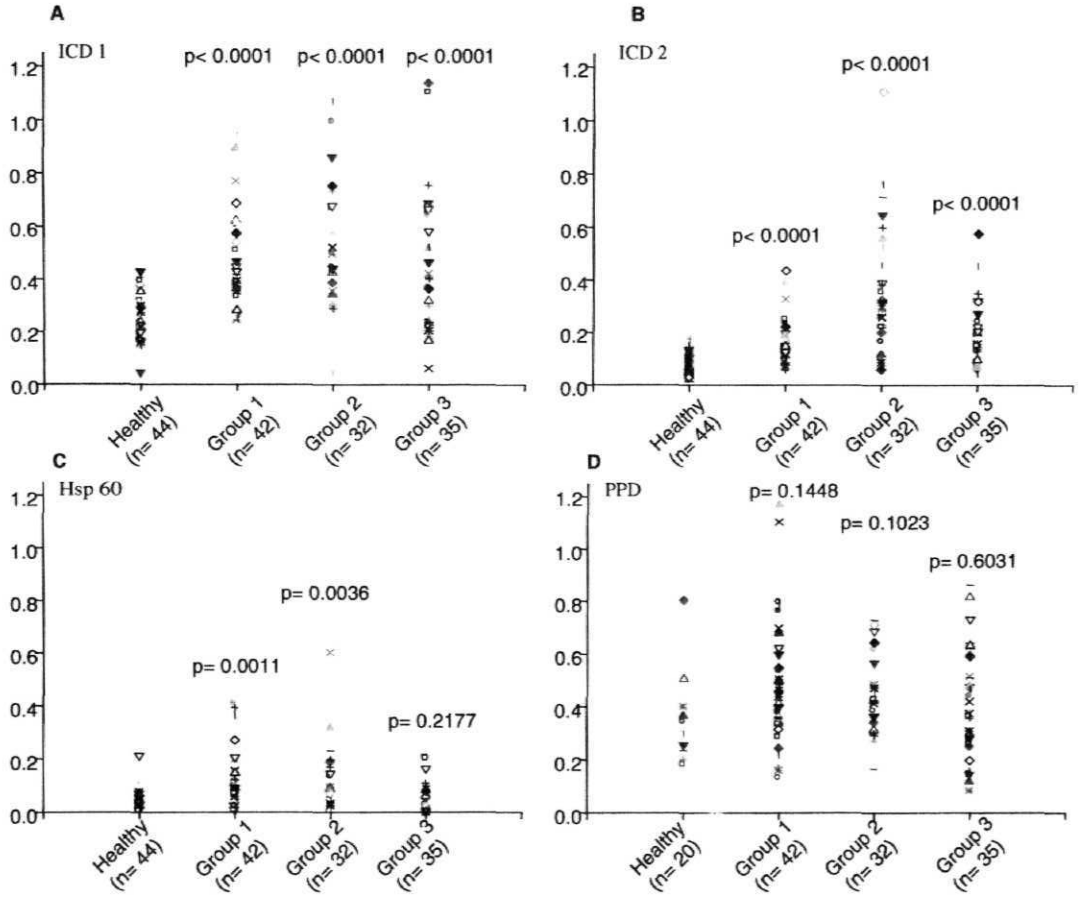


Figure 3.2: The humoral immune responses directed against the recombinant proteins, *M.tb* ICD-1 (3.2A) and *M.tb* ICD-2 (3.2B) and control antigens HSP 60 (3.2C) and PPD (3.2D) were compared between different categories of patients and healthy controls. Group 1: fresh infections, Group 2: relapsed infection and Group 3: extrapulmonary TB. The respective sample numbers and p values are given in the figure.

be significantly identified and categorized by *M.tb* ICDs and (iii) *M.tb* ICD-2 mounted a stronger antibody response in relapsed cases and could significantly discriminate them from Group 1 and Group 3 categories.

Table III: The categorywise comparison of mean ELISA reactivity as against the healthy controls with the respective standard errors of means (SEM) and p values for each antigen tested is tabulated below. ELISA reactivity of each group is compared against healthy controls for p value calculations.

Patient category	**Mean ELISA reactivity (absorbance at 492nm)**	**SEM**	**p-value**
***M. tb* ICD-1**			
Healthy Control	0.27631	0.01574	
Group 1	0.48047	0.03618	< 0.0001
Group 2	0.56052	0.04799	< 0.0001
Group 3	0.45738	0.04446	< 0.0001
Total Infected	0.49466	0.02464	< 0.0001
***M. tb* ICD-2**			
Healthy Control	0.08776	0.00595	
Group 1	0.16186	0.01391	< 0.0001
Group 2	0.36197	0.04157	< 0.0001
Group 3	0.18860	0.02177	< 0.0001
Total Infected	0.23239	0.01762	< 0.0001
HSP 60			
Healthy Controls	0.04795	0.00569	
Group 1	0.11322	0.01920	= 0.0011
Group 2	0.31022	0.09716	= 0.0036
Group 3	0.03449	0.00943	= 0.2177
Total Infected	0.14551	0.03218	= 0.0645
PPD			
Healthy Controls	0.36850	0.05244	
Group 1	0.47222	0.03333	= 0.1448
Group 2	0.47597	0.03444	= 0.1023
Group 3	0.41723	0.03053	= 0.6031
Total Infected	0.45051	0.01752	= 0.2301

Immunodominace of ICDs over HSP 60 and PPD

The immunogenicity of ICDs over HSP 60 was compared. Humoral response to HSP 60 in all the three categories of TB patients was tested and compared with those to ICDs (Figure 3.3). The data clearly indicate that the mean reactivity (represented by the horizontal bands in Figure 3.3) of HSP 60 in all the groups of patient sera was much lower than either ICD-1 or ICD-2 (Figure 3.3). Thus ICDs are immunodominant and serologically more sensitive than HSP 60. The mean values for ICD-1 in the Groups 1, 2 and 3 were 0.481, 0.565 and 0.457 respectively, while those for ICD-2 were 0.165, 0.362 and 0.188 respectively. It is therefore apparent that ICD-1 elicited a stronger response in all the three categories of patients tested than ICD-2. The data also confirm the discriminatory power of ICD-2 for relapsed case as compared to other categories.

As mentioned earlier, PPD is still used for tuberculin testing in many countries. To compare the immune potential of ICDs over PPD, humoral response to PPD in all the three categories of TB patients vis-a-vis healthy controls were tested (Figure 3.2D). The data clearly indicate that the difference in antibody titers (against PPD) between healthy and patient categories is not statistically significant. As apparent from Table III, PPD failed to distinguish amongst the patient categories and also failed to diffrentiate between BCG vaccinated population and TB patients ($p = 0.1448$ for group 1, $p = 0.1023$ for group 2 and $p = 0.6031$ for group 3). PPD therefore lacks the discriminatory power of either *M. tb* ICDs or HSP 60. In order to compare the serological sensitivity of *M. tb* ICDs over PPD, the data presented in Figure 3.2 were recalculated as percentage individuals showing absorbance above 0.40 (an approximate mean ELISA reactivity value) at 492 nm. In all the categories, *M. tb* ICD-1 mounted strong antibody responses of IgG type as compared to PPD (56.6% in group 1, 72.73% in group2 and 65.91% in group 3) (Figure 3.4). *M. tb* ICD-2 shows almost equal reactivity as that of PPD (Figure3.4). However, the serological importance of *M. tb* ICD-2 is in

M. tb ICDs are more immunogenic than Hsp 60

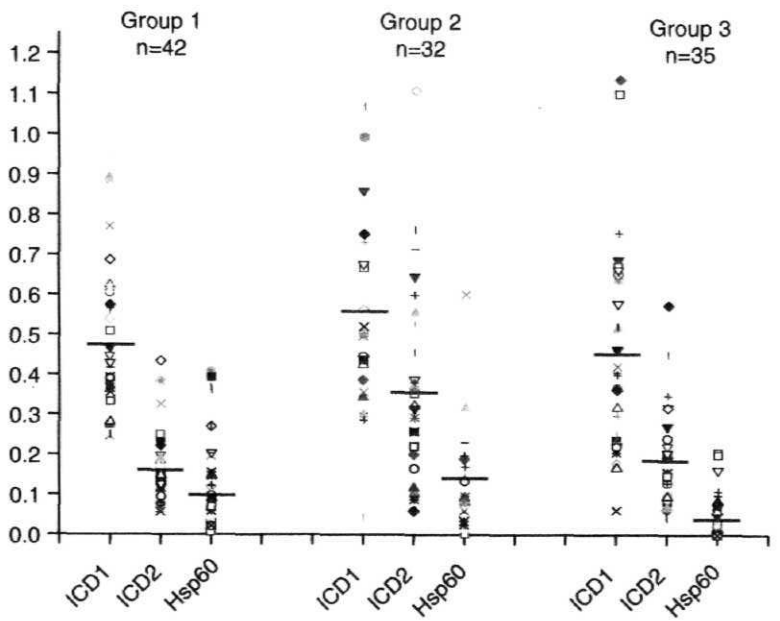


Figure 3.3: The ELISA reactivity to *M.tb* ICD-1, *M.tb* ICD-2 and control antigen HSP 60 were compared in different patient groups under the study. Horizontal bands represent the mean reactivity or average levels of humoral response in each category.

M. tb ICD-1 is more immunogenic than PPD

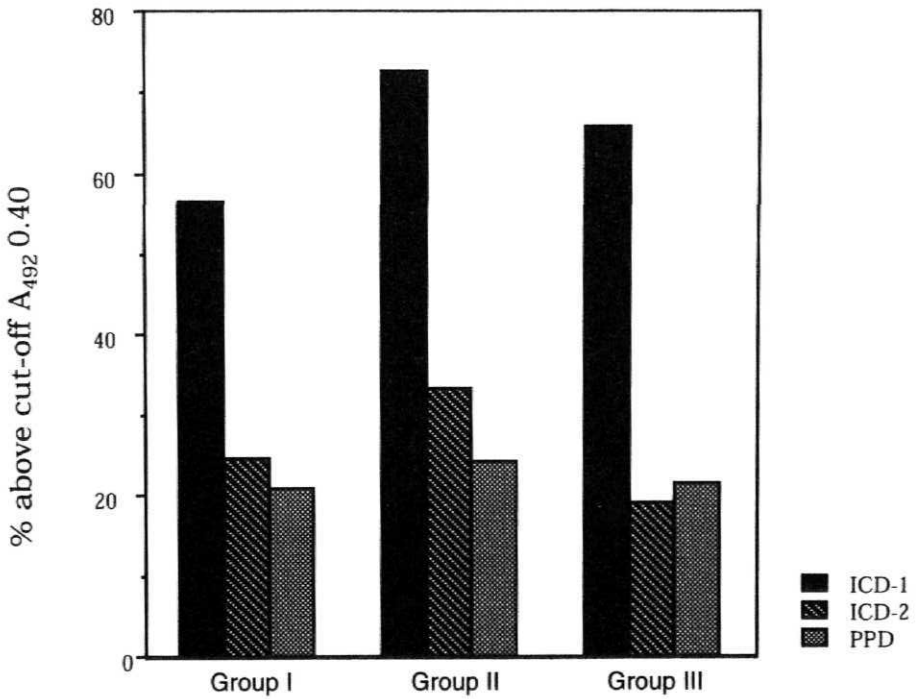


Figure 3.4: Comparative ELISA reactivity as percentage individuals showing absorbance above 0.40 at 492 nm for *M. tb* ICD-1, *M. tb* ICD-2 and PPD. The immunodominant nature of *M. tb* ICD-1 is evident from figure.

being able to statistically discriminate between BCG vaccinated healthy controls and TB patient sera which PPD cannot.

Immunospecificity of *M.tb* ICDs

The potential of *M.tb* ICDs to specifically distinguish between TB, NTMs and non-TB patient sera (those essentially culture negative for acid fast bacteria but harboring other pathogens) was tested by examining the cross-reactivity of the recombinant proteins with NTMs and non-TB patient sera. Thirty NTMs and thirty two non-TB patient sera were tested for their immunogenic response against *M.tb* ICD-1, *M.tb* ICD-2 and HSP 60. The data were statistically analyzed to check if ICDs could significantly distinguish between TB infected patients and NTMs or non-TB patients. Figure 3.5A and 3.5B show that ICDs could significantly distinguish TB-infected sera from NTMs ($p< 0.0001$) and non-TB ($p< 0.0001$ for ICD-1 and $p= 0.0008$ for ICD-2), thus ruling out any cross reactivity with non-tuberculous mycobacteria and other bacterial pathogen tested. HSP 60 appeared to react more broadly to the population under study and could not differentiate between TB infections and NTMs ($p= 0.4461$) or non-TB ($p= 0.3464$) significantly. This was apparent by calculating the average reactivity for each group; infected, NTMs and non-TB; where reactivity to HSP 60 remain almost the same (Figure 3.5A and 3.5B). The mean reactivity of HSP 60 is in contrast to mean humoral response to ICD-1 and ICD-2, where a distinctly higher response was seen in TB infected sera as compared to NTMs or non-TB cases.

DISCUSSION

The main objective of this study was to evaluate *M.tb* ICD-1 and *M.tb* ICD-2 in terms of their immune features as compared to the control antigens HSP 60 and PPD that are frequently used for diagnosis of tuberculosis. ICDs serve as marker of autolysis (Anderson *et al*, 1991; Ohman *et al*, 1996) and are amongst the secretory proteins released during late logrithmic phase. While earlier efforts have pointed to the antigenic potential of *M.tb* ICDs (Ohman *et al*, 1996),

M. tb ICDs could significantly distinguish TB-infected sera from NTMs and non-TB patient sera

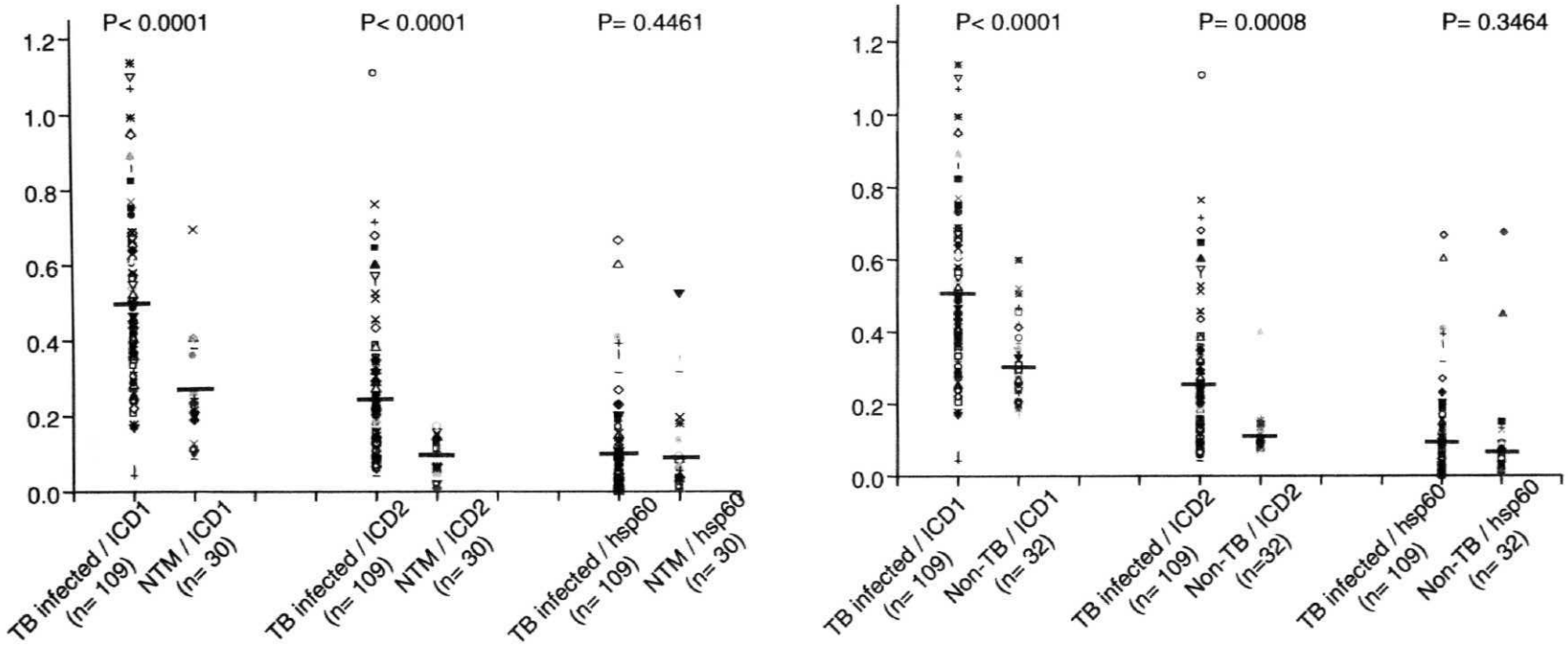


Figure 3.5: Recombinant *M.tb* ICD-1 and *M.tb* ICD-2 as well as HSP 60 were tested against sera of NTM (3.5A) and non-tuberculosis patients (3.5B). The respective humoral responses were compared to TB – infected sera, the p values for which are given in the figures. HSP 60 could not distinguish TB-infected patient from either NTM or non-TB significantly. Horizontal bands represent the mean reactivity in each category. The figure clearly demonstrates the immunospecificity of ICD-1 and ICD-2.

the present study is the first systematic investigation of their potential as an immune marker.

The cases of tuberculosis were identified and enrolled based on their history of treatment as fresh infections, relapsed cases and extrapulmonary infections. The categorization of patients largely depended upon the treatment history dictated by the patients or their family members. As evident from Figure 3.1, there is little doubt about the ability of either *M.tb* ICD-1 or *M.tb* ICD-2 to elicit a strong B-cell response, irrespective of patient category. When compared to *M.tb* ICD-2, *M.tb* ICD-1 was more antigenic (Figure 3.3 and 3.4). It would be interesting to explore this disparity. A comparative analysis of ELISA reactivities amongst different categories of patients for *M.tb* ICDs (Figure 3.2A and 3.2B) revealed higher reactivity in the Group 2 as compared to fresh (Group 1) and extrapulmonary (Group 3) infections. More specifically, the antigenic response in this category of patients to ICD 2 was significantly higher than that in Group 1 and Group 3. Since, these patients had undergone treatment earlier, high number of autolysed infected macrophages and autolysed pathogens could possibly explain the comparative high antibody response against *M.tb* ICDs in this category. Drugs, like isoniazid (INH), are known to affect the cell envelope architecture of mycobacteria and hence the increase in the production of the secreted proteins (Bardou *et al*, 1996).

Comparative immunoreactivity of *M.tb* ICD-1, *M.tb* ICD-2 and HSP 60 clearly indicates that the antigenic distinction between healthy and tuberculosis patients is statistically significant for both *M.tb* ICD-1 and *M.tb* ICD-2 ($p < 0.0001$), but not quite so for HSP 60 ($p = 0.0645$). Earlier reports have shown cross-reactive epitopes between microbial 60/65 and human HSP 60, which often serve as autoimmune targets in conditions like atherosclerosis (Perschinka *et al*, 2003). This probably could explain the broader reactivity of HSP 60 to healthy and infected sera. Since negligible antibody responses were obtained in the BCG-vaccinated healthy control group as

compared to TB infected patients and a statistically significant difference in the immunoreactivity between infected and healthy sera was observed, it can be argued that *M.tb* ICD-1 and *M.tb* ICD-2 can be used for diagnosis of *M. tuberculosis* infection even in areas where BCG vaccination is routinely followed. The poor performance of PPD can be attributed to its non-specific immune reaction in BCG-vaccinated healthy controls. Interestingly, the extrapulmonary TB cases (Group 3) in the study population did not show any significant humoral response to HSP 60 to distinguish them from healthy controls. On the other hand, Group 3 patients mounted a very significant B-cell response to ICD-1 and ICD-2, separating them from BCG vaccinated healthy controls.

Immunodominance is a parameter that was defined to compare the antibody titers against the tested proteins, i.e, ICD-1, ICD-2 and HSP 60, in the patient sera. The data presented clearly showed that ICD-1 was most antigenic and mounted a very strong B-cell response in all the patient categories, followed by ICD-2 and HSP 60 (Figure 3.3). Having shown that *M.tb* ICDs elicit a B-cell response much higher than HSP 60, immune specificity of these proteins was checked. Cross-reaction with sera of NTMs and non-TB patient is one of the critical parameters that needed to be checked before establishing any antigenic marker for possible serological studies in *M.tb*. *Mycobacterium tuberculosis* complex, including *M. bovis* and *M. africanum*, is responsible for more illness worldwide than any other bacteria. However, there are more than 82 recognized species of mycobacteria that occasionally infect mammalian hosts. These are referred to as nontuberculous mycobacteria (NTM). NTMs are omnipresent in the environment and most species are either non-pathogenic for humans or are rarely associated with disease, except a few like *M. avium*, that are opportunistic pathogens, more frequently associated with immunocompromised patients (Shiratsuch *et al* 2003, Hadad *et al*, 2004). The clinical significance of many NTM remains unclear, however it is important to check the crossreactivity of *M.tb* antigens with this group of mycobacteria.

These experiments could establish that *M.tb* ICDs do not cross-react with either NTMs or non-TB patient sera (Figure 3.5A and 3.5B).

The existing diagnostic tests for tuberculosis, even to this day, largely depends on tuberculin skin test and staining and culture techniques. These methods are slow, cumbersome and lack sensitivity and specificity in BCG vaccinated cases. As more and more recombinant antigens are being tested (Choudhary *et al*, 2003; Ramalingam *et al*, 2003; Brusasca *et al*, 2003; Maekura *et al*, 2003; Perkins *et al*, 2003; Maekura *et al*, 2001; Chakhaiyar *et al*, 2004) serological methods are likely to be favoured over others. ELISA *per se* is unlikely to replace the current tuberculosis diagnosis, however in combination or parallel with other rapid PCR based diagnostic techniques, ELISA can largely improve the accuracy and rapidity of tuberculosis diagnosis for an effective disease management. The results presented in this chapter, for the first time, reveal the antigenic potential of recombinant *M.tb* ICD-1 and also present a systematic study on immunogenicity of recombinant *M.tb* ICD-2. *M.tb* ICD-1 and *M.tb* ICD-2 can be further analyzed for their pathogen specific antigenic epitopes. Given their important role in the energy cycle, it would be interesting to evaluate these two enzymes of *M.tb* as possible drug targets. That such important enzymes can also have strong antigenic attributes, which enable them to significantly discriminate between BCG-vaccinated healthy controls and TB patients and at the same time TB from other pathogenic infections, is a very exciting and novel proposition possibly pointing to their immunomodulatory function.

Chapter 4

Characterization of *Mycobacterium tuberculosis* Aconitase as a Bifunctional Protein

INTRODUCTION

Intracellular iron homeostasis is a critical parameter both for the host cell and the pathogen during infection (Rouault *et al*, 1996; Fillebeen *et al*, 2002). The imperative relationship between intracellular iron concentration and pathogen survival has resulted in formulation of several mechanisms by the pathogen to maintain iron at a cellular threshold level. One such mechanism in response to iron deprivation and oxidative stress (Beinert *et al*, 1997; Kang *et al*, 2003) is post-transcriptional regulation of iron metabolism by iron regulatory proteins or IRPs. IRPs come under broad spectrum of iron-sulfur containing protein family, aconitases, and bind to RNA stem-loop structures, Iron Responsive Elements (IREs), either at 5' or 3' end of untranslated regions of mRNA. These proteins act as iron and oxygen sensors of the cells (Nunez *et al*, 2003; Pantopoulos *et al*, 1995). Aconitase family of proteins primarily include the energy cycle enzyme aconitase that by virtue of the same Fe-S cluster participates in electron transfer in catalyzing the reversible isomerization reaction of citrate and isocitrate via the dehydrated intermediate cis-aconitate (Beinert *et al*, 1996) in citric acid cycle. An eukaryotic cell, being a double-membrane system, has a partitioned cytosolic and mitochondrial aconitases (Philpott *et al*, 1994) that perform the functions of IRP and energy cycle enzyme respectively. The cytosolic aconitase function either as RNA binding protein that regulates the uptake, sequestration and utilization of iron or an enzyme that interconverts citrate and isocitrate depending on the status of [4Fe – 4S] cluster. The holoprotein with the intact [4Fe-4S] cluster functions as aconitase, whereas the apoprotein, [3Fe-4S] is an RNA-binding translational regulator (Rouault and Klausner 1997; Haile *et al*, 1992). Several bacteria, such as, *E. coli*, has two isoforms of aconitase, AcnA and AcnB, with differential expression and physiological properties (Jordan *et al*, 1999; Tang *et al*, 2002; Varghese *et al*, 2003), while prokaryotes like *Bacillus* or *Xanthomonas*, have only one copy of aconitase in their genome.

As mentioned earlier, *Mycobacterium tuberculosis* codes for a single copy of aconitase (*acn* or Rv1475c). Amino acid sequence comparisons of *Mycobacterial* aconitase (Swissprot ID: O53166) with *E.coli* aconitase AcnA (Swissprot ID: P25516) & AcnB (Swissprot ID: P36683) showed that mycobacterial aconitase shares 60% identity with AcnA of *E.coli* but only 20% identity with AcnB. As per earlier report, AcnA of *E.coli* is induced in stationary phase or during oxidative stress conditions and is less sensitive to oxygen mediated inactivation (Jordan *et al*, 1999; Tang *et al*, 2002). It, hence, can be proposed that *M. tb* being a strict aerobe would use the more stable AcnA / IRP aconitases for energy metabolism, a feature likely to be important for survival under oxidative stress. 4.74 folds down regulation in starvation model (Betts *et al*, 2002) and up regulation at high iron stress suggest (Wong *et al*, 1999) its involvement in energy production and iron homeostasis.

The work presented in this chapter documents for the first time the bifunctionality of the *M. tb* aconitase and shows that (i) *Mycobacterium tuberculosis* aconitase displays both enzymatic and RNA-binding activities depending on iron availability, (ii) aconitase activity is not required for RNA binding, (iii) cysteine residues are important for both enzyme activity and RNA binding. Further, the functional oligomeric state and basic biochemical properties of *M. tb* aconitase as a TCA cycle enzyme has been elucidated. The likely amino acid residues associated with the functional specificity of *M. tb* aconitase as a TCA cycle enzyme and the likely region of the protein to which IRE-RNA binds has been highlighted.

MATERIALS AND METHODS

PCR amplification, cloning and expression of *M. tb* aconitase

The 2.892 kb (ACN or Rv1475c) long ORF was amplified from H37Rv genomic DNA using proofreading *Taq* polymerase (AccuTaq™ LA DNA Polymerase, Sigma) using the primers having restriction sites *Xho*I and *Hind*III at 5' end (FP – ATTCCTCGAGATGTGACTAGCAAATCTGTG and RP –

ATGGAAGCTTTCAGCCTGACTTCAGTATGT). A 50µl PCR reaction contained 250ng of template DNA (genomic DNA of H37Rv), 200ng of each primer with 2.5U of *Taq* DNA polymerase (AccuTaq™ LA DNA Polymerase, Sigma), 500µ moles of dNTPs, 1µl DMSO and 1X PCR buffer supplied with the enzyme. The template DNA was amplified for 35 cycles with denaturation at 94°C, annealing at 50°C and elongation at 68°C in GeneAmp PCR system 9700 (PE Applied Biosystems, USA). The PCR products were gel extracted, purified and stored at -20°C.

The insert was cloned in the expression vector pRSET – C at *Xho*I / *Hind*III site followed by overexpression in the pRSET-C/ *E. coli* BL-21 (DE3) expression system. For the ligation reaction, 200ng of vector (pRSET C), 600ng of inserts (*acn*), 1 unit of T4 DNA ligase and 1X ligation buffer (30mM Tris-HCl pH 7.8, 10mM $MgCl_2$ 10mM DTT, 2 mM ATP, 5% PEG 8000) were mixed together and volume made to 20µl with distilled water. The tube was incubated at 16°C for 16 hours or overnight. The ligation mix was used to transform DH5α strain of *E. coli*. The clones were confirmed by single digestion with *Hind*III first and later by sequencing using the T7 promoter primer, on an ABI prism 377 DNA sequencer (PE Biosystems, USA). The generated constructs 'setAacn' was further transformed into the BL21 (DE3) strain of *E.coli*.

The overexpressed his-tagged recombinant protein was purified by Ni^{2+}-nitrilotriacetate affinity chromatography from soluble fraction under native conditions. The cells transformed with the construct plasmid were grown overnight in LB with 100 µg/ml ampicillin at 37°C for primary culture. 1% of the primary culture was used to grow a large scale secondary culture to an OD_{600} of 0.4 at 37°C, cooled to 27°C, induced with 0.1mM IPTG and grown overnight at 27°C in Terrific Broth. The cells were lysed by sonication, collected by centrifuging at 16000Xg for 30 minutes at 4°C and the clear lysate was loaded onto Ni^{2+}-NTA column. The column was washed with 50mM NaH_2PO_4, 300mM NaCl, 20mM imidazole, pH 8. The protein was eluted in the same buffer supplemented with 250mM

imidazole. The protein was 90-95% pure as checked on 10% SDS-PAGE followed by Commasie Blue staining. The purified protein was dialyzed against 20mM TrisCl pH 8, 100mM NaCl and 3% glycerol. The protein was concentrated using Centriplus Centrifugal filter units with 10Kda cut-off filter (Millipore, USA) and quantified using the Bradford method of protein quantification.

Reconstitution and inactivation of native protein

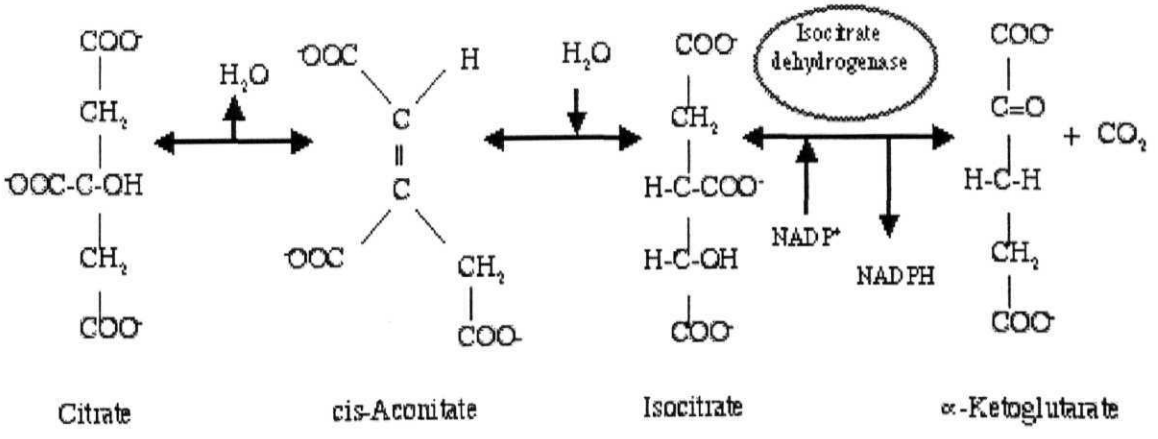
The protein purified under native condition was reconstituted by incubating with1mM ferrous ammonium sulphate and 10mM DTT in 50 mM TrisCl pH 8 at 25°C for 20 minutes. The reaction mix was centrifuged at 16000Xg at 4°C for 10 minutes to pellet the precipitated protein, if any. The reconstituted protein was dialyzed against 20mM TrisCl pH8 and100mM NaCl.

For inactivation of the purified enzyme 0.5mM of specific iron chelator, dipyridyl, or non-specific metal chelator, EDTA was used.

Biochemical assays for aconitase

Indirect Assay

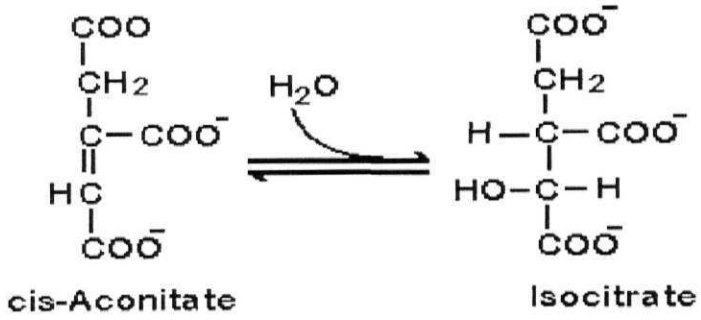
This assay is based on measurement of the concomitant formation of NADPH from $NADP^+$ in a two step reaction where citric acid is converted into isocitrate by aconitase and the second enzyme isocitrate dehydrogenase converts isocitrate into α-ketoglutaric acid in presence of NADP, which is reduced to NADPH. The formation of NADPH is monitored by the increase in absorbance at 340nm. The rate of NADPH production is proportional to aconitase activity.



Porcine heart isocitrate dehydrogenase (Sigma) was used for these reactions. Each 400µl reaction mix contained 20mM triethanolamine chloride buffer pH 7.5, 2mM NADP+, 2 mM citrate, 100mM NaCl, 10mM $MgCl_2$, 5µl porcine heart isocitrate dehydrogenase and 30-50 nM of *M. tb* aconitase. The reaction was tested both in absence and presence of *M. tb* aconitase to check for residual activity. The concentration of substrate used for Km and Vmax estimation varied from 1µM to 1mM.

Direct assay

Aconitase activity was measured spectrophotometrically by monitoring the time dependent conversion of isocitrate to cis-aconitate at 25°C in Unicam UV/Vis spectrometer at 240 nm, the absorbance maximum for cis-aconitate. The rate of cis-aconitate production is proportional to aconitase activity.



cis-Aconitate Isocitrate

Each standard assay solution of 400µL contained 20mM triethanolamine chloride buffer pH 8, 2mM DL-isocitrate, 100mM NaCl and 30 – 50nM of the enzyme. Environmental parameters for the enzyme was measured by altering the pH of the buffer (range 5 – 11), temperature (20 - 95°C), concentration of substrate (0.1mM - 4mM), metal ion (Mg^{++}, Zn^{++}, Mn^{++}; 1 – 10mM) and salt (0mM – 500mM) requirement. The pH dependence of the enzyme was measured using 20 mM of the following buffers: phosphate buffer (pH 5.7 to pH 7), TrisCl: pH 5-9; CAPS: pH9-11. The kinetic parameters were determined by altering the concentration of the

substrate (0.1mM – 4mM). The values were plotted and counter checked as Micheles Menten plot (V vs S), Hanes plot (S/V vs S) and Direct-linear plot (V vs S, with median values) for calculating Km and Vmax for this first order reaction.

Size exclusion chromatography: Size exclusion chromatography was performed at room temperature using FPLC equipped with Superdex-200 HR 10/30 column (Amersham Pharmacia Biotech). Calibration of the column was performed using protein molecular-mass standards for gel-filtration (Sigma, USA), namely thyroglobulin (669kDa), ferritin (443kDa), β-amylase (200kDa), alcohol dehydrogenase (150kDa), albumin (66kDa) and carbonic anhydrase (29kDa). The column was equilibrated with standard proteins in 50mM Tris-HCl, pH7.5 with desired concentration of NaCl and the elution volume (V_e) for each protein was determined. The void volume (V_o) was determined by running Blue Dextran on the column. The calibration curve was plotted as Ve/Vo versus log of molecular mass. No difference in the elution volume or the calibration curve was observed when different concentrations of NaCl (No NaCl; 150mM; 300mM; 1.5M) were used to calibrate the column. The column was equilibrated with three bed volumes of the elution buffer prior to each run. Protein elution was monitored at A_{280}. A 1.2 mg/ml (for both native and reconstituted protein) concentration of recombinant proteins were used for all gel filtration experiments. The buffers used for different experiments to determine the oligomeric state were 20 mM Tris pH8 without or with 100mm, 300mM, 500mM and 1M NaCl.

Sequence alignment and phylogenetic analysis: The amino acid sequence of *M. tb* aconitase was compared against the NCBI protein database. The sequences with the BLAST score upto e-153 or 65% identity were selected for construction of the phylogenetic tree. MegAlign, the multiple sequence alignment package of LASERGENE was used to determine the conserved residues across the aligned sequences. Only the conserved sequences were aligned and used to construct a phylogenetic tree using CLUSTALW program. The

sequence alignment file was used to construct a phylogenetic rooted tree applying DNA*star (Altschul *et al* 1990; Retief *et al* 2000).

Scanning of *M. tb* genome for IRE – like sequences

The *M. tb* genome was scanned using the pattern search program of TubercuList (http://genolist.pasteur.fr/TubercuList/) for the sequence C N N N N N C A G U G, with or without a single mismatch in the region CAGUG, located within 100 base pairs upstream of a start codon or 100 base pairs downstream of a stop codon. The selected sequences were subjected to secondary structure prediction (http://www.genebee.msu.su/services/rna2_reduced.html) for possible stem loop like configuration. The selected sequences were synthesized as complimentary oligonucleotides.

Table : Sequences of the oligonucleotides corresponding to the selected IRE like sequences of *M. tb* genome used for gel shift assays.

5' UTR of *M. tb trx*C	aattcGGCGATGCGCTGCGCTGTGGCGACC GCAGTGCGGCCGTCACCGAGATCCGGa
3' UTR of *M. tb acn*	aattcGCGGCTCGCCGCCGTCAGATCCTCG ACGGTGCGCGCCGTTGCTTTGCCGAAa
5' UTR of *M. tb ideR*	aattcGGTAGCAGACGGTATGCCCGCCGCG CCAGCGGCGGGCATACCGCTGCGGTGa
Control: 5' UTR of human ferritin	aattcGGGAGAGGATCCTGCTTCAACAGT GCTTGGACGGATCCa

***In-vitro* transcription**

Complimentary oligodeoxyribonucleotides of the selected sequences were cloned in pGEM-3Zf vector at *Eco*RI and *Hind*III sites. The vector was linearized using *Hind*III and used as a template for the *in-vitro* transcription using T7 RNA polymerase. The 10µl reaction mix contained 150ng of DNA template, 10mM DTT, 20-25 units of T7 RNA polymerase, 1X RNA polymerase reaction buffer and 0.5mM each of ATP, GTP, UTP with 0.5mM CTP, wherever unlabelled CTP

was used and 0.05mM CTP with 20μCi of α^{32} P CTP, wherever the probe was radiolabelled for gel shift assays. The reaction was incubated at 37°C for 1 hour followed by RNase-free DNase treatment (1 unit of DNase per reaction) for 15 minutes at 37°C. After phenol extraction, the RNA was precipitated with 0.5M ammonium acetate, 15-20 μg of yeast RNA and one volume of isopropanol at -20°C overnight. The RNA was dissolved in nuclease free water and checked on 7M denaturing 15% polyacrylamide gel. Both labeled and unlabeled RNA were denatured at 85°C for 5 minutes followed by renaturation by slow cooling to allow proper folding before using for gel shift assays.

