CLONING AND CHARACTERIZATION OF A PPE ANTIGEN OF *Mycobacterium tuberculosis*

Thesis submitted for the Degree of

Doctor of Philosophy
By

RAKESH KUMAR CHOUDHARY

CENTRE FOR DNA FINGERPRINTING AND DIAGNOSTICS
HYDERABAD

Department of Biochemistry
School of Life Sciences
University of Hyderabad
Hyderabad
March 2004
Registration no. 99LBPH04

CERTIFICATE

This is to certify that this thesis entitled "Cloning and Characterization of a PPE antigen of *Mycobacterium tuberculosis*" comprises the work done by Rakesh Kumar Choudhary under my guidance at the Centre for DNA Fingerprinting and Diagnostics. This work is original and has not been submitted in part or full for any degree or diploma of any University.

28/3/04

Dr A S Raghavendra
Dean, School of Life Sciences
University of Hyderabad

Seyed E Hasnain
Supervisor

Head, Department of Biochemistry
University of Hyderabad

DECLARATION

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

Seyed E Hasnain

Supervisor

Rakesh Kumar Choudhary

Rakesh Kumar Choudhary

Candidate

CONTENTS

Acknowledgements

I am greatly indebted to Dr. Seyed E. Hasnain, my research supervisor, for his constant support and encouragement throughout the course of this study, and for his empathy in many difficult situations.

I take this opportunity to thank Dr. Nasreen Z Ehtesham for her guidance and allowing me to utilize her laboratory facilities at the NIN.

My heartfelt thanks to Dr. Raghu Pullakhandham for his help in carrying out the biophysical experiments.

Thanks are due to Drs. Sangita Mukhopadhyay and KJR Murthy for their help in the serological experiments.

I will deeply cherish the memory of working at EGE, NII, and LMCB, CDFD. I was fortunate to have interacted with a wonderful set of lab colleagues, past and present, making my graduate experience not just intellectually rewarding, but also personally enriching.

I express my heartfelt thanks to all my friends at CDFD who have helped me all along in some way or the other towards the completion of this work.

Financial support from CSIR is acknowledged.

Finally I wish to thank my family members for providing me with the encouragement to complete this thesis and being patient and supportive through out these years. They were a constant source of love, inspiration and moral support on which I banked quite often.

Thank you for believing in me.

Rakesh Kumar Choudhary

ABBREVIATIONS

A	Angstrom
AIDS	Acquired immunodeficiency syndrome
BCG	Bacillus of Calmette-Guérin
Bp	Base pair(s)
CD	Circular Dichroism
CTAB	Cetyl trimethyl ammonium bromide
dATP	2'-deoxyadenosine-5'-triphosphate
dCTP	2'-deoxycytidine-5'-triphosphate
ddNTP	Dideoxynucleotide triphosphate
dGTP	2'-deoxyguanosine-5'-triphosphate
DNA	Deoxyribosenucleic acid
dNTP	Deoxynucleotide triphosphate
dTTP	2'-deoxythymidine-5'-triphosphate
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid (disodium salt)
ELISA	Enzyme linked Immunosorbent assay
gm	Gram(s)
HIV	Human immunodeficiency virus
Hsp	Heat shock protein
IPTG	Isopropyl-β-D-thiogalactopyranoside
IgA	Immunoglobulin A
IgE	Immunoglobulin E
IgG	Immunoglobulin G
IgM	Immunoglobulin M
IL	Interleukin
kb	Kilobase pair(s)
kDa	Kilodalton(s)
LJ	Lowenstein Jensen
M	Molar
M. bovis	*Mycobacterium bovis*
M.tb.	*Mycobacterium tuberculosis*
MDR TB	Multidrug resistant tuberculosis
mg	Milli gram (10^{-3} gram)
min	Minutes
ml	Millilitres (10^{-3} litres)
mM	Millimolar
MPTR	Major polymorphic tandem repeats
$MgCl_2$	Magnesium Chloride
$MgSO_4$	Magnesium Sulpate
NaCl	Sodium Chloride
OD	Optical density
PBS	Phosphate Buffered Saline

PCR	Polymerase chain reaction
PE	Proline - Glutamic acid
PPE	Proline - Proline - Glutamic acid
PGRS	Polymorphic GC-rich sequences
PPD	Purified Protein Derivative
RNA	Ribose nucleic acid
rpm	Revolutions per minute
RT-PCR	Reverse transcriptase polymerase chain reaction
SDS	Sodium dodecyl sulphate
Taq	*Thermus aquaticus*
TAE	TrisCl Acetic acid EDTA buffer
TBE	TrisCl Boric acid EDTA buffer
TE	TrisCl EDTA
TNT	Transcription and Translation
TST	Tuberculin Skin testing
TB	Tuberculosis
TNFα	Tumour Necrosis Factor alpha
Tris	Tris-(hydroxymethyl) aminomethane
TrisCl	Tris-(hydroxymethyl)aminomethane hydrochloride
U	Units
WHO	World Health Organization
µg	Micro gram (10^{-6} grams)

CHAPTER 1

INTRODUCTION

INTRODUCTION

Tuberculosis is an infectious disease caused by a bacillus – *Mycobacterium tuberculosis*, that spreads through the air. Left untreated, a single person with active TB can infect 10 to 15 others each year, creating a self-perpetuating pool of infection. Someone in the world is infected with TB every second, seven to eight million become sick and two million die from it every year. TB was first described by Hippocrates as Phthisis, meaning wasting away of the body, which remains a primary symptom of untreated TB and carries a lot of economic significance even today. Tuberculosis has existed in India since the earliest days. In 1500 BC, the Rig Veda described the illness as Rajayaksma, King of diseases. There is evidence that the disease existed in the ancient world as seen in 3000-year old mummies which show classical bone anomalies associated with TB (Salo *et al*, 1994). The disease assumed importance in feudal Europe, America, Asia, Africa and other pacific islands causing a great deal of mortality. The high mortality statistics of TB make it one of the most lethal infectious disease known to humans.

About one third of the world's population is infected by *Mycobacterium tuberculosis* (Bloom and Murray, 1992). The morbidity and mortality statistics of TB make it one of the most infectious

disease known to man. Global prevalence of MTB infection was 32% (1.86 billion people). Eighty percent of all incident TB cases were found in 22 countries, with more than half the cases occurring in 5 Southeast Asian countries. Nine of 10 countries with the highest incidence rates per capita were in Africa. The global burden of tuberculosis remains enormous, mainly because of poor control in Southeast Asia, Sub-Saharan Africa, and eastern Europe. Majority of TB cases in developing countries are in the economically productive age group (15-50 years). The main reasons for the increasing global TB burden are: 1) poverty and the widening gap between rich and poor in various populations; 2) neglect arising out of inadequate case detection, diagnosis and cure; 3) increase in population migration; and 4) HIV pandemic.

Historically the causative organism is supposed to have evolved into its human host *via* domesticated animals when humans started settling down and an agrarian order came into effect (Bates and Stead, 1993). Mycobacterium probably emerged from the soil to find a niche, first infesting various mammals and birds. Based on genetic relatedness *Mycobacterium bovis*, a common animal pathogen, and *Mycobacterium tuberculosi*s share a common ancestor. It is reasonable to infer that at some period of time *Mycobacterium bovis* underwent subtle changes to adapt to a parasitic role in humans

(Brosch *et al,* 2002). The tubercle bacilli scores over *Mycobacterium bovis* in that it has a natural reservoir in humans, diminished virulence potential in animals and exhibits airborne transmission. Although the two have different animal pathogenicity, and can easily be distinguished by biochemical or physiologic tests, both fall within the tuberculosis complex, a group of closely related types including *Mycobacterium microti* and *Mycobacterium africanum*. The extreme degree of DNA homology and immunologic properties suggest that *Mycobacterium tuberculosis* and *Mycobacterium bovis* should be regarded as subspecies rather than distinct species. This genetic concordance has prompted the suggestion that *Mycobacterium tuberculosis* evolved in the human organism from *Mycobacterium bovis* sometime after the domestication of cattle (Brosch *et al,* 2002). Evolution inherently proceeds in the direction of greater specialization. Because *Mycobacterium bovis* infects a broader variety of animals than does *Mycobacterium tuberculosis*, the latter is considered to be more evolved. Furthermore, the dynamic aspects of the coevolution of host and pathogen make it highly unlikely that mutations tying *Mycobacterium tuberculosis* so tightly to humans could have evolved in a nonhuman species. Although *Mycobacterium tuberculosis* and *Mycobacterium bovis* are equally capable of producing progressive tuberculosis in humans, the routes of infection and associated clinical patterns of disease are different.

1.1. Drug Resistance in *Mycobacterium tuberculosis*

The mycobacterial strains which are resistant to at least two of the frontline drugs rifampicin and isoniazid are considered to be multi-drug resistant (MDR). Rifampicin and isoniazid form the backbone of the short course therapy for treating tuberculosis. These drugs are the most effective, and resistance to them would necessitate using drugs which are more toxic coupled with a longer treatment period (Bloom and Murray, 1992).

The acquisition of resistance by the bacterium is a random event, and the chance of acquiring multidrug resistance is 10^{-14} (Harkin and Harris, 1995). These strains have been demonstrated all over the world. The frequency of resistance to multiple drugs varies geographically and acquired (secondary) resistance is more common than primary resistance. Delineation of the molecular mechanisms of antimicrobial agent resistance has resulted in the development of several strategies for early detection and treatment of drug resistant strains.

Resistance to drugs primarily arises due to mutations in genes coding for drug target proteins (Ramaswamy and Musser, 1998, Siddiqi *et al*, 1998, Siddiqi *et al*, 2002). However, this is not the

exclusive mechanism employed by the pathogen to evade killing by the drug. All mycobacteria, especially *Mycobacterium tuberculosis*, have a complex cell wall which blocks drug entry into the cell and thus prevents the drug from accessing it's cytoplasmic targets. A more recent mechanism which is well documented in other systems is the presence of membrane proteins which act as drug efflux pumps (Lewis, 1994, Siddiqi *et al*, 2004). The *Mycobacterium tuberculosis* genome has nearly 20 such pumps (Cole *et al*, 1998). However, none of these have been conclusively shown to operate in a clinical setting to give rise to drug resistance. Another mechanism for acquiring resistance that has been documented in other systems, but not yet in *Mycobacterium tuberculosis*, is drug alteration. Superimposed on the problem of MDR TB is the non compliance of therapy with some patients. Most recently, the problem of Multi Drug resistant tuberculosis has been aggravated by the deadly synergy of TB with the Human Immunodeficiency Virus.