Gel Shift Assays (GRAs)

The renatured radiolabeled RNA fragments (~ 1 pico mole) were allowed to bind with ~3 μg of purified *M. tb* aconitase in a biding reaction (20μl) containing 10mM TrisCl pH 8, 50mM KCl, 10% glycerol and 1 μg of total yeast RNA. The reaction was carried out at room temperature for 30 minutes. The reaction products were then loaded onto a 6% nondenaturing polyacrylamide gel and electrophoresed at room temperature and run in TGE buffer. 0.5 mM of didyridyl was used wherever *M. tb* aconitase was inactivated, with an incubation time of 15 minutes at 25 °C. Non-specific metal chelator EDTA was also used in some of the experiments. For competition experiments 75X of radiolabeled specific and non-specific RNA fragments were used.

Factor Xa digestion and protease protection assays

M. tb aconitase was allowed to bind with the selected IRE-RNA fragments and the complex was UV crosslinked after the binding. This was followed by digestion with Factor Xa by incubating at 22°C till 36 hours. Each 100 μl reaction contained 10mM TrisCl pH 8, 50mM KCl, 1 U of Factor Xa and 30 – 50μg of dipyridyl treated *M. tb* aconitase. Fractions were collected at the intervals of 18, 24 and 36 hours and were run on a denaturing 10% polyacrylamide gel followed by Commasie Blue staining to check for the RNA binding site of *M. tb*

aconitase. Both radiolabelled and unlabelled RNA were used for this experiment.

Chemical inactivation of cysteine residues

The reaction of 7-chloro-4-nitrobenzo-2-oxa-1,3-diazole (NBDCl, stock concentration in DMSO = 1.99mg/ml) with free thiol groups of *M. tb* aconitase (final concentration = 17µg/100µl) was initiated by pipetting 4 µl of the stock NBDCl into 100µl of the reaction solution containing 20mM TrisCl pH 8. The reaction was followed by monitoring absorbance at 420 nm over time as described by Yancey et al. (Yancey and Somero, 1979) to check the formation of complex between NBDCl and Cysteine residues of the protein. The NBDCl-treated protein was then used for gel shift assays.

RESULTS

A. Oligomeric assembly, Biochemical characterization and Phylogenetic affiliation

PCR amplification and cloning of *M.tb* in expression vector pRSET C

The insert, *M. tb acn*, was cloned in the expression vector pRSET – C at *Xho*I/*Hind*III site (Figure 4.1). The amplified PCR product corresponding to size ~2.89 kb was excised from 1.2% agarose gel and extracted (Figure 4.2A). The PCR product and the vector pRSET C were digested with *Xho* I and *Hind* III and ligated. The ligation mix was used to transform competent DH5α cells and the colonies were screened for the positive clone by single RE digestion with *Hind* III. The expected size of the vector with insert was ~5.7 kb (Figure 4.2B).

Expression and purification of protein coded by *M. tb* ORF Rv1475c

The insert cloned in the expression vector pRSET – C at *Xho*I/*Hind*III site was overexpressed in the *E. coli* BL-21 (DE3) cells. The transformed cells grown in Terrific Broth were induced by 0.1mM isopropyl-β-D-thiogalactopyranoside (IPTG) at 18°C for 16 hours for overexpression of *M. tb* aconitase. The over-expressed N-terminal His-tagged *M. tb* aconitase was then purified to 90% - 95%

Cloning of *M. tb* aconitase (*acn*) in expression vector pRSET C

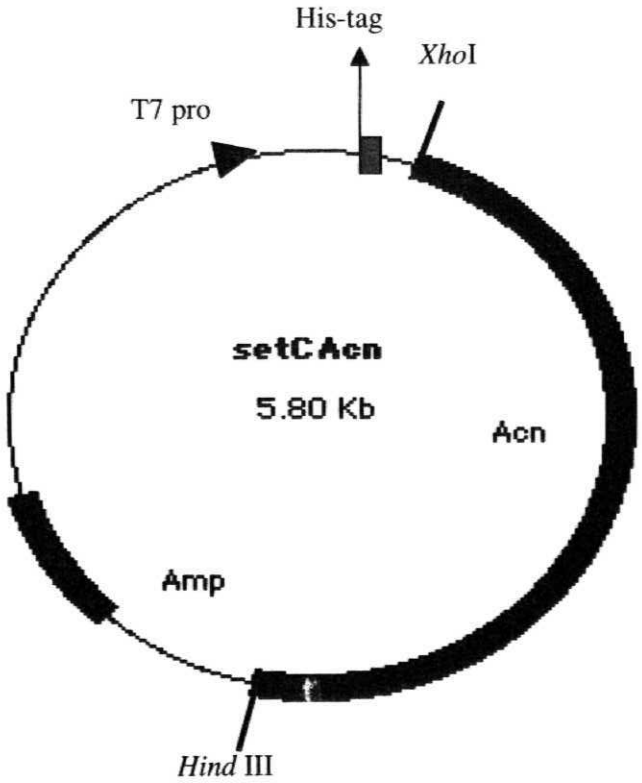


Figure: 4.1

PCR amplification and cloning of *M.tb* Aconitase (Rv1475c) in expression vector pRSET C

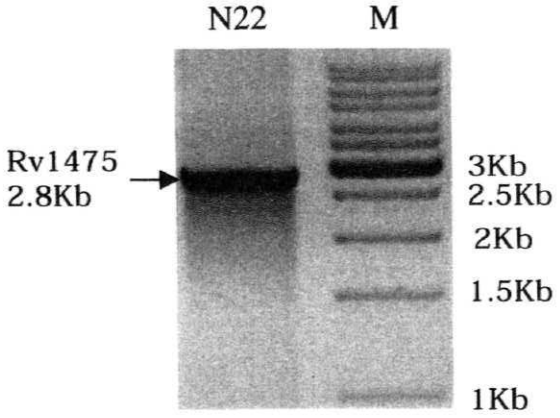


Figure 4.2A: PCR amplification of 2.8Kb fragment of Rv 1475 from H37Rv.

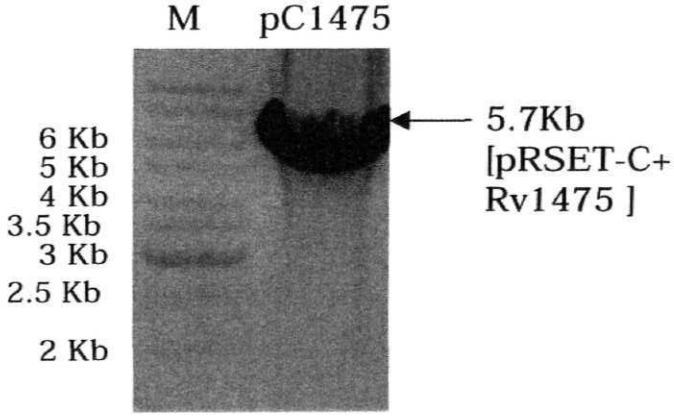


Figure 4.2B: Identification of the positive clone of Rv1475 in the pRSET-C vector by restriction digestion with *Hin*dIII

Purification of *M.tb* Aconitase by Ni-NTA affinity chromatography under native conditions

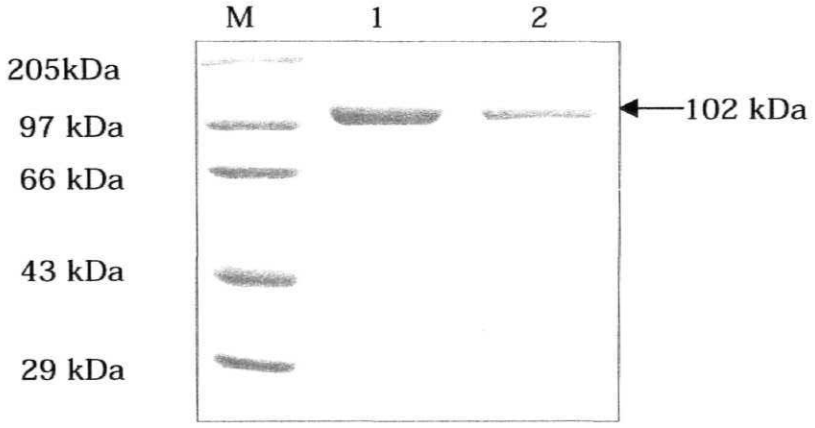


Figure 4.3: Histidine-tagged recombinant protein was purified by nickel column chromatography under native condition and stained with Coomassie Blue following electrophoresis on 10% SDS-PAGE. The different lanes are: lane M: protein molecular size markers (205 kDa, 97 kDa, 66 kDa, 43 kDa and 29 kDa); lanes 1 and 2: purified *M. tb* aconitase.

homogeneity using a Nickel affinity column (Figure 4.3). The molecular mass of the recombinant protein as determined by SDS-PAGE analyses was found to be ~ 102 kDa (Figure 4.3). The purification was carried out under native conditions from soluble fraction. ≥ 7mg of recombinant protein could be purified from 750ml of start culture.

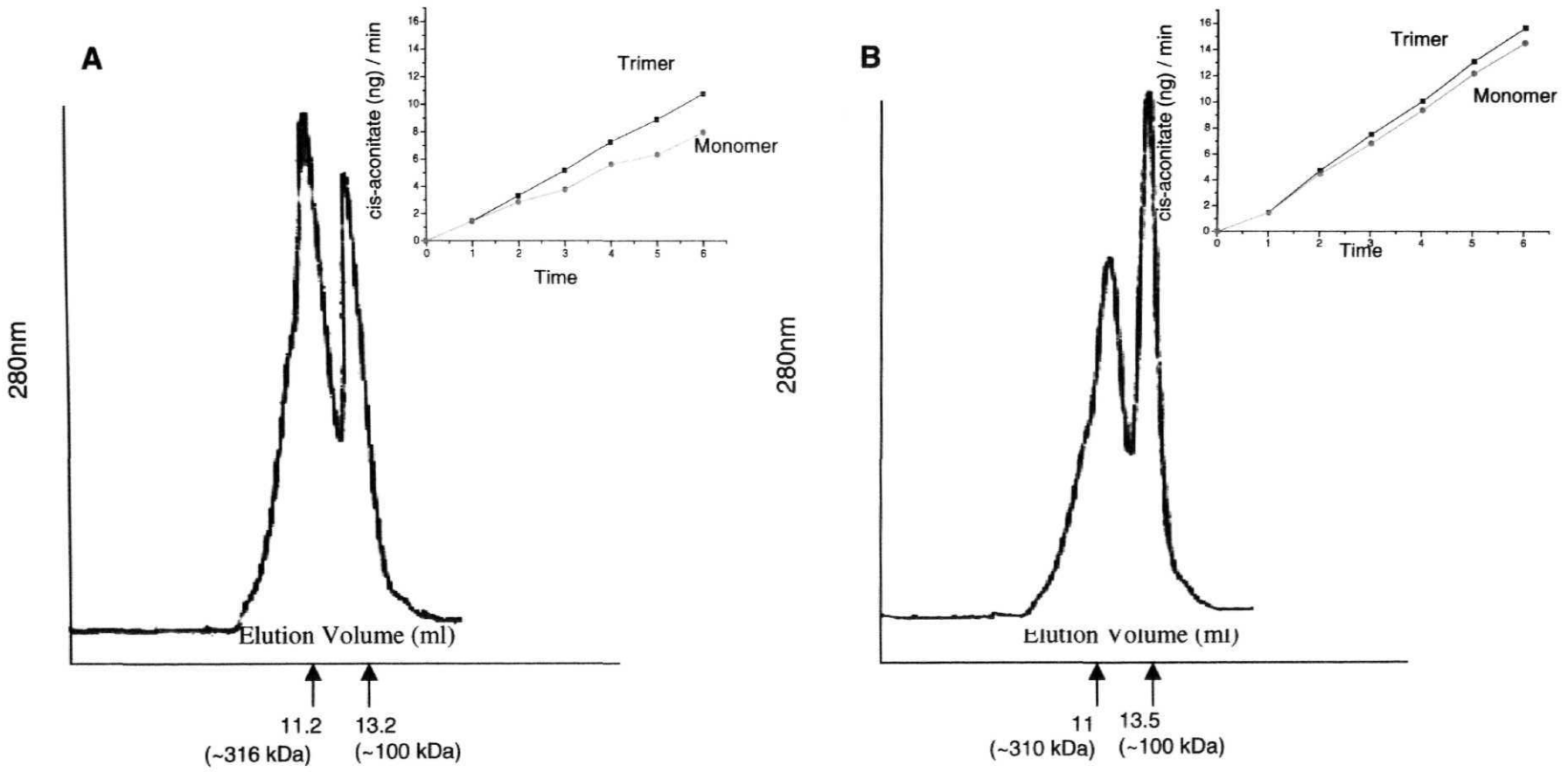
Oligomeric assembly

Recombinant *M. tb* aconitase is a functionally active TCA cycle enzyme and the functional form is a monomer

Recombinant *M. tb* aconitase could convert isocitrate to cis-aconitate in a reaction mix containing 2mM DL-isocitrate, 100mM NaCl and 30 – 50nM of the enzyme present in 20mM triethanolamine chloride buffer pH 8. The formation of cis-aconitate was monitored spectrophotometrically at 240 nm in a time dependent fashion at 25°C (Figure 4.4A, inset and 4.4B, inset). The active enzyme was used for further characterization.

FPLC profile of native protein (Figure 4.4A) and reconstituted protein (Figure 4.4B) showed presence of two oligomeric states corresponding to monomer and trimer in presence of 20mM TrisCl pH 8 and 100mM NaCl. Each fraction was collected separately, quantified and checked for activity (Figure 4.4A and 4.4B). The aconitase activities of purified protein under native condition were generally lower (Figure 4.4A, inset) than the reconstituted enzyme (Figure 4.4B, inset) indicating that a significant portion of purified enzyme is in the inactive [3Fe–4S] or apoenzyme form. Km for native enzyme were 1.2 mM ± 0.26 while that for reconstituted enzyme were 0.375 mM ± 0.17 indicating a 3.2 fold increase in the affinity for the substrate by the reconstituted protein. The inset graphs in the Figures 4.4A and 4.4B clearly show that both the forms were enzymatically active, hence it could not be judged how important the trimer is for the activity of the protein. The kinetic parameters of monomer and trimer were further calculated separately for native and reconstituted enzymes (Table IV). These experimental data established the facts that reconstituted aconitase had higher activity than native *M. tb* aconitase while the trimer of both reconstituted

Functional form of *M. tb* aconitase is a monomer



(A) Native aconitase

(B) Reconstituted aconitase

Figure 4.4A and B: The elution profiles of (A) native recombinant *M. tb* aconitase (B) reconsituted *M. tb* aconitase, on a Superdex-200 HR 10/30 column showed two peaks corresponding to trimer and monomer. Inset: Comparative activity of the collected fractions. [cont.]

Cont.

The trimer of native protein could be disrupted into dimer and monomer upon NaCl (500mM) treatment

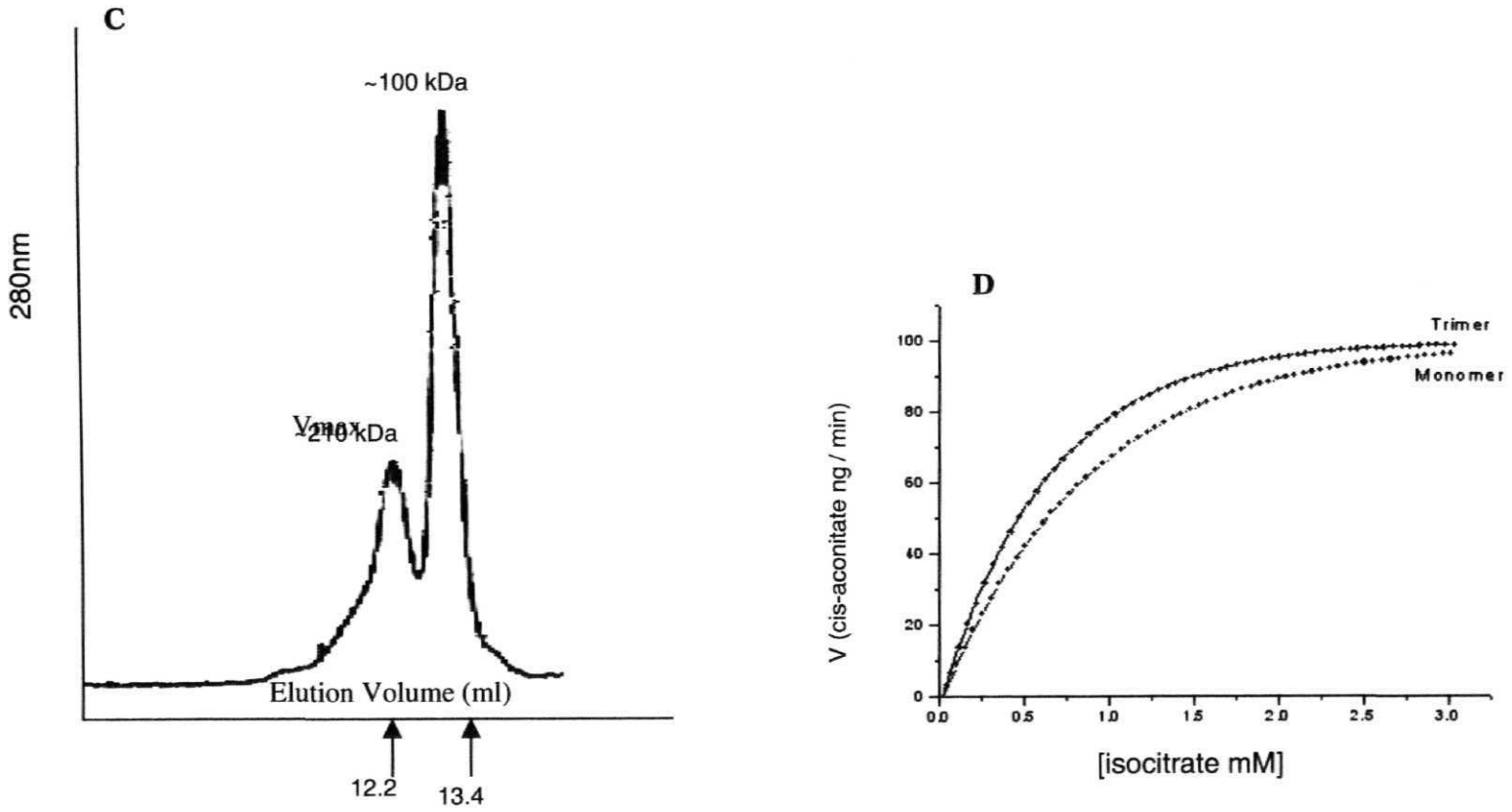


Figure 4.4C and D: (C) FPLC profile of *M. tb* aconitase after treatment with 500mM NaCl, showed two peaks corresponding to dimer and monomer. The trimer of native protein could be disrupted into dimer and monomer upon NaCl treatment. (D) Michaelis – Menten (V vs [S]) plot for trimer and monomer suggested insignificant substrate binding cooperativity.

and native enzymes displayed higher activity than monomer. This was nonetheless, not significant. Even though a trivial decrease (less than 2 fold) was observed in Km values, the specific activity of both the forms nearly remained same (Table IV). This indicated that trimer was probably an aggregate of monomers and not a separate oligomeric form.

Table IV: Kinetic parameters of different FPLC fractions of native and reconstituted recombinant *M. tb* aconitase

Oligomeric state		Km Isocitrate (mM)	Vmax cis-aconitate (ng/min)	Specific activity cis-aconitate (μg/min/mg)
Native	Trimer	0.5 ±0.09	133.3 ±5	111.1 ±4.17
	Monomer	0.9 ±0.03	125 ±6.5	100.3 ±5.42
Reconsti-tuted	Trimer	0.23 ±0.08	180.55 ±23.3	150.5 ±19.4
	Monomer	0.56 ±0.1	141.67 ±40	118.1 ±33.3

To further confirm that the trimer is not a stable functional form of *M. tb* aconitase and was just a loose aggregate, the protein was incubated with 300 mM, 500 mM and 1 mM of NaCl for 1 hour at room temperature followed by loading onto Superdex-200 column for fractionation. It could be clearly seen that the trimer is disrupted into dimer and monomer upon treatment with atleast 500 mM of NaCl (Figure 4.4C). This clearly showed that trimer is an aggregate formed due to ionic interactions between aconitase molecules and does not represent the stable functional form. A V vs [S] plot for trimer and monomer (Figure 4.4D) indicated substrate binding cooperativity that may explain the unsubstantial decrease in Km values for trimer. It can also be seen that the cooperativity implied by

the plot is also not very significant. These biochemical features clearly suggest that the functional form of *M. tb* aconitase is a monomer.

Biochemical characterization

pH profile of *M. tb* aconitase: Log_{10} Vmax vs pH plot suggests the involvement of histidine and cysteine side chains in ionization at the catalytic site of *M. tb* aconitase

The activity of *M. tb* aconitase as a function of pH was studied. Although the optimum pH for the activity of *M. tb* aconitase, both native and reconstituted as seen from the pH vs activity profile is pH 8, the enzyme however remains fairly active within a range of 7 – 10 (Figure 4.5A). The broader activity range of *M. tb* aconitase is in agreement with its apparent similarity to AcnA of *E. coli*. AcnA is a stationary phase enzyme when the intracellular pH may vary over a wider range than log phase (Jordan *et al*, 1999). Plotting experimental values as Log_{10} Vmax vs pH (Figure 4.5B) shows involvement of two ionization groups involved in catalytic function of the enzyme. Extrapolation of the linear portion of the plot gave the pKa values of the free amino acid residues, the side chains of which might be responsible for change in ionization state during catalysis (Price and Stevens). The plot reveals three distinct regions of slope (~ ±1 and 0). The point of intersection corresponds to the values 7.1 and 9.3. Since environment of an amino acid within an enzyme can be influenced by other amino acids, about 1.5 pH units on either side of the predicted values was therefore taken. With reference to the pKa values of free amino acids (Price and Stevens), the possibility of amino acids involved may be histidine (pKa 6), cysteine (pKa 8.3), tyrosine (pKa 10.1) or lysine (pKa 10.5). Further, when *M. tb* aconitase was aligned with bovine heart aconitase (P20004, Plank *et al*, 1988), human IRP-BP1 (NP_002188, Schneider *et al*, 2003) and IRP-BP2 (Schneider *et al*, 2003) human mitochondrial Acn (NM_001098, Mirel *et al* 1998), *E. coli* AcnA (P25516, Blattner *et al*, 1997) and AcnB (P36683, Blattner *et al*, 1997) to elucidate the strictly conserved catalytically important residues based on mutation

pH vs activity profile of *M. tb* aconitase

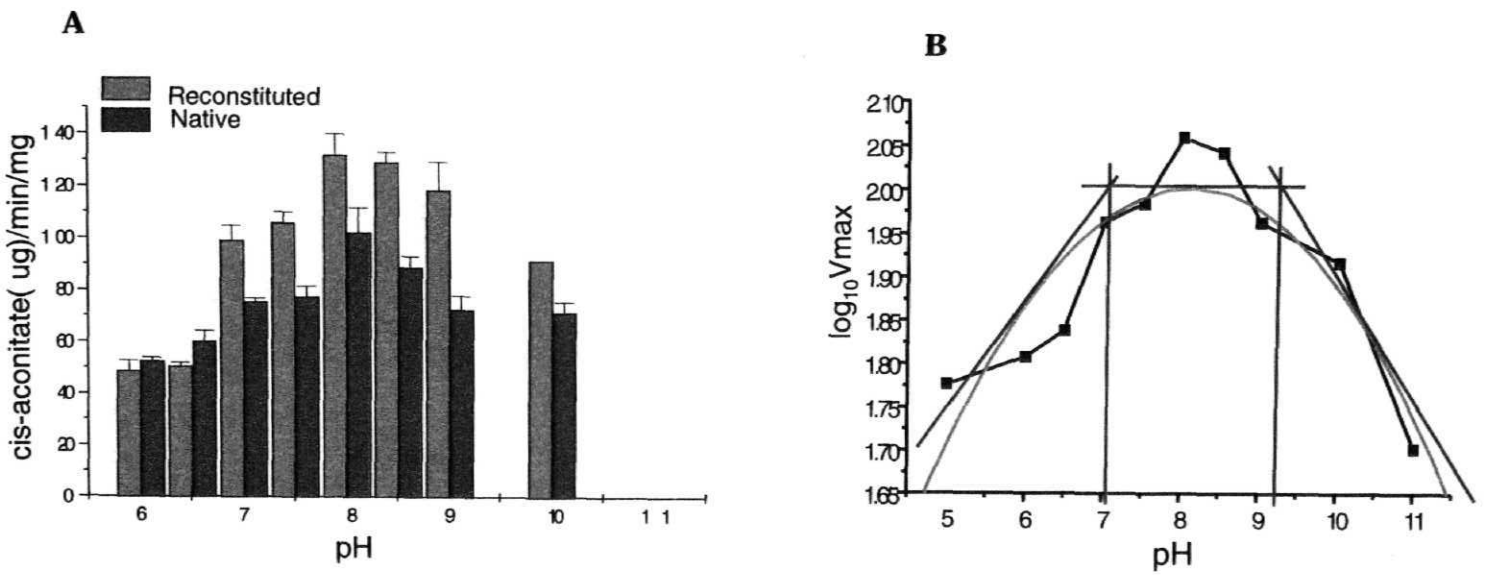


Figure 4.5: (A) The activity of *M. tb* aconitase, both native and reconstituted, was tested as a function of pH (6 to 10). 20 mM of the following buffers were used: phosphate buffer (pH 5.7 to pH 7), TrisCl: pH 5-9; CAPS: pH9-11. (B) Log_{10} Vmax vs pH plot for data set in (A) generated a parabolic curve that depicted involvement of atleast two ionization groups in catalysis of the enzyme. The intersection of the slopes (~ +1, 0 and ~ -1) correspond to the values 7.1 (pKa_1) and 9.3 (pKa_2). Log_{10} Vmax vs pH plot depicts involvement of **histidine** and **cysteine** side chains in ionization at the catalytic site of *M. tb* aconitase

His 208 and Cys 479, 545,548, residues are conserved across Acn A/IRP family of proteins

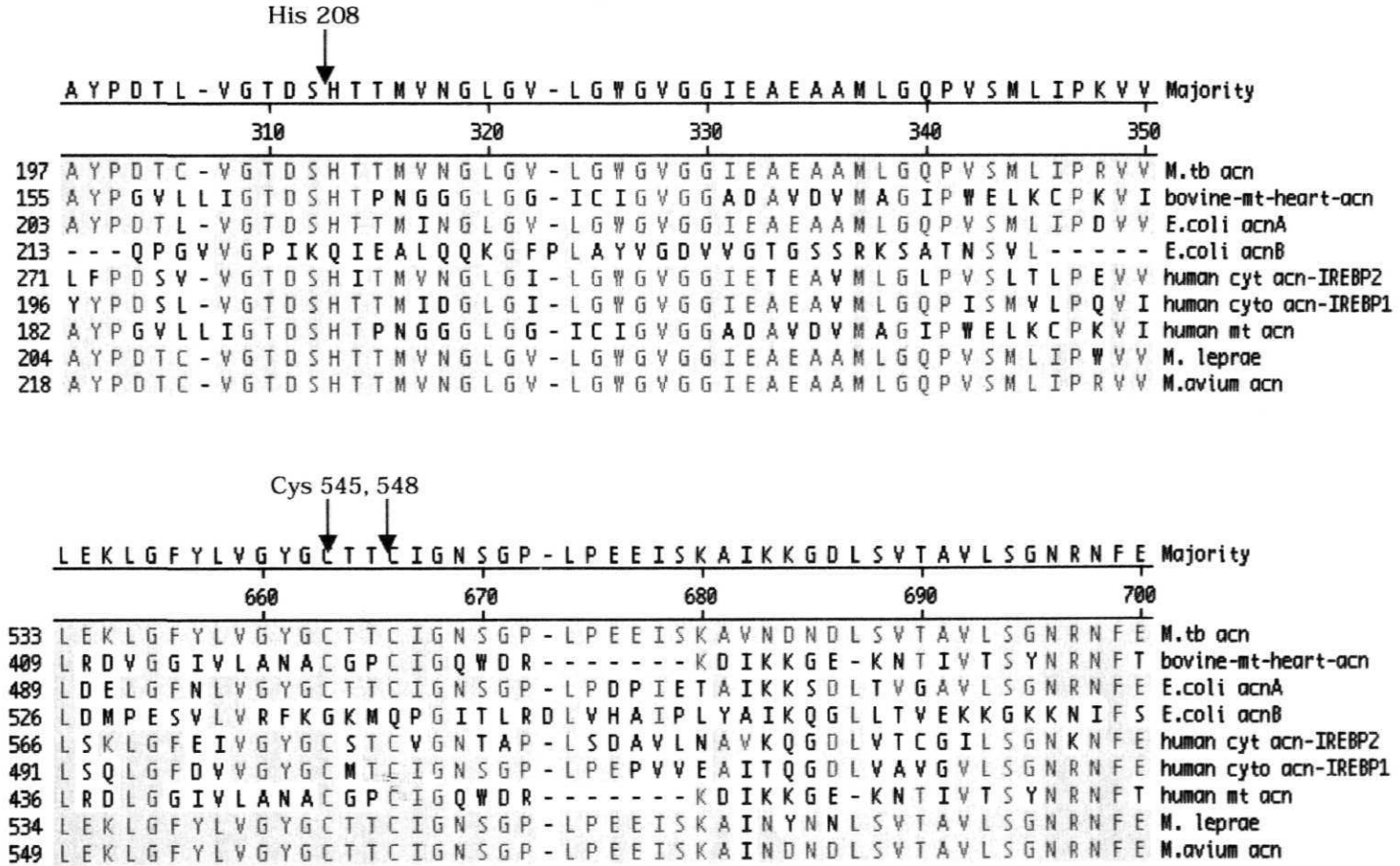


Figure 4.6: Discontinuous sample alignment report of *M. tb* aconitase to highlight a few catalytically important conserved amino acid residues

and crystallographic studies. Based on such alignment the corresponding residues in *M. tb* aconitase included Gln-87, Phe-89, Thr-90, His 127, Glu 226, Ser 207, His 208, Ser 478, Cys 479, 545,548, Ile 549, Asn 577, Arg-578, 583,745, 825 and Ser-823, 824. The alignment clearly showed conserved histidine
and cysteine but no lysine or tyrosine in the list of catalytically important residues. Discontinuous sample alignment report of *M. tb* aconitase to highlight a few catalytically important conserved amino acid residues has been presented in Figure 4.6. Therefore it would be reasonable to conclude that the side chains participating in ionization during catalysis could be His-127, 208 and Cys-479, 545, 548.

Optimum temperature:

***M. tb* aconitase exists in a single interconvertible form with a single activation energy irrespective of temperature change**

The activity of *M. tb* aconitase was studied as a function of temperature. Results clearly show that the *M. tb* aconitase is active over a broad range of temperature from 25°C to 45°C (Figure 4.7). The enzyme could be renatured after 30 minutes of incubation between 55°C to 60°C, where considerable activity could be regained. Thermal inactivation remains irreversible after incubation at 75°C for half an hour. The Arrhenius plot [ln(velocity) vs 1/T in Kelvin] shows a single peak (Figure 4.7), indicating presence of a single interconvertible form with a single activation energy irrespective profiles of both native and reconstituted protein remain by and large the same.

M. tb aconitase existed in a single interconvertible form irrelevant of temperature change

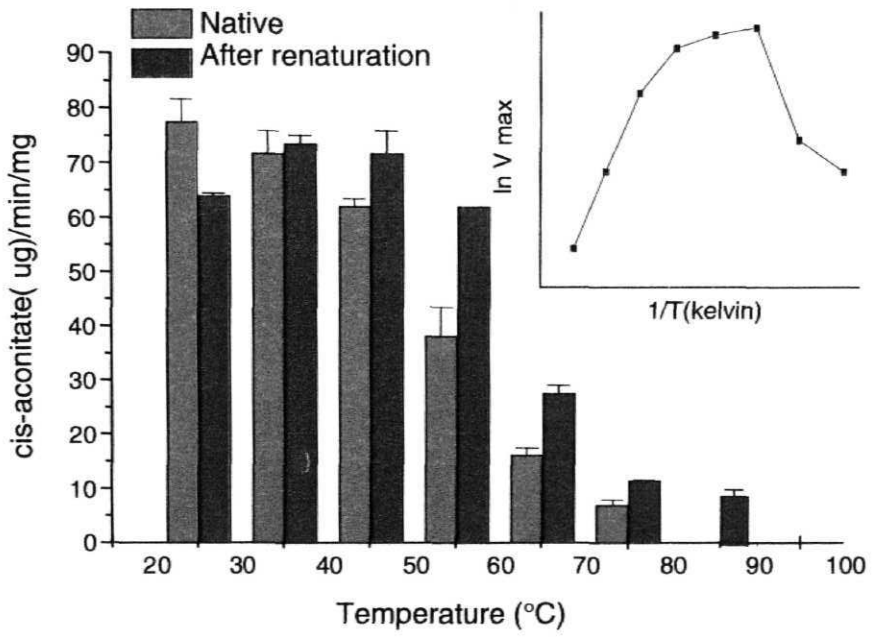


Figure 4.7: Temperature vs activity profile of native and renatured *M. tb* aconitase. Inset: The Arrhenius plot (ln(velocity) vs 1/T in Kelvin) showing a single peak, indicating presence of a single interconvertible form with a single activation energy irrespective of change in temperature

Metal ion as cofactors:

Divalent metal ions are not required as co-factors for *M. tb* aconitase activity

The enzyme (native and reconstituted) was tested for metal ion requirement with respect to four divalent metal ions, namely, Mg^{++}, Zn^{++}, Ca^{++} and Mn^{++} (Figure 4.8). It is apparent that there is no change in activity either in presence or absence of metal ions. Pre-incubation of isocitrate with metal ions also did not alter the rate of reaction (Figure 4.8).

Salt dependence of the *M.tb* aconitase suggested involvement of ionic interactions in stability and catalysis

The effect of NaCl on the stability as well as activity of the native as well as reconstituted enzyme was checked and it could be concluded that min 100mM – 200mM of salt is required for efficient activities in both cases (Figure 4.9). The absence of NaCl in the reaction decreased the activity of *M. tb* aconitase drastically. Higher concentrations of salt (above 300mM) proved to be detrimental for the activity. The decrease in the activity at higher salt concentration indicates involvement of ionic interactions during catalysis. The reversibility of the activity of the enzyme after removal of excess of salt, could not be checked as the enzyme precipitated upon prolonged exposure to room temperature.

Sequence alignment and phylogenetic analysis of *M. tb* aconitase: *M. tb* aconitase belongs to cytosolic aconitase (Acn A / IRP group)

Bacterial aconitase are classified into two main groups, those similar to eukaryotic iron regulatory proteins / cytosolic aconitases (the Acn A / IRP group) and the aconitase family found only in bacteria, the Acn B group. Recently one more group has been added that includes the aconitases closely related to mitochondrial aconitases. Multi-sequence alignment of primary protein structure of *M. tb* aconitase with other members of aconitase family followed by phylogenetic analysis puts *M.tb* Acn in AcnA/IRP group in broad classification of

Divalent metal ions were not required as co-factors for *M. tb* aconitase activity

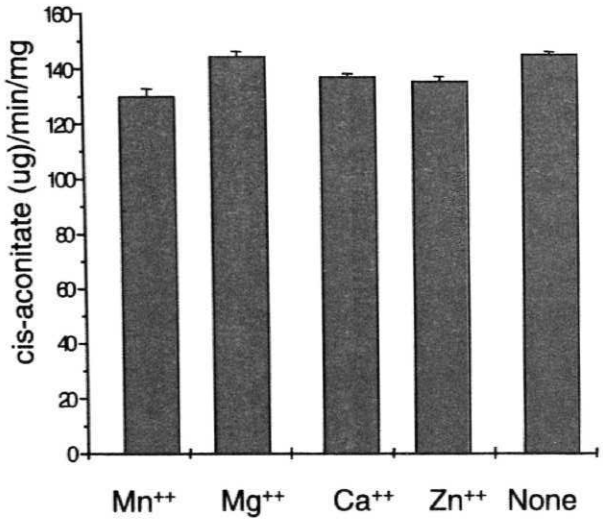


Figure 4.8: Activity profile of *M.tb* aconitase in absence and presence of Mn^{++} , Mg^{++} , Ca^{++} and Zn^{++} .

Salt dependence of the *M.tb* aconitase suggested involvement of ionic interactions in catalysis

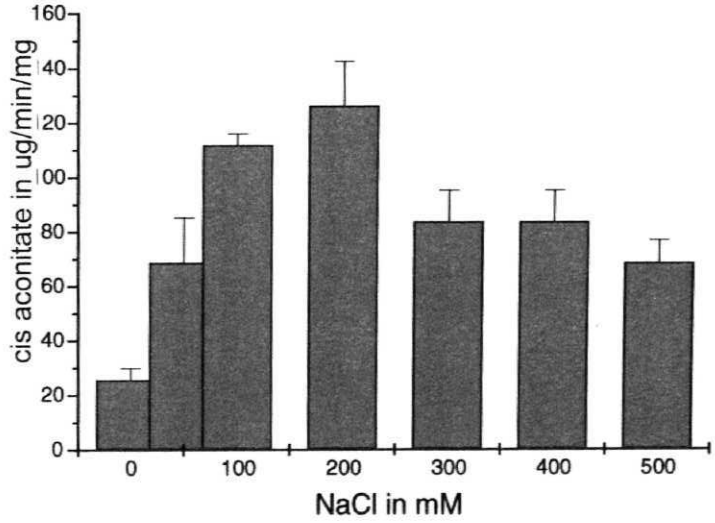


Figure 4.9: Activity profile of *M.tb* aconitase as a function of salt concentration

bacterial Acn family (Figure 4.10). Even though it could be observed that *M.tb* Acn clusters with aconitases of other *Mycobacterial* species, it is interesting to note the close proximity of *M.tb* Acn with those of bacteria that are mainly human pathogens such as *Corynebacterium diphtheria, Salmonella typhimurium, Shigella flexneri, Xanthomonas campestris* and *Xylella fastidiosa* etc (Figure 4.10). This suggests a possibility of lateral transmission of this gene over the evolutionary period of time. The distribution of sequences in phylogenetically divergent but ecologically similar microorganisms is ascribed to as lateral gene transfer (horizontal transfer). On the other hand vertical transmission of genes occurs within a species from one generation to the next through inheritance. Unlike eukaryotes, which evolve primarily through the modification of existing genetic information bacteria have obtained a significant proportion of genetic diversity through acquisition of genes form distantly related organisms. These lateral transfer produce extremely dynamic genomes and have effectively changed ecological and pathogenic character of bacterial species. As seen from the phylogenetic analysis the species specific region (Acn sequences) show high levels of protein sequence similarity from taxonomical families inferred to be very divergent suggesting the possibility of lateral gene transfer, however we strongly propose the possibility of a common ancestral gene.

The relatedness of *M.tb* Acn of *E.coli* is what puts *M. tb* in Acn A group The fact that *M.tb* Acn has higher identity to *E.coli* Acn A (60 %) than *E.coli* Acn B (20%) is of importance. AcnB of *E.coli* is the major aconitase for logarithmic growth in aerobic *E.coli* culture. Its synthesis is repressed in stationery phase and by anaerobiosis while the AcnA of *E.coli* is induced in stationery phase or during oxidative stress. This gives us a clue that *M.tb* aconitase might have a similar expression profile. Acn A is a more robust enzyme than Acn B and can withstand oxygen-mediated denaturation. Hence it is logical of *M. tb* Acn, which is strictly an aerobe, to be closely related to *E. coli* AcnA.

M. tb aconitase belongs to cytosolic aconitase (Acn A / IRP) group)

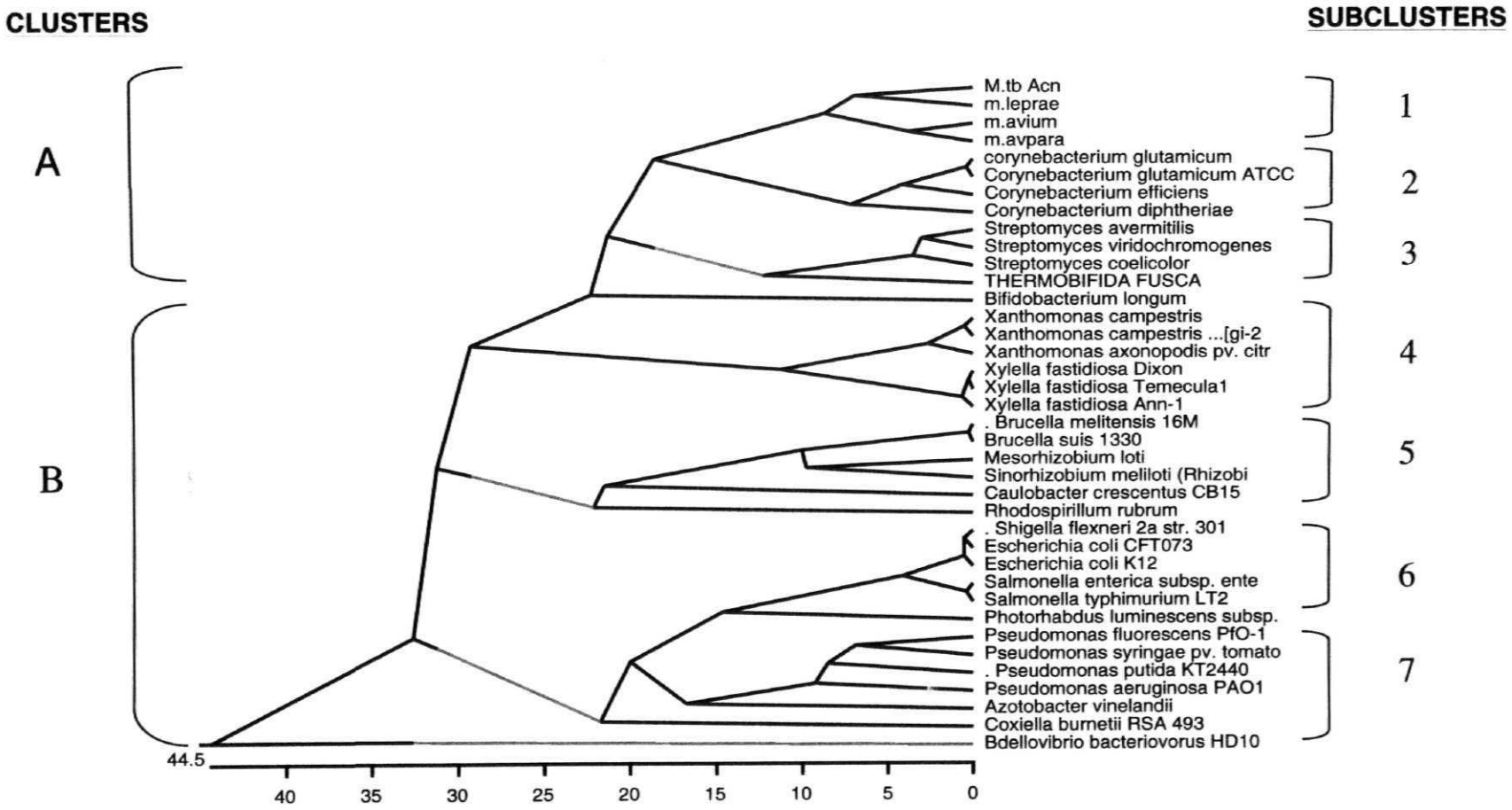


Figure 4.10: Phylogenetic tree of *M. tb* aconitase

B. Characterization of *M. tb* aconitase as a RNA binding protein

IRE like sequences in *M. tb* genome: As a first step to determine RNA binding activity of *M. tb* aconitase the *M. tb* genome was scanned for IRE like structures using the pattern search program of TubercuList (http://genolist.pasteur.fr/TubercuList/) for the sequence C N N N N N C A G U G, followed by secondary structure prediction (http://www.genebee.msu.su/services/rna2_reduced.html) for possible stem loop like configuration of typical iron responsive elements. Many candidates were enlisted after the pattern search for consensus IRE sequence. However, the selection of putative IRE-like sequences for the binding assays was based on their presence at untranslated regions (UTRs) of ORFs that were annotated to be involved in iron or oxygen metabolism of the cell.

5' IRE of *M. tb trxC* (thioredoxin) was thus selected because of its involvement in redox reactions. Cellular redox status is maintained by thioredoxin that participates in various redox reactions through the reversible oxidation of its active center dithiol, to a disulfide and catalyzes dithiol-disulfide exchange reactions. As mentioned earlier, aconitases, owing to their iron-sulphur cluster act both as iron and oxygen sensors (3, 4) and are affected by oxidative stress (14). Therefore it was postulated that *M. tb* aconitase might bind to IRE structure at 5' UTR of *trxC*. Other ORF that was selected was *M. tb acn* for the mere fact that aconitase is known to autoregulate its own expression by binding to 3' IRE like structures in other prokaryotes (31). As in *E. coli*, the 3' end of *M. tb acn* also showed multiple IRE – like structures.

IdeR is iron-dependent repressor and activator. 5' UTR containing IRE-like sequence was selected as it is directly linked to iron metabolism by acting as iron-binding repressor of siderophore biosynthesis and iron uptake.

***In vitro* transcription**

The complimentary oligodeoxyribonucleotides of the selected sequences were cloned in pGEM-3Zf vector at *Eco*RI and *Hind*III sites. The clones were confirmed by sequencing. The *in vitro* transcribed RNA after phenol extraction, was precipitated with 0.5M ammonium acetate, 15-20μg of yeast RNA and one volume of isopropanol at -20°C overnight. The RNA was dissolved in nuclease free water and checked on 7M denaturing 15% polyacrylamide gel (Figure 4.11).

***M. tb* aconitase binds to IRE-like RNA sequences with high specificity:** Having shown that *M. tb* aconitase has enzymatic activity, electrophoretic mobility shift assays were carried out to ascertain if *M. tb* aconitase indeed also has RNA binding properties. Binding of purified *M. tb* aconitase to selected IRE sequences was tested by assaying the interaction between the recombinant protein and *in-vitro* transcribed radiolabelled 57 – nucleotide RNA carrying IRE-like sequences (Figure 4.12). The RNA probes were denatured followed by slow cooling for attaining proper IRE-like configuration. IRE-RNA-aconitase protein complex is clearly evident when compared to the free probe lanes (Figure 4.12, lanes 1, 5, 9 and 12). The complex is seen with the control 5' human ferritin IRE (Figure 4.12, lanes 1- 4) as well as the selected IRE-like sequences of the *M. tb* genome; i.e.; 5' *M. tb trx*C (Figure 4.12, lanes 5- 8), 5'*M. tb Ide*R (Figure 4.12, lanes 9-11) and 3'*M. tb acn* (Figure 4.12, lanes 12-14) under these binding conditions. That these mobility shifts are specific is evident from homologous (specific) and heterologous (non-specific) cold competition experiments. The labeled IRE- like sequences is effectively competed out in the presence of 75X molar excess of specific unlabelled RNA corresponding to the ferritin probe (lane 4), 5' *M. tb trx*C (lane 8), 5'*M. tb Ide*R (lane11) and 3'*M. tb acn* (lane 14) confirming the specificity of the complex. Further more, the complex could not be competed out by
unlabelled non-specific RNA (Figure 4.12, lanes 3 and 7). Any kind of non-specific binding with the vector was ruled out by using a vector control where pGEM-3Zf vector was linearized with *Hind*III and

In-vitro transcription

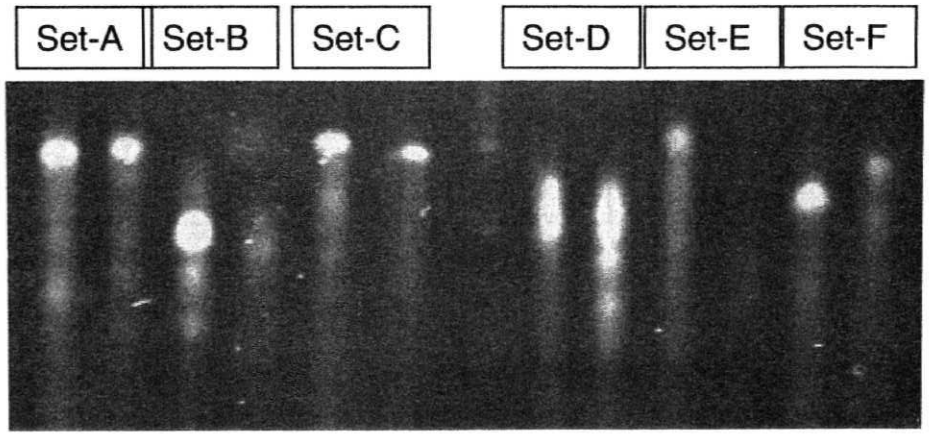


Figure 4.11: The complimentary oligodeoxyribonucleotides of the selected sequences were cloned in pGEM-3Zf vector at *EcoR*I and *Hind*III sites followed by in-vitro transcription. The RNA was checked on 7M denaturing 15% polyacrylamide gel. A and B: mammalian IREs;
C: 5' UTR of *M. tb ide*R; D: 3' UTR of *M. tb acn;* E: Bacillus *qou*X; F: 5' UTR of *M. tb trx*C.

M. tb aconitase binds to IRE-like RNA sequences with high specificity

IRE	human Ferritin				5' *M.tb trx*C					5' *M.tb ide*R			3' *M.tb acn*			Vector Control		
IRE Probe	+	+	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+
Rec. Aconitase	-	+	+	+		-	+	+	+	-	+	+	-	+	+	-	+	+
Specific Cold Comp.	-	–	-	+		-	–	-	+	-	–	+	-	–	+	-	–	+
Non-Specific Cold Comp.	-	-	+	-		-	-	+	-	-	-	-	-	-	-	-	-	-

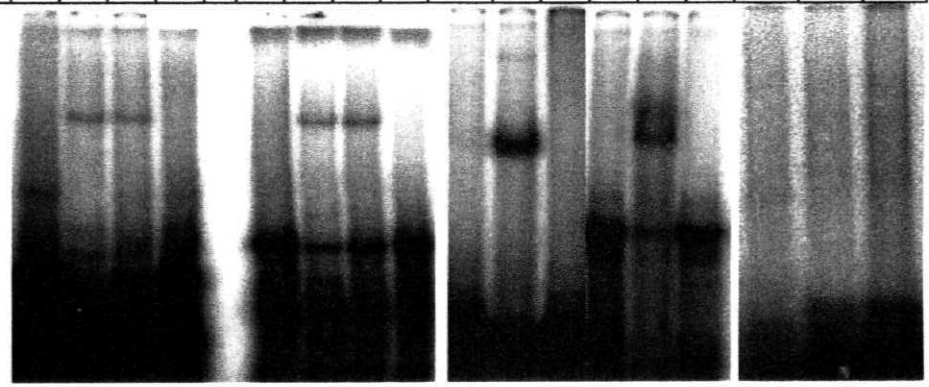


Figure 4.12: Gel retardation assays with the control 5' human ferritin IRE (lanes 1- 4), 5' *M. tb trx*C (lanes 5 - 8), 5'*M. tb Ide*R (lanes 9 -11), 3'*M. tb acn* (lanes 12 - 14) and vector control (lanes 15 - 17). The specificity of the shift was checked using 75X specific and non-specific competition

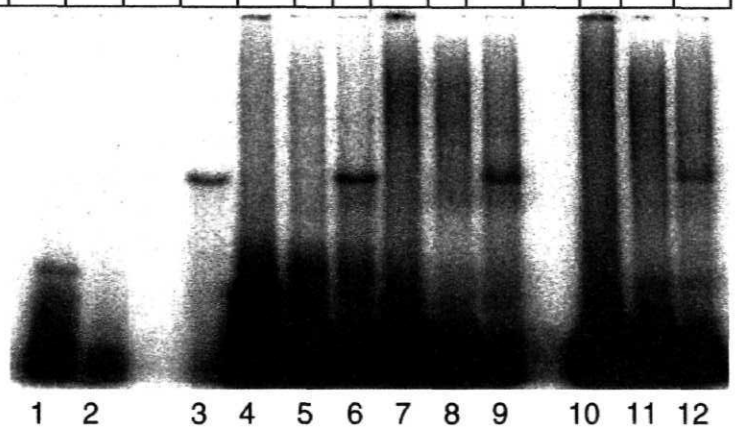
subjected to *in vitro* transcription and the labeled transcript was used as the vector control. *M. tb* aconitase did not bind to the vector RNA (Figure 4.12, lanes 15-17). In another control experiment the ability of these RNA to bind to the non-specific protein was rested. The selected IRE-like sequences did not bind to any of nonspecific proteins that were tested, namely, *M. tb* ICD-1, *M. tb* ICD-2 and BSA (data not shown). These experiments showed that *M. tb* aconitase is a RNA binding protein and binds to IRE-like sequences present in UTRs of 5' *M. tb trx*C, 3'*M. tb acn* and 5'*M. tb Ide*R mRNA.

RNA binding activity and aconitase activity of *M. tb* aconitase are mutually independent: Having shown the *M. tb* aconitase binds specifically to IRE elements present in the RNA, experiments were designed to investigate whether the enzymatic activity is independent of the RNA binding property. Recombinant purified aconitase was reconstituted with iron and inactivated with specific iron chelator dipyridyl. The results clearly shows that reconstitution of the purified protein with iron salt increases the enzymatic activity but decreases the IRE-binding ability of *M. tb* aconitase (Figure 4.13A, lanes 2, 5, 8 and 11). Further, depleting iron from the purified protein by dipyridyl inactivates the enzyme without affecting the RNA binding property of *M. tb* aconitase (Figure 4.13A, lanes 3, 6, 9 and12). The reconstituted and the inactivated proteins were checked for their enzyme activity by monitoring the concomitant formation of NADPH from $NADP^+$ in a two step reaction (Figure 4.13B). In the reaction, citric acid is converted into isocitrate by aconitase and the second enzyme isocitrate dehydrogenase converts isocitrate into α-ketoglutaric acid in presence of NADP, which is reduced to NADPH. The formation of NADPH is monitored by the increase in absorbance at 340nm. The rate of NADPH production is proportional to aconitase activity. Porcine heart isocitrate dehydrogenase (Sigma) was used for these reactions. Each 400 µl reaction mix contained 20mM triethanolamine chloride buffer pH 7.5, 2mM NADP+, 2mM citrate, 100mM NaCl, 10mM $MgCl_2$, 5µl porcine heart isocitrate dehydrogenase and100nM of *M. tb* aconitase. The reaction was tested both in absence and presence of *M. tb* aconitase to check for any

RNA binding activity of *M. tb* aconitase is independent of its aconitase activity

A

IRE	human Ferritin				5' *M.tb trx*C			3' *M.tb acn*				5' *M.tb ide*R		
IRE probe	+	+	-	+	+	+	+	+	+	+	-	+	+	+
Rec. Aconitase	-	+	-	+	-	+	+	-	+	+	-	-	+	+
Fe^{++}	-	+	-	-	-	+	-	-	+	-	-	-	+	-
Dipyridyl	-	-	-	+	-	-	+	-	-	+	-	-	-	+



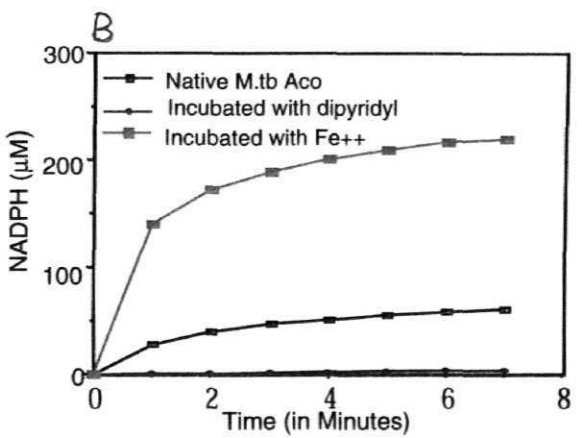


Figure 4.13: (A) Gel retardation assays of Fe^{++} and dipyridyl treated *M. tb* aconitase with 5' human ferritin IRE (lanes 1 – 3), 5' *M. tb trx*C (lanes 4 - 6), 5' *M. tb acn* (lanes 7 -9) and 3' *M. tb Ide*R (lanes 10 - 12). (B) Comparative activity curves of native, reconstituted and dipyridyl treated *M. tb* aconitase used for assays in (A)

residual activity. These results clearly demonstrated that the ability of *M. tb* aconitase to bind to RNA is independent of its enzymatic property.

Factor Xa protection assay show site specific binding of IREs to *M. tb* aconitase

Results presented above have shown that *in vitro* manipulations that assembles (through reconstitution of the native protein by Fe^{++}) and disassembles (through iron-specific chelator dipyridyl) the 4Fe-4S cluster of *M. tb* aconitase results in interconversion between two mutually exclusive functional states. Earlier reports on mammalian cytosolic aconitase which, like *M. tb* aconitase, also belongs to AcnA/IRP group, have shown through mutagenesis and genetic studies that the three cluster-ligating cysteine residues are essential for *in vivo* RNA-binding and aconitase activity (Philpott *et al*, 1994) and these cysteine residues are highly conserved throughout the aconitase/IRP family. The corresponding cysteine residues in *M. tb* as revealed by multiple sequence alignment are Cys 479, 545, 548. These residues may be important for 'Fe-S' cluster ligation and interaction with IRE-RNAs. Based on this information protease protection assays were performed to predict the possible region of the protein to which IREs bind in *M. tb*. The protease selected for the study was Factor Xa that has only two cleavage sites, at 449 and 583 amino acid residues, in *M. tb* aconitase. Therefore cleavage with Factor Xa would result in production of three truncated fragments of *M. tb* aconitase corresponding to ~48 kDa, ~41 kDa and ~12 kDa. All the predicted cysteine residues fall between these two Factor Xa cleavage sites. Therefore it was hypothesized that RNA binding to this region of the protein would confer protection against Factor Xa cleavage, resulting in partial digestion with formation of ~63 kDa and ~54 kDa fragments (Figure 4.14A).

M. tb aconitase, both bound to RNA and unbound, were subjected to Factor Xa digestion at 22°C for 24 hours, followed by fractionation of the digested fragments in 10% SDS-PAGE (Figure 4.14B). Protein without Factor Xa was also incubated at 22°C for 24 hours to check

Factor Xa protection assay show site specific binding of IREs to *M. tb* aconitase

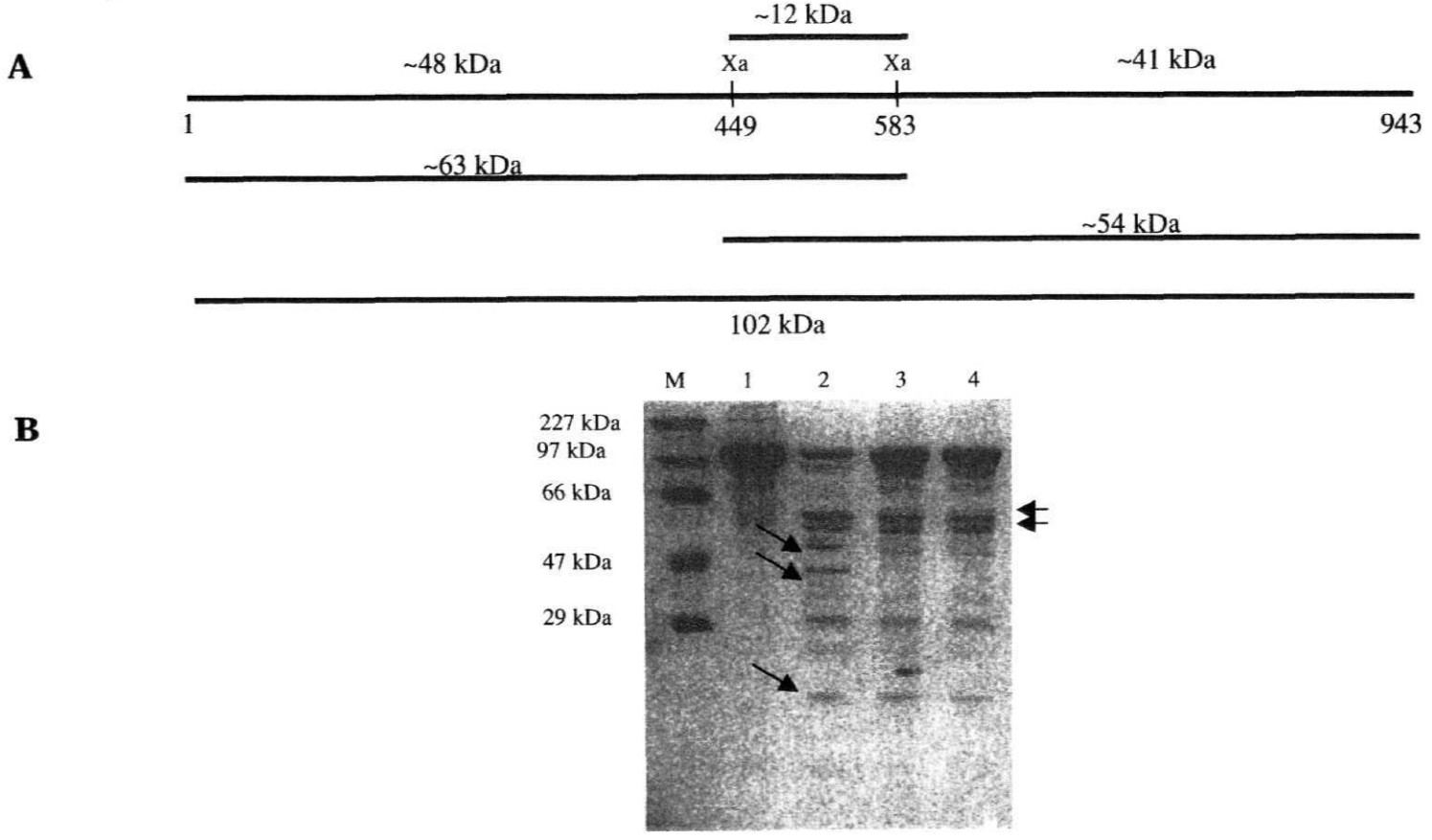


Figure 4.14: (A) FactorXa site in M. tb Acn and schematic representation of digestion pattern of M. tb Acn after factor Xa treatment. (B) *M. tb* aconitase, both bound to RNA and unbound, were subjected to Factor Xa digestion at 22°C for 24 hours, followed by fractionation of the digested fragments in 10% SDS-PAGE. Different lanes are: M: protein size marker; Lane1: Native protein at 22°C for 24hrs; Lane 2: Native protein without RNA at 22°C for 24hrs; Lane 3: Native protein bound to IRE1 at 22°C for 24hrs; Lane 4: Native protein bound to IRE2 at 22°C for 24hrs.

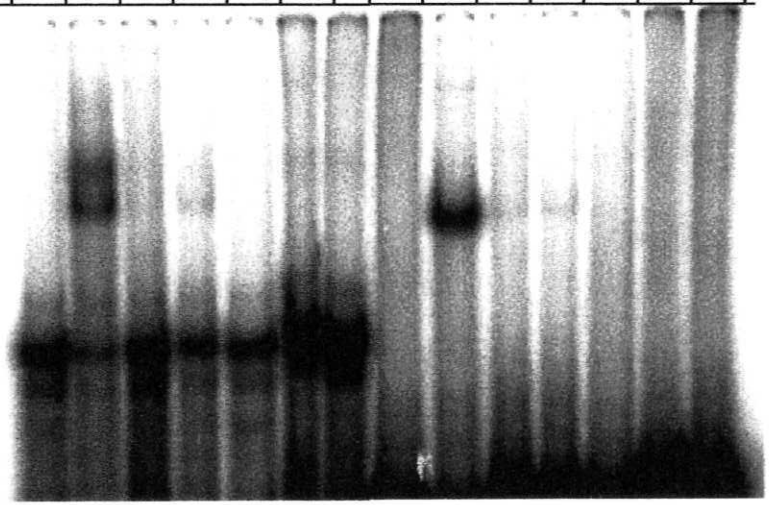
for degradation in that temperature (lane 1). The ~48 kDa and ~41 kDa fragments can be prominently seen in lane 2, which is due to complete digestion of the unbounded (without RNA binding to the protein) and hence unprotected protein. Lane 3 and 4 show digested fragments of *M.tb* aconitase bound to 5' *M. tb trx*C and 3'*M. tb acn* RNAs respectively, where only~63 kDa and ~54 kDa bands are conspicuous. Some background bands corresponding to ~48 kDa and ~12 kDa can be seen in lane 3 and 4. That is because binding of RNA to protein may not be 100% efficient. These result suggest that RNA bind to this region that also includes Cys 479, 545, 548 in *M. tb* aconitase.

Cysteine residues are important for IRP activity

The importance of the cysteine residues in RNA binding activity of *M. tb* aconitase was further investigated by chemical inactivation of cysteine residues by 0.4mM NBDCl (4-chloro-7-nitro-2,1,3-benzoxadiazole). NBDCl-treated recombinant aconitase along with reconstituted and dipyridyl-treated aconitase were used for gel-shift assays with respect to two predicted IRE elements of *M. tb* genome 3' *M.tb Acn and* 5' *M.tb Ide*R (Figure 4.15). The gel-picture clearly shows that NBDCl-treated aconitase could not bind to 3' *M.tb Acn* or 5' *M.tb Ide*R RNA (Figure 4.15, lanes 6 and lane 13 respectively). Also, consistent with earlier experiments, reconstituted aconitase did not bind to 3' *M.tb Acn and* 5' *M.tb Ide*R RNA (Figure 4.15, lane 5 and lane 12 respectively) either. Aconitase treated with dipyridyl that was expected to have a disassembled Fe-S cluster, alone could bind to these RNA sequences (Figure 4.15, lane 2 and lane 9). The experiment, thus, showed that cysteine residues are important for RNA binding activity of *M. tb* aconitase. It is to be noted that there are six cysteine residues in *M. tb* aconitase and NBDCl-treatment would indiscriminately inactivate all these residues. Therefore, inhibition of the RNA binding property of *M. tb* aconitase by cysteine residue inactivation may be more due to change in the structural conformation of the protein rather than inactivation of Cys 479, 545, 548 at the catalytic cleft.

Cysteine residues are involved in IRP activity of *M. tb* aconitase

IRE	3' *M.tb Acn*							5' *M.tb IdeR*						
IRE probe	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Dipyridyl- Aconitase	-	+	+	+	-	-	-	-	+	+	+	-	-	-
Cold comp. (75X)	-	-	+	-	-	-	-	-	-	+	-	-	-	-
Cold comp. (50X)	-	-	-	+	-	-	-	-	-	-	+	-	-	-
Activated ACN	-	-	-	-	+	-	-	-	-	-	-	+	-	-
NBDCl treated ACN	-	-	-	-	-	+	-	-	-	--	-	-	+	-
NBDC +DMSO	-	-	-	-	-	-	+	-	-	-	-	-	-	+



1 2 3 4 5 6 7 8 9 10 11 12 13 14

Figure 4.15: Chemical inactivation of cysteine residues by 0.4mM NBDCl (4-chloro-7-nitro-2,1,3-benzoxadiazole) inhibits the RNA binding property of *M. tb* Aconitase. Gel retardation assays of NBDCl-treated (lanes: 6 and 13), Fe^{++}- treated (lanes: 5 and 12) and dipyridyl treated (lanes: 2 and 9) *M. tb* aconitase with 5' *M. tb acn* (lanes 1-7) and 3' *M. tb IdeR* (lanes 8-14).

DISCUSSION

The property that makes iron indispensable to cell is also one that makes it potentially toxic to the cell. With its ability to readily exchange electrons, iron forms an important link in all fundamental cell functions including those that lead to formation of free radicals and ultimately oxidative stress and cell damage (Lill *et al*, 2000; Mühlenhoff *et al*, 2000; Mühlenhoff *et al*, 2002). This love-hate relationship is delicately balanced between 'bound' and toxic 'free-iron' inside cell. The aconitase superfamily includes aconitases and iron regulatory proteins. These proteins are closely related by virtue of their conserved residues (Walden, 2002). Post-transcriptional regulation by cytosolic aconitase or Iron Responsive Proteins in eukaryotes is a well established phenomenon (Gegout *et al*, 1999; Allerson *et al*, 2003; Loyevsky *et al*, 2001; Zhang *et al*, 2002). Prokaryotic aconitases are highly similar to eukaryotic IRPs (Walden, 2002; Baughn *et al*, 2002), but RNA binding property and regulatory function has been studied in a few prokaryotes, like *E. coli* (Tang *et al* 1999), *Bacillus subtilis* (Alén *et al*, 1999), *Xanthomonas campestris* (Wilson *et al*, 1998), *Pseudomonas aeruginosa* (Somerville *et al*, 1999) and a few others. These reports have also attempted to establish an indirect link between regulatory function of aconitase and virulence. This chapter presents evidence, for the first time, supporting bifunctional activity of *M. tb* aconitase. It was shown that purified, recombinant *M. tb* aconitase is functional both as a TCA cycle enzyme where it catalyses the reversible isomerization of citrate and isocitrate *via* the dehydrated intermediate cis-aconitate and as a RNA binding protein where it binds to selected IRE like sequences present within the UTRs of 5' *M. tb trx*C, 3'*M. tb acn* and 5'*M. tb Ide*R mRNA. Further it was demonstrated that these two functions are mutually exclusive.

Size exclusion chromatography followed by the assay for aconitase present in the eluted fractions presented some interesting features of the *M. tb* aconitase. A functional trimer form of aconitase was consistently present. That *M. tb* aconitase existed as a monomer was more than evident. The observed trimeric form was more likely due

to aggregation involving ionic interactions (Figure 4.4C). There was an insignificant decrease in Km while Vmax for both monomer and trimer remained the same. A V vs [S] plot suggested insignificant substrate binding cooperativity which could be the possible reason for increased substrate affinity of the trimer (Figure 4.4D).

Enzymatically, *M. tb* aconitase could tolerate a broad range of pH. This is expected given the fact that *M. tb* aconitase is closer to *E. coli* AcnA which is expressed during stationary phase when the intracellular pH is expected to vary over a wider range than log phase (Jordan *et al* 1999). The pH is an important parameter for any enzyme activity. The change in activity with change in pH indicates the change in ionization state of the side chains of amino acids involved. The amino acid residues involved could be roughly predicted by extrapolation of the linear portion of the plot $\log_{10}$ Vmax Vs pH. A parabolic curve (Figure 4.5B) implied involvement of at least two ionizing groups with pKa values corresponding to the points of intersection of the slopes. The nearest pKa values (± 1.5) corresponding to 7.2 and 9.3, suggested the involvement of histidine and cysteine side chains. Since the environment of the side chain in an enzyme can cause a shift in pKa values of free amino acids, it was further checked if these amino acid residues were conserved amongst the catalytically important residues along the phylogenetic line by multiple sequence alignment (Figure 4.6). His-127, 208 and Cys-479, 545, 548 were highly conserved in this family of proteins. Lysine and Tyrosine, whose pKa values lie within the ±1.5 range were not present amongst the conserved residues of catalytic site. This therefore adds evidence to the interpretation from $\log_{10}$ Vmax Vs pH plot that histidine and cysteine residues are involved in ionization during catalysis *of M. tb* aconitase.

The selection of probable IRE like sequences in *M. tb* genome was difficult because, even though IRPs bind to consensus IRE with high affinity, earlier studies have shown that IRPs also bind to alternate ligands with relative different efficiencies. Relative binding efficiency differs as per change in nucleotide within the stem-loop (Butt *et al*,

1996). Phylogenetic relativity, mutational analysis and comparisons among functional IREs in different transcripts have permitted the identification of the signatures of the IRE that are necessary for high affinity binding by the IRPs. The consensus structure of an IRE consists of a base-paired stem, interrupted by an unpaired 'C' five base pairs removed from a six member loop. The sequence of loop is always CAGUGN (even though not strictly) with atleast seven Watson-Crick base pairs to either side of the loop excluding the bulge (Butt *et al*, 1996). One or more mismatches in the loop region was opted while scanning the *M. tb* genome as it was evident from earlier reports that CAGUGN may not be the strict consensus. It is interesting to mention here that both 3' UTR of *M. tb acn* and 5' UTR of *IdeR* had two stem-loop like probable IREs each. Both the regions were tested for their ability to bind to *M. tb* aconitase, however only one region could bind. Presence of multiple stem-loop at UTRs is not unusual and could be seen in other prokaryotes and eukaryotes as well (Koeller *et al*, 1989; Tang *et al*, 1999).

Aconitase has 3 Fe atoms directly bound to the cysteine residues of the enzyme backbone, while the fourth Fe is ligated to the sulphur of the inactive (3Fe-4S) cluster. This iron-sulphur ligation provides a free co-ordination site that is involved in the binding of the substrate to the active site of the aconitase (Lauble *et al*, 1994). Increase in the affinity for the substrate by *M. tb* aconitase upon reconstitution, as implied by lowering of Km for the reconstituted protein as compared to the native protein (Table IV), pointed to the following status of the protein. Firstly, there is a loss of the 4^{th} labile iron during protein purification as a result of which a significant portion of purified enzyme is in the inactive [3Fe–4S] or apoenzyme form. The restoration of the labile 4^{th} iron after reconstitution increases its affinity for the substrate. Iron and dipyridyl (a chelator of Fe^{++}) had antagonistic effect on aconitase activity and RNA binding ability of *M. tb* aconitase. The reconstituted protein did not bind to IRE like sequences studied while inactivation by dipyridyl increased the binding affinity of *M. tb* aconitase (Figure 4.13). Thus *M. tb* aconitase

could act both as an enzyme and a RNA binding protein depending upon the availability of labile iron to the protein.

Many genes in *M. tb* have since been characterized after genome annotation. Some of these have been found to have additional properties than was expected by homology based predictions (Chakhaiyar *et al*, 2004; Choudhary *et al*, 2003; Siddiqi *et al*, 2004). The present work while documenting interesting biochemical properties of *M. tb* aconitase, also points to a hitherto unknown property of *M. tb* aconitase. Binding of *M. tb* aconitase to such unrelated IREs like that of mammalian ferritin validate earlier suggestions that such interactions are not species specific, even though the affinity by which interspecies IRE and IRP bind might differ (Alén *et al*, 1999; Wilson *et al*, 1998; Somerville *et al* 1999). The Acn system has so far been shown to autoregulate aconitase synthesis in prokaryotes (Tang *et al*, 1999). Presence of IRE-like sequences at 3' UTR of *M. tb acn* and binding of *M. tb* aconitase to these sequences makes it very tempting to speculate that *M. tb* aconitase may also be involved in autoregulation as in other prokaryotes. However, as it also binds to some other IRE-like sequences in *M. tb* genome, a more general role of regulating gene expression at post transcriptional level can not be ruled out. Even though this work has established *M. tb* aconitase as a RNA binding protein, the mechanism by which it regulates other ORFs and its own expression remains elusive. It would be important to characterize the Acn regulatory system of *M. tb* and its involvement in virulence, identify other Acn-regulated genes, the effects of oxidative stress and define the molecular basis of the *M. tb* aconitase and IRE- mRNA interactions.

Chapter 5

Summary

Importance of basic metabolic pathways in the development of a microbial disease became apparent with the discovery of involvement of fatty acid metabolic pathways and glyoxylate pathway in virulence in many pathogens (De voss *et al*, 2000; McKinney *et al*, 2000). This implies that if a pathogen is unable to synthesize the precursors of important metabolites required for growth, it is unlikely to proliferate or persist in host. Several adaptive alterations in house keeping metabolic pathways have been observed in diverse group of organism in response to internalization by macrophages or contact with host immune system. One of the prominent adaptations found amongst varied organisms from commensals to phytopathogens (McKinney *et al*, 2000; Manabe *et al*, 1999; Goldstein *et al*, 2001; Lorenz *et al*, 2002) is adjustments in physiology in response to nutrition starvation inside phagolysosome in macrophages. The microenvironment is usually devoid of complex carbon sources like glucose and certain essential amino acids. Pathogen in turn perceives simple carbon sources like acetyl Co-A or C2 sources or deficiency in essential metabolites like magnesium and iron in their environment and responds accordingly. Whole genome sequencing of *Mycobacterium tuberculosis* and microarray technology has presented a global picture of these adaptations in terms of controlled expression of selected genes. Often an additional, altered or unusual function of a gene has been noted under such conditions.

The work presented in this thesis is an extended investigation of the basic energy pathway enzymes, isocitrate dehydrogenase and aconitase and understanding their distinctive role in immune response and iron metabolism respectively as additional properties.