1.2. Tuberculosis and the Human Immunodeficiency Virus Infection

Infection with the human immunodeficiency virus (HIV) has transformed tuberculosis from an endemic disease into a world wide epidemic. During the past decade TB has graduated into becoming the major opportunistic infection compounding the HIV epidemic

worldwide, although other infections remain more common in certain geographic areas. In 1994 the Global Programme on AIDS of the World Health Organization estimated that the prevalence of HIV infection among adults worldwide was between 13 and 14 million (World Health Organization, Global Programme on AIDS, 1994). At the same time about 1.75 billion people were infected with *Mycobacterium tuberculosis*. Approximately 4 to 5.6 million people worldwide are coinfected. The converging epidemics of AIDS and TB have been complicated by the emergence of multidrug-resistant tuberculosis. That the TB bacillus can evolve differently based on the local environment and geographic pressures has given birth to the concept of geographic genomics (Majeed *et al*, 2004).

$CD4^+$ T lymphocytes and macrophages play a central role in immune response to mycobacterial infection. HIV infection results in progressive depletion and dysfunction of $CD4^+$ cells with defects in macrophage and monocyte function. As a result, HIV-infected patients have an increased risk of reactivation of latent TB as well as getting the disease from new infection. Macrophage based resistance against TB is due to its ability to control intracellular growth. $CD4^+$ lymphocytes prime macrophages to control intracellular growth by releasing cytokines such as interferon-ϒ interleukin-1 and interleukin-2, and tumour necrosis factor-alpha which activate blood

borne macrophages. An additional protective mechanism may be the generation of cytolytic T cells in response to mycobacteria. $CD4^+$ T cells and ϒ/δ T cells kill autologous infected macrophages (Kaufman, 1988; Kumararante *et al*, 1990; Munk *et al*, 1990). Forte and colleagues demonstrated decreased cytolytic T-cell activity in HIV-infected patients (Forte *et al*, 1992). The continued loss of $CD4^+$ cells in HIV-seropositive patients and their reduced production of IL-2 may be the cause of decreased cytolytic activity.

TB can accelerate the course of HIV infection. Infection with *Mycobacterium tuberculosis* stimulates release of mononuclear cytokines such as IL-1, IL-2, and TNF∝. These cytokines enhance replication of HIV *in vitro*. Taking into account the rapid emergence of multidrug resistant strains of TB coupled with the deadly synergy of TB with the Human Immunodeficiency virus, the onus for the treatment lies in the identification of novel immunodominant antigens which can be used as vaccine candidates.

1.3. Humoral response in Tuberculosis

The humoral responses to infection by *Mycobacterium tuberculosis* have been under investigation for several decades, primarily for the

purpose of devising specific and sensitive serodiagnoses for tuberculosis and for evaluating their use as a candidate vaccine. The task is complicated because the mycobacterial cell wall is made up of numerous proteins, carbohydrates, and lipids that exist independently; as well as closely associated lipopolysaccharides, glycolipids, peptidoglycolipids, glycoproteins, and so on. This leads to the existence of an enormous number of antigens and epitopes to which the immune system may respond. The first generation of investigations were performed by using nonpurified or crudely purified preparations of antigens, such as suspensions of *Mycobacterium tuberculosis*, old tuberculin, purified protein derivative (PPD), polysaccharide extracts, tri-chloroacetic acid-precipitated proteins from culture filtrates of *Mycobacterium tuberculosis* grown in various bacteriologic media, and concentrated culture filtrates. The immunologic methods used to evaluate the presence of antibodies to these antigen preparations included agglutination of bacterial suspensions, agglutination of erythrocytes sensitized with different bacterial antigens, complement fixation, agglutination of antibody-coated inert particles, radioimmunoassay, and gel diffusion. In general these studies revealed that a great majority, but not all patients with active tuberculosis, had antibodies to different antigens of *Mycobacterium tuberculosis*. Simultaneous studies aimed at characterization of antigens of *Mycobacterium*

tuberculosis identified the reasons underlying the cross-reactivity in antibody responses between diseased and control subjects. In view of the facts that the immune system recognizes both the shared as well as specific antigens, and that uninfected and infected but nondiseased individuals would nevertheless respond to environmental as well as the commensal bacteria, the overlap between humoral immune responses became inevitable.

The focus of investigations thus moved to obtaining specific antigens of *Mycobacterium tuberculosis* and evaluating their role in eliciting humoral responses in diseased individuals. Since *Mycobacterium tuberculosis* infections can remain latent in infected individuals for their lifetime, with only a small proportion of such individuals developing active disease years after the primary infection, and since millions of individuals in the world have been vaccinated with a closely related organism, Bacille Calmette-Guerin (BCG), efforts have been directed at identification of antigens that are associated with active disease.

A variety of approaches have been used to achieve these aims. One approach that identified several such potential antigens is to immunize mice with heat killed or sonicated or radiation killed mycobacteria and subsequently produce monoclonal antibodies from

them, thus obtaining reagents that are specific for individual antigens or epitopes (Kolk *et al,* 1984; Damiani *et al,* 1988; Ljungqvist *et al*, 1988; Mauch *et al,* 1988; Engers *et al,* 1986; Khanolkar-Young *et al,* 1992). These antibodies are then used to obtain the corresponding antigens in purified form by screening recombinant DNA expression libraries made from mycobacterial DNA. Some investigators used different biochemical approaches to purify important antigens of *Mycobacterium tuberculosis*. Daniel and coworkers obtained a semipurified antigen, called antigen 5, from cell-free culture filtrates of *Mycobacterium tuberculosis*. This antigen was reported to be a glycoprotein of a molecular mass of 35 kDa (Daniel *et al,* 1978; Daniel *et al,* 1979).

Actively growing mycobacteria release several proteins into their culture supernatants. Andersen and coworkers identified 33 such proteins of which the 23-kDa superoxide dismutase enzyme was the earliest to appear (Andersen *et al,* 1991). These proteins bind to fibronectin, a glycoprotein that participates in cell surface interactions between bacteria and eukaryotic cells. These proteins, also known as the antigen 85 complex are now known to be components of many different antigen preparations that were used for evaluation of humoral responses in tuberculosis patients (Wiker *et al,* 1992).

The humoral immune responses of humans to some of the antigens have been investigated. Antibodies to the 38-kDa antigen in patients, contacts, and control individuals have been assessed. Daniel and coworkers used the semipurified antigen 5 preparation in a direct enzyme-linked immunosorbent assay (ELISA) to detect antibodies to the 38-kDa antigen in serum (Daniel *et al*, 1979; Daniel and Debanne, 1987; Daniel, 1989). Others have used a 38 kDa antigen obtained by immunoaffinity purification based on murine monoclonal antibodies (Jackett *et al*, 1988; Bothamley *et al*, 1992; Bothamley *et al*, 1988). These groups reported that 85% of the sera from patients with advanced tuberculosis, whose sputum smear was positive for acid-fast bacilli, had antibodies to the 38kDa antigen and none of the control sera was positive. Only 15% of the smear negative tuberculosis patients made antibodies to the 38kDa antigen (Bothamley *et al*, 1988). Results indicate that in humans, antibodies to the 38kDa antigen are elicited during advanced tuberculosis when the bacterial load is high.

Of the secreted antigens, humoral responses to the fibronectin binding proteins have been extensively studied (Wiker *et al*, 1990; Van Vooren *et al*, 1991; Van Vooren *et al*, 1992). Antibodies to antigen 85A can be found in both tuberculosis and nontuberculosis sera, but only tuberculosis patients have antibodies to antigen 85B.

These antigens do not seem to be immunodominant during early stages of infection in mice with live *Mycobacterium tuberculosis* (Verbon *et al*, 1990).

During the past few years, some investigators attempted a direct analysis of the antibody responses elicited in the diseased humans. Thus, disrupted *Mycobacterium tuberculosis* cells or culture supernatants of actively growing bacteria have been fractionated by SDS-PAGE and the seperated proteins, which range in size from 10 to 100 kDa, have been electrophoretically transferred to nitrocellulose sheets, which were then probed with sera obtained from patients and control subjects. This technique has the advantage of being able to provide information on the total profile of the antigens that are recognized by antibodies to different antigens in a particular serum sample. Espitia and colleagues fractionated ammonium sulphate-precipitated proteins from culture supernatants of *Mycobacterium tuberculosis* H37Rv and found that the antigen preparation fractionated into 35 bands. Sera from different tuberculosis individuals recognized different sets of antigens, and no definite pattern was discernible (Espitia and Mancilla, 1989; Espitia *et al*, 1989). Havlir and coworkers performed a similar analysis of antigen profile of *Mycobacterium tuberculosis* culture supernatant antigens recognized in tuberculous and nontuberculous sera. They

found that although tuberculous sera had higher titers of antibodies, both tuberculous and normal sera reacted with many fractions on the Western blots, and it was not possible to identify any specific antibody response in tuberculosis (Havlir *et al,* 1991). Using sonicated bacterial pellets as the antigen source, Coates and associates reported that both tuberculosis and control sera had antibodies to many different antigens, and antigens of 14, 30, 38, 45, and 65kda were prominent among those that were recognized (Coates *et al,* 1989). Verbon and associates fractionated culture supernatant antigens both by SDS-PAGE and by gel permeation chromatography, and used the former in Western blots and the fractions obtained in the latter in an ELISA to study the reactivity of antibodies present in tuberculous and normal sera. They found that although both serum groups had antibodies to 65, 61, 58, 30, and 24 kDa antigens when tested by western blots, 16 of 25 tuberculosis sera, but no control sera, reacted with a 24 kDa protein on the ELISA. Thus the epitopes revealed by ELISA are different from those seen on the Western blots (Verbon *et al*, 1990; Verbon *et al*, 1992).

Thus, we now know that there are differences in the epitopes or antigens of *Mycobacterium tuberculosis* that are recognized by animal and human immune systems. This also has been observed in another mycobacterial disease, leprosy, which is caused by a closely

related organism - *Mycobacterium leprae*. Use of serum antibodies from leprosy patients enabled identification of a dominant *Mycobacterium leprae* antigen that was not recognized by either murine or rabbit antibodies (Meeker *et al*, 1989; Laal *et al*, 1991; Sela *et al*, 1991). In addition, recently, Wallis and associates fractionated the culture supernatant containing secreted antigens of *Mycobacterium tuberculosis* by two-dimensional gels, and found that this fraction contains about 150 different proteins (Wallis *et al*, 1993). In addition, many of the already known antigens of *Mycobacterium tuberculosis* have both cross-reactive and specific epitopes and this is expected to be true for other, as yet undefined antigens. Sera from normal as well as tuberculosis individuals would have antibodies to the epitopes that are cross-reactive between *Mycobacterium tuberculosis* and other environmental and commensal bacteria. In retrospect, it is easy to understand how a combination of these factors could have contributed to the failure to delineate specific humoral responses in tuberculosis. *Mycobacterium tuberculosis* presents hundreds of antigens, each with several epitopes, to the immune system. The human immune system, in turn, appears to respond to a variety of antigens by making antibodies to them. The challenge that still remains is to devise strategies to dissect the humoral immune response, in order to define the species specific antigens or epitopes of *Mycobacterium*

tuberculosis that are recognized only in individuals with active disease.

1.4. The PE and PPE protein families of *M.tuberculosis*

About 10% of the coding capacity of the genome of *Mycobacterium tuberculosis* is devoted to two large unrelated families of acidic, glycine-rich proteins, the PE and PPE families (Fig. 1 on page 45), whose genes are clustered and are often based on multiple copies of the polymorphic repetitive sequences referred to as PGRSs, and major polymorphic tandem repeats (MPTRs), respectively (Hermans *et al*, 1992; Poulet and Cole, 1995a; Poulet and Cole, 1995b). The names PE and PPE derive from the motifs Pro-Glu and Pro-Pro-Glu found near the N-terminus in most cases. The 99 members of the PE family all have a highly conserved N-terminal domain of ~110 amino acid residues that is predicted to have a globular structure, followed by a C-terminal segment that varies in size, sequence and copy number (Cole *et al*, 1998). Phylogenetic analysis separated the PE family into several subfamilies. The largest of these is the highly repetitive PGRS class, which contains 61 members; members of the other subfamilies, share very limited sequence similarity in their C-terminal domain. The predicted molecular weight of the PE proteins vary considerably as few members contain only the N-terminal

domain, whereas most have C-terminal extensions ranging in size from 100 to 1400 residues. The PGRS proteins have a high glycine content (upto 50%), which is the result of multiple tandem repetitions of Gly-Gly-Ala or Gly-Gly-Asn motifs, or variations thereof.

The 69 members of the PPE protein family also have a conserved N-terminal domain that comprises ~180 amino acid residues, followed by C-terminal segments that vary markedly in sequence and length. These proteins fall in three groups (Cole *et al*, 1998), one of which constitutes the MPTR class characterized by the presence of multiple, tandem copies of the motif Asn-X-Gly-X-Gly-Asn-X-Glu. The second subgroup contains a characteristic, well conserved motif around position 350, whereas the third contains proteins that are unrelated except for the presence of the common 180-residue PPE domain. The PE/PPE families of proteins could represent a principle source of antigenic variation and the glycine rich proteins might interfere with immune response by inhibiting antigen processing.

There are several parallels between the PGRS proteins and the Epstein-barr virus nuclear antigens (Levitskaya *et al*, 1995, Levitsky *et al*, 1996). Members of both polypeptide families are glycine-rich, contain extensive Gly-Ala repeats, and exhibit variation in the length

of the repeat region between different isolates. The Gly-Ala repeat region of the EBNA1 functions as a cis-acting inhibitor of the ubiquitin/proteasome antigen-processing pathway that generates peptides presented in the context of major histocompatibility complex (MHC) class 1 molecules (Levitskaya *et al*, 1995; Levitskaya *et al*, 1997). MHC class 1 knockout mice are very susceptible to *Mycobacterium tuberculosis*, underlining the importance of a cytotoxic T-cell response in protection against disease (Chan and Kaufman, 1994; Flynn *et al*, 1992, Flynn *et al*, 1993).

Several observations and results support the possibility of antigenic variation associated with the PE and PPE family of proteins. The PGRS member Rv1759 is a 55kDa fibronectin binding protein that elicits a variable antibody response, indicating either that individuals mount different immune responses or that this PGRS protein may vary between strains of *Mycobacterium tuberculosis* (Abou-Zeid *et al*, 1991). The latter possibility is supported by restriction fragment length polymorphisms for various PGRS and MPTR sequences in clinical isolates (Cole *et al*, 1998). There is evidence to show that some members of the PGRS family are variable surface antigens. Banu *et al* showed 10 of these genes to be expressed *in vitro* by reverse transcription-polymerase chain reaction (RT-PCR). Antibodies against five PE-PGRS proteins, raised in mice by DNA

vaccination, detected single proteins when the same plasmid constructs used for immunization were expressed in epithelial cells or in reticulocyte extracts, confirming that the PE-PGRS proteins are antigenic (Banu *et al,* 2002). As expected from the conserved repetitive structure, the antibodies cross-reacted with more than one PE-PGRS protein, suggesting that different proteins share common epitopes.

Subcellular fractionation studies and immunoelectron microscopy localized many PE-PGRS proteins in the cell wall and cell membrane of *Mycobacterium tuberculosis*. Brennan *et al* showed that a transposon insertion in a PE_PGRS gene Rv1818c found in *Mycobacterium bovis* BCG Pasteur, which is the BCG homologue of the *Mycobacterium tuberculosis* H37Rv gene Rv1818c, introduces new phenotypic properties to this BCG strain (Brennan *et al,* 2001). These properties include dispersed growth in liquid medium and reduced infection of macrophages. Complementation of the 1818(PE_PGRS)::Tn5367 mutant with the wild-type gene restores both aggregative growth (clumping) in liquid medium and re-establishes infectivity of macrophages to levels equivalent to those for the parent BCG strain (Brennan *et al,* 2001). Western blot analysis using antisera raised against the 1818(PE_PGRS) protein shows that PE_PGRS proteins are found in cell lysates of BCG and

Mycobacterium tuberculosis H37Ra and in the cell wall fractions of *Mycobacterium tuberculosis* H37Rv. Moreover, immunofluorescent labeling of mycobacteria indicates that certain PE_PGRS proteins are localized at the cell surface of BCG and *Mycobacterium tuberculosis* (Brennan *et al,* 2001). Furthermore, investigation of the immunological response to Rv1818c revealed that during infection of mice with *Mycobacterium tuberculosis*, a significant humoral immune response was observed against recombinant Rv1818c. Similarly, immunization with a Rv1818c PE_PGRS DNA construct induced antibodies directed against PE_PGRS protein. Dissection of the immune response using the full length Rv1818c and N-terminal region of Rv1818c revealed that antibody response is directed solely against the Gly-Ala-rich PGRS domain present in the C-terminal region (Delogu and Brennan, 2001). These results suggest that the Rv1818c can elicit an effective immune response and that immune recognition of the PE antigen is influenced by the Gly-Ala-rich PGRS domain (Delogu and Brennan, 2001). Singh *et al* showed that Rv3367, a member of the PE-PGRS family is recognized by pooled sera from TB patients but not from healthy controls, confirming its *in vivo* expression during active infection in humans. Rv3367 was also recognized by retrospective preclinical TB sera obtained, prior to the clinical manifestation of TB, from human immunodeficiency virus-TB patients (Singh *et al,* 2001).

Direct support for genetic variation within both the PE and PPE family was obtained by comparative DNA sequence analysis. The gene for PE-PGRS protein Rv0746 of BCG differs from that in H37Rv by the deletion of 29 codons and the insertion of 46 codons. Similar variation was seen for the PPE protein Rv0442 (Cole *et al*, 1998). As these differences were all associated with repetitive sequences they could have resulted from intergenic or intragenic recombinational events or, more probably, from strand slippage during replication (Poulet *et al*, 1995a). These mechanisms are known to generate antigenic variability in other bacterial species (Robertson and Meyer, 1992). Flores *et al* have reported differential expression of PE and PE_PGRS genes in *Mycobacterium tuberculosis* strains. The expression of the *Mycobacterium tuberculosis* PE, PE_polymorphic GC-rich sequences (PGRS) gene family encoding approximately 99 glycine-rich proteins was assayed by reverse-transcriptase polymerase chain reaction (RT-PCR) in *Mycobacterium tuberculosis* H37Rv, *Mycobacterium canettii* and two clinical isolates of *Mycobacterium tuberculosis*. Restriction analyses and sequencing of the RT-PCR products showed that all the strains expressed the PE Rv1172c gene while the PE_PGRS Rv3652 gene was only expressed by one of the *Mycobacterium tuberculosis* clinical isolates, and the PE_PGRS Rv0578c was not expressed by *Mycobacterium canettii.* It was also determined that the PE_PGRS Rv0278c and Rv0279c were

not expressed by any of the studied strains. The data presented in this report (Flores and Espitia, 2003) show that the PE, PE_PGRS genes are differentially expressed in *Mycobacterium tuberculosis* strains during *in vitro* growth. These findings suggest that PE, PE_PGRS genes may play a role in protein variation between *Mycobacterium tuberculosis* strains (Flores and Espitia, 2003). Lamichhane *et al* employed a postgenomic method for predicting essential genes of *Mycobacterium tuberculosis* at subsaturation levels of mutagenesis and found that the PE-PGRS family of genes, which are unique to mycobacteria, were disproportionately enriched in essential genes (Lamichhane *et al*, 2003). Saviola *et al*, employing the use of recombinase-based *in vivo* expression technology identified a PE-PGRS gene, Rv0834c, the expression of which was dependent on pH in both *Mycobacterium smegmatis* and *Mycobacterium tuberculosis* (Saviola *et al*, 2003). Fisher *et al* have used microarrays and real-time reverse transcription-PCR to analyze the global transcriptional response of *Mycobacterium tuberculosis* to low pH *in vitro*, which may mimic an environmental signal encountered by phagocytosed mycobacteria and found that the transcription of Rv1169c, a member of the PE-PGRS family was up-regulated at low pH (Fisher *et al*, 2002). Stewart *et al* used a combination of targeted mutagenesis and whole-genome expression profiling to characterize transcription factors responsible for control of genes encoding the

major heat-shock proteins of *Mycobacterium tuberculosis* and found that Rv1169c is one of the genes upregulated at high temperatures under the control of the regulon HrcA, indicating its probable role as a heat shock protein (Stewart *et al*, 2002).

As compared to the members of the PE-PGRS family, only a few reports are available about the PPE family proteins of *Mycobacterium tuberculosis*. Sampson *et al* have used tandem repeat regions of Rv1917c to determine the molecular basis for the gene polymorphism of this PPE ORF within the clinical isolates and found that Rv1917c is highly polymorphic in clinical isolates. RT-PCR analysis demonstrated that Rv1917c mRNA is expressed in liquid cultures of *Mycobacterium tuberculosis* H37Rv. Expression of the recombinant protein in *Mycobacterium smegmatis* and *Mycobacterium bovis* BCG was visualized by fluorescence microscopy and flow cytometry. A protein of the predicted size (166 kDa) was confirmed by Western blotting. Cell fractionation studies demonstrated that the recombinant protein is hydrophobic, suggestive of cell wall-association, while flow cytometric data derived from antibody binding experiments suggested that it is surface exposed (Sampson *et al*, 2001). Skeiky *et al* infected C57BL/6 mice with *Mycobacterium tuberculosis* and used this infection model in conjunction with direct T cell expression cloning to identify antigens involved with the early

control of the disease. A protective *Mycobacterium tuberculosis*-specific CD4 T cell line derived from mice at 3 week postchallenge was used to directly screen a *Mycobacterium tuberculosis* genomic expression library. This screen resulted in the identification of a genomic clone comprising two putative adjacent genes with predicted open reading frames of 10 and 41 kDa, MTB10 and MTB41, the products of Rv0916c and Rv0915c, respectively (Skeiky *et al*, 2000). MTB10 and MTB41 belong to the PE and PPE family of *Mycobacterium tuberculosis*. MTB41 was the antigen recognized by the cell line and by *Mycobacterium tuberculosis*-sensitized human PBMC. C57BL/6 mice immunized with MTB41 DNA developed both CD4- (predominantly Th1) and CD8-specific T cell responses to rMTB41 protein (Skeiky *et al*, 2000). More importantly, immunization of C57BL/6 mice with MTB41 DNA induced protection against infection with *Mycobacterium tuberculosis* comparable to that induced by bacillus Calmette-Guerin. Rindi *et al* employed an mRNA differential display (DD) assay to compare gene expression between *Mycobacterium tuberculosis* H37Rv and its avirulent mutant H37Ra and found that Rv2770c, a member of the PPE family was one of the genes expressed in H37Rv but not in H37Ra, pointing to its probable role in the pathogenesis of the bacterium (Rindi *et al*, 1999). However, it has not been well documented what consequence the PPE family proteins, unique in their protein sequence and possible

structure, may have on the immune system. Furthermore, a qualitative and quantitative immune response of PPE proteins in a clinical setting has not been shown. Since 180 amino acid residues in the N-terminal region of PPE proteins are conserved, it is interesting to speculate that the variation in the sequence and length in the C terminal region could represent a source of antigenic variability.