Two isoforms of isocitrate dehydrogenase, *M. tb icd*-1 and *M. tb icd*-2, have been identified in the *M. tb* genome. This study, for the first time, biochemically characterizes *M. tb* ICD-1 and ICD-2 in terms of their substrate and coenzyme specificity, metal ion requirement, pH, temperature, salt tolerance and oligomeric assembly. Differences in coenzyme affinity, oligomeric state, pH tolerance and phylogenetic affiliation of the two isoforms of *M. tb* ICD were evident. While *M. tb*

ICD-1 also accepts Zn^{+2} ion as metal ion apart from Mg^{+2} for catalytic reactions, *M. tb* ICD-2 did not show any activity in presence of Zn^{+2}. *M. tb* ICD-1 could tolerate a wider pH range than *M. tb* ICD-2 and showed 33-35% activity in an acidic pH upto 5. The optimum temperature for the enzyme activities was 37°C - 45°C. Both the enzymes required $NADP^+$ as the coenzyme, but remained inactive in presence of NAD^+. Km [isocitrate] for *M. tb* ICD-1 in presence of Mg^{++} or Zn^{++} was 10μM ± 5 and 22μM ± 7, respectively while that for *M. tb* ICD-2 was 20μM ± 1 in presence of Mg^{++}. *M. tb* ICD-1 is a homodimeric $NADP^+$ dependent isocitrate dehydrogenase, whereas, *M. tb* ICD-2, originally annotated as a monomeric protein, was found to exist in a dimeric state. Phylogenetic analysis revealed that unlike *M. tb* ICD-2 which groups with bacterial ICDs, *M. tb* ICD-1 exhibits a closer lineage to eukaryotic $NADP^+$ dependent ICDs.

A differential expression of *M. tb* ICD-1 and ICD-2 under different stages of growth of *Mycobacterial tuberculosis* in *in vitro* culture was further demonstrated. The experiment pointed to a transient expression of the two ICDs. ICD-1 was expressed during early log phase and reappeared along with ICD-2 during late log and stationary phase. However, ICD-1 was less expressed than ICD-2 on 12^{th} and 18^{th} day quantitatively. The tolerance to a broader range of pH by ICD-1 as shown by biochemical assays may be a logical explanation to its reappearance with ICD-2 expression during late log or stationary phases when pH inside the cell is much more variable than log phase.

Isocitrate dehydrogenase (ICD) was found to be amongst proteins that are released from *Mycobacterium tuberculosis* during late logarithmic growth phase. In addition to biochemical characterization, the immunogenic properties of the two isoforms of *M. tuberculosis* ICDs were evaluated and compared the same with the control antigen – HSP 60 as well as purified protein derivative (PPD). PPD lacks the sensitivity to distinguish between BCG vaccinated and TB-infected populations and therefore epidemiological relevance of PPD in BCG vaccinated regions is debatable. It was shown that *M.tb*

ICDs elicit a strong B-cell response in TB-infected population and can differentiate between healthy-BCG vaccinated population and those with tuberculosis. The study population (n=215) was categorized into different groups, namely, patients with fresh infection (n=42), relapsed TB cases (n=32), patients with extra-pulmonary tuberculosis (n=35), 44 clinically healthy donors, 30 NTMs and 32 non-TB patients (culture negative for acid fast bacteria but carrying other infections). The *M.tb* ICDs showed statistically significant antigenic distinction between healthy-BCG vaccinated controls and tuberculosis patients ($p<0.0001$) and those with other infections. Surprisingly, extrapulmonary infections could not be discriminated from healthy controls by HSP 60 ($p= 0.2177$) which *M. tb* ICDs could do significantly ($p< 0.0001$). These results highlight the immunodominant, immunosensitive and immunospecific nature of *M.tb* ICDs and point to a novel and an unusual property of this TCA cycle enzyme. A combinatorial study of biochemical and immunological properties, differential expression and phylogenetic analysis while reflecting an attempt to trace the evolution of intracellular pathogenicity, may additionally help understanding the adaptive role of isocitrate dehydrogenase in intracellular persistence of this pathogen.

Aconitase is the common enzyme between both TCA cycle and glyoxylate shunt. Also, *M. tb* Acn is a Fe-S cluster containing protein. Such proteins are known to act as iron regulatory and sensor proteins. Cellular iron levels are one of the critical nutritional alteration inside infected macrophages and hence is closely monitored by *Mycobacterium tuberculosis* for survival. Amongst the sensor proteins reported in eukaryotic systems and a few prokaryotes are iron responsive proteins (IRPs) that bind to iron responsive elements (IREs) and regulate the translation or stability of mRNAs encoding proteins involved in iron homeostasis. Sequence comparison at amino acid level showed that it carries the conserved residues of IRPs. Both the enzymatic and the IRP activity of *M. tb* Acn were evaluated in this study. *M. tb* Acn functionally existed as a monomer and was enzymatically active in converting isocitrate to cis-

aconitate at a broad pH range of 7-10 (optimum pH 8) with a single interconvertible transition state. Along with being enzymatically active, it was observed that *M. tb* Acn could successfully bind to IRE-like sequences of eukaryotes taken as control and also to predicted IRE–like sequences found in *M. tb* genome. *M. tb* genome was scanned for IRE-like elements and the selected sequences were subjected to RNA secondary structure prediction. Three such regions were selected which are found in either 3' or 5' untranslated regions of mRNAs of *trxC*, *ideR* and *acn* that have probable role in iron metabolism. IRE – binding activity of *M. tb* Acn was checked by gel retardation assays. It was observed that availability of iron affected both the IRP and the enzymatic activity of *M. tb* Acn *in-vitro*. *M.tb* Acn when reactivated with Fe^{++} acted as a TCA cycle enzyme, while when depleted of iron by specific iron chelator acted like an IRP binding to the selected IREs. It was also shown that both the activities of *M. tb* Acn are mutually independent. Factor Xa protection assay show site specific binding of IREs to *M. tb* aconitase. The importance of Cysteine residues for both enzymatic and IRP activity was demonstrated. These results support that *M. tb* Acn is a bifunctional protein and points to its role in iron homeostasis.

Summarizing the above work, I have biochemically characterized Isocitrate dehydrogenases and Aconitase of *Mycobacterium tuberculosis* and have pointed out the additional properties of these two proteins apart from them being TCA cycle enzymes. The results reveal that *M. tb* ICDs have additional property to elicit high B-cell response in TB-patients and can differentiate healthy BCG vaccinated population from TB infected population. This study also points to post transcriptional regulatory role of *M. tb* Aconitase as a RNA binding protein. Lineage of these two proteins was studied in detail in an attempt to trace the evolution of intracellular pathogenicity.

References

Ahmed N, Alam M, Rajender Rao K, Kauser F, Ashok Kumar N, Qazi NN, Sangal V, Sharma VD, Das R, Katoch VM *et al.* **2004**. Molecular genotyping of a large, multicentric collection of tubercle bacilli indicates geographical partitioning of strain variation and has implications for global epidemiology of *Mycobacterium tuberculosis*. J. Clin. Microbiol. 42: 3240-3247.

Ahmed N and Hasnain SE. **2004**. Genomics of *Mycobacterium tuberculosis*: Old threats and new trends. Indian J Med Res (In press).
16.

Ahmed N, Mohanty AK, Batish VK, Mukhopadhyay U and Grover S. **1998**. PCR-based Rapid detection of *Mycobacterium tuberculosis* in blood from immunocompetent patients with pulmonary tuberculosis. J. Clinical Microbiol. 36 : 3094-3095

Ahmed N, Caviedes L, Alam M, Rao KR, Sangal V, Sheen P, Gilman RH and Hasnain SE. **2003**. Distinctiveness of *Mycobacterium tuberculosis* genotypes from human immunodeficiency virus type 1-seropositive and seronegative patients in Lima, Peru. J. Clin. Microbiol. 41: 1712-1716.

Ahmed N, Alam M, Majeed AA, Rahman SA, Cataldi A, Cousins D and Hasnain, SE. **2003**. Genome sequence based, comparative analysis of the fluorescent amplified fragment length polymorphisms (FAFLP) of tubercle bacilli from seals provides molecular evidence for a new species within the *Mycobacterium tuberculosis* complex. Infection Genet. Evol. 2: 193-199.

Alén C and Sonenshein AL. **1999**. *Bacillus subtilis* aconitase is an RNA-binding protein Proc. Natl. Acad. Sci. USA.96: 10412-10417.

Allerson CR, Martinez A, Yikilmaz E and Rouault TA. **2003**. A high capacity RNA affinity column for the purification of human IRP1 and IRP2 overexpressed in Pichia pastoris. RNA. 9: 364-74.

Altschul SF, Gish W, Miller W, Myers EW and Lipman DJ. **1990**. Basic local alignment search tool. J. Mol. Biol. 215: 403-410.

Andersen P. 1994. Effective vaccination of mice against *Mycobacterium tuberculosis* infection with a soluble mixture of secreted mycobacterial proteins.Infect. Immun. 62: 2536–2544.

Anderson P, Askgaard D, Ljungqvist L, Bennedsen J and Heron I. **1991**. Proteins released from *Mycobacterium tuberculosis* during growth. Infect. Immun. 59: 1905-1910.

Armitige LY, Jagannath C, Wanger AR and Norris SJ. **2000.** Disruption of the genes encoding antigen 85A and antigen 85B of *Mycobacterium tuberculosis* H37Rv: effect on growth in culture and in macrophages.Infect. Immun. 68: 767–778.

Armstrong JA and Hart PDA. **1975**. Phagosome-lysosome Interactions in cultured macrophages infected with virulent tubercle bacilli. J. Exp. Med. 142:1–16.

Azad AK, Sirakova TD, Fernandes ND and Kolattukudy PE. **1997**. Gene knockout reveals a novel gene cluster for the synthesis of a class of cell wall lipids unique to pathogenic mycobacteria. J. Biol. Chem. 272:

Balamir A. 1983. NADP-linked isocitrate dehydrogenase from beef liver: a new method of purification and the effect of metal ion cofactor on its stability. Biochem. Med. 29: 194-206.

Bange FC, Brown AM and Jacobs WR Jr. **1996**. Leucine auxotrophyrestricts growth of *Mycobacterium bovis* BCG in macrophages. Infect. Immun. 64: 1794–1799.

Banu S, Honore N, Saint-Joanis B, Philpott D, Prevost MC and Cole ST. **2002**. Are the PE-PGRS proteins of *Mycobacterium tuberculosis* variable surface antigens? Mol. Microbiol. **44:** 9–19.

Bardou F, Quemard A, Dupont MA, Horn C, Marchal G and Daffe M. **1996**. Effects of isoniazid on ultrastructure of Mycobacterium aurum and *Mycobacterium tuberculosis* and on production of secreted proteins. Antimicrob. Agents Chemother. 40: 2459-2467.

Barnes DS. 2000. Historical perspectives on the etiology of tuberculosis. Microbes Infect. 2: 431–440.

Barnes PF, Chatterjee D, Abrams JS, et al.**1992**. Cytokines production induced by *Mycobacterium tuberculosis* lipoarabinomannon-relationship to chemical structure. J immunol. 149: 541-547

Bartowr EA and Mcmurray DN. **1990**. Cellular and humoral responses to recombinant mycobacterial peptides in guinea pigs infected by the respiratory route By *Mycobacterium tuberculosis*. FASEB. 4: 2192

Bates JH and Stead WW. **1993**. The history of tuberculosis as a global epidemic. Med. Clin. N. Am. 77: 1205-1217

Baughn AD and Malamy MH. **2002**. A mitochondrial-like aconitase in the bacterium *Bacteroides fragilis*: Implications for the evolution of the mitochondrial Krebs cycle. Proc Natl Acad Sci U S A 99: 4662-4667

Beaucher J, Rodrigue S, Jacques PE, Smith I, Brzezinski R and Gaudreau L. **2002**. Novel *Mycobacterium tuberculosis* anti-sigma factor antagonists control sigma F activity by distinct mechanisms. Mol. Microbiol. 45: 1527–1540.

Behr MA, Wilson MA, Gill WP, Salamon H, Schoolnik GK, Rane S and Small PM. **1999**. Comparative genomics of BCG vaccines by whole-genome DNA microarrays. Science 284: 1520–1523.

Beinert H and Kennedy MC. **1993**. Aconitase, a two-faced protein: enzyme and iron regulatory factor. FASEB J. 7: 1442-1449.

Beinert H, Holm RH and Munck E. **1997**. Iron-sulfur clusters: nature's modular, multipurpose structures. Science 277: 653–659

Beinert H, Kennedy MC and Stout CD. **1996** Aconitase as Iron-Sulfur Protein, Enzyme, and Iron-Regulatory Protein. Chem Rev. 96: 2335-2374.l

Belisle JT, Vissa VD, Sievert T, Takayama K, Brennan PJ and Besra GS. **1997**. Role of the major antigen of *Mycobacterium tuberculosis* in cell wall biogenesis. Science 276: 1420–1422.

Berthet F, Rauzier J, Lim EM, Philipp W, Gicquel B and Portnoi D. **1995**. Characterization of the *Mycobacterium tuberculosis erp gene* encoding a potential cell surface protein with repetitive structures. Microbiology 141: 2123–2130.

Berthet FX, Lagranderie M, Gounon P, Laurent-Winter C, Ensergueix D, Chavarot P, Thouron F, Maranghi E, Pelicic V, Portnoi D, Marchal G and Gicquel B. **1998**. Attenuation of virulence by disruption of the *Mycobacterium tuberculosis erp* gene. Science 282: 759–762.

Betts JC, Lukey PT, Robb LC, McAdam RA and Duncan K. **2002**. Evaluation of a nutrient starvation model of *Mycobacterium tuberculosis* persistence by gene and protein expression profiling. Mol Microbiol. 43: 717-31.

Betts JC, Lukey PT, Robb LC, McAdam RA, and Duncan K. **2002**. Evaluation of a nutrient starvation model of *Mycobacterium tuberculosis* persistence by gene and protein expression profiling. Mol Microbiol. 43: 717-31.

Bharadwaj VP, Katoch VM, Sharma VD, Kannan KB, Datta AK, Shivannavar CT. **1987**. Metabolic studies on mycobacteria IV. Assay of isocitrate lyase and malate synthase activity in M. leprae. Indian J Lepr. 59: 158-62.

Blattner FR, Plunkett III G, Bloch CA, Perna NT, Burland V, Riley M, Collado-Vides J, Glasner JD, Rode CK, Mayhew GF, et al. **1997**. The complete genome sequence of Escherichia coli K-12 Science. 277:1453-1474.

Bloom BR and Murray CJ L. **1992**. Tuberculosis: commentary on a reemergent killer. Science. 257: 1055-1064.

Boon C and Dick T. **2002**. *Mycobacterium bovis* BCG response regulator essential for hypoxic dormancy. J. Bacteriol. 184: 6760–6767.

Boon C, Li R, Qi R and Dick T. **2001**. Proteins of *Mycobacterium bovis* BCG induced in the Wayne dormancy model. J. Bacteriol. 183: 2672–2676.

Bradford MM. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Analyt. Biochem. 72: 248-252.

Braunstein M and Belisle J. **2000**. Genetics of protein secretion, p.203–220. *In* G. F. Hatfull and J. W. R. Jacobs (ed.), Molecular genetics of mycobacteria. American Society for Microbioogy, Washington, D.C.

Brosch R, Gordon SV, Marmiesse M, Brodin P, Buchrieser C, Eiglmeier K, Garnier T, Gutierrez C, Hewinson G, Kremer K, Parsons LM, Pym AS, Samper S, van Soolingen D and Cole ST. **2002**. A new evolutionary scenario for the *Mycobacterium tuberculosis* complex. Proc. Natl. Acad. Sci. USA 99: 3684–3689.

Brosch R, Gordon SV, Pym A, Eiglmeier K, Garnier T and Cole ST. **2000**. Comparative genomics of the mycobacteria. Int. J. Med. Microbiol. 290: 143–152.

Brusasca PN, Peters RL, Motzel SL, Klein HJ and Gennaro ML. **2003**. Antigen recognition by serum antibodies in non-human primates experimentally infected with *Mycobacterium tuberculosis*. Comp. Med. 53: 165-172.

Buchmeier N, Blanc-Potard A, Ehrt S, Piddington D, Riley L and Groisman EA. **2000**. A parallel intraphagosomal survival strategy shared by *Mycobacterium tuberculosis* and *Salmonella enterica.* Mol. Microbiol. 35: 1375–1382.

Butt J, Kim H-Y, Basilion JP, Cohen S, Iwai K, Philpott CC, Altschul S, Klausner RD and Rouault TA. **1996**. Differences in the RNA binding sites of iron regulatory proteins and potential target diversity PNAS 93: 4345 4349.

Camacho LR, Ensergueix D, Perez E, Gicquel B and Guilhot C. **1999**. Identification of a virulence gene cluster of *Mycobacterium tuberculosis* by signature-tagged transposon mutagenesis. Mol. Microbiol. 34: 257–267.

Chakhaiyar P, Nagalakshmi Y, Aruna B, Murthy KJR, Katoch VM and Hasnain SE. **2004**. Regions of high antigenicity within the hypothetical PPE_MPTR Rv2608 ORF show a differential humoral response but similar T cell response in various categories of TB patients. J. Infect. Dis. (In Press).

Chakhaiyar P and Hasnain SE. **2004**. Defining the Mandate of Tuberculosis Research in a Postgenomic Era. Med Princ Pract. 13: 177-184

Chan J, Ran X, Hunter SW, Brennan PJ and Bloom BR. **1991**. Lipoarabinomannan, a possible virulence factor involved in persistence of *Mycobacterium tuberculosis* within macrophages. Infect. Immun. 59: 1755–1761.

Chatterjee D, Roberts AD, Lowell K et al.**1992**. Structural basis of capacity of liparrabinomannon to induce secretion of tumour necrosis factor. Infect Immun. 60: 1249-1253

Chen P, Ruiz RE, Li Q, Silver RF and Bishai WR. **2000**. Construction and characterization of a *Mycobacterium tuberculosis* mutant lacking the alternate sigma factor gene, sigma F. Infect. Immun. 68: 5575–5580.

Chen R and Yang H. **2000**. A highly specific monomeric isocitrate dehydrogenase from *Corynebacterium glutamicum*. Arch. Biochem. Biophys. 383: 238-245.

Choudhary RK, Pullakhandam R, Ehtesham NZ, Hasnain SE. **2004**. Expression and characterization of Rv2430c, a novel immunodominant antigen of *Mycobacterium tuberculosis*. Protein Expr Purif. 36: 249-53.

Choudhary RK, Mukhopadhyay S, Chakhaiyar P, Sharma N, Murthy KJR, Katoch VM and Hasnain SE **2003**. PPE antigen Rv2430c of *Mycobacterium tuberculosis* induces a strong B-Cell response. Infection Immunity. 71: 6338-6343.

Chowgule RV and Deodhar L. **1998**. Pattern of secondary acquired drug resistance to antituberculosis drug in Mumbai, India 1991-1995. Indian J. Chest Dis. Allied Sci. 40: 23-31

Clemens DL and Horwitz MA. **1995**. Characterization of the *Mycobacterium tuberculosis* phagosome and evidence that phagosomal maturation is inhibited. J. Exp. Med. 181: 257-270.

Cohn DL, Bustreo F and Raviglione MC. **1997**. Drug-resistant tuberculosis: Review of the worldwide situation and the WHO/IULATD global surveillance project. Clin. Infect. Dis., 24 suppl. 1: S121-30

Cole ST and Smith DR. **1994**. Toward mapping and sequencing the genome of *Mycobacterium tuberculosis*, p. 227–238. In B. R. Bloom (ed.), Tuberculosis: pathogenesis, protection, and control. American Society for Microbiology, Washington, D.C.

Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, Gordon SV, Eigenmeir K, Gas S, Barry III CE, et al.**1998**.

Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. Nature 393: 537–544.

Cole ST, Eiglmeier K, Parkhill J, James KD, Thomson NR, Wheeler PR, Honore N, Garnier T, Churcher C, Harris D, et al. **2001**. Massive gene decay in the leprosy bacillus. Nature 409: 1007–1011.

Cooper AM, Dalton DK, Steward TA et al. **1993.** Dissimalated tuberculosis in interferon-gamma gene- disrupted mice. J exp med. 178: 2243-2247

Coronado VG, Beck-Sague CM, Hutton MD, Davis BJ, Nicholas T, Willartal C, Woodley CL, Kilburn JO, Crawford JT, Frieden TR, Sinkowitz RL, and Jarvis WR. **1993**. Transmission of multidrug-resistant *Mycobacterium tuberculosis* among persons with human immunodeficiency virus infection in an urban hospital: epidemiologic and restriction fragment length polymorphism analysis. J. Infect. Dis. 168:1052-1055

Cox JH, Knight BC and Ivanyi J.**1989**. Mechanism of recrudescence of *Mycobacterium bovis* BCG infection in mice. Infect immun. 57: 1719-1724

Cox JS, Chen B, MacNeil M and Jacobs WR. **1999**. Complex lipid determines tissue-specific replication of *Mycobacterium tuberculosis* in mice. Nature 402: 79–83.

Crowle AJ, Dahl R, Ross E et al. **1991.** Evidence that vesicles containing living virulent *Mycobacterium tuberculosis* or *Mycobacterium avium* in cultured human macrophages are not acidic. Infect immun. 59: 1823-1831

Das Gupta SK, Bashyam MD and Tyagi AK. **1993**. Cloning and assessment of mycobacterial promoters by using a plasmid shuttle vector. J. Bacteriol. 175: 5186–5192.

Das S, Paramasivan CN, Lowrie DB, Prabhakar R and Narayanan PR. **1995**. IS6110 restriction fragment length polymorphism typing of clinical isolates of *Mycobacterium tuberculosis* from patients with pulmonary tuberculosis in Madras, south India. Tuber. Lung Disease. 76: 550-4

De Voss JJ, Rutter K, Schroeder BG, Su H, Zhu Y and Barry III CE. **2000**. The salicylate-derived mycobactin siderophores of *Mycobacterium tuberculosis* are essential for growth in macrophages. Proc. Natl. Acad. Sci. USA 97: 1252–1257.

Dellagostin OA, Esposito G, Eales LJ, Dale JW and McFadden J. **1995**. Activity of mycobacterial promoters during intracellular and extracellular growth. Microbiology 141: 1785–1792.

DeMaio J, Zhang Y, Ko C and Bishai WR. **1997**. *Mycobacterium tuberculosis sigF* is part of a gene cluster with similarities to the

Bacillus subtilis sigF and *sigB* operons. Tubercule Lung Dis. 78: 3–12.

DeMaio J, Zhang Y, Ko C, Young DB and Bishai WR. **1996**. A stationary-phase stress-response sigma factor from *Mycobacterium tuberculosis*. Proc. Natl. Acad. Sci. USA 93: 2790–2794.

Dhariwal KR and Venkitasubramanian TA. **1987**. NADP-specific isocitrate dehydrogenase of *Mycobacterium phlei* ATCC 354: purification and characterization. J. Gen. Microbiol. 133: 2457-2460.

Dubey VS, Sirakova TD and Kolattukudy PE. **2002**. Disruption of msl3 abolishesthe synthesis of mycolipanoic and mycolipenic acids required for polyacyltrehalose synthesis in *Mycobacterium tuberculosis* H37Rv and causes cell aggregation. Mol. Microbiol. 45: 1451–1459.

Dubnau E, Chan J, Raynaud C, Mohan VP, Laneelle MA, Yu K, Quemard A, Smith I and Daffe M. **2000**. Oxygenated mycolic acids are necessary for virulence of *Mycobacterium tuberculosis* in mice. Mol. Microbiol. 36: 630–637.

Dubnau E, Fontan P, Manganelli R, Soares-Appel S and Smith I. **2002**. *Mycobacterium tuberculosis* genes induced during infection of human macrophages. Infect. Immun. 70: 2787–2795.

Dubnau E, Laneelle M-A, Soares S, Benichou A, Vaz T, Prome D, Daffe M and Quemard A. **1997**. *Mycobacterium bovis* BCG genes involved in the biosynthesis of cyclopropyl keto- and hydroxy-mycolic acids. Mol. Microbiol. 23: 313–322.

Dussurget O, Rodriguez GM and Smith I. **1996**. An *ideR* mutant of *Mycobacterium smegmatis* has a derepressed siderophore production and an altered oxidative-stress response. Mol. Microbiol. 22: 535–544.

Dussurget O, Stewart G, Neyrolles O, Pescher P, Young D and Marchal G. **2001**. Role of *Mycobacterium tuberculosis* copper-zinc superoxide dismutase. Infect. Immun. 69: 529–533.

Dye C, Espinal, MA, Watt CJ, Mbiaga C and Williams BG. **2002**. Worldwide incidence of multidrug-resistant tuberculosis. J. Infect. Dis. 185: 1197-1202.

Dye C, Scheele S, Dolin P, Pathania V and Raviglione MC. **1999**. Global burden of tuberculosis: estimated incidence, prevalence, and mortality by country. JAMA. 282: 677-686.

Edlin BR, Tokars JI, Grieco MH, Crawford JT, Williams J, Sordillo EM, Ong KR, Kilburn JO, Dooley SW, Castro KG, Jarvis WR and Holmberg SD. **1992**. An outbreak of multidrug-resistant tuberculosis

among hospitalized patients with the acquired immunodeficiency syndrome. New Eng. J. Med. 326: 1514-1521

Edwards KM, Cynamon MH, Voladri RK, Hager CC, DeStefano MS, Tham KT, Lakey DL, Bochan MR and Kernodle DS. **2001**. Iron-cofactored superoxide dismutase inhibits host responses to *Mycobacterium tuberculosis*. Am. J. Respir. Crit. Care Med. 164: 2213–2219.

Ewann F, Jackson M, Pethe K, Cooper A, Mielcarek N, Ensergueix D, Gicquel B, Locht C and Supply P. **2002**. Transient requirement of the PrrA-PrrB two-component system for early intracellular multiplication of *Mycobacterium tuberculosis*. Infect. Immun. 70: 2256–2263.

Fillebeen C and Pantopoulos K. **2002**. Redox control of iron regulatory proteins. Redox Rep. 7: 15–22.

Florio W, Bottai D, Batoni G. Esin S, Pardini M, Maisett G and Campa M. **2002**. Identification, molecular cloning, and evaluation of potential use of isocitrate dehydrogenase II of *Mycobacterium bovis* BCG in serodiagnosis of tuberculosis.Clin. Diagn. Lab. Immunol. 9: 846-851.

Flynn JA, Chan J, Triebold Kj et al. **1993.** An essential role for IFN-gamma in resistance to *M. tuberculosis* infection. J Exp. Med. 178: 2249-2253

Frishman D and Hentze MW. **1996**. Conservation of aconitase residues revealed by multiple sequence analysis. Implications for structure/function relationships. Eur. J. Biochem. 239: 197-200

Fuangthong M, Herbig AF, Bsat N and Helmann JD. **2002**. Regulation of the *Bacillus subtilis fur* and *perR* genes by PerR: not all members of the PerR regulon are peroxide inducible. J. Bacteriol. 184: 3276–3286.

Gegout V, Schlegl J, Schlager B, Hentze MW, Reinbolt J, Ehresmann B, Ehresmann C and Romby P. **1999**. Ligand-induced structural alterations in human iron regulatory protein-1 revealed by protein footprinting. J Biol Chem. 274: 15052-8.

Ghosh S, Rasheedi S, Rahim SS, Banerjee S, Choudhary RK, Chakhaiyar P, Ehtesham NZ, Mukhopadhyay S, Hasnain SE. **2004**. A novel method for enhancing solubility of the expressed proteins in *E. coli*. Biotechniques. [In Press]

Glickman MS, Cox JS and Jacobs Jr. WR. **2000**. A novel mycolic acid cyclopropane synthetase is required for cording, persistence, and virulence of *Mycobacterium tuberculosis*. Mol. Cell 5: 717–727.

Gold B, Rodriguez GM, Marras SA, Pentecost M and Smith I. **2001**. The *Mycobacterium tuberculosis* IdeR is a dual functional regulator that controls transcription of genes involved in iron acquisition, iron storage and survival in macrophages. Mol. Microbiol. 42: 851–865.

Goldstein AL and McCusker JH. **2001.** Development of sacchromyces cerevisiae as a model pathogen. A system for the genetic identification of gene productsrequired for survival in the mammalian host environment. Genetics 159: 499-513.

Gomez JE and Bishai WR. **2000**. *whmD* is an essential mycobacterial gene required for proper septation and cell division. Proc. Natl. Acad. Sci. USA 97: 8554–8559.

Gomez M and Smith I. **2000**. Determinants of mycobacterial gene expression, p. 111–129. *In* G. F. Hatfull and W. R. Jacobs, Jr (ed.), Molecular genetics of mycobacteria. American Society for Microbiology, Washington. D.C.

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

Graham JE and Clark-Curtiss JE. **1999**. Identification of *Mycobacterium tuberculosis* RNAs synthesized in response to phagocytosis by human macrophages by selective capture of transcribed sequences (SCOTS). Proc. Natl Acad. Sci. USA 96: 11554–11559.

Gray NK, Pantopoulos K, Dandekar T, Ackrell BAC and Hentze MW. **1996**. Translational regulation of mammalian and Drosophila citric acid cycle enzymes via iron-responsive elements Proc Natl Acad Sci U S A. 93: 4925-4930.

Groisman EA. 2001. The pleiotropic two-component regulatory system PhoP-PhoQ. J. Bacteriol. 183:1835–1842.

Hadad DJ, Palaci M, Pignatari AC, Lewi DS, Machado MA, Telles MA, Martins MC, Ueki SY, Vasconcelos GM and Palhares MC. **2004**. Mycobacteraemia among HIV-1-infected patients in Sao Paulo, Brazil: 1995 to 1998. Epidemiol. Infect. 132: 151-155.

Haile DJ, Rouault TA, Tang CK, Chin J, Harford JB and Klausner RD. **1992**. Reciprocal Control of RNA-Binding and Aconitase Activity in the Regulation of the Iron-Responsive Element Binding Protein: Role of the Iron-Sulfur Cluster Proc. Natl. Acad. Sci. U. S. A. 89: 7536–7540.

Hasnain SE. 2003. Molecular epidemiology of infectious diseases: a case for increased surveillance. *Bulletin of the W.H.O.* 81: 474.

Helmuth L. 2000. A Weak Link in TB Bacterium Is Found. Science. 289: 1123-1125.

Hentze MW, Caughman SW, Casey JL, Koeller DM, Rouault TA, Harford JB and Klausner RD. **1988**. A model for the structure and functions of iron-responsive elements. *Gene*. 72:201-8.

Heym B, and Cole ST. **1996**. Tuberculosis. ed. T.M. Shinnick. Springer, Germany.

Heym B, Stavropoulos E, Honore N, Domenech P, Saint-Joanis B, Wilson TM, Collins DM, Colston MJ and Cole ST. **1997**. Effects of overexpression of the alkyl hydroperoxide reductase AhpC on the virulence and isoniazid resistance of *Mycobacterium tuberculosis*. Infect. Immun. 65:1395–1401.

Holms H. 1996. Flux analysis and control of the central metabolic pathways in *Escherichia coli*. FEMS. Microbiol. Rev. 19: 85-116.

Holms H. 2001. Flux analysis: a basic tool of microbial physiology. Adv Microb Physiol.45: 271-340.

Holms WH. 1987. Control of flux through the citric acid cycle and the glyoxylate bypass in *Escherichia coli*. Biochem. Soc. Symp. 54: 17-31.

Hondalus MK, Bardarov S, Russell R, Chan J, Jacobs Jr. WR, and Bloom BR. **2000**. Attenuation of and protection induced by a leucine auxotroph of *Mycobacterium tuberculosis*. Infect. Immun. 68: 2888–2898.

Horwitz MA, Harth G, Dillon BJ and Maslesa-Galic S. **2000**. Recombinant bacillus Calmette-Guerin (BCG) vaccines expressing the *Mycobacterium tuberculosis* 30-kDa major secretory protein induce greater protective immunity against tuberculosis than conventional BCG vaccines in a highly susceptible animal model. Proc. Natl. Acad. Sci. USA 97: 13853–13858.

Horwitz MA, Lee BW, Dillon BJ and Harth G. **1995**. Protective immunity against tuberculosis induced by vaccination with major extracellular proteins of *Mycobacterium tuberculosis*. Proc. Natl. Acad. Sci. USA 92: 1530-1534.

Hunter SW, Gaylord H and Brennan PJ. **1986**. Structure and antigenicity of the phosphorylated lipopolysaccharide antigens from the leprosy and tubercle bacilli. J. Biol. Chem. 261: 12345–12351.

Hussein S, Curtis J, Akuffo H and Turk JL.**1987**. Dissociation between delayed-type hypersensitivity and resistance to pathogenic mycobacteria demonstrated by T cell clones. Infect Immun. 55; 564-569 Immunity. 71: 6338-6343.

Iseman MD and Sabarbaro JA. **1992**. The increasing prevalence of resistance to antituberculosis chemotherapeutic agents: Implications for global tuberculosis control. Curr. Trop. Infect. Dis. 12: 188-204.

Jackson M, Phalen SW, Lagranderie M, Ensergueix D, Chavarot P, Marchal G, McMurray DN, Gicquel B and Guilhot C. **1999**. Persistence and protective efficacy of a *Mycobacterium tuberculosis* auxotroph vaccine. Infect. Immun. 67: 2867–2873.

Jennings GT, Minard KI and McAlister-Henn L. 1997. Expression and mutagenesis of mammalian cytosolic NADP+-specific isocitrate dehydrogenase. Biochemistry. 36: 13743-13747.

Jensen-Cain DM and Quinn FD. **2001**. Differential expression of *sigE* by *Mycobacterium tuberculosis* during intracellular growth. Microb. Pathog. 30: 271–278.

Jordan PA, Tang Y, Bradbury AJ, Thomson AJ and Guest JR.**1999**. Biochemical and spectroscopic characterization of Escherichia coli aconitases (AcnA and AcnB). Biochem J. 344: 739-46.

Kanao T, Kawamura M, Fukui T, Atomi H and Imanaka T. **2002**. Characterization of isocitrate dehydrogenase from the green sulfur bacterium *Chlorobium limicola*, A carbon dioxide-fixing enzyme in the reductive tricarboxylic acid cycle. Eur. J. Biochem. 269: 1926-1931.

Kang DK, Jeong J, Drake SK, Wehr NB, Rouault TA and Levine RL. **2003**. Iron regulatory protein 2 as iron sensor. Iron dependent oxidative modification of cysteine. *J Biol Chem.*, 278: 14857-14864.

Kannan KB, Katoch VM, Bharadwaj VP, Sharma VD, Datta AK and Shivannavar CT. **1985**. Metabolic studies on mycobacteria--II. Glyoxylate by-pass (TCA cycle) enzymes of slow and fast growing mycobacteria Indian J Lepr. 57: 542-8.

Katoch VM, Sharma VD, Kannan KB, Datta AK, Shivannavar CT, Bharadwaj VP.**1987**. Metabolic studies on mycobacteria. III. Demonstration of key enzymes of TCA cycle in M. leprae. Indian J Lepr. 59: 152-7.

Kaufmann SHE. **1989**. In-vitro analysis of the cellular mechanism involved in immunity to tuberculosis. Rev. infect. dis. 11:448-454.

Koeller DM, Casey JL, Hentze MW, Gerhardt EM, Chan LN, Klausner RD and Harford JB. **1989**. A cytosolic protein binds to structural elements within the iron regulatory region of the transferrin receptor mRNA. Proc Natl Acad Sci U S A. 86: 3574-3578.

Koch R. 1882. Die Aetiologie der Tuberculose. Berl. Klin. Wchnschr., xix: 221-230. In Milestones in Microbiology: 1556 to 1940, translated and edited by Thomas D. Brock, ASM Press. 1998.

Kornberg HL. 1966. The role and control of the glyoxylate cycle in Escherichia coli. Biochem. J. 99: 1-11.

Laidlaw M. 1989. In Practical Medical Microbiology, eds. Colle JG, Duguid JP, Fraser Ag and Marimon BP. (New York, Churchill Livingstone), pp. 399-416.

Lakshmi TM and Helling RB. **1978**. Acetate metabolism in Escherichia coli. Can. J. Microbiol. 24: 149-153.

Lamichhane G, Zignol M, Blades NJ, Geiman DE, Dougherty A, Grosset J, Broman KW and Bishai WR. **2003**. A postgenomic method for predicting essential genes at subsaturation levels of mutagenesis: application to *Mycobacterium tuberculosis*. Proc Natl Acad Sci U S A. 100: 7213-8.

Lauble H, Kennedy MC, Beinert H and Stout CD. **1994**. Crystal structures of aconitase with trans-aconitate and nitrocitrate bound. J Mol Biol. 237: 437-51.

Lee B-Y and Horwitz MA. **1995**. Identification of macrophage and stress-induced proteins of *Mycobacterium tuberculosis.* J. Clin. Investig. 96: 245–249.

Lewis KN, Liao R, Guinn KM, Hickey MJ, Smith S, Behr MA and Sherman DR. **2003**. Deletion of RD1 from *Mycobacterium tuberculosis* mimics bacille Calmette-Guerin attenuation. J. Infect. Dis. 187: 117–123.

Lill R and Kispal G. **2000**. Maturation of cellular Fe/S proteins: The essential function of mitochondria. Trends Biochem. Sci. 25: 352-356.

Litwin CM and Calderwood SB. **1993**. Role of iron in regulation of virulence genes. Clin. Microbiol. Rev. 6: 137–149.

Lorencz MG and Wackernagel W. **1994**. Bacterial gene transfer by natural genetic transformation in the environment. Microbial Reviews 58: 563-602.

Lorenz MC and Fink GR. **2002.** Life and Death in a Macrophage: Role of glyoxylate cycle in virulence. Eukaryotic Cell. 1: 657-662.

Louise R, Skjot V, Agger EM and Andersen P. **2001**. Antigen discovery and tuberculosis vaccine development in the post-genomic era. Scand. J.Infect. Dis. 33: 643–647.

Lounis N, Truffot-Pernot C, Grosset J, Gordeuk VR and Boelaert JR. **2001**. Iron and *Mycobacterium tuberculosis* infection. J. Clin. Virol. 20: 123–126.

Loyevsky M, LaVaute T, Allerson CR, Stearman R, Kassim OO, Cooperman S, Gordeuk VR and Rouault TA. **2001**. An IRP-like protein from Plasmodium falciparum binds to a mammalian iron-responsive element. Blood. 98: 2555-2562.

Maekura R, Kohno H, Hirotani A, Okuda Y, Ito M, Ogura T and Yano I. **2003**. Prospective clinical evaluation of the serologic tuberculous glycolipid test in combination with the nucleic acid amplification test J. Clin. Microbiol. 41: 1322-1325.

Maekura R, Okuda Y, Nakagawa M, Hiraga T, Yokota S, Ito M, Yano I, Kohno H, Wada M, Abe C, *et al.* **2001**. Clinical evaluation of anti tuberculous glycolipid immunoglobulin G antibody assay for rapid serodiagnosis of pulmonary tuberculosis. J. Clin. Microbiol. 39: 3603 3608.