Keeping the above observations in mind, the present study was undertaken to identify and characterize immunodominant antigen(s) present in the PPE family. The third sub-group members of the PPE family are unrelated except for the presence of the common PPE domain. An *in silico* approach was employed to identify candidate ORFs within this subgroup having a high antigenic index, coupled with the data available from Microarray studies suspecting an association of the candidate PPE ORFs to the pathogenesis of *Mycobacterium tuberculosis*. Using both approaches an ORF Rv2430c was shortlisted, the gene was cloned and expressed and the recombinant protein was evaluated for its use in generating a B cell response from a panel of sera obtained from four well classified categories of patients. The different categories were: Category 1 (n=32) which comprised patients who had contracted TB for the first time and had no history of TB treatment, Category 2 (n=30) which

comprised patients with relapsed TB, i.e., who were treated earlier for TB but the symptoms resurfaced after the completion of treatment. Category 3 (n=32) which comprised patients with extrapulmonary TB in which case the disease was confirmed by tissue biopsy and Category 4 were patients with MDR TB. Clinically healthy patients were taken as controls. In order to evaluate the diagnostic ability of Rv2430c, the reactivity of Rv2430c was compared with the sera obtained from non TB patients and the four infected groups. To elucidate the structure of the recombinant protein coded by Rv2430c, the protein was purified from inclusion bodies of *E.coli*. Initial attempt to refold the protein using dialysis led to protein precipitation. An on-column refolding strategy (Pullakhandham *et al*, 2004) was therefore employed to generate soluble protein. The soluble protein was further characterized by Circular Dichroism and Flourescence Spectroscopy. In order to assess the local environment of aromatic residues, the purified protein was incubated in the presence and absence of 8M urea.

In summary, the work presented in this thesis describe, for the first time, the immunological and structural characteristics of a protein belonging to the PPE family of *Mycobacterium tuberculosis*.

CHAPTER 2

MATERIALS AND METHODS

2.1. In silico analyses of PPE family ORFs

In silico analysis of a few of the PPE ORFs belonging to the third subgroup of the PPE family was carried out using the protein analysis software (Protean 4.0, Lasergene Navigator, DNASTAR Inc., Madison, WI) namely Kyte-Doolittle Hydrophilicity plot, Jameson-Wolf antigenic index and the Emini surface probability plot.

2.2. RNA Extraction

RNA was extracted from liquid cultures of H37Rv. The strain was grown in MiddleBrooks 7H9 liquid broth (Difco, USA). RNA was extracted using the Qiaquick Total RNA extraction kit (Qiagen, USA). The manufacturer's protocol was modified to get better yields of RNA from mycobacteria. All the reagents were diethyl pyrocarbonate (DEPC) treated to inhibit RNAse activity. A loop full of cells (~10^9) was transferred from the liquid media to 1 ml TE (10 mM Tris-HCl, 1 mM EDTA pH 8.0). The cells were centrifuged at 10,000 rpm for 10 minutes to collect the cell pellet. The soup was discarded and 100 µl of fresh TE containing 10 mg/ml lysozyme was added and the tube incubated for 15 minutes at 37^0C. The cells were lysed by adding 350 µl buffer RLT (containing 4 M guanidine iso-thiocynate, 0.25 M sodium citrate, 0.1% sarcosine, 0.2 M sodium acetate, 0.7% β-mercaptoethanol). The sample was vortexed vigorously and incubated at 65^0C for 1 hour. The cell debris was centrifuged for 2

minutes at 12,000 rpm in an eppenndorf rotor using a table top Heraeus centrifuge. The supernatant was transferred to a fresh tube. 250 µl of absolute ethanol was added and mixed by inverting the tube a couple of times. This was then put into the spin column (containing the silica matrix). The tube was centrifuged at 8,000 rpm for 1 minute. The flow through was discarded and 700 µl of buffer RW1 (as per the instructions given in the RNeasy Mini Handbook, Qiagen, USA) was added. The tube was centrifuged at 8,000 rpm for 15 seconds, the flow through was discarded and 500 µl of buffer RPE (containing ethanol for wash) was added. The tube was centrifuged at 8,000 rpm for 15 seconds. This step was repeated to wash out all the impurities. A final centrifugation at 10,000 rpm for 1 minute ensured that the silica matrix was dry and no carry over of Buffer RPE took place. The tube was transferred to a clean 1.5 ml tube and 50 µl of nuclease free water was added into the tube. The tube was centrifuged at 8,000 rpm for a minute to collect the dissolved RNA and stored at -70^{0}C till further use.

2.3. Reverse Transcriptase-Polymerase Chain Reaction (RT-PCR)

Primers designed on the basis of the published sequence of Rv2430c were used for the RT-PCR reaction. Promega Access RT-PCR kit was used for the amplification of RNA. A 50 µl reaction contained ~100 ng

of RNA, 1 μM each of the oligonucleotide primers specific to the corresponding gene, 0.2 mM each of the dNTPs, 1X of the provided reaction buffer, 1mM of $MgSO_4$ and 0.1 U/μl each of AMV Reverse Transcriptase and *Tfl* DNA Polymerase. For the first strand synthesis the tube was incubated at 48^0C for 45 minutes followed by incubation at 96^0C for 2 minutes to inactivate the reverse transcriptase. The second strand synthesis and PCR was performed using the following conditions; 25 cycles of 94^0C-1 minute, 48^0C-1 minute, and 68^0C-2 minute. One final hold at 68^0C for 10 minutes ensured the synthesis of complete length amplicons. In the control series two reactions were set up. One had no reverse transcriptase to check for DNA contamination. The second had no template to check for aerosol contamination. The other reaction conditions were identical to sample RT-PCR conditions. All other routine precautions of handling RNA and setting up of PCR were observed to limit chances of contamination. The PCR products were resolved by electrophoresis on 2% agarose gel.

2.4. Genomic DNA Extraction from H37Rv

Conventional Protocol

The *M.tuberculosis* cells (10^7cells) were scraped from Lowenstein-Jensen agar slants and suspended in 1 ml of TE (10 mM TrisCl, 1

mM EDTA pH 8.0). The cells were killed by three cycles of boiling at 100°C followed by freezing at -70°C. The cells were pelleted by centrifugation at 12,000 rpm for 10 minutes. The pellet was resuspended in 500 µl of fresh TE. Dead cells were lysed by incubation at 37°C with 100 µg/ml lysozyme for 30 minutes. The cells were vortexed for one minute and 2% Sodium dodecyl sulphate (SDS) and 15 µg/ml Proteinase K were added and the tubes were incubated at 65°C for 2 hour. The proteins, carbohydrates and other contaminants were selectively precipitated out with 10% Cetyl trimethyl ammonium bromide (CTAB) and 0.7M NaCl. After addition of CTAB-NaCl the tubes were incubated at 65°C for 15 minutes. The tubes were then cooled on ice for 10 minutes. This was followed by centrifugation at 10,000 rpm for 10 minutes. The supernatant was collected in a fresh tube and equal amount of phenol chloroform isoamyl alcohol (25:24:1) was added. The centrifugation step was repeated and the aqueous phase was collected in a fresh tube. To precipitate the DNA, 0.7 volumes of isopropanol was added and mixed by inverting the tubes a couple of times. The tubes were kept at 4°C for 2 hours. The DNA was pelleted by centrifugation at 10,000 rpm for 10 minutes. The supernatant was discarded and the DNA pellet was washed with 70% ethanol to remove salt. The DNA was finally resuspended in TE (10 mM TrisCl, I mM EDTA pH 8.0).

2.5. Polymerase Chain Reaction (PCR) amplification of Rv2430c

The Rv2430c gene was PCR amplified from the genomic DNA of H_{37}Rv using upstream (5' –GGATCCATGCATTTCGAAGCGTAC- 3') and downstream (5' –AAGCTTCTAAGTGTCTGTACGCGATGA- 3') primers. *Bam*HI and *Hin*dIII sites were incorporated in the 5' and 3' of the primers respectively, to facilitate directional cloning. For amplification, the *Taq* DNA polymerase (Promega, USA) was used. The PCR was set up in a 50 µl volume containing 100 ng of template (genomic DNA from the isolate), 1X Reaction buffer, 0.2 mM each of dNTPs, 1 unit of *Taq* DNA polymerase, 1.5 mM $MgCl_2$, 100 ng each of the primers and distilled water to make up the volume. The template was added to initiate the reaction and the tubes transferred to the GeneAmp PCR system 9700 thermal cycler (Applied Biosystems, USA). The amplification reaction was carried out as per the following conditions:

Denaturation	94°C - 2 min.
Annealing	48°C - 90 sec.
Extension	72°C - 2 min.

These conditions were repeated for 30 cycles and the resulting PCR product (amplicon) was subsequently used for cloning purposes.

2.6. DNA Purification

Once the amplicon was checked on the agarose gel, it was purified either by Qiagen's Qiaquick PCR purification kit or Qiaquick gel extraction kit. The former was used in cases where there was no non-specific amplification. In cases where there were more than one bands the gel extraction protocol was applied. These kits employ the technique of binding DNA to a silica matrix in the presence of high salt buffers. While the DNA stays bound to the matrix, the impurities are washed out by 70% ethanol. This DNA can then be eluted by TE or water. The DNA so obtained is of high quality with the O.D. A260/A280 in the range of 2.0.

2.7. PCR purification

The PCR product was diluted in 5 volumes of Buffer PB. The spin column was placed in a 2 ml collection tube and the diluted mix added to the column. The Buffer PB provides the high salt conditions for DNA binding to the silica matrix in the column. The spin column was centrifuged at 8,000 rpm for a minute. The flow through was discarded and the column was put back in the collection tube. The bound DNA was given two washes with 0.5 ml Buffer PE (ethanol containing buffer). To completely dry the tube to make it free of any

solutions a final spin for 2 minutes was given and the spin column was transferred to a fresh 1.5 ml tube. The DNA was eluted by adding 50 µl of distilled water to the column and the tube centrifuged at 8000 rpm for a minute. The DNA was checked for quantity and quality by running an aliquot (5 µl) on a 2% agarose gel. The tube was stored at -20^{0}C till further use. It was important to dissolve the DNA in water and not in TE as the EDTA in the buffer inhibits the PCR sequencing step. All the samples that needed to be sequenced were dissolved in distilled water.

2.8. Gel Extraction

For gel extraction the remaining (usually 40 µl) sample was loaded on 2% agarose gel and electrophoresed in 1X TAE buffer. The DNA was resolved and the specific band (based on the amplicon's molecular weight) was excised. DNA was extracted from the gel slices using the QIAquick Gel Extraction kit (Qiagen, USA). The gel slice was weighed and 3 volumes of Buffer QG (for solubilizing the gel piece and for providing the right condition for DNA binding to the silica matrix in the spin column) were added to it. The tube was incubated at 55^{0}C for 15 minutes with gentle agitation. 1 volume of isopropanol was added to the mix and vortexed. This was then transferred to the spin column placed in a 2 ml collection tube. To bind the DNA the column

was given a brief spin at 8,000 rpm for a minute. The flow through was discarded and 0.5 ml of Buffer PE (ethanol containing buffer) was added to the column. The column was placed in the collection tube and centrifuged as above. This wash step was repeated once more to remove unwanted primers, salts, enzymes, unincorporated nucleotides etc. The spin column was then given a dry run that is without any buffers. This ensured that the column was free of any solutions and the resin within the column was completely dry. The spin column was transferred to a fresh 1.5 ml tube and 50 µl of distilled water was added. The tube was spun at 8,000 rpm for a minute to elute the DNA. The tube was transferred to -20^{0}C after an aliquot (5 µl) of the eluate was electrophoresed on an agarose gel to confirm quantity and quality of DNA.

2.9. Cloning of the PPE gene Rv2430c

The purified fragment was ligated into the pGEMT-easy vector (Promega Inc., USA) and the ligation mixture was then transformed into competent DH5α cells using standard protocols. The transformed cells were plated on LB agar with 100 µg/ml ampicillin and the resulting colonies were screened for the insert. The insert was then sequenced from the vector using T7 promoter primer.

2.10. Automated DNA Sequencing

Sanger's Dideoxy method was used for sequencing the DNA samples. The technique depends on the requirement of the growing chain of single strand DNA for a 3' hydroxyl group to form a phosphodiester bridge with the incoming nucleotide. DNA polymerases are capable of incorporating analogues of nucleotide bases. So instead of a 2', 3' deoxynucleotide (dNTP) 2', 3' dideoxynucleotide (ddNTP) is added which does not have the hydroxyl group needed to form the phosphodiester bond with the incoming nucleotide. This leads to chain termination at this point. In automated fluorescent sequencing, fluorescent dye labels are incorporated into DNA extension product using 3'-dye labeled dideoxy nucleotide triphosphates (ddNTP). Each of the four different ddNTP (A, G, C, T) is labeled with a different dye that emits fluorescence at different wavelength. So the ddNTP that terminates the chain can be determined by the colour it emits. This varying positional colour-code is read on the automated sequencer and the DNA sequence obtained. Cycle sequencing is a simple method in which successive rounds of denaturation, annealing, and extension in a thermal cycler result in linear amplification of the extension products. The advantages of the technique are- the sequencing reaction can be done in a single tube, there is requirement of less DNA template and more samples can be

sequenced per gel (instead of four well per sample in radiolabeled DNA sequencing, here one sample is loaded in one well of the gel). The sequence lengths of 600 bases can be read with an accuracy of 95%. Both strands were sequenced to ensure the accuracy of sequencing.

2.11. Cycle Sequencing Reaction

For each sample two sequencing reactions, using forward or reverse primer were set up. The sequencing primers were identical to the ones used in PCR amplification of the different loci. PCR sequencing was carried out using the BigDye terminator kit (ABI Prism, USA) according to the manufacturer's instructions. Each sequencing reaction of 20 µl volume consisted of 50 ng template, 5 pmoles primer, 8 µl sequencing mix (AmpliTaq DNA polymerase, FS, with thermally stable pyrophosphate, the 4 dNTPs, MgCl2, Tris-HCl buffer pH 9.0, 4 ddNTPs labeled with 4 different dyes) and distilled water to 20 µl. The cycling parameters were as follows; 30 cycles of 96^0C for 30 seconds; 45^0C-60^0C for 30 seconds (depending on the Tm of primer used); 60^0C for 3 minutes. The labeled extension product was purified to remove unincorporated labeled ddNTPs. This was done by adding 0.1 volumes of 3 M sodium acetate pH 4.5 and 2.5 volumes of absolute ethanol. The solutions were mixed by vortexing and the

tube was left at room temperature for 10 minutes. The tube was centrifuged at 10,000 rpm for 10 minutes to pellet the DNA. The soup was discarded and the pellet was washed twice with 70% ethanol. After the final wash the tubes were air-dried. The pellet was dissolved in 5 µl of loading dye (90% Formamide, containing 0.1% blue dextran) and heated at 95^0C for 3 minutes to denature the DNA. The tube was kept on ice till loading and the left over was stored at-20^0C. 1.5-2.0 µl of the sample was loaded onto the 5% polyacrylamide gel (19:1 acrylamide-bis acrylamide, 1X TBE buffer) and electrophoresed at 3000 Volts for 3 hours in 1X TBE buffer (89 mM Tris base, 89 mM boric acid, 2 mM EDTA, pH 8.3). Sequencing of the amplicons was carried out using the ABI Prism Automated DNA Sequencer 377 sequencer (ABI Prism, USA). The automated sequencer was set to collect data with filter set E as recommended by the manufacturer. The software package Sequencing Analysis 3.3™ was used for analyzing the gel information. Sequences generated by the program were compared with their respective wild type sequences using MegAlign software (Lasergene, DNASTAR, Inc USA).

2.12. Subcloning of Rv2430c in PQE30 expression vector

After confirmation of the sequence the insert was extracted from the pGEMT-easy vector by digestion with *Bam*HI and *Hin*dIII enzymes

and then subcloned into the pQE30 expression vector (Qiagen Inc., USA), downstream of a 6X histidine sequence tag, at the *Bam*HI and *Hin*dIII sites to generate the plasmid construct called PQERv2430c. The ligation mix contained 100ng each of the vector and the insert, 1U DNA ligase, 1X reaction buffer and the volume was made up with water. The reaction was allowed to proceed at 14°C for 12 hours and the ligation mix was transformed in DH5α cells using the standard protocols. The resulting construct PQERv2430c was purified and was finally transformed into M15pREP4 strain of *E.coli* for expression of the recombinant protein.

2.13. Expression and purification of the recombinant protein coded by Rv2430c

A single colony of M15pREP4 cell line harbouring the construct was inoculated in 5 ml of LB broth with the appropriate antibiotics, as mentioned above, and grown overnight at 37°C with constant agitation. 100 µl of this overnight culture was inoculated into 5 ml of LB broth with the appropriate antibiotics and grown till a cell density corresponding to an absorbance value of 0.6 (A_{590}= 0.6) was achieved. The culture was then induced with 1 mM IPTG. A separate aliquot of uninduced culture was kept as a control. Cells were harvested 3 hours post induction, suspended in 1X SDS sample buffer and

denatured by heating at 100°C for 10 min. The samples were resolved in a 12% SDS polyacrylamide gel, confirming that the expected 23 kDa protein was expressed.

The recombinant protein was then purified to homogeneity using the QIAExpressionist kit (Qiagen, Inc., USA). Cells harvested from 10 ml of induced culture were resuspended in lysis buffer containing 100 mM NaH_2PO_4, 10 mM Tris.Cl and 8M urea (pH 8.0). The lysate was loaded onto a Ni-NTA column pre-equilibriated with the lysis buffer. The column was washed with wash buffer containing 100 mM NaH_2PO_4, 10 mM Tris.Cl and 8 M urea (pH 6.3). Finally, the protein was eluted with elution buffer containing 100 mM NaH_2PO_4, 10 mM Tris.Cl and 8 M urea (pH 4.5), and resolved by electrophoresis in a 12% SDS polyacrylamide gel. A single 23 kDa protein band was observed upon staining with Coomassie Brilliant Blue dye, indicating the purified protein. The protein was then dialyzed against 1 X PBS, pH 7.5 using a 10 kDa cut-off dialysis membrane and the dialyzed protein was quantified using the PIERCE Micro BCA Protein Assay Reagent Kit, according to the manufacturer's instructions.

2.14. Human Study Population

Serum samples were obtained from 101 TB patients who reported to the outpatient department of Mahavir Hospital and Research Centre,

Hyderabad. These 101 patients belonged to four well classified categories of patients. Category 1 (n=32) comprised patients who had contracted TB for the first time and had no history of TB treatment. Category 2 (n=30) comprised patients with relapsed TB, i.e., who were treated earlier for TB but the symptoms resurfaced after the completion of treatment. Category 3 (n=32) comprised patients with extrapulmonary TB in which case the disease was confirmed by tissue biopsy and category 4 (n=9) were patients with MDR TB. In the case of category 1 and category 2 patients, diagnosis was confirmed by the examination of the sputum (acid-fast bacilli smear positive). 10 clinically healthy donors, which were *M.bovis* BCG vaccinated, were included in the study. 15 non TB patients, i.e patients which harbored any pathogen but were culture negative for *Mycobacterium tuberculosis* were also included in the study. The study was carried out after approval from the Institute Bioethics Committee.

2.15. Serological characterization of the recombinant protein

Serological characterization of the recombinant protein was carried out using ELISA (Enzyme linked immunosorbent assay). ELISAs were performed in 96-well microtitre plates (Corning, Costar) coated with the recombinant Rv2430c protein. After overnight incubation at 4°C, the plates were washed three times with phosphate-buffered saline

(PBS) buffer and blocked with 200µl of blocking buffer (PBS containing 1% bovine serum albumin) for 1h at 37°C. The plates were then washed with PBS-Tween wash buffer (0.05% Tween 20 in 1X PBS, pH 8.0) and incubated for 1h at 37°C with human sera (1:200 dilution in blocking buffer). The plates were washed with PBS-Tween and further incubated with either anti-human immunoglobulin G (IgG)- horseradish peroxidase (HRP) or anti-human IgM-HRP (Sigma). HRP activity was detected using a chromogenic substance, o-phenylenediamine tetrahydrochloride (Sigma) in citrate-phosphate buffer (pH 5.4) and 1µ of H_2O_2/ml. The reactions were terminated using 1 N H_2SO_4 and the absorbance values were measured at 492nm in an ELISA reader (Bio-Rad, USA).