Mahairas GG, Sabo PJ, Hickey MJ, Singh DC and Stover CK.**1996**. Molecular analysis of genetic differences between *Mycobacterium bovis* BCG and virulent *M. bovis*. J. Bacteriol. 178:1274–1282.

Majeed AA, Ahmed N, Rao KR, Ghousunnissa S, Kausar F, Bose B, Nagarajaram HA, Katoch VM, Cousins DV, Sechi LA, Gilman RH and Hasnain SE. **2004**. AmpliBASE MTTM: A *Mycobacterium tuberculosis* diversity knowledgebase. Bioinformatics. 20: 989-992.

Manabe YC, Saviola BJ, Sun L, Murphy JR and Bishai WR. **1999**. Attenuation of virulence in *Mycobacterium tuberculosis* expressing a constitutively active iron repressor. Proc. Natl. Acad. Sci. USA 96:12844–12848.

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

Manganelli R, Voskuil MI, Schoolnik GK, Dubnau E, Gomez M and Smith I. **2002**. Role of the extracytoplasmic-function sigma Factor sigma H in *Mycobacterium tuberculosis* global gene expression. Mol. Microbiol. 45: 365–374.

Marcinkeviciene JA, Magliozzo RS and Blanchard JS. **1995**. Purification and characterization of the *Mycobacterium smegmatis* catalase-peroxidase involved in isoniazid activation. J. Biol. Chem. 38: 22290–22295.

Mariani F, Cappelli G, Riccardi G and Colizzi V. **2000**. *Mycobacterium tuberculosis* H37Rv comparative gene-expression analysis in synthetic medium and human macrophage. Gene 253: 281–291.

McAdam RA, Weisbrod TR, Martin J, Scuderi JD, Brown AM, Cirillo JD, Bloom BR and Jacobs Jr. WR. **1995**. In vivo growth characteristics of leucine and methionine auxotrophic mutants of *Mycobacterium bovis* BCG generated by transposon mutagenesis. Infect. Immun. 63: 1004– 1012.

McDonough KA, Kress Y and Bloom BR. **1993**. Pathogenesis of tuberculosis: interaction of *Mycobacterium tuberculosis* with macrophages. Infect. Immun. 61: 2763-2773.

McKinney JD, Bentrup KHZ, Munoz-Elias EJ, Miczak A, Chen B, Chan WT, Swenson D, Sacchettini JC, Jacobs Jr. WR and Russell DG. **2000**. Persistence of *Mycobacterium tuberculosis* in macrophages and mice requires the glyoxalate shunt enzyme isocitrate lyase. Science 406: 735–738.

Mcmurray DN and Bartow. **1992.** Immunosuppression and alteration of resistance to pulmonary tuberculosis in guinea pigs by proteins undernutrition. J Nutr., 738-743

Mirel DB, Marder K, Graziano J, Freyer G, Zhao Q, Mayeux R and Wilhelmsen KC. **1998**. Characterization of the human mitochondrial aconitase gene (ACO2) Gene. 213: 205-218

Molle V, Palframan WJ, Findlay KC and Buttner MJ. **2000**. WhiD and WhiB, homologous proteins required for different stages of sporulation in *Streptomyces coelicolor* A3(2). J. Bacteriol. 182: 1286–1295.

Moncrief MB and Maguire ME. **1998**. Magnesium and the role of MgtC in growth of *Salmonella typhimurium.* Infect. Immun. 66: 3802–3809.

Mori T, Sakatani M, Yamagishi F, Takashima T, Kawabe Y, Nagao K, Shigeto E, Harada N, Mitarai S, Okada M, *et al.* **2004**. Specific detection of tuberculosis infection: an interferon-gamma-based assay using new antigens. Am. J. Respir. Crit. Care Med. **0**, 200402179-0 (In Press).

Mühlenhoff U and Lill R. **2000**. Biogenesis of iron-sulfur proteins in eukaryotes: A novel task of mitochondria that is inherited from bacteria. Biochem. Biophys. Acta 1459: 370-382.

Mühlenhoff U, Richhardt N, Ristow M, Kispal G and Lill R. **2002**. The yeast frataxin homologue Yfh1p plays a specific role in the maturation of cellular Fe/S proteins. Hum. Mol. Genet. 11: 2025-2036.

Muro-Pastor MI and Florencio FJ. **1992**. Purification and properties of NADP-isocitrate dehydrogenase from the unicellular cyanobacterium *Synechocystis sp.* Eur. J. Biochem. 203: 99-105.

Mustafa AS. 2002. Development of new vaccines and diagnostic reagents against tuberculosis. Mol. Immunol. 39: 113–119.

Nimmo HG, Borthwick AC, el-Mansi EM, Holms WH, MacKintosh C and Nimmo GA. **1987**. Regulation of the enzymes at the branchpoint between the citric acid cycle and the glyoxylate bypass in *Escherichia coli.* Biochem. Soc. Symp. 54: 93-101.

Noss EH, Pai RK, Sellati TJ, Radolf JD, Belisle J, Golenbock DT, Boom WH and Harding CV. **2001**. Toll-like receptor 2-dependent inhibition of macrophage class II MHC expression and antigen processing by 19-kDa lipoprotein of *Mycobacterium tuberculosis.* J. Immunol. 167: 910–918.

Nunez MT, Nunez-Millacura C, Tapia V, Munoz P, Mazariegos D, Arredondo M, Munoz P, Mura C and Maccioni RB. **2003**. Iron-activated iron uptake: a positive feedback loop mediated by iron regulatory protein 1.Biometals. 16: 83-90.

Nystrom T and Gustavsson N.**1998**. Maintenance energy requirement: what is required for stasis survival of *Escherichia coli*? Biochim. Biophys. Acta. 1365: 225- 231.

Nystrom T, Larsson C and Gustafsson L. **1996**. Bacterial defense against aging: role of the *Escherichia coli* ArcA regulator in gene expression, readjusted energy flux and survival during stasis EMBO. J. 15: 3219-3228.

Ohman R and Ridell M. **1996**. Purification and characterisation of isocitrate dehydrogenase and malate dehydrogenase from *Mycobacterium tuberculosis* and evaluation of their potential as suitable antigens for the serodiagnosis of tuberculosis. Tuber. Lung Dis. 77: 454-461.

Ohman R and Ridell M. **1995**. Enzymatic and antigenic analyses of strains of *Mycobacterium bovis*, *M. bovis* BCG, and *M. tuberculosis*. Curr. Microbiol. 30: 161-165.

Orme IM. 1988. Characteristics and specificity of accquired immunology memory to *Mycobacterium tuberculosis* infection. J immunol. 140: 3589-3593.

Pancholi P, Mirza A, Bhardwaj N and Steinman RM.**1993**. Sequesteration from immune CD4+ T cells of mycobacteria growing in human macrophages. Science. 260: 984-987.

Pantopoulos K and Hentze MW. **1995**. Rapid responses to oxidative stress mediated by iron regulatory protein EMBO J.14: 2917-24.

Parish T, Gordhan BG, McAdam RA, Duncan K, Mizrahi V and Stoker NG. **1999**. Production of mutants in amino acid biosynthesis genes of *Mycobacterium tuberculosis* by homologous recombination. Microbiology145: 3497–3503.

Perez E, Samper S, Bordas Y, Guilhot C, Gicquel B and Martin C. **2001**. An essential role of *phoP* in *Mycobacterium tuberculosis* virulence. Mol. Microbiol. 41:179–187.

Perkins MD, Conde MB, Martins M and Kritski AL. **2003**. Serologic diagnosis of tuberculosis using a simple commercial multiantigen assay. Chest 123: 107-112.

Perschinka H, Mayr M, Millonig G, Mayerl C, van der Zee R, Morrison SG, Morrison RP, Xu Q and Wick G. **2003**. Cross-reactive B-cell epitopes of microbial and human heat shock protein 60/65 in atherosclerosis. Arterioscler Thromb Vasc Biol. 23: 1060-1065.

Pethe K, Alonso S, Biet F, Delogu G, Brennan MJ, Locht C and Menozzi FD. **2001**. The heparin-binding haemagglutinin of *M. tuberculosis* is required for extrapulmonary dissemination. Nature 412: 190–194.

Philpott CC, Klausner RD and Rouault TA. **1994**. The bifunctional iron-responsive element binding protein/cytosolic aconitase: the role of active-site residues in ligand binding and regulation. Proc Natl Acad Sci U S A. 91: 7321-7325.

Piddington DL, Fang FC, Laessig T, Cooper AM, Orme IM and Buchmeier NA. **2001.** Cu,Zn superoxide dismutase of *Mycobacterium tuberculosis* contributes to survival in activated macrophages that are generating an oxidative burst. Infect. Immun. 69: 4980–4987.

Pitchenik AE, Burr J, Laufer M, Miller G, Cacciatore R, Bigler WG, Witte JJ, and Cleary T. **1990**. Outbreaks of drug-resistant tuberculosis at AIDS centre. Lancet. 336: 440-441.

Plank DW and Howard JB. **1988**. Identification of the reactive sulfhydryl and sequences of cysteinyl-tryptic peptides from beef heart aconitase. J Biol Chem. 263: 8184-8189.

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

Prabhakaran K. 1986. Biochemical studies on *Mycobacterium leprae*. J Basic Microbiol.26: 117-26.

Predich M, Doukhan L, Nair G and Smith I. **1995**. Characterization of RNA polymerase and two factor genes from *Mycobacterium smegmatis*. Mol. Microbiol. 15: 355–366.

Price NC and Stevens L, Fundamentals of Enzymology, third edition, oxford university press, Nicholas C. Price and Lewis Stevens, 3.3, the determination of aminoacid composition, page 56, 4.3, page 135.

Primm TP, Andersen SJ, Mizrahi V, Avarbock D, Rubin H and Barry III CE. **2000**. The stringent response of *Mycobacterium tuberculosis* is required for long-term survival. J. Bacteriol. 182:4889–4898.

Pym AS, Brodin P, Brosch R, Huerre M and Cole ST. **2002**. Loss of RD1 contributed to the attenuation of the live tuberculosis vaccines

Mycobacterium bovis BCG and *Mycobacterium microti*. Mol. Microbiol. 46: 709–717.

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

Quadri LEN, Sello J, Keating TA, Weinreb PH, and. Walsh CT.**1998**. Identification of a *Mycobacterium tuberculosis* gene cluster encoding the biosynthetic enzymes for assembly of the virulence-conferring siderophore mycobactin. Chem. Biol. 5: 631–645.

Ramalingam B, UmaDevi KR and Raja A. **2003**. Isotype-specific anti-38 and 27 kDa (mpt 51) response in pulmonary tuberculosis with human immunodeficiency virus coinfection. Scand. J. Infect. Dis. 35: 234-239.

Raman LA, Siddiqi N, Shamim M, Deb M, Mehta G and Hasnain SE. **2000**. Molecular characterization of *Mycobacterium abscessus* strains isolated from a hospital outbreak. Emerging Infectious Diseases 6: 561-562.

Ratledge C and Ewing M. **1996**. The occurrence of carboxymycobactin, the siderophore of pathogenic mycobacteria, as a second extracellular siderophore in *Mycobacterium smegmatis*. Microbiology 142: 2207–2212.

Raynaud C, Guilhot C, Rauzier J, Bordat Y, Pelicic V, Manganelli R, Smith I, Gicquel B and Jackson M. **2002**. Phospholipases C are involved in the virulence of *Mycobacterium tuberculosis*. Mol. Microbiol. 45: 203–217.

Reeves HC, Daumy GO, Lin CC and Houston M. **1972**. $NADP^+$-specific isocitrate dehydrogenase of *Escherichia coli*.I. Purification and characterization. Biochim. Biophys. Acta. 258: 27-39.

Renshaw PS, Panagiotidou P, Whelan A, Gordon SV, Hewinson RG, Williamson RA and Carr MD. **2002**. Conclusive evidence that the major T-cell antigens of the *Mycobacterium tuberculosis* complex ESAT-6 and CFP-10 form a tight, 1:1 complex and characterization of the structural properties of ESAT-6, CFP-10, and the ESAT-6_CFP-10 complex. Implications for pathogenesis and virulence. J. Biol. Chem. 277: 21598–21603.

Retief JD. 2000. Phylogenetic analysis using PHYLIP. Methods. Mol. Biol. 132: 243-258.

Rindi L, Fattorini L, Bonanni D, Iona E, Freer G, Tan D, Deho G, Orefici G and Garzelli C. **2002**. Involvement of the *fadD33* gene in the growth of *Mycobacterium tuberculosis* in the liver of BALB/c mice. Microbiology 148: 3873–3880.

Roche PW, Triccas JA, Avery DT, Fifis T, Billman-Jacobe H and Britton WJ. **1994**. Differential T cell responses to mycobacteria-secreted proteins distinguish vaccination with bacille Calmette-Guerin from infection with *Mycobacterium tuberculosis*. J. Infect. Dis. 170: 1326-1330.

Rodriguez GM, Gold B, Gomez M, Dussurget O and Smith I. **1999**.Identification and characterization of two divergently transcribed iron regulated genes in *Mycobacterium tuberculosis.* Tubercle Lung Dis. 79: 287–298.

Rodriguez GM, Voskuil MI, Gold B, Schoolnik GK and Smith I. **2002**. *ideR*, An essential gene in *Mycobacterium tuberculosis:* role of IdeR in iron-dependent gene expression, iron metabolism, and oxidative stress response. Infect. Immun. 70: 3371–3381.

Rouault T and Klausner R. **1997**. Regulation of iron metabolism in eukaryotes.Curr. Top. Cell Regul. 35: 1–19.

Rouault TA and Klausner RD. **1996**. Iron-sulfur clusters as biosensors of oxidants and iron Trends Biochem. Sci. 21: 174–177

Saiman L. 2004. The mycobacteriology of non-tuberculous mycobacteria. Paediatr. Respir. Rev. 5: 221-3.

Sambandamurthy VK, Wang X, Chen B, Russell RG, Derrick S, Collins FM, Morris SL and Jacobs Jr. WR. **2002**. A pantothenate auxotroph of *Mycobacterium tuberculosis* is highly attenuated and protects mice against tuberculosis. Nat. Med. 8:1171–1174.

Sassetti CM, Boyd DH and Rubin EJ. **2003**. Genes required for mycobacterial growth defined by high density mutagenesis. Mol Microbiol. 48: 77-84.

Saunders BM and Cooper AM. **2000**. Restraining mycobacteria: role of granulomas in mycobacterial infections. Immunol. Cell. Biol. 78: 334- 341.

Saunders BM, Frank AA. and Orme IM. **1999**. Granuloma formation is required to contain bacillus growth and delay mortality in mice chronically infected with *Mycobacterium tuberculosis*. Immunology. 98: 324-328.

Schlesinger LS and Horwitz MA. **1991.** Phagocytosis of *Mycobacterium leprae* by human monocyte-derived macrophages is mediated by compliment receptors CR1 (CD35) CR3 (CD11b/CD18) and CR4 (CD11c/Cd18) and interferon-gamma activation inhibits compliment receptor function and pkagocytosis of this bact erium.J immunol. 147: 1983-1994.

Schlesinger LS, Bellinger-kawahara CG, Payne NR et al.**1990**. Phagocytosis of *Mycobacterium tuberculosis* mediated by human

monocyte complement receptor and compliment component. J. immunol., 144: 2771-2780

Schlesinger LS. 1993. Macrophage phagocytosis of virulent but not attenuated strains of *Mycobacterium tuberculosis* is mediated by mannose receptors in addition to complement receptors. J. Immunol. 150: 2920–2930.

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

Schnarrenberger C and Martin W. **2002**. Evolution of the enzymes of the citric acid cycle and the glyoxylate cycle of higher plants. A case study of endosymbiotic gene transfer. Eur. J. Biochem. 269: 868-883.

Schneider BD and Leibold EA. **2003**. Effects of iron regulatory protein regulation on iron homeostasis during hypoxia. Blood. 102: 3404-3411

Schoolnik GK. 2002. Microarray analysis of bacterial pathogenicity. Adv. Microb. Physiol. 46: 1–45.

Segal W and Bloch H. **1956**. Biochemical differentiation of *Mycobacterium tuberculosis* grown in vivo and in vitro. J. Bacteriol 72: 132–141.

Segal W and Bloch H. **1957**. Pathogenic and immunogenic differentiation of *Mycobacterium tuberculosis* grown *in vivo* and *in vitro*. Am. Rev. Tuberc. Pulm. Dis. 75: 495–500.

Sharma VD, Katoch VM, Datta AK, Kannan KB, Shivannavar CT and Bharadwaj VP. **1985**. Metabolic studies on mycobacteria-I. Demonstration of key enzymes of glycolysis and tricarboxylic acid cycle on polyacrylamide gels. Indian J Lepr. 57: 534-41.

Sherman DR, Sabo PJ, Hickey MJ, Arain TM, Mahairas GG, Yuan Y, Barry CE and Stover CK. **1995**. Disparate responses to oxidative stress in saprophytic and pathogenic mycobacteria. Proc. Natl. Acad. Sci. USA 92: 6625–6629.

Sherman DR, Voskuil M, Schnappinger D, Liao R, Harrell MI and Schoolnik GK. **2001**. Regulation of the *Mycobacterium tuberculosis* hypoxic response gene encoding alpha-crystallin. Proc. Natl. Acad. Sci. USA 98: 7534–7539.

Shi L, Jung YJ, Tyagi S, Gennaro ML and North RJ. **2003**. Expression of Th1-mediated immunity in mouse lungs induces a *Mycobacterium tuberculosis* transcription pattern characteristic of nonreplicating persistence. Proc. Natl. Acad. Sci. USA 100: 241–246.

Shiratsuch H and Basson MD. **2003**. Caspase activation may be associated with *Mycobacterium avium* pathogenicity. Am. J. Surg. 186: 547-551.

Siddiqi N, Das R, Pathak N, Banerjee S, Ahmed N, Katoch VM and Hasnain SE. **2004**. *Mycobacterium tuberculosis* isolate with a distinct genomic identity overexpresses a TAP-like efflux pump. Infection. 32: 109-111.

Siddiqi N, Shamim M, Hussain S, Choudhary, RK, Ahmed N, Prachee, Banerjee S, Savithri GR, Alam M, Pathak N, *et al.* **2002**. Molecular characterization of multidrug-resistant isolates of *Mycobacterium tuberculosis* from patients in North India Antimicrob. Agents Chemother. 46: 443-450.

Siddiqi N, Shamim MD, Amin AA, Chauhan DS, Ram Das, Srivastava K, Singh D, Sharma VD, Katoch VM, Sharma SK, Hanief M and Hasnain SE. **2001**. Typing of drug resistant isolates of *Mycobacterium tuberculosis* from India using the IS6110 element reveals substantive polymorphism. Infection Genetics and Evolution 1: 109-116.

Siddiqi N, Shamim M, Jain NK, Rattan A, Amin A, Katoch VM, Sharma SK, Hasnain SE. **1998**. Molecular genetic analysis of multi-drug resistance in Indian isolates of *Mycobacterium tuberculosis*. Mem Inst Oswaldo Cruz. 93: 589-594.

Singh SK, Miller SP, Dean A, Banaszak LJ and LaPorte DC. **2002**. *Bacillus subtilis* Isocitrate Dehydrogenase. a substrate analogue for *Escherichia coli* isocitrate dehydrogenase kinase/phosphatase. J. Biol. Chem. 277: 7567- 7573.

Skjot RLV, Oettinger T, Roswnkrands I, Ravn P, Brock I, Jacobsen S and Andersen P. **2000**. Comparative evaluation of low-molecular-mass proteins from *Mycobacterium tuberculosis* identifies members of the ESAT-6 family as immunodominant T-cell antigens. Infect. Immun. 68: 214–220.

Smith DA, Parish T, Stoker NG and Bancroft GJ. **2001**. Characterization of auxotrophic mutants of *Mycobacterium tuberculosis* and their potential as vaccine candidates. Infect. Immun. 69: 1142–1150.

Smith DW and Weigheshaus EH. **1989.** What animal models can teach us about the pathogenesis of tuberculosis. Rev infect dis. 11: 385-393

Smith I. 2003. *Mycobacterium tuberculosis* pathogenesis and molecular determinants of virulence, Clin Microbiol Rev. 16: 463-96

Soliveri JA, Gomez J, Bishai WR and Chater KF. **2000**. Multiple paralogous genes related to the *Streptomyces coelicolor* developmental regulatory gene whiB are present in *Streptomyces* and other actinomycetes. Microbiology 146: 333–343.

Somerville G, Mikoryak CA, Reitzer L. **1999**. Physiological characterization of Pseudomonas aeruginosa during exotoxin A synthesis: glutamate, iron limitation, and aconitase activity. J Bacteriol. 181: 1072-1078.

Sonnenberg MG and Belisle JT. **1997**. Definition of *Mycobacterium tuberculosis* culture filtrate proteins by two-dimensional polyacrylamide gel electrophoresis, N-terminal amino acid sequencing, and electrospray mass spectrometry. Infect. Immun. 65: 4515–4524.

Soundar S, Danek BL and Colman RF. **2000**. Identification by mutagenesis of arginines in the substrate binding site of the porcine NADP-dependent isocitrate dehydrogenase. J. Biol. Chem. 275: 5606-5612.

Steen IH, Madern D, Karlstrom M, LienT, Ladenstein R, and Birkeland NK. **2001**. Comparison of isocitrate dehydrogenase from three hyperthermophiles reveals differences in thermostability, cofactor specificity, oligomeric state, and phylogenetic affiliation. J. Biol. Chem. 276: 43924-43931.

Stewart GR, Snewin VA, Walzl G, Hussell T, Tormay P, O'Gaora P, Goyal M, Betts J, Brown IN and Young DB. **2001**. Overexpression of heatshock proteins reduces survival of *Mycobacterium tuberculosis* in the chronic phase of infection. Nature Med. 7: 732–737.

Stewart GR, Wernisch L, Stabler R, Mangan JA, Hinds J, Laing KG, Young DB and Butcher PD. **2002**. Dissection of the heat-shock response in *Mycobacterium tuberculosis* using mutants and microarrays. Microbiology 148: 3129–3138.

Stokes RW, Haidl ID, Jefferies WA et al. **1993.** Mycobacteria-macrophage interactions. J immunol. 151: 7067-7076.

Stryer L. 1995. Biochemistry, 4th Ed, W. H. Freeman and Co., USA.

Sturgill-Koszycki S, Schlesinger PH, Chakraborty P, Haddix PL, Collins HL, Fok AK, Allen RD, Gluck SL, Heuser J and Russell DG. **1994**. Lack of acidification in Mycobacterium phagosomes produced by exclusion of the vesicular proton-ATPase. Science 263: 678-681.

Tang Y and Guest JR. **1999**. Direct evidence for mRNA binding and post-transcriptional regulation by Escherichia coli aconitases. Microbiology. 145: 3069-79.

Tang Y, Quail MA, Artymiuk PJ, Guest JR and Green J. **2002**. Escherichia coli aconitases and oxidative stress: post-transcriptional regulation of sodA expression. Microbiology. 148: 1027-37.

Thoma-Uszynski S, Stenger S, Takeuchi O, Ochoa MT, Engele M, Sieling PA, Barnes PF, Rollinghoff M, Bolcskei PL, Wagner M, et al. **2001**. Induction of direct antimicrobial activity through mammalian toll-like receptors. Science 291:1544–1547.

Trajkovic V, Natarajan K and Sharma P. **2004**. Immunomodulatory action of mycobacterial secretory proteins. Microbes Infect. 6: 513-519.

Tullius MV, Harth G and Horwitz MA. **2001**. High extracellular levels of *Mycobacterium tuberculosis* glutamine synthetase and superoxide dismutase in actively growing cultures are due to high expression and extracellular stability rather than to a protein-specific export mechanism. Infect.Immun. 69: 6348–6363.

van Crevel R, Ottenhoff TH and van der Meer JW. **2002**. Innate immunity to *Mycobacterium tuberculosis.* Clin. Microbiol. Rev. 15: 294–309.

Varghese S, Tang Y and Imlay JA. **2003**. Contrasting sensitivities of Escherichia coli aconitases A and B to oxidation and iron depletion. J Bacteriol.185: 221-30.

Vasquez B and Reeves HC. **1981**. NADP-specific isocitrate dehydrogenase from *Escherichia coli.* V. Multiple forms of the enzyme. Biochim. Biophys. Acta. 660: 16-22.

Walden WE. 2002. From bacteria to mitochondria: Aconitase yields surprises. Proc Natl Acad Sci U S A 99: 4138-4140

Wayne LG and Liu KY. **1982**. Glyoxalate metabolism and adaptation of *Mycobacterium tuberculosis* to survival under anaerobic conditions. Infect.Immun. 37: 1042–1049.

Wayne LG. 1994. Dormancy of *Mycobacterium tuberculosis* and latency of disease. Eur. J. Clin. Microbiol. Infect. Dis. 13: 908–914.

Weber I, Fritz C, Ruttkowski S, Kreft A and Bange FC. **2000**. Anaerobic nitrate reductase (*narGHJI*) activity of *Mycobacterium bovis* BCG *in vitro* and its contribution to virulence in immunodeficient mice. Mol. Microbiol. 35: 1017–1025.

Wei J, Dahl JL, Moulder JW, Roberts EA, O'Gaora P, Young DB and Friedman RL. **2000**. Identification of a *Mycobacterium tuberculosis* gene that enhances mycobacterial survival in macrophages. J. Bacteriol. 182: 377–384.

Wernisch L, Kendall SL, Soneji S, Wietzorrek A, Parish T, Hinds J, Butcher PD and Stoker NG. **2003**. Analysis of whole-genome microarray replicates using mixed models. Bioinformatics 19: 53–61.

Wilson T, de Lisle GW, Marcinkeviciene JA, Blanchard JS and Collins DM. **1998.** Antisense RNA to *ahpC*, an oxidative stress defence gene involved in isoniazid resistance, indicates that AhpC of *Mycobacterium bovis* has virulence properties. Microbiology 144: 2687–2695.

Wilson TJ, Bertrand N, Tang JL, Feng JX, Pan MQ, Barber CE, Dow JM, Daniels MJ. **1998**. The rpfA gene of Xanthomonas campestris pathovar campestris, which is involved in the regulation of pathogenicity factor production, encodes an aconitase. Mol Microbiol. 28: 961-70.

Wong DK, Lee BY, Horwitz MA and Gibson BW. **1999**. Identification of fur, aconitase, and other proteins expressed by *Mycobacterium tuberculosis* under conditions of low and high concentrations of iron by combined two-dimensional gel electrophoresis and mass spectrometry. Infect.Immun. 67: 327–336.

Xu J and Gordon JI. **2003**. Honor thy symbionts. Proc Natl Acad Sci U S A. 100: 10452-10459.

Xu S, Cooper A, Sturgill-Koszycki S, van Heyningen T, Chatterjee D, Orme I, Allen P and Russell DG.**1994**. Intracellular trafficing in *Mycobacterium tuberculosis* and *Mycobacterium avium*-infected macrophages. J immunol. 2568-2578

Yancey PH and Somero GN. **1979**. Counteraction of urea destabilization of protein structure by methylamine osmoregulatory compounds of elasmobranch fishes. Biochem. J. 183, 317-323

Young D, Lathigra R, Hendrix R, Sweetser D and Young RA. **1988**. Stress proteins are immune targets in leprosy and tuberculosis. Proc. Natl. Acad. Sci. USA 85:4267–4270.

Yuan Y and Barry III CE. **1996**. A common mechanism for the biosynthesis of methoxy and cyclopropyl mycolic acids in *Mycobacterium tuberculosis.* Proc. Natl. Acad. Sci. USA 93: 12828–12833.

Yuan Y, Crane DD, and Barry III CE. **1996**. Stationary phase-associated protein expression in *Mycobacterium tuberculosis:* function of the mycobacterial alpha-crystallin homolog. J. Bacteriol. 178: 4484–4492.

Yuan Y, Crane DD, Simpson RM, Zhu Y, Hickey MJ, Sherman DR and Barry III CE. **1998**. The 16-kDa α-crystallin (Acr) protein of *Mycobacterium tuberculosis* is required for growth in macrophages. Proc. Natl. Acad. Sci. USA 95: 9578–9583.

Zahrt TC and Deretic V. **2001**. *Mycobacterium tuberculosis* signal transduction system required for persistent infections. Proc. Natl. Acad. Sci. USA 98: 12706–12711.

Zahrt TC, Song J, Siple J and Deretic V. **2001**. Mycobacterial FurA is a negative regulator of catalase-peroxidase gene *katG*. Mol. Microbiol. 39: 1174–1185.

Zhang D, Dimopoulos G, Wolf A, Minana B, Kafatos FC and Winzerling JJ. **2002**. Cloning and molecular characterization of two mosquito iron regulatory proteins. Insect Biochem Mol Biol. 32: 579-89.

SHARMISTHA BANERJEE
Laboratory of Molecular and Cellular Biology
Centre for DNA Fingerprinting and Diagnostics,
Hyderbad 500 076
India

Date of Birth : January 7, 1973

Educational Qualifications

- **Ph.D** (Submitted to University of Hyderabad), 2004
  Centre for DNA Fingerprinting and Diagnostics.
  Title of thesis "Characterization of Mycobacterium tuberculosis TCA cycle enzymes isocitrate dehydrogenase and aconitase".
- **Master of Science (MSc)**, 1998
  MSc in Zoology (with specialization in Ecology) from the Jamshedpur Co-operative College, Ranchi University, Jharkhand, India. Division - 1st
- **Bachelor of Science (BSc)**, 1995
  BSc in Zoology (Honours) with Botany Chemistry, English and Hindi from The Graduate School, College for Women, Ranchi University Jharkhand, India. Division - 1st

Publications

- **Banerjee S**, Nandyala A, Podili R, Katoch VM, Murthy KJR and Hasnain SE. *Mycobacterium tuberculosis* (*Mtb*) isocitrate dehydrogenases show strong B cell response and distinguish vaccinated controls from TB patients ***Proc. Nat. Acad. Sci.***, USA.2004; **101**: 12652-12657
- Siddiqi N, Das R, Pathak N, **Banerjee S**, Ahmed N, Katoch VM and Hasnain SE. *Mycobacterium tuberculosis* isolate with a distinct genomic identity overexpresses a TAP like efflux pump. ***Infection***. 2004; **32**: 93-95
- Ghosh S, Rasheedi S, Rahim SS, **Banerjee, S**, Choudhary RK, Chakhaiyar P, Ehtesham NZ, Mukhopadhyay S, Hasnain SE. A novel method for enhancing solubility of the expressed proteins in *E. coli*. ***Biotechniques***. 2004 (In Press)
- Siddiqi N, Shamim M, Hussain S, Choudhary RK, Ahmed N, Prachee, **Banerjee S**, Savithri GR, Alam M, Pathak N, Amin A, Hanief M, Katoch VM, Sharma SK and Hasnain SE. Molecular Characterization of Multi-Drug Resistant Isolates of *Mycobacterium tuberculosis* From Patients in North India. ***Antimicrob. Agents Chemother.*** 2002; **46**:443-450.

Other Publications

- Niyaz Ahmed, Abhijit Bal, Aleem A. Khan, Mahfooz Alam, Anju Kagal, Vidya Arjunwadkar, Amarsingh Rajput, Ahmed A. Majeed, Syed Asad Rahman, **Sharmistha Banerjee**, Suvarna Joshi, Renu Bharadwaj. Whole genome fingperprinting and genotyping of multiple drug resistant (MDR) isolates of *Pseudomonas aeruginosa* from endophthalmitis patients in India. ***Infection, Genetics and Evolution.*** 2002; **1**:1-6.

Mycobacterium tuberculosis (*Mtb*) isocitrate dehydrogenases show strong B cell response and distinguish vaccinated controls from TB patients

Sharmistha Banerjee*, Ashok Nandyala*, Raviprasad Podili*, V. M. Katoch†, K. J. R. Murthy‡, and Seyed E. Hasnain*§¶

*Laboratory of Molecular and Cellular Biology, Centre for DNA Fingerprinting and Diagnostics, Nacharam, Hyderabad 500076, India; †Laboratory of Microbiology, Central JALMA Institute for Leprosy, Agra 282001, India; ‡Mahavir Hospital and Research Centre, Hyderabad 500004 India; and §Jawaharlal Nehru Centre for Advanced Scientific Research, Jakkur, Bangalore 560012, India

Communicated by G. Balakrish Nair, International Centre for Diarrhoeal Disease Research, Dhaka, Bangladesh, June 24, 2004 (received for review January 20, 2004)

Proteins released from *Mycobacterium tuberculosis* (*Mtb*) during late logarithmic growth phase are often considered candidate components of immunogenic or autolysis markers. One such protein is isocitrate dehydrogenase (ICD), a key regulatory enzyme in the citric acid cycle. We have evaluated the immunogenic properties of two isoforms of *Mtb* ICD and compared them with the control antigens heat-shock protein 60 and purified protein derivative (PPD). PPD lacks the sensitivity to distinguish between bacillus Calmette–Guérin (BCG)-vaccinated and tuberculosis (TB)-infected populations, and, therefore, epidemiological relevance of PPD in BCG-vaccinated regions is debatable. We show that *Mtb* ICDs elicit a strong B cell response in TB-infected populations and can differentiate between healthy BCG-vaccinated populations and those with TB. The study population (n = 215) was categorized into different groups, namely, patients with fresh infection (n = 42), relapsed TB cases (n = 32), patients with extrapulmonary TB (n = 35), clinically healthy donors (n = 44), nontuberculous mycobacteria patients (n = 30), and non-TB patients (culture negative for acid-fast bacteria but carrying other infections, n = 32). The *Mtb* ICDs showed statistically significant antigenic distinction between healthy BCG-vaccinated controls and TB patients ($P < 0.0001$) and those with other infections. Although extrapulmonary infections could not be discriminated from healthy controls by heat-shock protein 60 ($P = 0.2177$), interestingly, the *Mtb* ICDs could significantly ($P < 0.0001$) do so. Our results highlight the immunodominant, immunosensitive, and immunospecific nature of *Mtb* ICDs and point to an unusual property of this tricarboxylic acid energy cycle enzyme.

Tuberculosis (TB), caused by *Mycobacterium tuberculosis* (*Mtb*), remains a major threat to the human population, being responsible for ≈2–3 millions deaths every year worldwide (1–3). The secret of the pathogen's success is its ability to escape the host immune system and remain undetected in lungs for decades. In only 10% of infected people, the number being higher in immunocompromised patients, does TB erupt as a full-blown disease (4). Delay in diagnosis and treatment impedes the downstream management and control of the disease. With the increasing emergence of multidrug resistant strains and coinfection with HIV, the problem is further compounded (5–7). Early diagnosis, therefore, is a matter of utmost concern not just for TB disease management but also for epidemiological investigations (8). Current diagnostic tools for TB often lack sensitivity and can be time consuming. TB diagnosis in developing countries largely banks on tuberculin skin tests and staining and culture methods. The epidemiological relevance of the tuberculin test with purified protein derivative (PPD) is questionable in areas where bacillus Calmette–Guérin vaccination is compulsory because PPD is not sensitive enough to distinguish between vaccinated and infected individuals (9). Microscopic determination of the bacilli in the sputum samples is a direct way of examining pulmonary TB (5). This method, however, requires high titers of bacilli (5,000–10,000 per ml) in sputum, a condition seen only in full-blown TB patients. Culture techniques can detect very low titers but are time-consuming, taking ≈3–6 weeks (10).

The importance of the major extracellular proteins of the pathogen as candidate components of a subunit vaccine has been reported earlier (11). Discovery of the RD1 locus in the *Mtb* genome, encoding mainly the proteins actively secreted by mycobacteria into the culture medium, such as CFP-10 and ESAT-6, has further encouraged immunological tests as an adjunct to conventional diagnosis (12–15). Proteins that are released from *Mtb* during late logarithmic growth phase, such as superoxide dismutase and isocitrate dehydrogenase (ICD), are used as autolysis markers (16). The use of ICD as a potential antigen for serodiagnosis along with malate dehydrogenase has been suggested (17, 18). The *Mtb* genome carries two isoforms of ICD, *Mtb* ICD-1 and *Mtb* ICD-2. Multiple sequence alignment revealed a closer similarity of *Mtb* ICD-1 to eukaryotic $NADP^+$-dependent ICDs, whereas *Mtb* ICD-2 groups with bacterial ICDs (unpublished data). We have evaluated the utility of ICDs as immunogenic markers for TB through the detection of anti-*Mtb* ICD antibodies in sera of different well characterized categories of TB patients through enzyme-linked immunosorbent assays. We describe the sensitivity and specificity of ICDs to distinguish TB patients from those vaccinated with bacillus Calmette–Guérin and from those patients infected with nontuberculous mycobacteria (NTM) or other pathogens by means of the conventional antigens, heat-shock protein 60 (HSP 60) (19) and PPD.

Materials and Methods

Cloning, Expression, and Purification of *Mtb* ICD-1 and *Mtb* ICD-2. The ORFs corresponding to *Mtb* ICD-1 (Rv3339c, 1.230 kb) and *Mtb* ICD-2 (Rv0066c, 2.238 kb) were PCR-amplified from the genomic DNA of H37Rv. *Bam*HI and *Hin*dIII restriction sites were incorporated in the 5′ end of forward and reverse primers, respectively, for both *Mtb* ICD-1 and *Mtb* ICD-2. The primers and parameters for thermal cycle amplification have been tabulated in Table 1. The amplicons carrying the full-length *Mtb* ICD-1 and *Mtb* ICD-2 were cloned at the *Bam*HI and *Hin*dIII sites of the expression vector pRSET-A (Invitrogen) with six N-terminal histidine sequence tags. The generated constructs setAicd1 and setAicd2 were further transformed into the BL21 (DE3) strain of *Escherichia coli*. The clones were confirmed by

Freely available online through the PNAS open access option.

Abbreviations: *Mtb*, *Mycobacterium tuberculosis*; TB, tuberculosis; ICD, isocitrate dehydrogenase; PPD, purified protein derivative; HSP 60, heat-shock protein 60; NTM, nontuberculous mycobacteria.