2.16. Localization of Rv2430c using an *in-vitro* coupled transcription and translation system

Localization of Rv2430c was checked by using a coupled transcription-translation system. The ORF Rv2430c was cloned in pET23a expression vector under the control of T7 promoter to generate a construct pET23aRv2430c. The coupled transcription and translation reaction was carried out using the TNT Coupled Reticulocyte Lysate Systems as per the suppliers instructions (Promega, USA). Briefly, 1µg of the plasmid (pET23aRv2430c), 0.5µl of

amino acid mix, 1µl of RNAsin, 1µl of ^{35}S Methionine, 0.5µl of T7 RNA polymerase, 13µl of Rabbit Reticulocyte Lysate, 1µl of reaction buffer and 6µl of water was incubated at 30°C for 90 minutes. A separate reaction with the above reaction mix plus Canine Pancreatic Microsomal membrane was also set up. Both the reaction products were treated with Trypsin and analysed on SDS-PAGE.

2.17. Statistical analysis

Student's t-test was used for analysis of statistical significance (p-value). Graphpad Quickcalcs (Online t-test calculator) was used for this purpose. http://www.graphpad.com/quickcalcs/ttest1.cfm.

2.18. Secondary structure prediction of Rv2430c

In silico analysis of the Rv2430c was carried out using the Protein analysis software (Protean 4.0, Lasergene Navigator, DNASTAR Inc., Madison, WI). Predict protein (http://cubic.bioc.columbia.edu/predictprotein/) and PSIPRED (http://bioinf.cs.ucl.ac.uk/psiform.html) both of which were web based.

2.19. On-column refolding of recombinant Rv2430c

A 100ml culture of M15pREP4 strain carrying the recombinant plasmid PQERv2430c was grown till log phase and induced with 1mM IPTG. Cells were harvested 3 hours post induction and suspended in Buffer A (25mM Tris-Cl, pH 8.0 containing 8M urea and 0.9% NaCl) and incubated on an end-to-end shaker for 30 minutes at room temperature for lysis. The lysate was centrifuged at 13000 rpm for 30 minutes and the supernatant was then incubated with pre-equilibriated Ni-NTA slurry (Qiagen Inc., USA) for 15-20 minutes with gentle agitation to maximize the binding of the recombinant protein. The protein bound to slurry was then packed into a column. The bound protein was then subjected to on-column refolding by using a 250 ml gradient of buffer A and Buffer B (25mM Tris, pH 8.0, 5mM imidazole, 1mM glutathione and 0.1M L-Arginine hydrochloride) at a flow rate of 1ml/min using Acta-Prime chromatographic unit (Pharmacia Biotech). At the end of the gradient, the column was further washed with 50 ml of buffer B and then eluted with 25mM Tris, pH 8.0 containing 500 mM imidazole. The homogeneity of the eluted protein was confirmed by 12% SDS-PAGE and the purified protein was dialyzed extensively at 4°C against 25mM Tris HCl pH 8.0 containing 0.9% NaCl. Protein was quantified by Pierce Micro BCA Protein Assay Reagent kit (Pierce,

USA) and was subsequently used for spectroscopic analyses. The dialyzed protein was further centrifuged at 13000 rpm for 15 minutes to check for the presence of any visible aggregation.

2.20. Circular Dichroism (CD) Spectroscopy

CD measurements were carried out on a spectropolarometer (JASCO-715, JAPAN) using a 0.02-cm cell at 0.2-nm intervals and a two-nanometer bandwidth. Spectra were signal averaged by adding atleast 4 accumulations. The base line was corrected by subtracting the spectra of respective buffer blank obtained under identical conditions. Percentage of secondary structure was calculated using the web based programme K2D (http://www.embl-heidelberg.de/~andrade/k2d/).

2.21. Fluorescence spectroscopy

Purified recombinant Rv2430c protein was incubated in the presence or absence of 8M urea for 2h at room temperature and the fluorescence emission spectra (300-400nm) were recorded by exciting the protein at 280nm using Perkin-Elmer LS-3B spectrofluorimeter. The slit width was 10nm and the scan speed was 50nm/sec.

CHAPTER 3

RESULTS

RESULTS

3.1 : In silico analyses

3.1.1. In silico and web based analyses of the PPE family ORFs: The members of PE and PPE family have a conserved N-terminal domain of 110 and 180 amino acid residues respectively as evident from Fig. 1. The 69 members of the PPE family (Cole *et al*, 1998) display significant sequence and length variation in their C-terminal region. These 69 members of the PPE protein family have a conserved N-terminal domain that comprises ~180 amino acids followed by C-terminal segments that vary markedly in sequence and length. Based on our pattern search analysis of the Tuberculist database (http://genolist.pasteur.fr/TubercuList/) these proteins were categorised into three groups. Subgroup 1, represented by 20 members, constitutes the MPTR class characterized by the presence of multiple, tandem copies of the motif Asn-X-Gly-X-Gly-Asn-X-Gly. The second subgroup, comprising of 21 members, contains a characteristic well conserved motif Gly-X-X-Ser-Val-Pro-X-X-Trp around position 350, and the third subgroup proteins, with 28 members, are unrelated except for the presence of the common PPE domain (Table 1). ORFs belonging to the third subgroup with coding capacity equal to or less than 200 amino acids were shortlisted. ORFs Rv2430c and Rv3425 displayed major antigenic stretches with

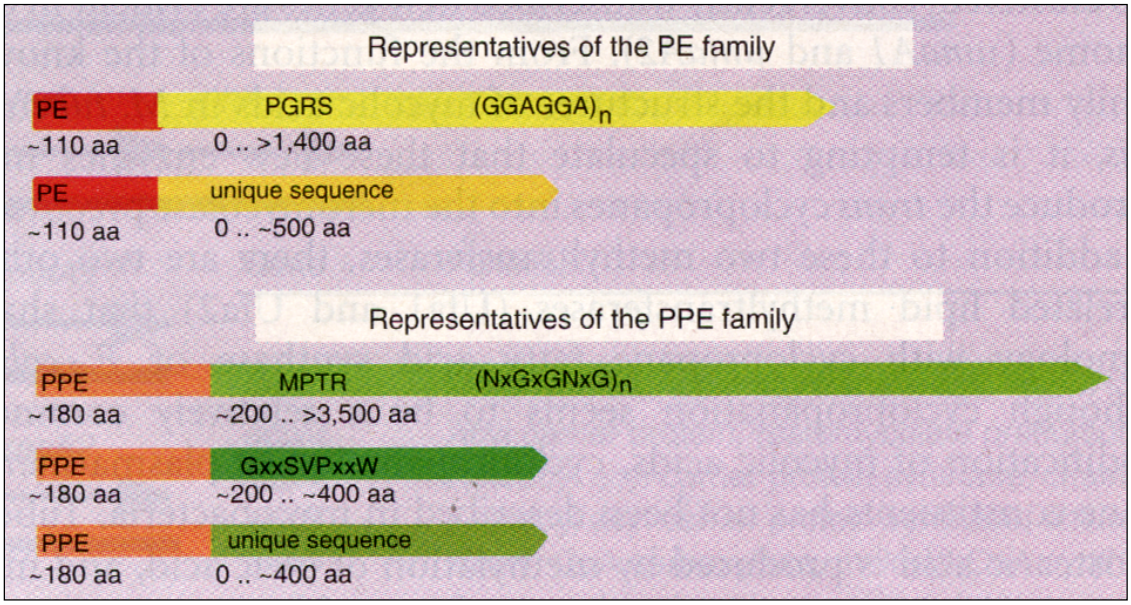


Fig. 1 Schematic representation of the PE and the PPE families of *Mycobacterium tuberculosis* indicating the variation in sequence and length in the C-terminal region.

	Sub Group 1	Sub Group 2	Sub Group 3
Members	20	21	28
Motif	Asn-X-Gly-X-Gly-Asn-X-Gly	Gly-X-X-Ser-Val-Pro-X-X-Trp	Common PPE domain

Table 1. Classification of the members of the PPE family based on the pattern search analysis of the Tuberculist database (http://genolist.pasteur.fr/TubercuList/).

peak values ≥ 3.0 (Fig. 2). This shortlist was further narrowed down based on two very important criteria – namely, antigenic profile and the association of the ORF with pathological conditions as evident from DNA Microarray expression data (Rodriguez *et al*, 2002). Comparative analysis of the PPE members, belonging to the third subgroup, less than 200 amino acids revealed that Rv2430c has a high antigenic index. Furthermore, DNA Microarray results demonstrated that Rv2430c was one of the genes induced in IdeR mutant of *Mycobacterium tuberculosis*, suggesting its possible role in pathogenesis. Rv2430c was accordingly shortlisted for the present study and was evaluated for its role as an antigen in a clinical setting.

3.1.2. *In silico* analysis predicts a high content of alpha helices within PPE ORF Rv2430c: *In silico* and web based analysis of the PPE family comprising of all three groups was carried out (Table 2). It appears that while most of them have an irregular or a random coil structure, the remaining ones display α helical structure. Given the fact that microarray data indeed pointed to the likely importance of a few members of this family under stress conditions, their function however remains largely unknown. Detailed computational analyses of Rv2430c predicted regions of high antigenic index with corresponding hydrophilicity and surface probability.

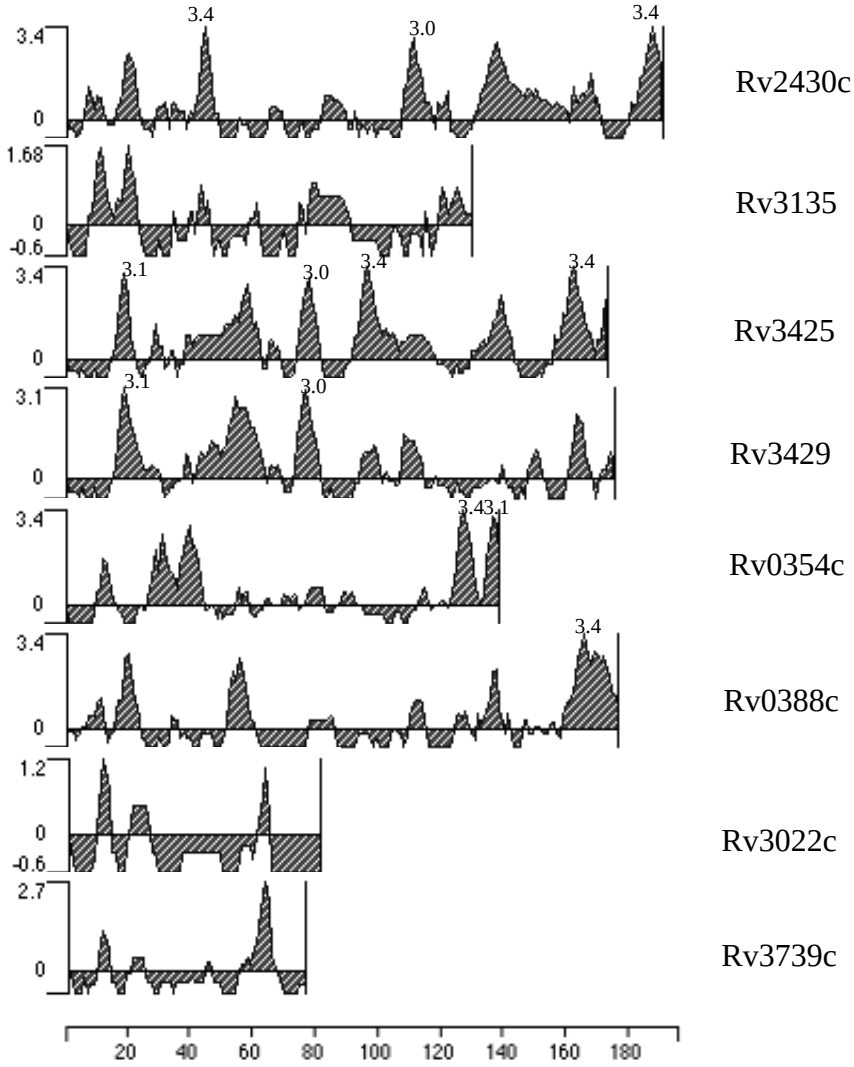


Fig. 2. *In silico* analysis of the selected PPE ORFs belonging to the third sub group. The analysis reveals a variable antigenic index. Relative peak heights are shown to the left. The scale at the bottom depicts the length of the ORFs in amino acids.

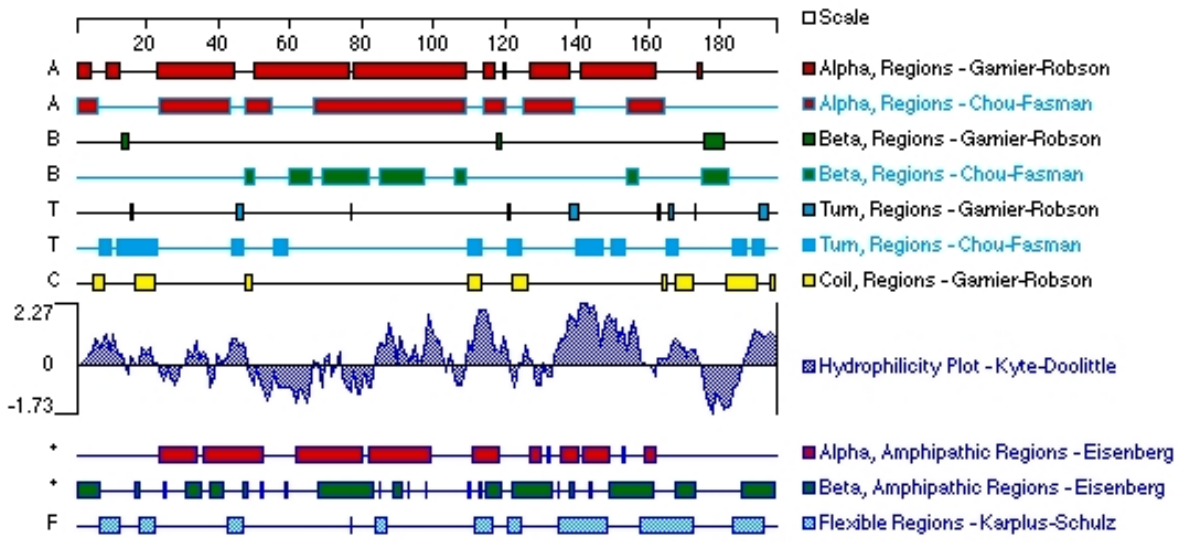


Fig. 3. *In silico* analysis of Rv2430c employing Protean 4.0, DNASTAR reveals a very high content of α helical structure in Rv2430c. The Garnier-Robson and Chou-Fasman methods of secondary structure prediction revealed similar results.

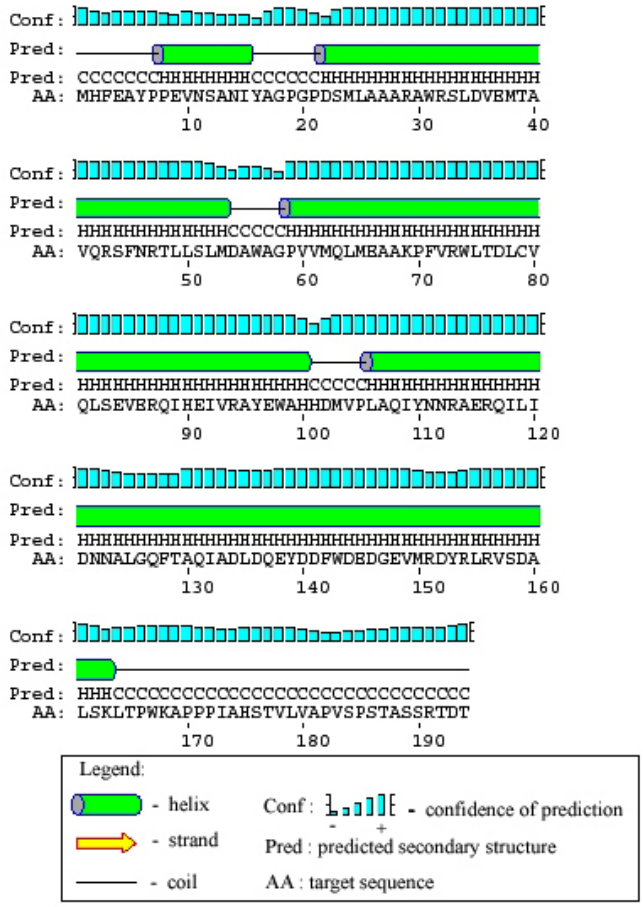


Fig. 4. Secondary structure analysis of Rv2430c using the web based server PSIPRED showed a predominant α helical composition with a high prediction score.

Table 2 : Results of *in silico* and web based analysis of the PPE ORFs belonging to all the three groups.

	Amino acids	Microarray data[a]	*In silico* prediction [b]	Suggested function
MPTR (Subgroup1)				
Rv0304c	2204		R	Unknown
Rv0305c	963		Alpha	Unknown
Rv0355c	3300		R	Unknown
Rv0442c	487		Alpha	Unknown
Rv0755c	645		Alpha	Unknown
Rv0878c	443		Alpha	Unknown
Rv1135c	618		Alpha	Unknown
Rv1548c	678		R	Unknown
Rv1753c	1053		R	Unknown
Rv1917c	1459		R	Unknown
Rv1918c	987		R	Unknown
Rv2353c	354		R	Unknown
Rv2356c	615		Alpha	Unknown
Rv2608	580		Alpha	Unknown
Rv3159c	590	mRNA identified in starvation model	Alpha	Unknown
Rv3343c	2523		R	Unknown
Rv3347c	3157		R	Unknown
Rv3350c	3716		R	Unknown

Rv3533c	582		R	Unknown
Rv3558	552		R	Unknown
Motif around position 350 (Subgroup2)				
Rv0915c	423		R	Unknown
Rv1039c	391	mRNA identified in starvation model	R	Unknown
Rv1168c	346		R	Unknown
Rv1196	391	mRNA identified in starvation model	R	Unknown
Rv1361c	396		Alpha	Unknown
Rv1706c	394		R	Unknown
Rv1787	365		R	Unknown
Rv1789	393		R	Unknown
Rv1790	353		R	Unknown
Rv1801	423	mRNA identified in starvation model	R	Unknown
Rv1802	463		Alpha	Unknown
Rv1807	399		R	Unknown
Rv1808	409	mRNA identified in starvation model	R	Unknown
Rv1809	468	mRNA identified in starvation model	R	Unknown
Rv2352c	391		R	Unknown

Rv2768c	394		R	Unknown
Rv2770c	382		R	Unknown
Rv2892c	408		Alpha	Unknown
Rv3125c	391		R	Unknown
Rv3136	380	mRNA identified in starvation model	R	Unknown
Rv3532	406		R	Unknown
Others (Subgroup3)				
Rv0096	463		R	Unknown
Rv0256c	556	mRNA identified in starvation model	R	Unknown
Rv0280	536		Alpha	Unknown
Rv0286	513		Alpha	Unknown
Rv0354c	141	mRNA identified in starvation model	R	Unknown
Rv0388c	180		Alpha	Unknown
Rv0453	518		R	Unknown
Rv1387	539	mRNA identified in starvation model and in response to acidic conditions	Alpha	Unknown
Rv1705c	385		R	Unknown
Rv1800	655		Alpha	Unknown
Rv2108	243		Alpha	Unknown
Rv2123	473		R	Unknown

Rv2430c	194		Alpha	Unknown
Rv3018c	434		R	Unknown
Rv3021c	358		R	Unknown
Rv3022c	82		R	Unknown
Rv3135	132		R	Unknown
Rv3144c	409		R	Unknown
Rv3425	176		Alpha	Unknown
Rv3426	232		Alpha	Unknown
Rv3429	178		Alpha	Unknown
Rv3478	393		R	Unknown
Rv3539	479		Alpha	Unknown
Rv3621c	413		Alpha	Unknown
Rv3738c	315		R	Unknown
Rv3739c	77		R	Unknown
Rv3873	368		R	Unknown
Rv3892c	399		R	Unknown

[a]Microarray data were obtained from Tuberculist (http://genolist.pasteur.fr/TubercuList/). [b]*In silico* analysis was carried out using the Protein analysis software, PROTEAN, DNASTAR. (R indicates irregular or random structure).

Secondary structure prediction employing Protean 4.0 reveals a very high content of α helical structure in Rv2430c. This is evident from both the Garnier-Robson and Chou-Fasman methods of secondary structure prediction (Fig. 3). Web based analysis of Rv2430c using PSIPRED similarly revealed a predominant α helical composition with a high prediction score (Fig. 4). Similar results were obtained with Predict protein (76% α helix, data not shown).

3.2: Expression of PPE Rv2430c

3.2.1. The PPE ORF Rv2430c is transcribed in liquid cultures of *Mycobacterium tuberculosis* : Rv2430c was listed as a hypothetical protein (Cole *et al*, 1998) and therefore it was important to check whether this hypothetical Rv2430c indeed represented a functional gene. To address this primary question, mRNA extracted from *in vitro* cultured *Mycobacterium tuberculosis* H_{37}Rv cells was used as a template for reverse transcription followed by PCR. The RT-PCR reaction product was fractionated by electrophoresis on a 1% agarose gel. A 597 bp band was observed upon staining with ethidium bromide indicating the expression of this ORF at the mRNA level in the liquid cultures of *Mycobacterium tuberculosis* (Fig. 5).

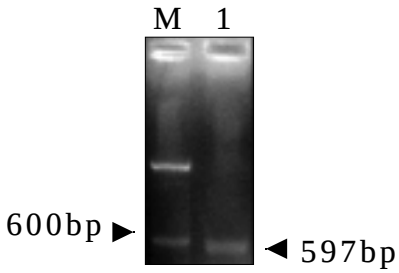


Fig. 5. The hypothetical PPE ORF Rv2430c is transcribed in liquid cultures of *Mycobacterium tuberculosis*. RNA extracted from the virulent H37Rv laboratory strain of *Mycobacterium tuberculosis* was used in an RT-PCR reaction. A 597 bp RT-PCR product was observed after electrophoresis in a 1% agarose gel (lane 1). Lane M depicts the 100 bp DNA molecular size marker run alongside. Arrowhead indicates the position of the 597bp PCR product.