¶To whom correspondence should be addressed at: Centre for DNA Fingerprinting and Diagnostics, Nacharam, Hyderabad 500076, India. E-mail: ehtesham@cdfd.org.in.

www.pnas.org/cgi/doi/10.1073/pnas.0404347101

Table 1. PCR primers and thermal cycle parameters for amplification of *Mtb* ICD-1 and ICD-2

Primer	Sequence	PCR parameters	Amplicon size
Mtb icd-1			
forward	ggatccATGTCCAACGCACCCAAGATA	94°C for 2 h	≈1.2 kb
reverse	aagcttCTAATTGGCCAGCTCCTTTTC	35 cycles:	
		94°C for 30 min,	
		50°C for 1 h,	
		72°C for 3 h,	
		72°C for 7 h	
Mtb icd-2			
forward	AGCTTggatccATGAGCGCCGAACAGCC	94°C for 2 h	≈2.23 kb
reverse	CATGGaagcttTCAGCCTTGGACAGCCT	10 cycles:	
		94°C for 30 min,	
		50°C for 30 min,	
		72°C for 3.5 h	
		25 cycles:	
		94°C for 30 min,	
		72°C for 3.5 h,	
		72°C for 7 h	

sequencing with the T7 promoter primer on an Applied Biosystems Prism 377 DNA sequencer.

The genes were overexpressed in the pRSET-A/*E. coli* BL-21 (DE3) expression system. The overexpressed His-tagged recombinant protein was purified by Ni^{2+}-nitrilotriacetate affinity chromatography. The cells transformed with the constructs were grown in Terrific Broth containing ampicillin (100 μg/ml) to an OD_{600} of 0.4–0.5 at 37°C, cooled to 27°C, induced with 0.1 mM isopropyl β-D-thiogalactoside, and grown overnight at 27°C. The cells were lysed by sonication, followed by centrifugation at 16,000 × *g* for 30 min at 4°C. The clear lysate was loaded onto a Ni^{2+}-nitrilotriacetic acid column, which was then washed with 50 mM NaH_2PO_4/300 mM NaCl/20 mM imidazole, pH 8. The protein was eluted in the same buffer supplemented with 200 mM imidazole. The proteins were 90–95% pure as seen on 10% SDS/PAGE followed by Coomassie blue staining (Fig. 1). The purified recombinant proteins were dialyzed against 20 mM Tris·HCl, pH 7.5, with 100 mM NaCl and 3% glycerol and quantified by using Bradford reagent (20).

Human Sera. The study population (n = 215) comprised the *Mtb*-infected human sample population reporting to the Mahavir Hospital and Research Centre and the Central JALMA Institute for Leprosy. These populations were categorized into three groups, namely Group 1 (n = 42 patients), Group 2 (n = 32 patients), and Group 3 (n = 35 patients). In addition to the above, 44 clinically healthy donors, 30 NTM cases, and 32 non-TB patients who were proven culture-negative for acid-fast bacteria were also included as controls in this study. Group 1 comprised patients with fresh infection with no history of TB treatment. Group 2 comprised patients with relapsed cases, i.e., those who were treated earlier for TB but whose symptoms reemerged after the completion of the treatment. Group 3 included patients with extrapulmonary TB. Patients from Groups 1 and 2 were diagnosed by sputum examination (acid-fast bacillus smear positive and negative), whereas the extrapulmonary cases were confirmed by tissue biopsy. Clinically healthy donors were *Mycobacterium bovis* bacillus Calmette–Guérin-vaccinated and had no symptoms of TB at the time of sera collection. Randomly picked individuals from the population of healthy controls were subjected to a PCR test for TB and were found to be PCR-negative. Mycobacteria other than *Mycobacterium leprae* that are not included in the *Mtb* complex are referred to as NTMs (21). However, the group referred to as NTMs in this study included sera collected from patients infected with NTM species (n = 14), such as *Mycobacterium avium*, *Mycobacterium xenopi*, and *Mycobacterium fortuitum*, as well as sera from patients with *M. leprae* infection (n = 16). The non-TB patient category included infected individuals who were tested negative for acid-fast bacteria by staining and culture-based techniques. These patients were also negative for HIV and hepatitis B virus. These randomly picked patients were suffering from pneumonia, lower respiratory infections, septicemia, urinary tract infections, gastrointestinal infections, cirrhosis, or fever of unknown origin. The study population had no sex or age bias. This study was approved by the Institutional Ethics Committee.

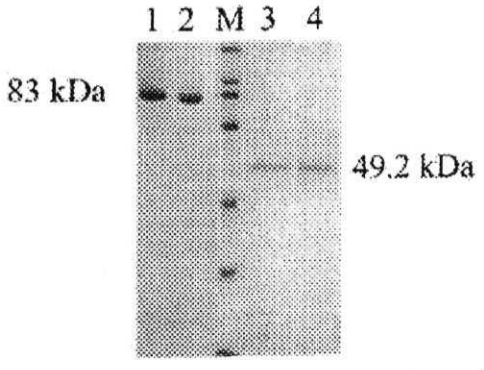


Fig. 1. Affinity purification of *Mtb* ICD-1 and *Mtb* ICD-2. His-tagged recombinant protein was purified by nickel column chromatography under native conditions and stained with Coomassie blue after 10% SDS/PAGE. Lanes: 1 and 2, *Mtb* ICD-2; M, protein molecular mass markers (200, 116, 97, 66, 45, 31, and 21.5 kDa); 3 and 4, *Mtb* ICD-1.

Immunosorbent Assays. ELISAs were performed to check the B cell immune response in humans to the *Mtb* ICD-1 and ICD-2 proteins and control antigens HSP 60 and PPD. The HSP 60 used was *Mtb* HSP65/GroEL. In brief, the 96-well microtiter plates (Costar) were coated with ≈500 ng of either control antigens or recombinant *Mtb* ICD-1 and *Mtb* ICD-2. The plates were incubated overnight at 4°C, washed three times with PBS, and blocked with 100 μl of blocking buffer (2% BSA in PBS) for 2 h at 37°C. The plates were then washed three times with wash buffer PBST (0.05% Tween 20 in 1× PBS). The *Mtb*-infected human sera belonging to different clinical groups were diluted 200 times in blocking buffer (1% BSA in PBS). Serum (50 μl) was added to antigen-coated wells and then incubated for 1 h at

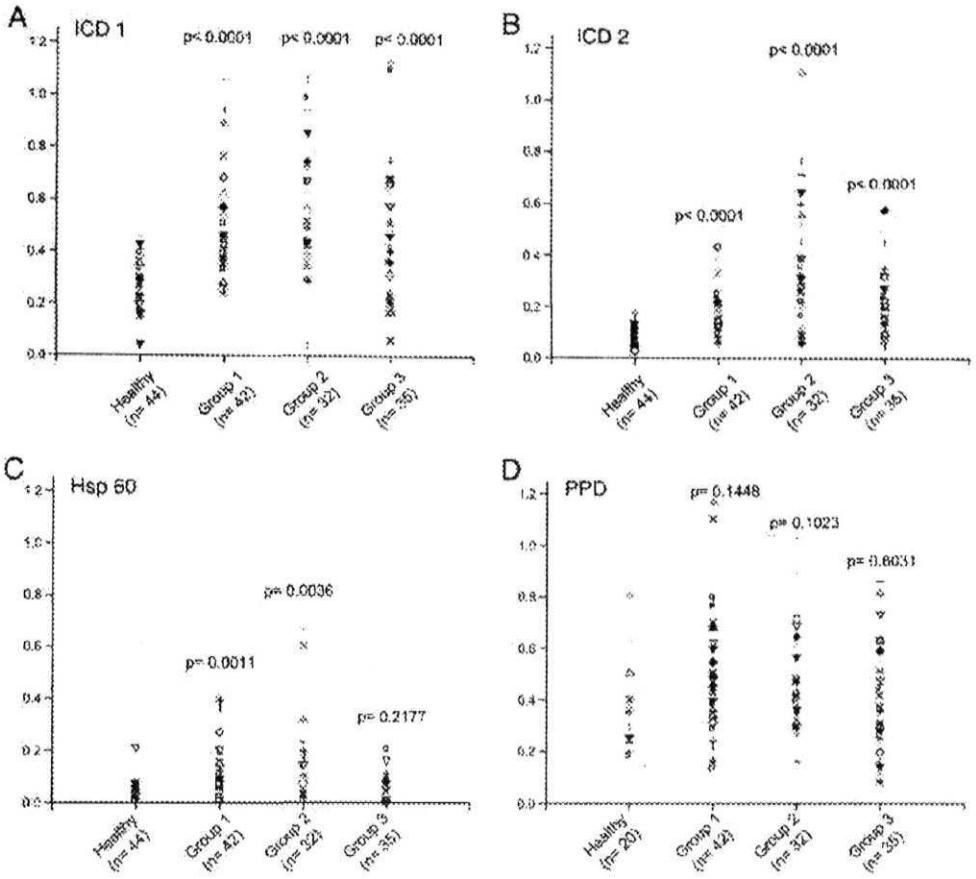


Fig. 2. *Mtb* ICD-1 and ICD-2 show high B cell reactivity to sera from TB-infected patients from different groups as opposed to bacillus Calmette–Guérin-vaccinated healthy controls. The humoral immune responses directed against the recombinant proteins *Mtb* ICD-1 (*A*) and *Mtb* ICD-2 (*B*) and control antigens HSP 60 (*C*) and PPD (*D*) were compared among different categories of patients and healthy controls. Group 1, fresh infections; Group 2, relapsed infection; Group 3, extrapulmonary TB. The respective sample numbers and *P* values are shown.

37°C. The plates were thoroughly washed with PBST and further incubated with anti-human IgG-horseradish peroxidase (Sigma) at 37°C for 1 h. Horseradish peroxidase activity was detected by using a chromogenic substance, *o*-phenylenediamine tetrahydrochloride (Sigma), in citrate-phosphate buffer (pH 5.4) and H_2O_2 (Qualigens Fine Chemicals, Ahmedabad, India) as 1 μl/ml. The reactions were terminated by using 0.5 M H_2SO_4, and the absorbance values were measured at 492 nm in an ELISA reader (Bio-Rad). Each ELISA was repeated at least twice with some randomly picked serum samples tested three times for confirmation, with and without replicates for each sample within individual ELISA.

Data Analysis. Student's *t* test was performed to compare the means of two variable groups, healthy and infected classes, by using the online scientific calculator of GraphPad (www.graphpad.com/quickcalcs/ttest1.cfm) to calculate means, SEM, and *P* values.

Results

Expression and Purification of *Mtb* ICD-1 and *Mtb* ICD-2. The overexpressed N-terminal His-tagged *Mtb* ICD-1 was purified to 95% homogeneity on a nickel affinity column (Fig. 1). The molecular mass of the recombinant ICD-1 was determined to be 49.2 kDa. The purification was carried out under native conditions from soluble fractions with an yield of 3.25 mg of protein per 500 ml of start culture. Similarly *Mtb* ICD-2, an 83-kDa protein, was purified to 90–95% homogeneity (Fig. 1) with a yield of ≈20.4 mg per 1,000 ml of start culture.

***Mtb* ICD-1 and *Mtb* ICD-2 Show High Reactivity to Patient Sera As Opposed to Bacillus Calmette–Guérin-Vaccinated Healthy Controls.** Humoral immune responses directed against the *Mtb* ICD-1 and *Mtb* ICD-2 were compared among patients with TB and bacillus Calmette–Guérin-vaccinated healthy controls (Fig. 2 *A* and *B*). The recombinant proteins were used to screen the infected and healthy sera by ELISA using anti-human IgG-horseradish peroxidase as a conjugate. The sera were also tested against *Mtb* HSP 60 and the PPD (Fig. 2 *C* and *D*). The immunoreactivity of ICD-1, ICD-2, HSP 60, and PPD was statistically analyzed and compared with respect to both infected and healthy sera. These data demonstrate that sera of all of the infected patients mounted a statistically significant ($P < 0.0001$) antibody response against recombinant *Mtb* ICD-1 and *Mtb* ICD-2 proteins as compared with that of the healthy controls. PPD, on the other hand, reacted against both healthy and TB-infected sera. It is interesting to note that, compared with ICD-1 and ICD-2 ($P < 0.0001$), the difference in the reactivity of PPD to total infected and healthy sera was negligible and statistically insignificant ($P = 0.2301$). Because PPD, a mixture of proteins, showed statistically insignificant discrimination between healthy populations and different categories of infected populations (Fig. 2*D*), the reactivities of the recombinant proteins were compared with the B

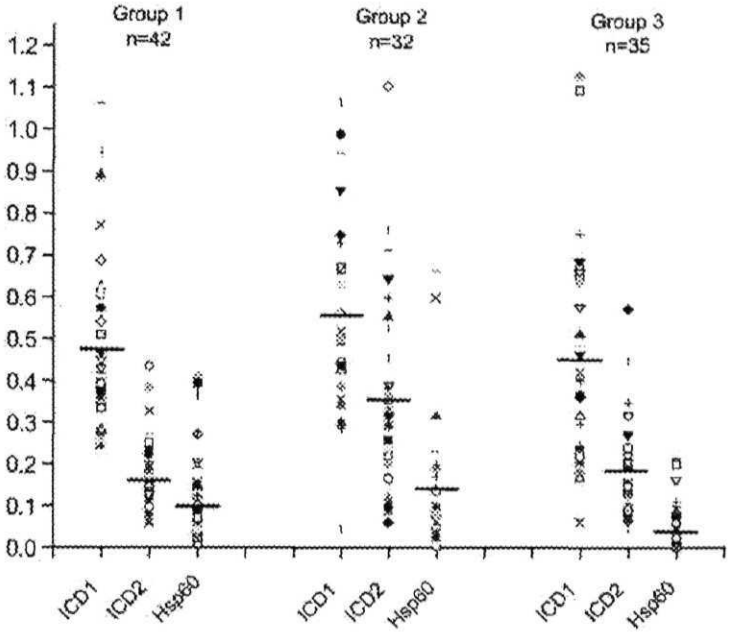


Fig. 3. *Mtb* ICDs are more immunogenic than HSP 60. The ELISA reactivity to *Mtb* ICD-1, *Mtb* ICD-2, and control antigen HSP 60 was compared in different patient groups. Horizontal bands represent the mean reactivity or average levels of humoral response in each category.

cell response to *Mtb* HSP 60 (Fig. 2*C*) in different categories of patients. The difference between reactivity to HSP 60 between TB patients and healthy controls was statistically not quite significant ($P = 0.0645$).

A correlation between reactivity against *Mtb* ICD-1 and *Mtb* ICD-2 in patient sera with the state of disease, fresh or relapse, was attempted by comparing the antibody responses to *Mtb* ICD-1 and *Mtb* ICD-2 of various clinical categories (Fig. 2 *A* and *B*, respectively). *Mtb* ICD-1 failed to discriminate between fresh, relapsed, and extrapulmonary TB cases because no significant differences in immunoreactivity in different patient groups were observed (Fig. 2*A*). Yet compared with bacillus Calmette–Guérin-vaccinated healthy controls each category yielded P values <0.0001, indicating that *Mtb* ICD-1 can differentiate substantially between bacillus Calmette–Guérin-vaccinated healthy populations and any category of *Mtb*-infected patients, pulmonary or nonpulmonary. *Mtb* ICD-2, on the other hand, could also discriminate relapsed cases from both fresh infections ($P < 0.0001$) and extrapulmonary infections ($P = 0.0003$). Like *Mtb* ICD-1, *Mtb* ICD-2 could also distinguish substantially between bacillus Calmette–Guérin-vaccinated healthy population and any category of *Mtb*-infected patients. Surprisingly, HSP 60, although it could discriminate Groups 1 and 2 from healthy controls ($P = 0.0011$ and 0.0036, respectively), failed to distinguish the extrapulmonary infections from bacillus Calmette–Guérin-vaccinated healthy controls ($P = 0.2177$). These results demonstrate that (*i*) recombinant *Mtb* ICD-1 and ICD-2 proteins could differentiate sera from TB-infected patients by means of healthy bacillus Calmette–Guérin-vaccinated controls, (*ii*) the extrapulmonary infections that could not be distinguished from healthy controls by HSP 60 could be significantly identified and categorized by *Mtb* ICDs, and (*iii*) *Mtb* ICD-2 mounted a stronger antibody response in relapsed cases and could significantly discriminate them from Group 1 and Group 3. These proteins, which have an apparently important metabolic role, are thus able to elicit a strong B cell response as a function of the TB infection.

Immunodominace of ICDs over HSP 60. We compared the immunogenicity of ICDs over HSP 60. Humoral response to HSP 60 in all of the three categories of TB patients was tested and compared with those to ICDs (Fig. 3). The data clearly indicate that the mean reactivity (represented by the horizontal bands in Fig. 3) of HSP 60 in all of the classes of patient sera was much lower than in either ICD-1 or ICD-2 (Fig. 3). Thus, ICDs are immunodominant and serologically more sensitive than HSP 60. The mean values for ICD-1 in Groups 1, 2, and 3 were 0.481, 0.565, and 0.457, respectively, whereas those for ICD-2 were 0.165, 0.362, and 0.188, respectively. It is therefore apparent that ICD-1 elicited a stronger response in all of the three categories of patients tested than ICD-2. The data also confirm the discriminatory power of ICD-2 for relapsed case as compared with other categories.

Immunospecificity of *Mtb* ICDs. The potential of *Mtb* ICDs to specifically distinguish between TB, NTMs, and non-TB patient sera (those essentially culture-negative for acid-fast bacteria but harboring other pathogens) was tested by examining the cross-reactivity of the recombinant proteins with NTMs and non-TB patient sera. Thirty NTMs and 32 non-TB patient sera were tested for their immunogenic response against *Mtb* ICD-1, *Mtb* ICD-2, and HSP 60. The data were statistically analyzed to check whether ICDs could significantly distinguish between TB-infected patients and NTMs or non-TB patients. Fig. 4 shows that ICDs could significantly distinguish TB-infected sera from NTMs ($P < 0.0001$) and non-TB patients ($P < 0.0001$ for ICD-1 and $P = 0.0008$ for ICD-2), thus ruling out any crossreactivity with NTMs and other bacterial pathogens tested. HSP 60 appeared to react more broadly to the population under study and could not differentiate between TB infections and NTMs ($P = 0.4461$) or non-TB cases ($P = 0.3464$) significantly. That HSP 60 reacted broadly was apparent by calculating the average reactivity for each group (infected, NTMs, and non-TB), for which reactivity to HSP 60 remained almost the same (Fig. 4). The mean reactivity for HSP 60 is in contrast to the mean

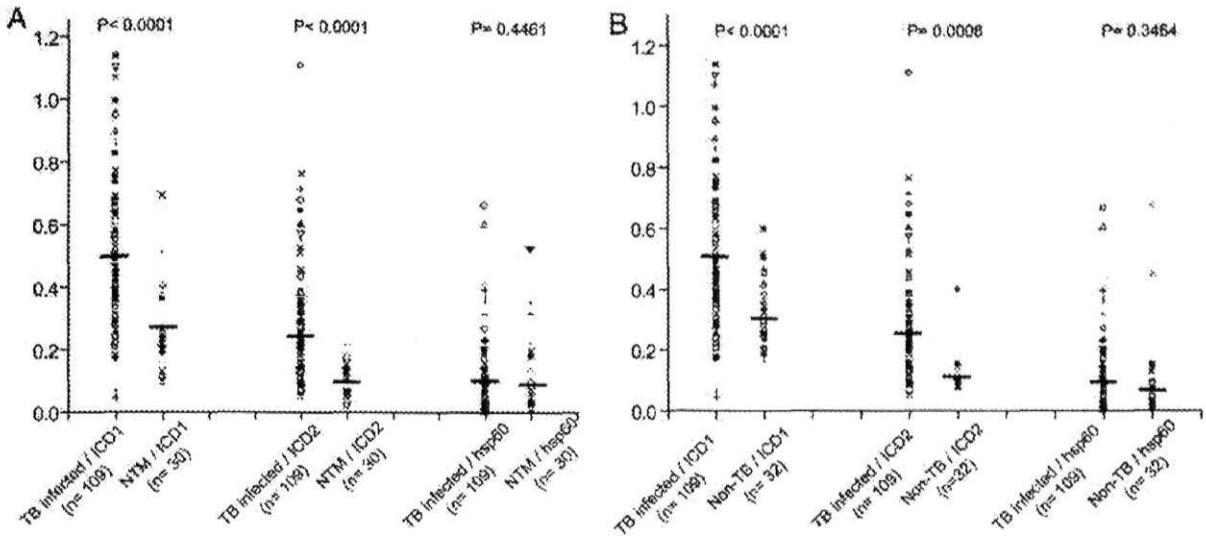


Fig. 4. *Mtb* ICDs could significantly distinguish TB-infected sera from NTMs and non-TB patient sera. Recombinant *Mtb* ICD-1 and *Mtb* ICD-2 and HSP 60 were tested against sera of NTM (*A*) and non-TB (*B*) patients. The respective humoral responses were compared with those for TB-infected sera, the *P* values for which are shown. HSP 60 could not distinguish TB-infected patients from either NTM or non-TB patients significantly. Horizontal bands represent the mean reactivity in each category.

humoral response to ICD-1 and ICD-2, for which a distinctly higher response was seen in TB-infected sera as compared with NTMs or non-TB cases.

Discussion

The main objective of our study was to evaluate *Mtb* ICD-1 and *Mtb* ICD-2 in terms of their immune features as compared with the control antigens, HSP 60 and PPD, that are frequently used for diagnosis of TB. ICDs serve as markers of autolysis (16, 17) and are among the secretory proteins released during the late logarithmic phase. Although earlier efforts have pointed to the antigenic potential of *Mtb* ICDs (17), the present study is systematic investigation of their potential as an immune marker.

The cases of TB were identified and enrolled based on their history of treatment as fresh infections, relapsed cases, and extrapulmonary infections. The categorization of patients largely depended on the treatment history dictated by the patients or their family members. As evident from Fig. 2, there is little doubt about the ability of either *Mtb* ICD-1 or *Mtb* ICD-2 to elicit a strong B cell response, irrespective of patient category. When compared with *Mtb* ICD-2, *Mtb* ICD-1 was more antigenic (Figs. 2 and 3). It would be interesting to explore this disparity. A comparative analysis of ELISA reactivities among different categories of patients for *Mtb* ICDs (Fig. 2 *A* and *B*) revealed higher reactivity in Group 2 as compared with fresh (Group 1) and extrapulmonary (Group 3) infections. More specifically, the antigenic response in this category of patients to ICD 2 was significantly higher than that in Group 1 and Group 3. Because these patients had undergone treatment earlier, a high number of autolysed infected macrophages and autolysed pathogens could possibly explain the comparative high-antibody response against *Mtb* ICDs in this category. Drugs, like isoniazid, are known to affect the cell envelope architecture of mycobacteria and, hence, the increase in the production of the secreted proteins (22).

Comparative immunoreactivity of *Mtb* ICD-1, *Mtb* ICD-2, and HSP 60 clearly indicates that the antigenic distinction between healthy and TB patients is statistically significant for both *Mtb* ICD-1 and *Mtb* ICD-2 ($P < 0.0001$) but not quite so for HSP 60 ($P = 0.0645$). Earlier reports have shown crossreactive epitopes between microbial HSP 60/65 and human HSP 60, which often serve as autoimmune targets in conditions like atherosclerosis (19). The crossreactive epitopes probably could explain the broader reactivity of HSP 60 to healthy and infected sera. Because negligible antibody responses were obtained in the bacillus Calmette–Guérin-vaccinated healthy control group as compared with TB-infected patients and because a statistically significant difference in the immunoreactivity between infected and healthy sera was observed, it can be argued that *Mtb* ICD-1 and *Mtb* ICD-2 can be used for diagnosis of *Mtb* infection even in areas where bacillus Calmette–Guérin vaccination is routinely followed. The poor performance of PPD can be attributed to its nonspecific immune reaction in bacillus Calmette–Guérin-vaccinated healthy controls. Interestingly, the extrapulmonary TB cases (Group 3) in the study population did not show any significant humoral response to HSP 60 to distinguish them from healthy controls. On the other hand, Group 3 patients mounted a very significant B cell response to ICD-1 and ICD-2, separating them from bacillus Calmette–Guérin-vaccinated healthy controls.

Immunodominance is a parameter that we defined to compare the antibody titers against the tested proteins, i.e., ICD-1, ICD-2, and HSP 60, in the patient sera. Our data clearly showed that ICD-1 was most antigenic and mounted a very strong B cell response in all of the patient categories, followed by ICD-2 and HSP 60 (Fig. 3). Having shown that *Mtb* ICDs elicit a B cell response much higher than HSP 60, we checked for immune specificity of these proteins. Crossreaction with sera of NTMs and non-TB patients is one of the critical parameters that needed to be checked before establishing any antigenic marker for possible serological studies in *Mtb*. *Mtb* complex, including *M. bovis* and *Mycobacterium africanum*, is responsible for more illness worldwide than any other bacteria. However, there are >82 recognized species of mycobacteria that occasionally infect mammalian hosts. These species, referred to as NTMs, are omnipresent in the environment, and most species are either nonpathogenic for humans or are rarely associated with disease, except for a few, like *M. avium*, that are opportunistic pathogens, more frequently associated with immunocompromised patients (23, 24). The clinical significance of many NTMs remains unclear; however, it is important to check the crossreactivity of

Mtb antigens with this group of mycobacteria. Our experiments could establish that *Mtb* ICDs do not crossreact with either NTMs or non-TB patient sera (Fig. 4).

The existing diagnostic tests for TB, even to this day, largely depend on tuberculin skin tests and staining and culture techniques. These methods are slow, cumbersome, and lack sensitivity and specificity in bacillus Calmette–Guérin-vaccinated cases. As more and more recombinant antigens are being tested (25–31), serological methods are likely to be favored over others. ELISA *per se* is unlikely to replace the current TB diagnosis; however, in combination or parallel with other rapid PCR-based diagnostic techniques, ELISA can largely improve the accuracy and rapidity of TB diagnosis for an effective disease management. Our data reveal the antigenic potential of recombinant *Mtb* ICD-1 and also present a systematic study on immunogenicity of recombinant *Mtb* ICD-2. *Mtb* ICD-1 and *Mtb* ICD-2 can be further analyzed for their pathogen-specific antigenic epitopes. Given their important role in the energy cycle, an evaluation of these two *Mtb* enzymes as possible drug targets is necessary. That such important enzymes can also have strong antigenic attributes that enable them to significantly discriminate between bacillus Calmette–Guérin-vaccinated healthy controls and TB patients and at the same time TB from other pathogenic infections is a very exciting proposition, possibly pointing to their immunomodulatory function.

We thank Dr. Shekhar Mande (Centre for DNA Fingerprinting and Diagnostics) for providing purified recombinant *Mtb* HSP 65/GroEL. This work was supported by research grants from the Council of Scientific and Industrial Research and the Department of Biotechnology, Government of India (to S.E.H.). S.B. was supported by a Senior Research Fellowship from the Council of Scientific and Industrial Research.

1. Dye, C., Scheele, S., Dolin, P., Pathania, V. & Raviglione, M. C. (1999) *J. Am. Med. Assoc.* **282**, 677–686.
2. Bloom, B. R. & Murray, C. J. L. (1992) *Science* **257**, 1055–1064.
3. Chakhaiyar, P. & Hasnain, S., E. (2004) *Med. Prin. Prac.* **13**, 177–184.
4. Helmuth, L. (2000) *Science* **289**, 1123–1125.
5. Dye, C., Espinal, M. A., Watt, C. J., Mbiaga, C. & Williams, B. G. (2002) *J. Infect. Dis.* **185**, 1197–1202.
6. Siddiqi, N., Shamim, M., Hussain, S., Choudhary, R. K., Ahmed, N., Prachee, Banerjee, S., Savithri, G. R., Alam, M., Pathak, N., *et al.* (2002) *Antimicrob. Agents Chemother.* **46**, 443–450.
7. Ahmed, N., Caviedes, L., Alam, M., Rao, K. R., Sangal, V., Sheen, P., Gilman, R. H. & Hasnain, S. E. (2003) *J. Clin. Microbiol.* **41**, 1712–1716.
8. Ahmed, N., Alam, M., RajenderRao, K., Kauser, F., Ashok Kumar, N., Qazi, N., N., Sangal, V., Sharma, V., D., Das, R., Katoch, V., M., *et al.* (2004) *J. Clin. Microbiol.* **42**, 3240–3247.
9. Roche, P. W., Triccas, J .A., Avery, D. T., Fifis, T., Billman-Jacobe, H. & Britton, W. J. (1994) *J. Infect. Dis.* **170**, 1326–1330.
10. Laidlaw, M. (1989) in *Practical Medical Microbiology*, eds. Colle, J. G., Duguid, J. P., Fraser, A. G. & Marimon, B. P. (Churchill Livingstone, New York), pp. 399–416.
11. Horwitz, M. A., Lee, B. W., Dillon, B. J. & Harth, G. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 1530–1534.
12. Mustafa, A. S. (2002) *Mol. Immunol.* **39**, 113–119.
13. Louise, R., Skjot, V., Agger, E. M. & Andersen, P. (2001) *Scand. J. Infect. Dis.* **33**, 643–647.
14. Trajkovic, V., Natarajan, K. & Sharma, P. (2004) *Microbes Infect.* **6**, 513–519.
15. Mori, T., Sakatani, M., Yamagishi, F., Takashima, T., Kawabe, Y., Nagao, K., Shigeto, E., Harada, N., Mitarai, S., Okada, M., *et al.* (2004) *Am. J. Respir. Crit. Care Med.* **170**, 59–64.
16. Anderson, P., Askgaard, D., Ljungqvist, L., Bennedsen, J. & Heron, I. (1991) *Infect. Immun.* **59**, 1905–1910.
17. Ohman, R. & Ridell, M. (1996) *Tuberc. Lung Dis.* **77**, 454–461.
18. Fiorio, W., Bottai, D., Batoni, G., Esin, S., Pardini, M., Maisetta, G. & Campa, M. (2002) *Clin. Diagn. Lab. Immunol.* **9**, 846–851.
19. Perschinka, H., Mayr, M., Millonig, G., Mayerl, C., van der Zee, R., Morrison, S. G., Morrison, R. P., Xu, Q. & Wick, G. (2003) *Arterioscler. Thromb. Vasc. Biol.* **23**, 1060–1065.
20. Bradford, M. M. (1976) *Analyt. Biochem.* **72**, 248–252.
21. Saiman, L. (2004) *Paediatr. Respir. Rev.* 221–223.
22. Bardou, F., Quemard, A., Dupont, M. A., Horn, C., Marchal, G. & Daffe, M. (1996) *Antimicrob. Agents Chemother.* **40**, 2459–2467.
23. Shiratsuch, H. & Basson, M. D. (2003) *Am. J. Surg.* **186**, 547–551.
24. Hadad, D. J., Palaci, M., Pignatari, A. C., Lewi, D. S., Machado, M. A., Telles, M. A., Martins, M. C., Ueki, S.Y., Vasconcelos, G. M. & Palhares, M. C. (2004) *Epidemiol. Infect.* **132**, 151–155.
25. Choudhary, R. K. Mukhopadhyay, S., Chakhaiyar, P., Sharma, N., Murthy, K. J. R., Katoch, V. M. & Hasnain, S. E. (2003) *Infect. Immun.* **71**, 6338–6343.
26. Ramalingam, B., UmaDevi, K. R. & Raja, A. (2003) *Scand. J. Infect. Dis.* **35**, 234–239.
27. Brusasca, P. N., Peters, R. L., Motzel, S. L., Klein, H. J. & Gennaro, M. L. (2003) *Comp. Med.* **53**, 165–172.
28. Maekura, R., Kohno, H., Hirotani, A., Okuda, Y., Ito, M., Ogura, T. & Yano, I. (2003) *J. Clin. Microbiol.* **41**, 1322–1325.
29. Perkins, M. D., Conde, M. B., Martins, M. & Kritski, A. L. (2003) *Chest* **123**, 107–112.
30. Maekura, R., Okuda, Y., Nakagawa, M., Hiraga, T., Yokota, S., Ito, M., Yano, I., Kohno, H., Wada M., Abe, C., *et al.* (2001) *J. Clin. Microbiol.* **39**, 3603–3608.
31. Chakhaiyar, P., Nagalakshmi, Y., Aruna, B., Murthy, K., J., R, Katoch, V., M. & Hasnain, S., E. (2004) *J. Infect. Dis.*, in press.

Mycobacterium tuberculosis Isolate with a Distinct Genomic Identity Overexpresses a Tap-Like Efflux Pump

N. Siddiqi, R. Das, N. Pathak, S. Banerjee, N. Ahmed, V. M. Katoch, S.E. Hasnain

Abstract

Background: One mechanism proposed for drug resistance in *Mycobacterium tuberculosis* (MTB) is by efflux of the drugs by membrane located pumps. We report a novel and definite association between drug resistance and transcription levels of a tap-like pump (Rv1258c) in a multi-drug resistant MTB patient isolate (ICC154) which possesses a unique genotypic signature.

Materials and Methods: The isolate ICC154 was tested for drug sensitivity. Over-expression of Rv1258c as a function of drug pressure was analyzed by RT-PCR and the strain was typed using fluorescent amplified fragment length polymorhism (FAFLP).

Result: In the presence of rifampicin and ofloxacin, this isolate shows increased transcription of the gene Rv1258c. Genotypic fingerprinting revealed the presence of unique FAFLP markers.

Conclusion: A clear association between drug resistance and overexpression of an efflux protein is evident from our studies. The presence of specific markers has implications in rapid identification of MDR clinical isolates and consequent disease management.

Infection 2004; 32: 109–111
DOI 10.1007/s15010-004-3097-x

Introduction

Drug efflux pumps are major contributors to drug resistance in human pathogens and cancer cells. Drug resistance in bacteria, including some species of *Mycobacterium*, has been associated with membrane-located efflux pumps that prevent cytosolic accumulation of drugs. The presence of pumps in *Mycobacterium smegmatis* [1], *Mycobacterium fortuitum* [2] and *Mycobacterium tuberculosis* (MTB) [3] has been reported. The MTB strain H37Rv genome carries 20 such putative efflux proteins [4] but there is no direct evidence to date of the involvement of these proteins in drug resistance in MTB clinical isolates. In many instances it has not been possible to relate the level of resistance in multi-drug resistant (MDR) strains to mutations in the target genes [5], pointing to the possible alternative, the existence of such MDR pumps. In their paper *Ainsa* et al. [2] have shown that the Rv1258c gene of MTB encodes a protein that is similar to the tap protein in *M. fortuitum* and confers weak resistance to tetracycline. We report a definite association between drug resistance and transcription levels of the Rv1258c gene in an MDR MTB patient isolate (ICC154) collected from the National Mycobacterial Repository at the Central JALMA Institute for Leprosy, Agra, India. This isolate was resistant to the following drugs: rifampicin (MIC 40 μg/ml), ofloxacin (MIC 4 μg/ml), isoniazid (MIC 2 μg/ml) and minomycin (MIC 2 μg/ml) as determined by the proportionate method [6]. Another interesting feature of this strain was the specific genomic signature obtained with fluorescent amplified fragment length polymorphism (FAFLP) based microrestriction analysis [7].

Materials and Methods

The isolate was cultured in Middlebrooks 7H9 medium with and without the following drugs: rifampicin (MIC 10 μg/ml), isoniazid (MIC 0.5 μg/ml) and ofloxacin (MIC 1 μg/ml). Thus, only sub-inhibitory concentrations of the drugs (1/4 of the MICs) were used to check the transcription profile. The H37Rv strain was included as a control. RNA was extracted using an RNA extraction kit (Qiagen, USA). Deoxyribonucleotide primers corresponding to the Rv1258c gene were designed from the sequence available in the Genebank (acct. no. AL123456) to generate a 456 bp amplicon upon RT-PCR. A control pair of primers was designed to amplify a 245 bp region of the housekeeping gene Rv1437 coding for phosphoglycerate kinase. The RT-PCR products were resolved in a 2% agarose gel and the products' band intensities determined using 'Quantity One (4.1.1)' gel documentation software (BioRad, USA).

N. Siddiqi
Centre for DNA Fingerprinting and Diagnostics, Nacharam, Hyderabad and National Institute of Immunology, New Delhi, India
R. Das, V.M. Katoch
Central JALMA Institute for Leprosy, Tajganj, Agra, India
N. Pathak, S. Banerjee, N. Ahmed, S.E. Hasnain *(corresponding author)*
Centre for DNA Fingerprinting and Diagnostics, ECIL Road, Nacharam, Hyderabad, 500076, India; Phone: (+91/40) 7155-610, Fax: -479, e-mail: ehtesham@cdfd.org.in

Received: June 12, 2003 • Revision accepted: December 1, 2003

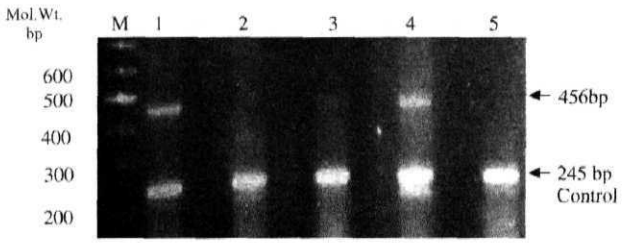


Figure 1. The RT-PCR products were fractionated in a 2% agarose gel. *Lane 5 represents the control strain H37Rv. Lanes 1 to 4 show the amplicon generated from RNA extracted from ICC 154 when the isolate was grown in the presence of rifampicin, isoniazid, no drug and* ofloxacin, respectively. Lane M is the 100 bp DNA size marker.

Whole genome micro-restriction fingerprinting with EcoRI/MseI enzymes using fluorescence tagged primer pairs EcoRI+A/MseI+0 and EcoRI+G/MseI+0 [7–9] was performed for ICC154 and H37Rv genomes. The profiles for each were then generated by electrophoretic separation on a 5% acrylamide gel and scoring of the fluorescent markers using an automated DNA analysis workstation (ABI Prism 377 DNA sequencer).

Results

The RT-PCR based analysis revealed a tenfold increase in the Rv1258c transcript level in the presence of rifampicin, and a sixfold increase in the presence of ofloxacin. The 456 bp amplicon band intensity of Rv1258c relative to the intensity of the 245 bp control band provided the means to compare the RNA transcript level of the Rv1258c gene. The normalized band intensities in all the lanes were subsequently compared to each other. It is clear that the levels of Rv1258c RNA vary in the presence or absence of different drugs (Figure 1). The control H37Rv strain in lane 5 shows hardly any RT-PCR product corresponding to Rv1258c RNA. The drug-resistant isolate ICC154 shows an appreciable increase in the corresponding RNA levels when grown in the presence of rifampicin (lane 1) or ofloxacin (lane 4) as compared to the isolate grown in the absence of any drug (lane 3).

Whole genome microrestriction fingerprinting with EcoRI/MseI enzymes using primers EcoRI+A/MseI+0 revealed marked differences (Figure 2, shown in boxes) in the profiles obtained with ICC154 and H37Rv genomes. One of the markers absent in ICC 154 was mapped to a hypothetical protein gene Rv2998 in H37Rv and its homolog MT 3076 in the CDC 1551 genome. In another analysis with primer pair EcoRI+G/MseI+0, a distinct genotypic pattern for ICC154 was observed. It comprised four distinct markers of sizes 408, 433, 465, 487 and 494 base pairs. These markers were found to be clearly absent in the case of the H37Rv genotype. The most remarkable observation was, however, the characteristic absence of the 383/384 base pair marker amplified by EcoRI+T/MseI+0 primers in the case of ICC 154 and all other 52 clinical isolates (data not shown) with a documented ofloxacin resistance pattern *in vitro*. This marker coincides with the *pst* C2 gene of H37Rv and its homolog in the CDC 1551 genome.

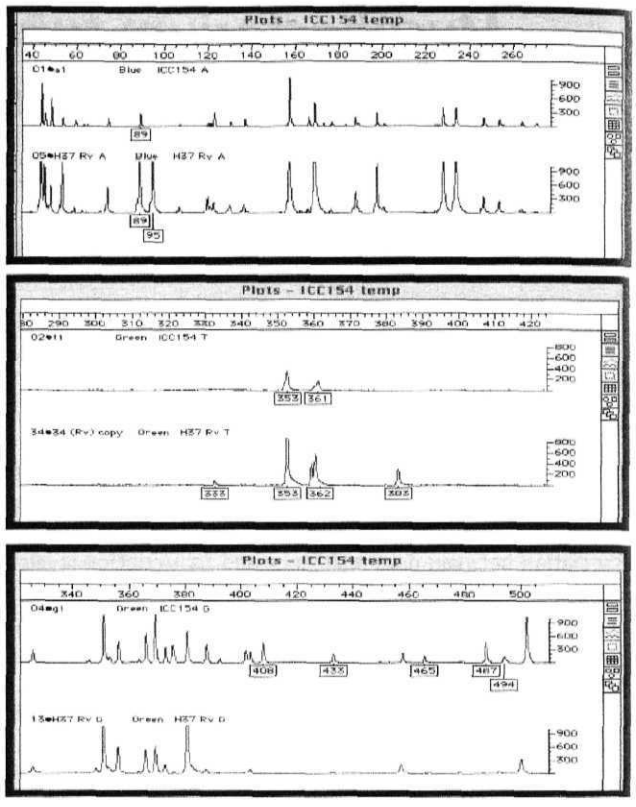


Figure 2. FAFLP analysis [8] results comparing ICC154 to H37Rv generated using primers EcoRI+A/MseI+0 (upper panel), EcoRI+T/MseI+0 (middle panel) and EcoRI+G/MseI+0 (lower panel). Horizontal scale represents marker sizes in base pairs, while vertical scale shows relative peak heights signifying quantitative variation in PCR products.

Discussion

A clear association between drug resistance and overexpression of a tap-like efflux pump under drug pressure in a clinical setup is evident from our studies. Densitometric quantification of the amplicon in RT-PCR based analyses revealed a tenfold increase in the Rv1258c transcript level when the isolate was grown in the presence of rifampicin, and a sixfold increase in the presence of ofloxacin. The MDR loci *viz rpoB, katG* and *gyrA* in ICC154 were amplified and sequenced [5], which revealed a deletion of the *katG* gene. Even though the observed mutations in the *gyrA* and *rpoB* loci were found to be those generally associated with drug resistance [10], the high resistance to rifampicin and ofloxacin could further be a reflection of overexpression of the Rv1258c gene as evident from the RT-PCR result.

The presence of the unique markers as revealed by the FAFLP is interesting, especially the characteristic absence

of the 383/384 base pair marker in ICC 154 as well as all the other 52 clinical isolates (data not shown) with a documented ofloxacin resistance pattern *in vitro*. It is likely that strain ICC154, and other related isolates that might be circulating in North India with an MDR phenotype may have distinct genotypic signatures. While this observation is exciting, nonetheless it is too early to suggest that FAFLP typing could be a replacement for susceptibility testing for direct detection of MDR strains. Clearly, a large well-controlled study showing correlation of antibiotic resistance with FAFLP signatures would be required before the latter could be used as a surrogate for actual susceptibility.

Acknowledgement

This study was supported by a research grant to SEH from the New Millennium Indian Technology Leadership Initiative of the Council of Scientific and Industrial Research, Government of India.

References

1. Takiff HE, Cimino M, Musso MC, Weisbrod T, Martinez R, Delgado MB, Salazar L, Bloom BR, Jacobs WR: Efflux pump of the proton antiporter family confers low level fluoroquinolone resistance in *Mycobacterium smegmatis*. Proc Nat Acad Sci USA 1996; 93: 362–366.
2. Ainsa JA, Blokpoel MCJ, Otal I, Young DB, Smet KA, Martin C: Molecular cloning and characterization of Tap, a putative multidrug efflux pump present in *Mycobacterium fortuitum* and *Mycobacterium tuberculosis*. J Bacteriology 1998; 180: 5836–5843.
3. Doran JL, Pang Y, Mdluli KE, Moran AJ, Victor TC, Stokes RW, Mahenthiralingam E, Kreiswirth BN, Butt JL, Baron GS, Treit JD, Kerr VJ, Van Helden PD, Roberts MC, Nano FE: *Mycobacterium tuberculosis* Efpa encodes an efflux protein of the qaca transporter family. Clin Diagn Lab Immunol 1997; 4: 23–32.
4. Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, Gordon SV, Eiglmeier K, Gas S, Barry CE 3rd, Tekaia F, Badcock K, Basham D, Brown D, Chillingworth T, Connor R, Davies R, Devlin K, Feltwell T, Gentles S, Hamlin N, Holroyd S, Hornsby T, Jagels K, Barrell BG: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. Nature 1998; 393: 537–544. Erratum in: Nature 1998; 396: 190.
5. Siddiqi N, Shamim M, Hussain S, Choudhary RK, Ahmed N, Prachee, Banerjee S, Savithri GR, Alam M, Pathak N, Amin A, Hanief M, Katoch VM, Sharma SK, Hasnain SE: Molecular characterization of multidrug-resistant isolates of *Mycobacterium tuberculosis* from patients in North India. Antimicrob Agents Chemother 2002; 46: 443–450.
6. Canetti G, Fox W, Khomenko A, Mahler HT, Menon NK, Mitchison DA, Rist N, Smelev NA: Advances in techniques of testing mycobacterial drug sensitivity, and the use of sensitivity tests in tuberculosis control programmes. Bull World Health Organ 1969; 41: 21–43.
7. Ahmed N, Alam M, Majeed AA, Rahman SA, Cataldi A, Cousins D, Hasnain SE: Genome sequence based, comparative analysis of the Fluorescent Amplified Fragment Length Polymorphisms (FAFLP) of tubercle bacilli from seals provides molecular evidence for a new species within the *Mycobacterium tuberculosis* complex. Infection Genet Evol 2003; 2: 193–199.
8. Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M: AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 1995; 23: 4407–4414.
9. Ahmed N, Caviedes L, Alam M, Rao KR, Sangal V, Sheen P, Gilman RH, Hasnain SE: Distinctiveness of *Mycobacterium tuberculosis* genotypes from human immunodeficiency virus type 1-seropositive and seronegative patients in Lima, Peru. J Clin Microbiol 2003; 41: 1712–1716.
10. Siddiqi N, Shamim M, Jain NK, Rattan A, Amin A, Katoch VM, Sharma SK, Hasnain SE: Molecular genetic analysis of multidrug resistance in Indian isolates of *Mycobacterium tuberculosis*. Mem Inst Oswaldo Cruz 1998; 93: 589–594.

ANTIMICROBIAL AGENTS AND CHEMOTHERAPY, Feb. 2002, p. 443–450
0066-4804/02/$04.00+0 DOI: 10.1128/AAC.46.2.443–450.2002

Vol. 46, No. 2

Molecular Characterization of Multidrug-Resistant Isolates of *Mycobacterium tuberculosis* from Patients in North India

Noman Siddiqi,[1,2] Mohammed Shamim,[2] Seema Hussain,[2] Rakesh Kumar Choudhary,[1] Niyaz Ahmed,[1] Prachee,[1] Sharmistha Banerjee,[1] G. R. Savithri,[1] Mahfooz Alam,[1] Niteen Pathak,[1] Amol Amin,[2] Mohammed Hanief,[3] V. M. Katoch,[4] S. K. Sharma,[5] and Seyed E. Hasnain[1,2,6]*

Centre for DNA Fingerprinting and Diagnostics, Hyderabad 500076,[1] National Institute of Immunology,[2] New Delhi TB Centre,[3] and Department of Medicine, A.I.I.M.S.,[5] New Delhi, Central Jalma Institute of Leprosy, Agra,[4] and Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore 560064,[6] India

Received 17 April 2001/Returned for modification 29 June 2001/Accepted 1 November 2001

The World Health Organization has identified India as a major hot-spot region for *Mycobacterium tuberculosis* infection. We have characterized the sequences of the loci associated with multidrug resistance in 126 clinical isolates of *M. tuberculosis* from India to identify the respective mutations. The loci selected were *rpoB* (rifampin), *katG* and the ribosomal binding site of *inhA* (isoniazid), *gyrA* and *gyrB* (ofloxacin), and *rpsL* and *rrs* (streptomycin). We found known as well as novel mutations at these loci. Few of the mutations at the *rpoB* locus could be correlated with the drug resistance levels exhibited by the *M. tuberculosis* isolates and occurred with frequencies different from those reported earlier. Missense mutations at codons 526 to 531 seemed to be crucial in conferring a high degree of resistance to rifampin. We identified a common Arg463Leu substitution in the *katG* locus and certain novel insertions and deletions. Mutations were also mapped in the ribosomal binding site of the *inhA* gene. A Ser95Thr substitution in the *gyrA* locus was the most common mutation observed in ofloxacin-resistant isolates. A few isolates showed other mutations in this locus. Seven streptomycin-resistant isolates had a silent mutation at the lysine residue at position 121. While certain mutations are widely present, pointing to the magnitude of the polymorphisms at these loci, others are not common, suggesting diversity in the multidrug-resistant *M. tuberculosis* strains prevalent in this region. Our results additionally have implications for the development of methods for multidrug resistance detection and are also relevant in the shaping of future clinical treatment regimens and drug design strategies.

Recent years have witnessed a dramatic upsurge in cases of drug-resistant *Mycobacterium tuberculosis* infections. The acquisition of resistance by the bacterium is a random event, and in a given mycobacterial population, 1 in 10^6 bacteria mutates to develop isoniazid resistance, while 1 in 10^8 mutates to develop rifampin resistance (8). The chance that a bacterium will acquire multidrug resistance (defined as resistance to at least rifampin and isoniazid) is thus 10^{-14} (8). The drug-resistant phenotype may get selected due to single-drug therapy, poor patient adherence, and improper diagnosis. With the AIDS pandemic fuelling increasing numbers of multidrug-resistant (MDR) strains of *M. tuberculosis*, urgent measures need to be taken to contain this scourge (2). A recently published World Health Organization report reviewing the global status of tuberculosis has pointed to an increasing incidence of drug-resistant tuberculosis (5). The highest rates of MDR tuberculosis have been reported in Nepal (48.0%), Gujarat, India (33.8%), New York City (30.1%), Bolivia (15.3%), and Korea (14.5%). Furthermore, the report points to the alarming increase in the number of tuberculosis patients in the Indian subcontinent, with India being singled out as having the greatest burden of tuberculosis patients. Three different studies from North and Northwest India indicate an increasing incidence of acquired MDR tuberculosis (9, 12, 15). Furthermore, the incidence of primary MDR tuberculosis in North India was put at 3.3% in one of the studies (12).

While there is lot of literature on the molecular epidemiology and characterization of MDR isolates from the United States and Europe, the same is not true for the Indian strains. The prevalence of drug-resistant tuberculosis in North India is known, but no serious efforts have been made to identify the drug resistance genotypes or their prevalence in the community. The present study was undertaken to characterize mutations prevalent in patient isolates of *M. tuberculosis* from North India with respect to a few of these drug target loci. We have chosen to look at the drug target genes for the drugs rifampin, isoniazid, streptomycin, and fluoroquinolones, which are commonly prescribed for the treatment of tuberculosis in North India. The first three drugs are the frontline drugs in tuberculosis chemotherapy, while fluoroquinolones are prescribed for drug-resistant cases. The loci studied were *rpoB* (RNA polymerase B subunit), *katG* (catalase-peroxidase), *inhA* (enoyl coenzyme A reductase), *rpsL* (ribosomal protein S12), *rrs* (16S rRNA), and *gyrAB* (DNA gyrases A and B). The present study, in combination with the molecular epidemiology of the drug-resistant strains, will help track the routes of infection and the extent of drug-resistant tuberculosis in this region. The elucidation of common and novel mutations in these loci could form the basis for the creation of new diagnostic tools and the

* Corresponding author. Mailing address: C.D.F.D., ECIL Rd., Nacharam, Hyderabad-500 076, India. Phone: 91-040 7155604 or 91-040 7155605. Fax: 91-040 7155610 or 91-040 7155479. E-mail: director@www.cdfd.org.in.

TABLE 1. Primers used in the study to amplify and sequence the different loci, amplicon sizes, annealing temperatures, and amplicon positions on the respective genes

Gene (accession no.)	Primer	Sequence	Annealing temp (°C)	Position (nt)	Amplicon size (bp)
rpoB (L27989)	Forward	GGG AGC GGA TGA CCA CCC	60	2266	350
	Reverse	GCG GTA CGG CGT TTC GAT GAA C		2615	
katG (X68081)	Forward	GCC CGA GCA ACA CCC	60	3	237
	Reverse	ATG TCC CGC GTC AGG		239	
	Forward	CGA GGA ATT GGC CGA CGA GTT	55	1187	414
	Reverse	CGG CGC CGC GGA GTT GAA TGA		1600	
inhA regulator sequence	Forward	CCT CGC TGC CCA GAA AGG GA	45	Upstream of *inhA* gene	248
	Reverse	ATC CCC CGG TTT CCT CCG GT			
gyrA (L27512)	Forward	CAG CTA CAT CGA CTA TGC GA	45	2383	320
	Reverse	GGG CTT CGG TGT TAC CTC AT		2702	
gyrB (L27512)	Forward	CCA CCG ACA TCG GTG GAT T	55	1538	428
	Reverse	CTG CCA CTT GAG TTT GTA CA		1965	
rpsl (X70995)	Forward	GGC CGA CAA ACA GAA CGT	54	5′ noncoding region	505
	Reverse	GTT CAC CAA CTG GGT GAC		S7 gene	
rrs (Z83862)	Forward	TTG GCC ATG CTC TTG ATG CCC	54	141	1140
	Reverse	TGC ACA CAG GCC ACA AGG GA		1280	

development of novel strategies that can be used to combat the menace of drug-resistant *M. tuberculosis*.

MATERIALS AND METHODS

Sources of *Mycobacterium* isolates. *Mycobacterium* isolates were collected from patients reporting to the outpatient departments of hospitals in northern India, primarily New Delhi and its neighboring regions. Another source of samples was the National Mycobacterial Repository at the Central Jalma Institute for Leprosy, Agra, India. The samples collected over a 3-year period from 1995 to 1998 were included in the present study. A large number of the patients (75%) had histories of previous treatment and were on antitubercular treatment at the time of collection of their sputa. Most of these patients had been through various degrees of antitubercular drug therapy during the previous 20 months. Rifampin and isoniazid were the most common drugs used in these regimens. Sputum samples collected from patients reporting with pulmonary tuberculosis were processed by standard methods and were streaked onto Lowenstein-Jensen slants. Most of them were coded with ICC numbers (ICC01, ICC201, etc.). The samples were biochemically characterized as belonging to the *M. tuberculosis* complex by nitrate reduction, niacin production, and BACTEC NAP tests. Drug susceptibility profiles were evaluated by the proportion method. The drugs tested were rifampin (Lupin, India), isoniazid (Lupin), ofloxacin (Ranbaxy, India), and streptomycin (Lupin). The MICs at which the isolates were considered resistant were as follows: 10 μg/ml for rifampin, 1 μg/ml for isoniazid, 2 μg/ml for ofloxacin, and 2 μg/ml for streptomycin. The numbers of drug-resistant isolates included in the study were as follows: for rifampin, $n = 94$; for isoniazid, $n = 74$; for streptomycin, $n = 14$; and for ofloxacin, $n = 68$. A total of 126 isolates were tested. Thirty-six isolates were resistant to a single drug, 66 isolates were resistant to two drugs, 22 isolates were resistant to three drugs, and 4 isolates were resistant to four drugs.

DNA isolation and PCR. The isolates were cultured on Lowenstein-Jensen slants. The colonies were scraped, resuspended in 500 μl of TE (10 mM Tris, 1 mM EDTA [pH 8]), and killed by freezing at −70°C followed by heating at 80°C. This cycle was repeated thrice to kill all the bacteria. The DNA was isolated (by tretament with cetyltrimethylammonium bromide in the presence of 0.7 M sodium chloride) and amplified by standardized protocols as reported previously (21).

Table 1 lists the sequences of the different primers used and their positions on the corresponding genes. It also lists the amplicon sizes generated and the annealing temperatures used for PCR cycling. The temperatures used for all cycles were identical for all PCRs except for that for annealing, the temperature of which varied for each primer pair. Briefly, 35 cycles of 94°C for 1 min, 45 to 60°C for 1 min, and 72°C for 2 min were used to amplify the loci. The samples were resolved in a 2% agarose gel, and the specific bands were excised. DNA was extracted from the gel slices with a QIAquick gel extraction kit (Qiagen, Chatsworth, Calif.) according to the manufacturer's instructions. The purified DNA was resuspended in sterile double-distilled water and was used for the sequencing studies.

DNA sequencing. Sequencing of the amplicons was carried out with an ABI Prism 377 automated DNA sequencer (ABI Prism). PCR sequencing was carried out with a BigDye terminator kit (ABI Prism) according to the manufacturer's instructions. The Sequencing Analysis (version 3.3) software package was used to analyze the gel information. The sequences generated with the program were compared to their respective wild-type sequences by using MegAlign software (Lasergene; DNASTAR, Inc., Madison, Wis.).

RESULTS

Mutations in the hot-spot regions of various loci were characterized. The results are summarized in Table 2. On the basis of the drug susceptibility profile for an isolate, the corresponding loci (representing the drug target gene) were amplified and sequenced. The largest number of samples was obtained from New Delhi, followed by Chandigarh, Ahmedabad, Agra, Bangalore, and Shimla, with a few samples coming from Jaipur and Chennai. Except for Chennai and Bangalore, all the cities are located in North India. We could establish a previous treatment history for patients from whom 94 of the 126 isolates

TABLE 2. Characteristics of *M. tuberculosis* isolates from patients

Strain no.	Geographic location	Treatment history[a]	Drug susceptibility[b]	Polymorphism[c]			
				rpo	*katG* or *inhA*	*gyrA*	*rpsL*
ICC14	New Delhi	+	R^r, I^r, O^r	D516V	N35D, NA at second locus	S95T	
ICC19	New Delhi	+	R^r, I^r, O^r	L511L, S531L	R463L	S95T	
ICC23	New Delhi	+	R^r, I^r	L511L, S531L	NA		
ICC98	New Delhi	−	I^r, O^r		R463L, Inh (C/T)	S95T	
ICC100	New Delhi	+	R^r, I^r, O^r, S^r	S531L	R463L	S95T	NM
ICC101	New Delhi	+	R^r, O^r	S531L		S95T	
ICC102	New Delhi	+	R^r, I^r, O^r	S531L	NM	S95T	
ICC103	New Delhi	+	R^r, I^r, O^r	L511L, N518T	NM	A90A, S95T	
ICC104	New Delhi	+	R^r, I^r, O^r	D516V	Δ30C, R463L	S95T	
ICC105	New Delhi	+	R^r, I^r	K527N	R463L		
ICC107	New Delhi	−	R^r, O^r, S^r	N518T, R528P		S95T	NM
ICC109	New Delhi	+	I^r		NM		
ICC111	New Delhi	+	R^r, I^r, S^r	S531W	Insertion 185C		K121K
ICC114	New Delhi	+	I^r		R463L		
ICC115	New Delhi	+	R^r, I^r	S531W	Insertion 98A, R463L		
ICC123	New Delhi	+	R^r, I^r, O^r	R528P	NA	S95T	
ICC124	New Delhi	+	R^r, I^r	H526Y	R463L		
ICC125	New Delhi	+	R^r	L511L, S531L			
ICC128	New Delhi	+	R^r, O^r, S^r	H526Y, R528H		S95T	K121K
ICC129	New Delhi	+	R^r, I^r, S^r	R528P	R463L		K121K
ICC147	New Delhi	−	R^r, I^r, O^r	S531W	Δ109G, R463L	S95T	
ICC203	New Delhi	+	R^r	D516V			
ICC204	New Delhi	+	R^r, I^r	L521L, K527N	R463L		
ICC205	New Delhi	+	R^r, I^r	D516V	NA		
ICC206	New Delhi	+	R^r, I^r	S531W	NM		
ICC208	New Delhi	−	R^r	D516V	R463L		
ICC209	New Delhi	−	R^r	D516V			
ICC210	New Delhi	+	R^r, O^r	L511V, N518T		S95T	
ICC211	New Delhi	+	I^r, O^r		R463L	S95T	
ICC212	New Delhi	−	R^r	S531L			
ICC213	New Delhi	+	R^r, I^r	R528H, S531W	R463L, Inh (C/T)		
ICC214	New Delhi	+	I^r		R463L		
ICC215	New Delhi	+	R^r	S531L			
ICC216	New Delhi	+	R^r, S^r	S531L			K121K
ICC217	New Delhi	+	R^r, I^r, O^r	S531L	R463L	S95T	
ICC218	New Delhi	+	R^r, I^r, O^r	S522Q	R463L	S95T	
ICC219	New Delhi	+	I^r		T12P, R463L		
ICC220	New Delhi	+	R^r, I^r, O^r, S^r	S531W	R463L, Inh (T/A)	D94G, S95T	NM
ICC221	New Delhi	−	R^r, O^r	L521L		D94A, S95T	
ICC222	New Delhi	−	O^r			S95T	
ICC223	New Delhi	−	R^r, O^r	H526Y		S95T	
ICC225	New Delhi	+	R^r, O^r	S531L		S95T	
ICC237	New Delhi	−	R^r, O^r	D516G		S95T	
ICC239	New Delhi	+	R^r, I^r	D516V	R463L		
ICC240	New Delhi	+	R^r, I^r	D516V	R463L		
ICC242	New Delhi	+	R^r, O^r	L511V		S95T	
ICC244	New Delhi	+	R^r, I^r, O^r	S531L	A61T, R463L	S95T	
ICC246	New Delhi	+	I^r, O^r		Insertion 185C, R463L	S95T	
ICC275	New Delhi	+	R^r, O^r	H526Y		S95T	
ICC277	New Delhi	+	R^r, I^r, O^r, S^r	H526Y	Δ30C	D94A, S95T	K121K
ICC284	New Delhi	−	O^r			NM	
ICC286	New Delhi	+	R^r, I^r	D516G	NM		
ICC287	New Delhi	+	R^r, I^r	H526Y	NM		
ICC325	New Delhi	+	I^r, O^r, S^r		NM	NM	NM
ICC326	New Delhi	+	R^r, I^r, S^r	H526L	Insertion 98A, R463L		NM
ICC327	New Delhi	+	R^r, I^r, S^r	S509R	R463L		K121K
ICC328	New Delhi	−	O^r			NM	
ICC408	New Delhi	+	R^r, I^r, O^r, S^r	Q510H, S531W	R463L	S95T	K121K
ICC425	New Delhi	+	R^r, I^r	S531L	R463L		
F4	New Delhi	+	R^r, O^r	H526Y		D94G, S95T	
F5	New Delhi	+	R^r, O^r	S531L		S95T	
F7	New Delhi	+	R^r, O^r	S531L		A90V, S95T	
F8	New Delhi	+	R^r, O^r	S531L		A90V, S95T	
F9	New Delhi	+	R^r, O^r	N518T		S91P, S95T	
N31	New Delhi	+	I^r		R463L		

Continued on following page

TABLE 2—*Continued*

Strain no.	Geographic location	Treatment history[a]	Drug susceptibility[b]	Polymorphism[c]			
				rpo	*katG* or *inhA*	*gyrA*	*rpsL*
N33	New Delhi	+	I^r		R463L		
N34	New Delhi	+	I^r		R463L		
N35	New Delhi	+	I^r		D73N, R463L		
N36	New Delhi	−	I^r		R463L		
ICC32	Ahmedabad	+	R^r, I^r	S531L	NA		
ICC33	Ahmedabad	+	R^r, I^r, O^r	S509R	R463L	S95T	
ICC36	Ahmedabad	+	I^r, O^r		R463L	S95T	
ICC37	Ahmedabad	+	R^r, I^r	D516G	Insertion 98A, R463L		
ICC131	Ahmedabad	+	R^r, I^r	H526Y, R528P	R463L		
ICC132	Ahmedabad	−	O^r			S95T	
ICC133	Ahmedabad	+	R^r, I^r	H526R	Δ30C, R463L		
ICC134	Ahmedabad	−	R^r, O^r	S522Q		S95T	
ICC136	Ahmedabad	+	R^r, I^r	H526R	Insertion 98A, R463L		
ICC137	Ahmedabad	−	O^r			S95T	
ICC138	Ahmedabad	−	O^r			S95T	
ICC226	Ahmedabad	+	R^r, I^r	L511L, H526R	Δ109G, R463L		
ICC233	Ahmedabad	+	R^r, I^r	D516V, H526Y	Δ30C, R463L		
ICC151	Chandigarh	−	O^r			S95T	
ICC154	Chandigarh	−	R^r, I^r, O^r	S531L	Δ30C, R463L	S95T	
ICC155	Chandigarh	+	I^r		Δ30C, R463L		
ICC159	Chandigarh	+	R^r, O^r	H526Y		S95T	
ICC161	Chandigarh	+	R^r, O^r	S522Q		S95T	
ICC162	Chandigarh	−	R^r, O^r	S522Q		S95T	
ICC164	Chandigarh	−	O^r			S95T	
ICC165	Chandigarh	−	O^r			S95T	
ICC166	Chandigarh	−	O^r			S95T	
ICC167	Chandigarh	+	I^r, O^r		R463L	S95T	
ICC168	Chandigarh	+	O^r			S95T	
ICC169	Chandigarh	+	R^r, I^r	N518T	R463L		
ICC170	Chandigarh	−	O^r			S95T	
ICC171	Chandigarh	+	R^r, I^r	S531L	R463L		
ICC172	Chandigarh	+	R^r, O^r	S522Q		S95T	
ICC173	Chandigarh	−	R^r, S^r	H526L			K121K
ICC174	Chandigarh	+	O^r			S95T	
ICC175	Chandigarh	−	R^r, O^r	H526Y		S95T	
ICC247	Chandigarh	−	R^r, O^r	D516V		S95T	
ICC248	Chandigarh	−	R^r, O^r	D516V		S95T	
ICC249	Chandigarh	−	O^r			S95T	
ICC251	Chandigarh	−	O^r			NM	
ICC254	Chandigarh	+	R^r	S531L			
ICC255	Chandigarh	−	R^r, O^r	N518T		NM	
ICC256	Chandigarh	+	R^r, I^r	H526Y	NM		
ICC257	Chandigarh	+	R^r, O^r	Q510H, L511L		S95T	
ICC262	Chandigarh	+	R^r, I^r, O^r	D516V	R463L	S95T	
ICC95	Bangalore	−	O^r			S95T	
ICC96	Bangalore	+	R^r, O^r	S531L		S95T	
ICC399	Bangalore	+	R^r, O^r	S531W		NM	
ICC524	Bangalore	+	R^r, I^r	S531L	R463L		
ICC525	Bangalore	+	R^r, I^r	S531L	R463L		
ICC143	Shimla	−	O^r			NM	
ICC144	Shimla	+	I^r		R463L		
ICC145	Shimla	−	O^r			NM	
A3	Agra	+	R^r, I^r, S^r	S531L	R463L		NM
A4	Agra	+	R^r, I^r	D516V	R463L		
A9	Agra	+	R^r, I^r	S531L	R463L		
A11	Agra	+	R^r, I^r	D516G	T11A, R463L		
A12	Agra	+	R^r, I^r	D516V	N35D, R463L		
A13	Agra	+	R^r, I^r	D516V, N518T	R463L		
A14	Agra	+	R^r, I^r	D516V	R463L		
A15	Agra	+	R^r, I^r	S531L	R463L		
ICC332	Jaipur	+	R^r, I^r	H526R	Insertion 185C, R463L		
ICC337	Jaipur	+	R^r, I^r	S531L	R463L	S95T	
ICC85	Chennai	+	R^r, I^r, O^r	H526R	NA	S95T	

[a] History of treatment in the previous 20 months.
[b] R^r, rifampin resistant; I^r, isoniazid resistant; O^r, ofloxacin resistant; S^r, streptomycin resistant.
[c] NA, no amplification; NM, no mutation; Inh, mutation in the *inhA* ribosome binding site; Δ, deletion at the indicated nucleotide position.

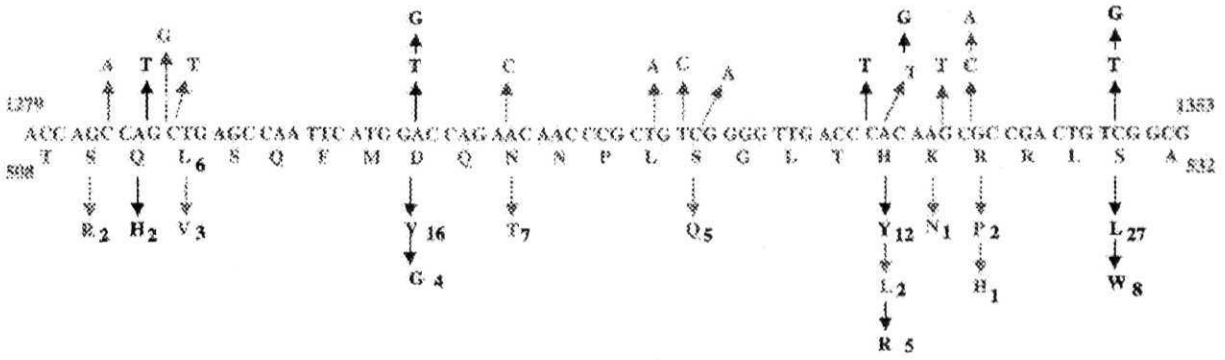


FIG. 1. Summary of mutations at codons 508 to 532 in the *rpoB* gene. The wild-type sequence and amino acids are shown in the middle frame. Nucleotide changes are marked with arrows in the top frame, and the corresponding amino acid changes are denoted in the bottom frame. The amino acids are subscripted with numbers that indicate the number of isolates harboring the change. Changes marked with orange lines (dotted arrows) are novel mutations; silent mutations are marked with blue lines (dashed arrows). Codons 531, 526, and 516 exhibit high degrees of polymorphism. Codons 509, 511, 522, 527, and 528 show novel mutations.

were recovered. These isolates probably represent those with acquired resistance, as the patients had at some time point been given antitubercular drug therapy.

A stretch of 30 amino acids at the center of the amplicon for the *rpoB* locus was studied. Amino acids 432 to 458 comprised the hot-spot region for mutations. For the sake of comparison, we used the corresponding *Escherichia coli* numbering, which is amino acids 507 to 533. We identified previously reported mutations as well as certain novel mutations. Codon 531 seemed to be the most vulnerable to mutations, as most rifampin-resistant isolates had this mutation (Fig. 1). Of the 93 rifampin-resistant strains in our study, 28 had the missense mutation Ser531Leu and 8 had the substitution Ser531Trp. The next most common mutations were the amino acid substitutions Asp516Val or Asp516Gly (20 isolates) and His526Tyr, His526Leu, or His526Arg (19 isolates). We found two isolates with Gln510His changes. While all these mutations have been reported earlier, we also found mutations that have not been reported previously. These included Ser509Arg (isolate ICC33), Leu511Val (isolate ICC242), Asn518Thr (isolate ICC107), Ser522Gln (isolate ICC172), Lys527Asn (isolate ICC105), Arg528Pro (isolate ICC129), and Arg528His (isolate ICC213). Most of these mutations occurred less frequently, comprising about 24% of the total mutations in the 94 isolates studied. Other mutations identified in our study were silent mutations at amino acids Leu511 and Leu521. Interestingly, the mutation at position 511 never occurred alone and was present only in isolates with more than one mutation at the *rpoB* locus.

An important outcome of these studies is the direct correlation of certain mutations to high MICs. Table 3 lists the isolates, their mutations, and the corresponding MICs at which they remained resistant. Mutations in codons 516 and 521 conferred low-level resistance (MIC, <40 μg/ml) to rifampin, whereas mutations in codons 510, 526, 527, 528, and 531 were seen to confer high levels of resistance (MICs, ≥64 μg/ml). Amino acids 526 to 531 appear to be very important in drug target interactions, and mutations in them result in MICs in the range of 64 μg/ml and above. In a few cases (e.g., for isolates ICC204, ICC257, and ICC128), double mutations were found to have an additive effect on the degree of resistance.

Insertion, deletion, and substitution mutations were mapped in the *katG* locus in 24 isoniazid-resistant isolates. In the present study we looked for mutations in the 5′ region (nucleotides [nt] 3 to 239) and the midregion (nt 1187 to 1600) of the *katG* gene, corresponding to amino acid positions 2 to 77 and 395 to 533, respectively. The results are summarized in Fig. 2. A C nucleotide at position 30 was deleted in six of the isolates. This deletion results in chain termination, thereby generating only a short polypeptide of 26 amino acids. Another deletion of a single nucleotide, a G residue at position 109, was observed in two isolates; this deletion would result in the production of

TABLE 3. Correlation of specific mutations with rifampin MICs[a]

Strain	Rifampin MIC (μg/ml)	Mutation	Mutation type	Amino acid change
ICC221	10	G1317A	Novel	L521L
ICC208, ICC205	10	A1304T	Reported	D516V
ICC37	10	A1304G	Reported	D516G
ICC204	40	G1317A	Novel	L521L
ICC204	40	G1336T	Novel	K527N
ICC105	40	G1336T	Novel	K527N
ICC129	40	G1338C	Novel	R528P
ICC131	40	C1331T	Reported	H526Y
ICC131	40	G1338C	Novel	R528P
ICC123	64	G1338C	Novel	R528P
ICC100	64	C1349T	Reported	S531L
ICC213	64	G1340A	Novel	R528H
ICC213	64	C1349G	Reported	S531W
ICC218	64	T1321C	Novel	S522Q
ICC218	64	C1322A	Novel	S522Q
ICC257	64	G1287T	Novel	Q510H
ICC257	64	C1288T	Novel	L511L
ICC275	64	C1333T	Reported	H526Y
ICC220	64	C1349G	Reported	S531W
ICC128	128	C1331T	Reported	H526Y
ICC128	128	G1338A	Novel	R528H

[a] Missense mutations in the RpoB protein at amino acid positions 510, 511, 522, 526, 527, 528, and 531 confer higher levels of resistance (MICs, ≥40 μg/ml) than those at positions 509, 516, and 521 (MICs, ≤10 μg/ml).

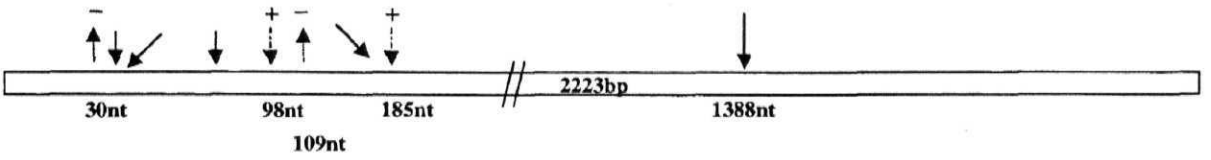


FIG. 2. Summary of mutations in the *katG* gene. Deletions are indicated by lines with a minus sign, while insertions are depicted by dashed lines with a plus sign. Solid lines show the substitutions. Codon 463 exhibited the highest degree of polymorphism, followed by the deletion at nucleotide 30.

a 45-amino-acid truncated polypeptide. Insertions were also observed at nt positions 98 (an A nucleotide) and 185 (a C nucleotide) in four and three isolates, respectively. Both of these insertions cause aberrant chain termination. Ala61Thr, Thr12Pro, Thr11Ala, Asp73Asn, and Asn35Asp missense mutations were observed in this locus in a few of the isolates. These are novel observations, as there are no reports of such mutations occurring in isoniazid-resistant strains from other parts of the world. We were unable to amplify this locus in six of the isolates (isolates ICC14, ICC23, ICC32, ICC85, ICC123, and ICC205), indicating a partial deletion of the gene. A common mutation in all these isolates was Arg463Leu. However, this mutation has been shown to have no direct consequence for drug resistance. To confirm this we sequenced this locus for all 126 isolates included in the study. It was found that the majority of the isolates carried this change. It has been argued previously that this polymorphism in the *katG* locus might be more important as a marker of evolution than as a marker of resistance (22). Three isoniazid-resistant isolates carried mutations in the ribosomal binding site upstream of the *inhA* gene. While two isolates showed a C-to-T transition, one had a T-to-A transversion. These mutations have previously been reported by other groups. The present understanding of these mutations is that they probably confer resistance by a drug titration effect.

Sixty-eight ofloxacin-resistant isolates were analyzed. The hot-spot region of the *gyrA* gene spanning codons 89 to 95 was sequenced to identify mutations. Most of the isolates showed a single mutation corresponding to the amino acid change Ser95Thr (Fig. 3). The second most common mutation, observed in four isolates, was Asp94Gly or Asp94Ala. Two isolates had an Ala90Val substitution, while one had a silent mutation at this codon. Seven isolates had double mutations, with the S95T change being common to all seven. These mutations were present in MDR isolates for which the MICs of the drugs were high, including the frontline drugs used in antituberculosis therapy. All strains were also checked for mutations in the *gyrB* locus, which is associated with low levels of resistance. However, we found no mutations in the *gyrB* loci of these isolates. It has been argued that the S95T mutation does not correlate with drug resistance (22). It therefore appears that the isolates have acquired resistance to ofloxacin via other mechanisms.

We tested 14 isolates resistant to streptomycin for mutations in the *rpsL* and *rrs* loci. In eight strains we found a novel silent mutation at amino acid position 121 in the *rpsL* locus, where the codon AAA (Lys) was changed to AAG (Lys), but we found no mutations in the *rrs* genes. To our knowledge, there are no reports of this mutation. The reported mutations at the *rpsL* locus are generally Leu43Arg, Leu43Thr, or Lys88Arg. We are still not clear about how a mutation at this locus leads to the development of streptomycin resistance. The remaining isolates probably acquired resistance by other means, such as by the development of a permeability barrier or by the production of drug-altering enzymes.

A point to be kept in mind is that the majority of isolates included in the present study were from North India. Our data are therefore inherently biased toward drug-resistant strains from this region and should not be seen as representative for isolates from the whole of India.

DISCUSSION

The mycobacterium uses various mechanisms to evade killing by drugs, including mutations in genes that code for drug target proteins (20), a complex cell wall which blocks drug entry, and membrane proteins that act as drug efflux pumps (6, 14). The objective of the present study was to identify muta-

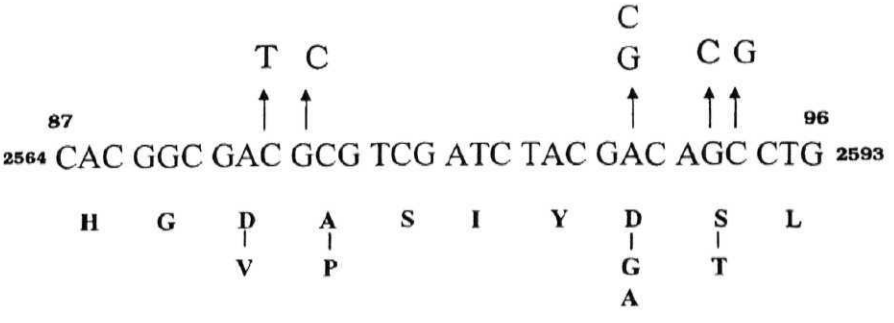


FIG. 3. Summary of missense mutations in the *gyrA* locus. Nucleotide changes are indicated on top of the wild-type sequence, and the corresponding amino acid changes are shown at the bottom. The most common mutation in this locus is Ser95Thr.

tions in drug target loci in Indian strains of *M. tuberculosis* and to identify the different drug resistance genotypes. As in all such studies, the aim was to generate information about the markers associated with drug resistance, polymorphisms in the drug target genes, the association of the level of resistance with particular mutations, etc. Our findings of mutations in the *rpoB*, *katG*, and *rpsL* loci are similar to those reported from other parts of the world, especially the common mutations, which reflect a global pattern (20). Rifampin resistance is often regarded as an excellent surrogate marker for MDR tuberculosis (4, 10), and our study corroborates this hypothesis. The mutation frequency of codon 531 (*rpoB*) was similar to that reported earlier (13, 17, 19, 20, 21, 25). Significantly, the frequency of mutations (relative to those of other mutations) was higher at codon 516 and lower at codon 526 in Indian isolates compared to those reported elsewhere. We found novel mutations that broaden the range of known mutations at this locus. When taken together, these mutations were detected in a significant number of drug-resistant isolates, a fact that needs to be considered when designing tools for the detection of MDR *M. tuberculosis*. We found a definite correlation between MICs and the type of mutation in many isolates. As reported by previous investigators (24), mutations at positions 528 and 531 are important in the development of high MICs. Our findings further strengthen the belief that the degree of resistance to rifampin exhibited by an isolate is related to the type of mutation in the *rpoB* locus.

In isoniazid-resistant isolates, significantly more deletion and insertion mutations than substitution mutations were found, of which a few have been reported previously (11, 19). We observed that almost all isolates studied carried the Arg463Leu substitution, which is also present in isolates that were sensitive to isoniazid. This is in concordance with a report from Sreevatsan et al. (22), who argue that polymorphism at this residue does not contribute to resistance per se but is an important marker for evolutionary genetics. The insertions and deletions in the *katG* locus invariably resulted in chain truncation and termination, leading to the generation of dysfunctional polypeptides. We found changes in the putative ribosomal binding site of the *inhA* gene in three isolates. While the exact mechanism of how these mutations confer resistance to isoniazid is not clear, reports (1, 18, 20) indicate that they probably increase the levels of enoyl-acyl carrier protein reductase which in turn leads to resistance via a drug titration mechanism. In isolates with no mutations in the hot-spot region of the gene, the complete sequencing of the gene is being done. However, resistance to isoniazid can also be due to mutations in the *ahpC-oxyR* and *kasA* gene loci (7, 16).

Fluoroquinolones comprise the secondary drug regimen in the treatment of tuberculosis. A large number of isolates were resistant to ofloxacin, which could be due in part to the inaccurate diagnosis of tuberculosis as a bacterial infection and fluoroquinolone overuse in the population. Codons 89, 90, 91, 94, and 95 in the *gyrA* gene have been shown to be polymorphic (20, 23, 26). The most common mutation in ofloxacin-resistant isolates in the present study was Ser95Thr, which reportedly has no direct role in the development of drug resistance, as it also occurs in drug-sensitive strains (22). It seems likely that ofloxacin resistance possibly results due to mutations elsewhere in the gene or the presence of drug efflux pumps.

Mutations in codons 43 and 88 of the *rpsL* gene generally result in high levels of resistance to streptomycin, while mutations in the loop at codon 530 or the region at codon 915 of the *rrs* locus are associated with low levels of resistance (3). We did not find any of these mutations in the 14 streptomycin-resistant isolates included in our study. However, we did observe a silent mutation at codon 121 that has not been reported by any other group.

Our study provides valuable data on the different kinds of mutations occurring at various target loci in Indian clinical isolates of *M. tuberculosis* that enhance our understanding of the molecular mechanisms of drug resistance. The diversity of the polymorphisms exhibited at these loci by the drug-resistant strains indicates the prevalence of a large numbers of drug-resistant strains in this region. Additionally, our data will also assist in the process of designing new molecular biology-based techniques for the diagnosis of MDR tuberculosis. Such methods promise faster detection rates compared to those achieved by methods based solely on culture of the isolates.

ACKNOWLEDGMENTS

This study was supported by research grants from the Department of Biotechnology, Government of India.

We thank Sunder S. Bisht and Mohammed Iliyas Ghazi, who helped with the generation of the sequencing data. We are also grateful to Ram Das, Kiran Srivastava, and V. Chauhan of the Central Jalma Institute of Leprosy, who provided us with the isolate DNAs.

REFERENCES

1. **Basso, L. A., R. Zheng, J. M. Musser, W. R. Jacobs, Jr., and J. S. Blanchard.** 1998. Mechanism of isoniazid resistance in *Mycobacterium tuberculosis*: enzymatic characterization of enoyl reductase mutants identified in isoniazid-resistant clinical isolates. J. Infect. Dis. **178**:769–775.
2. **Bloom, B. R., and J. L. Murray.** 1992. Tuberculosis: commentary on a reemergent killer. Science **257**:1055–1064.
3. **Bottger, E. C.** 1994. Resistance to drugs targeting protein synthesis in mycobacteria. Trends Microbiol. **2**:416–421.
4. **Centers for Disease Control and Prevention.** 1993. Initial therapy for tuberculosis in the era of multidrug resistance. Recommendations of the Advisory Council for the Elimination of Tuberculosis. Morb. Mortal. Wkly. Rep. **42**(RR-7).
5. **Cohn, D. L., F. Bustreo, and M. C. Raviglione.** 1997. Drug-resistant tuberculosis: review of the worldwide situation and the W.H.O./IULATD global surveillance project. Clin. Infect. Dis. **24**(Suppl. 1):S121–S130.
6. **Cole, S. T., R. Brosch, J. Parkhill, T. Garnier, C. Churcher, D. Harris, et al.** 1998. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. Nature **393**:537–544. (Erratum, **396**:190–198.)
7. **Collins, D. M., and T. M. Wilson.** 1996. *ahpC*, a gene involved in isoniazid resistance of the *Mycobacterium tuberculosis* complex. Mol. Microbiol. **19**: 1025–1034.
8. **Harkin, T. J., and H. W. Harris.** 1995. Treatment of multidrug resistant tuberculosis, p. 843–850. *In* N. W. Rom and S. Garay (ed.), Tuberculosis. Little, Brown & Company, Boston, Mass.
9. **Harris, K. A., Jr., U. Mukundan, J. M. Musser, B. N. Kreiswirth, and M. K. Lalitha.** 2000. Genetic diversity and evidence for acquired antimicrobial resistance in Mycobacterium tuberculosis at a large hospital in South India. Int. J. Infect. Dis. **4**:140–147.
10. **Hasnain, S. E., A. Amin, N. Siddiqi, M. Shamim, N. K. Jain, A. Rattan, V. M. Katoch, and S. K. Sharma.** 1998. Molecular genetics of multiple drug resistance (MDR) in *Mycobacterium tuberculosis*, p. 35–40. *In* R. L. Singhal and O. P. Sood (ed.), Drug resistance: mechanism and management. Proceedings of the Fourth Annual Ranbaxy Science Foundation Symposium. Ranbaxy Science Foundation, New Delhi, India.
11. **Heym, B., P. M. Alzari, N. Honore, and S. T. Cole.** 1994. Missense mutations in the catalase-peroxidase gene, *katG*, are associated with isoniazid resistance in *Mycobacterium tuberculosis*. Mol. Microbiol. **15**:235–245.
12. **Janmeja, A. K., and B. Raj.** 1998. Acquired drug resistance in tuberculosis in Haryana, India. J. Assoc. Physicians India **46**:194–198.
13. **Kapur, V., L.-L. Li, S. Iordanescu, M. C. Hamrick, A. Wanger, B. N. Kreiswirth, and J. M. Musser.** 1994. Characterization by automated DNA sequencing of mutations in the gene (*rpoB*) encoding the RNA polymerase β

subunit in rifampin-resistant *Mycobacterium tuberculosis* strains from New York City and Texas. J. Clin. Microbiol. **32**:1095–1098.
14. **Lewis, K.** 1994. Multidrug resistance pumps in bacteria: variations on a theme. Trends Biochem. Sci. **19**:119–123.
15. **Mathur, M. L., P. K. Khatri, and C. S. Base.** 2000. Drug resistance in tuberculosis patients in Jodhpur District. Indian J. Med. Sci. **54**:55–58.
16. **Mdluli, K., R. A. Slayden, Y. Zhu, S. Ramaswamy, X. Pan, D. Mead, D. D. Crane, J. M. Musser, and C. E. Barry III.** 1998. Inhibition of *Mycobacterium tuberculosis* β-ketoacyl ACP synthase by isoniazid. Science **280**:1607–1610.
17. **Miller, L. P., J. T. Crawford, and T. M. Shinnick.** 1994. The *rpoB* gene of *Mycobacterium tuberculosis*. Antimicrob. Agents Chemother. **38**:805–811.
18. **Morris, S., G. Han Bai, P. Suffys, L. P. Gomez, M. Fairchok, and D. Rouse.** 1995. Molecular mechanisms of multiple drug resistance in clinical isolates of *Mycobacterium tuberculosis*. J. Infect. Dis. **171**:954–960.
19. **Musser, J. M.** 1995. Antimicrobial agent resistance in mycobacteria: molecular genetic insights. Clin. Microbiol. Rev. 8:496–514.
20. **Ramaswamy, S., and J. M. Musser.** 1998. Molecular genetic bases of antimicrobial agent resistance in *Mycobacterium tuberculosis*: 1998 update. Tuber. Lung Dis. **79**:3–29.
21. **Siddiqi, N., M. Shamim, N. K. Jain, A. Rattan, A. Amin, V. M. Katoch, S. K. Sharma, and S. E. Hasnain.** 1998. Molecular genetic analysis of multidrug resistance in Indian isolates of *Mycobacterium tuberculosis*. Mem. Inst. Oswaldo Cruz **93**:589–594.
22. **Sreevatsan, S., X. Pan, K. E. Stockbauer, N. Connell, B. N. Kreiswirth, T. S. Whittam, and J. M. Musser.** 1997. Restricted structural gene polymorphism in the *Mycobacterium tuberculosis* complex indicates evolutionary recent global dissemination. Proc. Natl. Acad. Sci. USA **94**:9869–9874.
23. **Takiff, H. E., L. Salazar, C. Guerrero, W. Philipp, W. M. Huang, B. Kreiswirth, S. T. Cole, W. R. Jacobs, Jr., and A. Telenti.** 1994. Cloning and nucleotide sequence of the *Mycobacterium tuberculosis gyrA* and *gyrB* genes and characterization of quinolone resistance mutations. Antimicrob. Agents Chemother. **38**:773–780.
24. **Taniguchi, H., H. Aramaki, Y. Nikaido, Y. Mizuguchi, M. Nakamura, T. Koga, and S. Yoshida.** 1996. Rifampicin resistance and mutations of the *rpoB* gene in *Mycobacterium tuberculosis*. FEMS Microbiol. Lett. **144**:103–108.
25. **Telenti, A., P. Imboden, F. Marchesi, D. Lowrie, S. T. Cole, M. J. Colston, L. Matter, K. Schopfer, and T. Bodmer.** 1993. Detection of rifampicin-resistance mutations in *Mycobacterium tuberculosis*. Lancet **341**:647–650.
26. **Xu, C., B. N. Kreiswirth, S. Sreevatsan, J. M. Musser, and K. Drlica.** 1996. Fluoroquinolone resistance associated with specific gyrase mutations in clinical isolates of multidrug-resistant *Mycobacterium tuberculosis*. J. Infect. Dis. **174**:1127–1130.

Method for enhancing solubility of the expressed recombinant proteins in *Escherichia coli*

Sudip Ghosh[1,2], Sheeba Rasheedi[2], Sheikh Showkat Rahim[2], Sharmistha Banerjee[2], Rakesh Kumar Choudhary[2], Prachee Chakhaiyar[2], Nasreen Z. Ehtesham[1], Sangita Mukhopadhyay[2] and Seyed E. Hasnain[2,3]

[1]*National Institute of Nutrition (ICMR), Hyderabad,* [2]*CDFD, Nacharam, Hyderabad, and* [3]*Jawaharlal Nehru Centre for Advanced Scientific Research, Jakkur, Bangalore, India*

BioTechniques 37: _-_ (September 2004)

The production of correctly folded protein in Escherichia coli *is often challenging because of aggregation of the overexpressed protein into inclusion bodies. Although a number of general and protein-specific techniques are available, their effectiveness varies widely. We report a novel method for enhancing the solubility of overexpressed proteins. Presence of a dipeptide, glycylglycine, in the range of 100 mM to 1 M in the medium was found to significantly enhance the solubility (up to 170-fold) of the expressed proteins. The method has been validated using mycobacterial proteins, resulting in improved solubilization, which were otherwise difficult to express as soluble proteins in* E. coli. *This method can also be used to enhance the solubility of other heterologous recombinant proteins expressed in a bacterial system.*

INTRODUCTION

Escherichia coli is the most widely used system for the rapid and economical production of recombinant proteins because of its very well-characterized genetics and rapid growth rate in inexpensive culture media. One major disadvantage of *E. coli* is that heterologous proteins are often expressed as insoluble aggregates of folding intermediates known as inclusion bodies. Expression in soluble fraction is paramount for the expressed protein to be biologically active. In order to recover soluble proteins from the inclusion bodies, the inclusion bodies are solubilized in the presence of strong denaturants such as urea or guanidinium hydrochloride, followed by the removal of the denaturants under optimal conditions that favor refolding. Although considerable progress has been made for efficient refolding of proteins (1), specific folding conditions differ greatly from protein to protein. Even under optimal conditions of refolding, quite a large number of proteins are found to be recalcitrant to refolding, and the yield of renatured protein is relatively low.

Several general and protein-specific methods are available for increased solubility of expressed proteins in *E. coli*. One approach is the coexpression of molecular chaperones, which assists in the correct folding of the heterologous protein (2,3). Similarly, concomitant overexpression of thioredoxin (TrxA) is known to improve the solubility of the expressed proteins (4). Another approach that has gained considerable success in recent years is the use of gene fusion (5). Fusion partners such as glutathione-S-transferase (GST) and maltose binding protein (MBP) are known to impart solubility of many heterologous proteins in addition to serving as a tag for affinity purification. Sometimes, soluble expression can also be enhanced by supplying essential cofactors necessary for the activity of the protein in question. For example, soluble expression of human cystathionine β-synthase, a heme-containing protein, could be increased over 8-fold by the addition of the heme precursor δ-aminolevulinic acid (δ-ALA) to the culture medium (6). In some instances, coexpression of nuclear receptor partners is also found to increase the solubility of nuclear receptors expressed in *E. coli* (7). In yet other cases, specific substitution of some amino acid residues was found to enhance the solubility of the expressed proteins (8). In this report, we describe a novel method of enhancing the solubility of expressed proteins by inducing recombinant protein expression in the presence of the dipeptide glycylglycine. We deliberately chose as an example mycobacterial proteins that are known to be difficult to express as soluble proteins in *E. coli*. The solubilization of these proteins was enhanced up to 170-fold.

MATERIALS AND METHODS

Preparation of Recombinant Constructs

The open reading frames (ORFs) Rv0256c, Rv2430c (both are the members of the PPE gene family), Rv3339c (isocitrate dehydrogenase-1), and Rv1609 (anthranilate synthase) of *Mycobacterium tuberculosis* were amplified from the genomic DNA of the strain H37Rv by PCR. Oligodeoxyribonucleotide primers were chemically synthesized (Microsynth GmbH, Balgach, Switzerland), with appropriate restriction sites suitable for in-frame cloning into expression vectors, with N-terminal 6x-histidine tags. The primers used for amplification are shown in Table 1. The recombinant proteins were expressed as N-terminal His-tagged fusions. The positive clones were confirmed by DNA sequencing.

Cell Culture

Competent BL21(DE3)pLys cells (Novagen, Abingdon, UK) were transformed with pRSET256, pRSET1609, and pRSET3339 plasmid DNA, and the colonies were grown overnight on Luria Bertani (LB) plates (9) containing 100 µg/mL ampicillin. For pQE2430, competent M15(pREP4) cells (Qiagen GmBH, Hilden, Germany) were used and grown in LB plates containing ampicillin (100 µg/mL) and kanamycin (50 µg/mL). Fresh colonies were first inoculated into 5 mL LB media containing appropriate antibiotics and grown overnight at 37°C with shaking. These

Table 1. Sequence of the Primers, Restriction Sites, and Vectors Used for Expressing Different Mycobacterial Proteins

Gene	Primers with Restriction Enzyme Sites	Restriction Enzyme Sites	Cloning Vector
Rv0256c	5′-CGAGATCTATGACCGCCCCGATCTGGAT-3′	*Bgl*II	pRSET B
	5′-GCGAATTCTCACTCCACCCGGGTCG CTGA-3′	*Eco*RI	
Rv2430c	5′-GGATCCATGCATTTCGAAGCGTAC-3′	*Bam*HI	pQE30
	5′-AAGCTTCTAAGTGTCTGTACGCCATGA-3′	*Hind*III	
Rv1609	5′-AATCTCGAGGTGTCCGAGCTCAGCGT-3′	*Xho*I	pRSET A
	5′-AATCCATGGCTGGCGTGCAACCAGATAA-3′	*Nco*I	
Rv3339c	5′-AGGATCCATGTC CAACGCACCCAAGATA-3′	*Bam*HI	pRSET A
	5′-TAAGCTTCTAATTGGCCAGCTCCTTTTC-3′	*Hind*III	

Restriction enzyme sites have been underlined.

overnight cultures were diluted 10-fold into 10 mL Terrific Broth (TB) medium (9), containing different concentrations (0, 50, 200, 500, and 1000 mM) of glycylglycine (Amersham Biosciences, Buckinghamshire, UK) and further grown at 37°C in an orbitory shaker till the absorbance $(A)_{600} = 1$. The cultures were cooled to room temperature, and protein expression was induced with 0.5 mM isopropyl-β-D-thiogalactoside (IPTG) for 14–16 h at 27°C.

Preparation of Cell Lysate and Protein Solubility Analysis

The induced cells were centrifuged at 10,000× *g*, and the cell pellet was resuspended in extraction buffer (50 mM phosphate buffer, pH 8.0, 300 mM NaCl) with lysozyme to a final concentration of 1 mg/mL and incubated on ice for 30 min, followed by sonication 5 times with a burst duration of 15 s each. The sonicated lysates were centrifuged at 13,000 rpm for 10 min, and the supernatants containing the soluble proteins were collected into fresh tubes. The concentration of total protein in supernatant was estimated using Bio-Rad DC Protein Assay Kit (Bio-Rad Laboratories, Hertfordshire, UK). About 60 µg of protein from each supernatant were electrophoresed in a 12% sodium dodecyl sulfate (SDS)-polyacrylamide gel (10) and transblotted onto a nitrocellulose membrane (Hybond-ECL; Amersham Biosciences) at 300 mA for 2 h at 4°C in a transfer buffer (25 mM Tris-HCl, 190 mM glycine, 20% methanol). The membrane was blocked for 1 h in phosphate-buffered saline (PBS; Reference 9) containing 4% non-fat dry milk and then incubated with 1:200 diluted monoclonal anti-his tag antibodies (Santa Cruz Biotechnologies) in PBS. The membrane was washed thrice with excess PBST (PBS containing 0.1% Tween® 20) for 15 min each. Goat anti-mouse immunoglobulin G (IgG)-horseradish peroxidase (HRP) conjugate (Amersham Biosciences) was used at 1:10,000 dilution as the second antibody. The membrane was again washed thrice for 5 min each with PBST. The reactive bands were developed by chemiluminescence with luminol reagents (Santa Cruz Biotechnologies). All the experiments were performed at least three times, and the representative blots are presented.

Densitometric Analyses

Densitometric analyses were performed using National Institutes of Health (NIH)-Image software, available in the public domain, developed by Wayne Rasband for Macintosh® computers.

Biochemical Activity Assays

The isocitrate dehydrogenase-1 activity of the purified Rv3339c solubilized in the presence of 500 mM glycylglycine in the medium was measured spectrophotometrically by monitoring the time-dependent reduction of $NADP^+$ to NADPH at 25°C at 340 nm. The standard assay solution contained 20 mM triethanolamine chloride buffer, pH 7.5, 2 mM $NADP^+$, 0.03 mM DL-isocitrate, 10 mM $MgCl_2$, 100 mM NaCl, and 100 pmol of the enzyme in a total reaction volume of 400 µL.

RESULTS AND DISCUSSION

All the ORFs, Rv256c, and Rv2430c belonging to the PPE family of proteins and Rv1609 (anthranilate synthase) and Rv3339c (isocitrate dehydogenase 1) of *M. tuberculosis* were found to have enhanced solubility in the presence of the dipeptide glycylglycine in TB media. Because these proteins were expressed with an N-terminal histidine tag, the relative amount of protein going into the soluble fractions was determined by subjecting equal amounts of soluble proteins to SDS-polyacrylamide gel electrophoresis (PAGE) and subsequent detection by Western blot analysis using monoclonal anti-his tag antibodies. In the case of Rv256c, the amount of soluble protein was very low when grown in the absence of glycylglycine (Figure 1A). However, with the increasing concentrations of glycylglycine, the amount of soluble protein dramatically increased, with 1 M glycylglycine being the most effective. In the presence of 1 M glycylglycine, there was a more than 170-fold increase in solubility. Use of glycylglycine higher than 1 M was found to affect the growth of the bacteria (data not shown), Therefore, concentrations greater than 1 M were not used.

For the other proteins, Rv2430c, Rv1609 (anthranilate synthase), Rv3339c (isocitrate dehydrogenase-1), our initial attempts to express these proteins in soluble form employing various other strategies such as low temperature induction, induction at low IPTG concentration, etc., proved to be futile. However, when we induced these proteins in the presence of glycylglycine, soluble proteins were readily detected as compared to the undetectable level in experiments without glycylglycine (Figure 2, A–C). For Rv3339c, the maximum soluble yield was in the presence of 500 mM glycylglycine (Figure 2A), whereas for Rv1609, the maximum soluble expression was achieved in presence of 1 M glycylglycine (Figure 2B). Similarly, in the case of Rv2430c, the soluble expression was

Table 2. Relative Levels of Soluble Proteins Expressed in the Presence of Glycylglycine for Different Mycobacterial Proteins

Protein	Solubility in Glycylglycine (mM)						*Escherichia coli* Strain Used
	0	50	100	200	500	1000	
Rv0256C	-/+	+	+	+	+++	++++++++	BI21(DE3)pLys
Rv2430c	-/+	-/+	-/+	++	++	-/+	M15(pREP4)
RV1609	-	-	-	-	+	++	BI21(DE3)pLys
Rv3339c	-	-	-	+	++	+	BI21(DE3)pLys

- = not detectable; + = detectable

maximum in the presence of 200 mM glycylglycine (Figure 2C). The results are summarized in Table 2.

We next tested whether or not the recombinant proteins solubilized in the presence of glycylglycine possessed biological activity. For this, we purified the soluble protein coded by the Rv3339c ORF, encoding isocitrate dehydrogenase-1, solubilized in the presence of 500 mM glycylglycine in the medium, and assayed its ability to reduce $NADP^+$ to NADPH. The purified solubilized protein was found to possess enzymatic activity following a typical Michaelis-Menten reaction kinetics (Figure 3) and was stable over 1 week at 4°C.

We provide a simple and novel way of enhancing solubility of difficult proteins that are otherwise expressed in a nonproductive fashion into inclusion bodies. Inclusion body formation may be a consequence of the rate of protein translation exceeding the capacity of the cell to fold the newly synthesized protein correctly (11). This phenomenon has been defined as secretory load for the baculovirus insect cell system (12) and has been addressed by reducing the transcription level using a weaker promoter or by allowing more time for the insect cells to process the recombinant protein (13). Therefore, decreasing the rate of protein production is one of the major strategies to overcome this problem. Some general approaches to achieve this have been to induce the cells at lower temperature (14), use low IPTG concentration (15), use weak promoters (16), etc. Yet other methods utilize various "compatible solutes" to induce osmotic stress (17,18). Improved solubility has also been reported by the use of a specific host strain in which heterologous proteins can be expressed upon osmotic induction with high salt (19). Although the method was initially developed to enhance the soluble expression of some of the mycobacterial proteins, this method is also applicable to other proteins refractory to soluble expression in *E. coli* system

The mechanism of glycylglycine-mediated enhanced solubilization remains to be understood. One of the possibilities could be the increased osmolarity of the medium by the dipeptide. Osmotic stress induces the expression of heat shock proteins with chaperone-like activity to assist correct folding. Another possibility is the direct interaction of glycylglycine with the expressed protein by acting as a chemical chaperone (20,21). *E. coli* is known to possess specific transporters for dipetides and oligopeptides. These in turn are of particular advantage to the bacteria, which thrives in the peptide-rich gut lumen environment (22). Glycylglycine transport behaves similar to other shock-sensitive transport systems requiring ATP for its transport (23). In the presence of higher concentrations of glycylglycine in the media, the bacteria probably

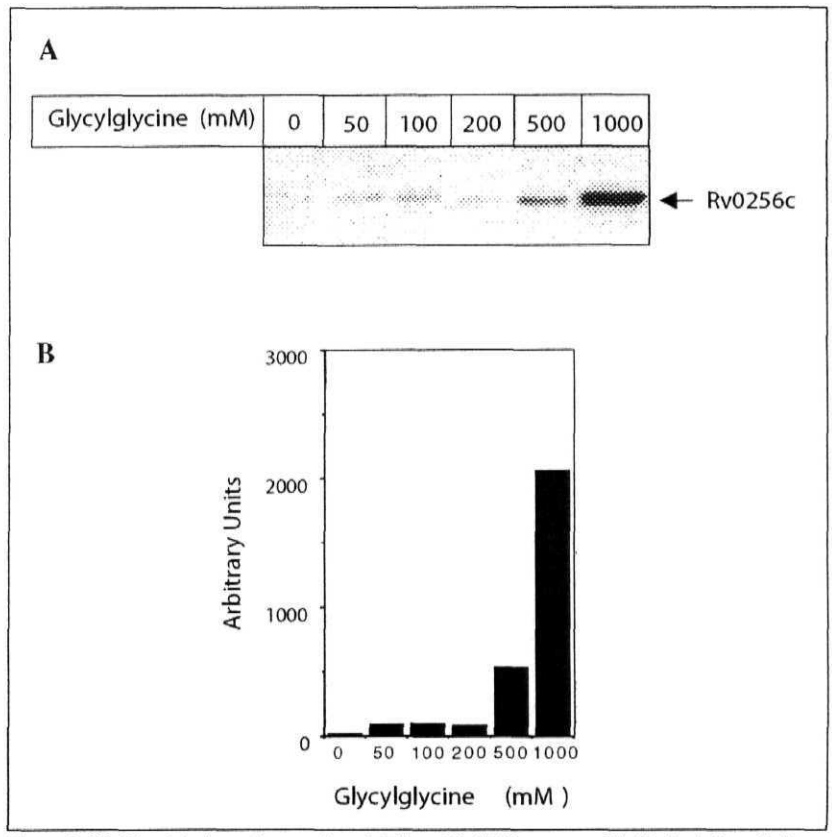


Figure 1. (A) Glycylglycine enhances the solubility of Rv0256c expressed in *Escherichia coli*. About 60 µg of soluble protein from the sonicated extract were loaded for each lane and transferred onto the nitrocellulose membrane. Upon Western transfer, the blot was probed with monoclonal anti-6xhistidine antibodies and developed by ECL reagents. The concentrations of glycylglycine used are indicated at the top of each lane. (B) Densitometric analyses of the same Western blot, in which the density of the individual bands were plotted against the glycylglycine concentrations used in the culture medium.

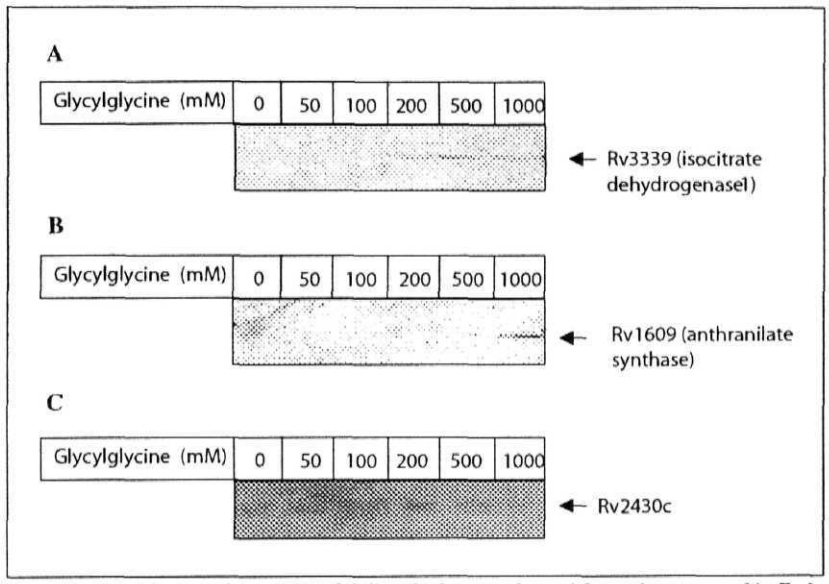


Figure 2. Glycylglycine enhances the solubility of other mycobacterial proteins expressed in *Escherichia coli*. About 60 µg of soluble protein from the sonicated extract were loaded in each lane and transferred onto the nitrocellulose membrane. The membrane was probed with monoclonal anti-6xhistidine antibodies and developed by ECL reagents. The concentrations of glycylglycine used are indicated at the top of each lane. (A) Western blot of Rv3339. (B). Western blot of Rv1609. (C). Western blot of Rv2430c.

ends up spending considerable energy in active glycylglycine transport, thus slowing down the overall metabolic rate including the rate of translation. This probably allows more time for the expressed proteins to be folded correctly. However, it will be interesting to study how a dipeptide can actually help proteins to be folded in its native condition.

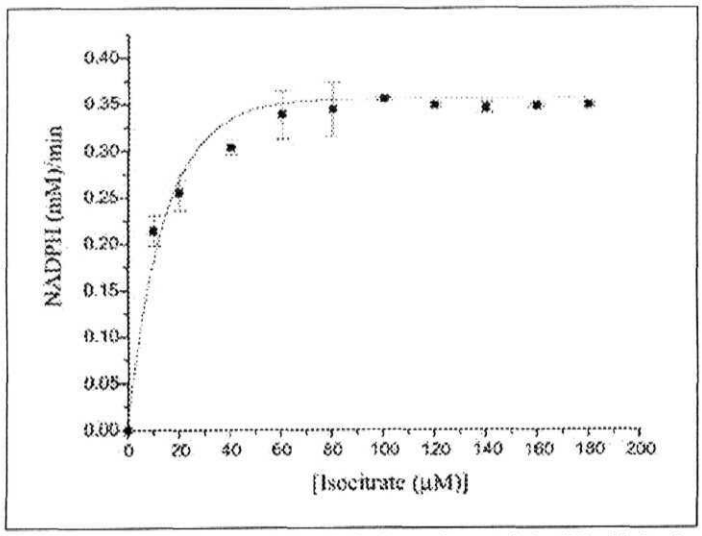


Figure 3. The recombinant protein purified using media containing 500 mM glycylglycine and was biochemically active. *Mycobacterium tuberculosis* isocitrate dehydrogenase-1 activity was measured spectrophotometrically by monitoring the time-dependent reduction of $NADP^+$ to NADPH at 25°C at 340 nm. The standard assay solution per 400 µL contained 20 mM triethanolamine chloride buffer, pH 7.5, 2 mM $NADP^+$, 0.03 mM DL-isocitrate, 10 mM $MgCl_2$, 100 mM NaCl, and 100 pmol of the enzyme.

ACKNOWLEDGMENTS

This work was supported by core support to CDFD from the Department of Biotechnology, Government of India. S.G. was supported by a Post-Doctoral Fellowship from the Department of Biotechnology, while S.R., S.B., R.K.C., P.C., and S.S.R. were the recipients of Research Fellowships from the Council of Scientific and Industrial Research, New Delhi, India.

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing interests.

REFERENCES

1. **Lilie, H., E. Schwarz, and R. Rudolph.** 1998. Advances in refolding of proteins produced in *E. coli*. Curr. Opin. Biotechnol. 9:497-501.
2. **Amrein, K.E., B. Takacs, M. Stieger, J. Molnos, N.A. Flint, and P. Burn.** 1995. Purification and characterization of recombinant human p50csk protein-tyrosine kinase from an Escherichia coli expression system overproducing the bacterial chaperones GroES and GroEL. Proc. Natl. Acad. Sci. USA *92*:1048-1052.
3. **Goenka, S. and C.M. Rao.** 2001. Expression of recombinant zeta-crystallin in Escherichia coli with the help of GroEL/ES and its purification. Protein Expr. Purif. *21*:260-267.
4. **Yasukawa, T., C. Kanei-Ishii, T. Maekawa, J. Fujimoto, T. Yamamoto, and S. Ishii.** 1995. Increase of solubility of foreign proteins in *Escherichia coli* by coproduction of the bacterial thioredoxin. J. Biol. Chem. *270*:25328-25331.
5. **LaVallie, E.R. and J.M. McCoy.** 1995. Gene fusion expression systems in *Escherichia coli*. Curr. Opin. Biotechnol. *6*:501-506.
6. **Kery, V., D. Elleder, and J.P. Kraus.** 1995. Delta-aminolevulinate increases heme saturation and yield of human cystathionine beta-synthase expressed in Escherichia coli. Arch. Biochem. Biophys. *316*:24-29.
7. **Li, C., J.W. Schwabe, E. Banayo, and R.M. Evans.** 1997. Coexpression of nuclear receptor partners increases their solubility and biological activities. Proc. Natl. Acad. Sci. USA *94*:2278-2283.
8. **Du Bois, G.C., S.P. Song, I. Kulikovskaya, J.L. Rothstein, M.W. Germann, and C.M. Croce.** 2000. Purification and characterization of recombinant forms of murine Tcl1 proteins. Protein Expr. Purif. *18*:277-285.
9. **Sambrook, J., E.F. Fritsch, and T. Maniatis.** 1989. Molecular Cloning: A laboratory Manual. CSH Laboratory Press, New York.
10. **Laemmli, U.K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature *227*:680-685.

11. **Kiefhaber, T., R. Rudolph, H.H. Kohler, and J. Buchner.** 1991. Protein aggregation *in vitro* and *in vivo*: a quantitative model of the kinetic competition between folding and aggregation. Biotechnology (NY) *9*:825-829.
12. **Sridhar, P. and S.E. Hasnain.** 1993. Differential secretion and glycosylation of recombinant human chorionic gonadotropin (beta hCG) synthesized using different promoters in the baculovirus expression vector system. Gene *131*:261-264.
13. **Sridhar, P., A.K. Panda, R. Pal, G.P. Talwar, and S.E. Hasnain.** 1993. Temporal nature of the promoter and not relative strength determines the expression of an extensively processed protein in a baculovirus system. FEBS Lett. *315*:282-286.
14. **Schein, C.H. and M.H.M. Noteborn.** 1988. Formation of soluble recombinant proteins in *Escherichia coli* is favored by low growth temperature. BioTechnology *6*:291-294.
15. **Sawyer, J.R., J. Schlom, and S.V. Kashmiri.** 1994. The effects of induction conditions on production of a soluble anti-tumor sFv in *Escherichia coli*. Protein Eng. *7*:1401-1406.
16. **Swartz, J.R.** 2001. Advances in *Escerichia coli* production of therapeutic proteins. Curr. Opin. Biotechnol. *12*:195-201.
17. **Barth, S., Huhn, B. Matthey, A. Klimka, E.A. Galinski, and A. Engert.** 2000. Compatible-solute-supported periplasmic expression of functional recombinant proteins under stress conditions. Appl. Environ. Microbiol. *66*:1572-1579.
18. **Blackwell, J.R. and R.A. Horgan.** 1991. Novel strategy for production of a highly expressed recombinant protein in an active form. FEBS Lett. *295*:10-12.
19. **Bhandari, P. and J. Gowrishankar.** 1997. An *Escherichia coli* host strain useful for efficient overproduction of cloned gene products with NaCl as the inducer. J. Bacteriol. *179*:4403-4406.
20. **Gekko, K. and S. Koga.** 1983. Increased thermal stability of collagen in the presence of sugars and polyols. J. Biochem. (Tokyo) *94*:199-205.
21. **Ou, W.B., Y.D. Park, and H.M. Zhou.** 2002. Effect of osmolytes as folding aids on creatine kinase refolding pathway. Int. J. Biochem. Cell Biol. *34*:136-147.
22. **Sussman, A.J. and C. Gilvarg.** 1971. Peptide transport and metabolism in bacteria. Annu. Rev. Biochem. *40*:397-408.
23. **Cowell, J.L.** 1974. Energetics of glycylglycine transport in *Escherichia coli*. J. Bacteriol. *120*:139-146.

Received 23 October 2003; accepted 8 June 2004.

Address correspondence to Seyed E. Hasnain, Centre for DNA Fingerprinting and Diagnostics, Nacharam, Hyderabad, 5000076, India. e-mail: ehtesham@www.cdfd.org.in