3.2.2. Expression of the recombinant PPE protein Rv2430c: The PPE ORF Rv2430c was cloned in PQE30 expression vector and the expression of the recombinant protein was induced by 1mM IPTG. High level expression of 6x His-tagged proteins in *E.coli* using PQE vectors is based on the T5 promoter transcription-translation system. PQE plasmids belong to the pDS family of plasmids (Bujard *et al,* 1987) and were derived from plasmids pDS56/RBS11 and pDS781/RBS11- DHFRS (Stüber *et al,* 1990). These low copy plasmids have an optimized promoter-operator element consisting of phage T5 promoter (recognized by the *E.coli* RNA polymerase) and two lac operator sequences, which increase *lac* repressor binding and ensure efficient repression of the powerful T5 promoter, synthetic ribosomal binding site, RBS11, for high translation rates, 6x Histidine-tag coding sequence either 5' or 3' to the cloning region, multiple cloning site and translational stop codons in all reading frames for convenient preparation of expression constructs. Two strong transcription terminators : t_o from phage lambda, and t1 from the rrnB operon of *E.coli,* are present to prevent read through transcription and ensure stability of the expression construct, alongwith β-lactamase gene confering resistance to ampicillin and ColE1 origin of replication.

E.coli host strain containing both the expression and the repressor plasmid was used for the production of recombinant Rv2430c protein. For this *E.coli* strain M15[pREP4] which permits high level expression and is easy to handle was used. M15 strain is derived from *E.coli* K12 and has the phenotype Nal^s, Str^s, Rif^s, Thi^-, Lac^-, Ara^+, Gal^+, Mtl^-, F^-, $RecA^+$, Uvr^+, Lon^+. The expression of Rv2430c protein was achieved by inducing *E.coli* M15, transformed with the recombinant plasmid, by the addition of 1mM IPTG.

3.2.3. Purification of the recombinant PPE protein Rv2430c : Recombinant Rv2430c protein, folded on-column, does not aggregate upon dialysis

Overexpression of the PPE protein in the M15 strain of *E.coli* led to its localization in the inclusion bodies. Hence the protein was extracted under denaturing conditions using 8M urea. Immobilized metal affinity chromatography was employed to purify the recombinant protein from *E.coli*. The 6x Histidine affinity tag facilitated the binding of the recombinant protein to Ni-NTA (Nitrilotriacetic acid). Nitrilotriacetic acid is a tetradentate chelating adsorbent that occupies four of the six ligand binding sites in the coordination sphere of the Nickel ion, leaving two sites free to interact with the 6x Histidine tag.

The on-column refolding strategy has been earlier used to facilitate minimal protein aggregation and precipitation. Chemical chaperones such as L-arginine together with glutathione provide reducing equivalents during folding. This strategy resulted in about 2mg pure protein per 100 ml culture. Fractions eluted from the purification column were analysed on 12% SDS-PAGE (Fig. 6) which confirmed the homogeneity (>95%) of the recombinant protein. No visible precipitation of the protein was observed upon extensive dialysis and high speed centrifugation (13000 rpm, 15 minutes) pointing to the stable conformation of the refolded protein.

3.3 : Biophysical characterization of the on column folded PPE protein

3.3.1. Rv2430c displays a significant α helical composition : The secondary structure of the on-column refolded protein was characterized by Circular Dichroism (CD) spectroscopy. The observed CD spectrum was characteristic of a helical protein (Fig. 7). Data obtained from the CD spectra recorded in 200 – 250 nm range were used to calculate the secondary structure composition of the protein using the web based programme K2D. The recombinant protein displays 81% α helical content and 19% random coil structure (Inset-Fig. 7).

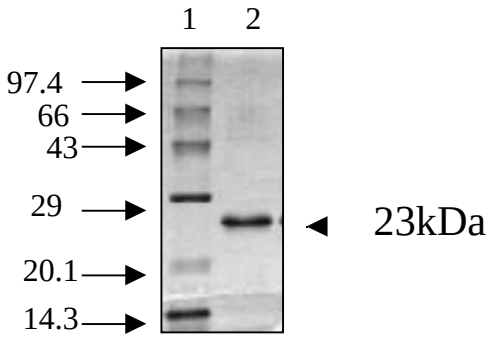


Fig. 6. Expression and purification of the recombinant PPE protein. The recombinant protein was expressed in *E.coli* M15pREP4 strain and was purified to homogeneity using the NiNTA protein purification kit. The gel was stained by Coomassie Brilliant blue. Lane 1 is the protein molecular size marker, Lane 2 is the purified protein. Arrowhead indicates the position of the 23kDa protein.

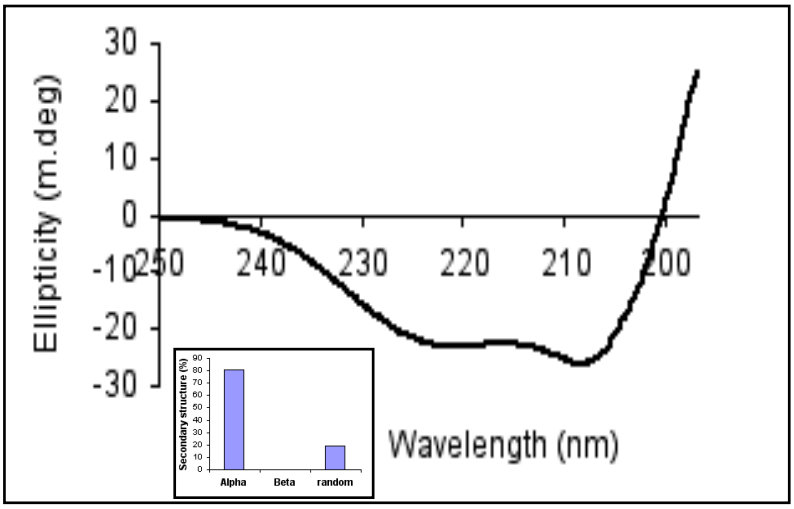


Fig. 7. Rv2430c has alpha helical structure as revealed from far-UV CD spectra analyses. Far-UV CD spectrum was recorded in 25mM Tris-Cl, pH 8.0 with 0.9% NaCl in the range of 190 to 250nm. Cell of path length 0.02cm was used for recording the Far-UV spectra. Ellipticity was represented in millidegrees. The spectrum obtained is characteristic of alpha helical protein. Figure in Inset shows the percentage composition of different secondary structures in Rv2430c. Data obtained from the CD spectra were calculated to get the percentage composition of secondary structure. Rv2430c shows a very high content (81%) of alpha helical structure.

While these data confirm the proper folding of the purified protein, they are also in agreement with the *in silico* predictions that Rv2430c is a predominantly α helical protein.

3.3.2. The flourescence spectra of the folded protein differ significantly from the unfolded protein : In order to assess the local environment of the aromatic residues, the folded protein was subjected to urea denaturation followed by fluorescence emission spectroscopy. The refolded PPE protein displayed the emission maximum at 340nm suggesting that the aromatic amino acid residues are buried in the protein, indicative of a folded protein. A significant red-shift in the absorption maxima and increased emission intensity was observed from 340 nm to 350 nm upon denaturation of protein with 8M urea (Fig. 8). These data suggest that in the native conformation of the protein, aromatic residues are present in a hydrophobic environment, which become exposed upon protein denaturation.

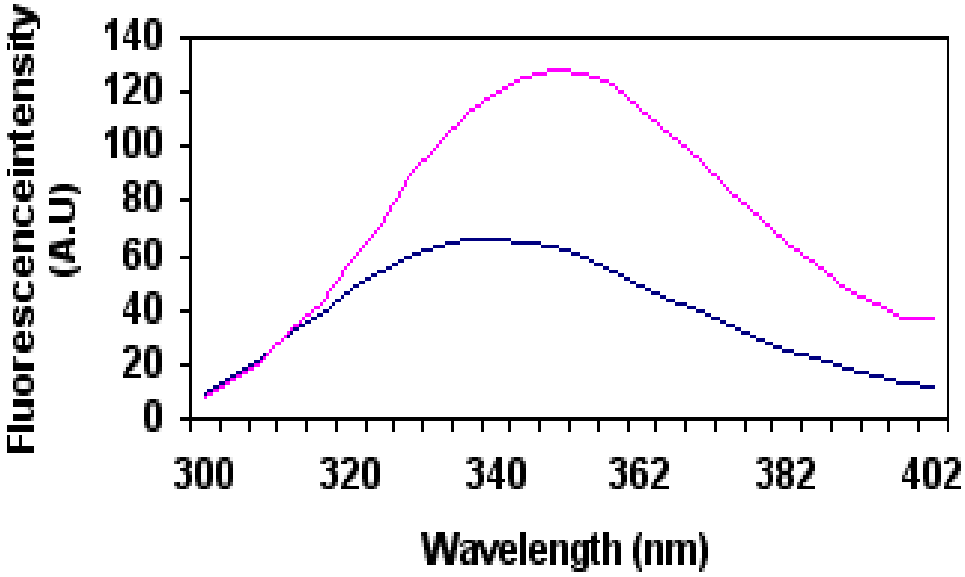


Fig. 8. The aromatic residues in Rv2430c are present in a hydrophobic region. Purified PPE protein was incubated in the presence and absence of 8M urea for 2hr at room temperature and the fluorescence emission spectra were recorded by exciting the protein at 280nm and recording the emission spectra in the range of 300nm-400nm. Pink line represents the unfolded protein whereas the blue line denotes the folded protein. A significant red shift in absorption was observed.

3.4 : Immunological characterization of the recombinant PPE protein Rv2430c

3.4.1. A strong B-cell response is elicited against Rv2430c in patients infected with TB: Having shown that the ORF encoded by the PPE family of *Mycobacterium tuberculosis* was expressed at the mRNA level, experiments were designed to evaluate the expression of this protein in patients infected with TB. This was achieved by assaying the immune response of TB patients to recombinant Rv2430c. For this the recombinant Rv2430c protein was used to screen TB patient sera by enzyme linked immunosorbant assay (ELISA) using anti-human IgG-HRP and anti-human IgM-HRP as conjugates. The humoral immune responses directed against the recombinant protein were compared between patients with tuberculosis and healthy controls. The data (Fig. 9) reveal that sera of all the infected patients mounted a significantly higher antibody responses against Rv2430c as compared to that of the healthy controls ($p<0.0001$). Since negligible antibody responses were obtained in the healthy control group, it is likely that this protein is expressed during the course of *Mycobacterium tuberculosis* infection and may be associated with disease manifestation and progression.

3.4.2. Immunodominant nature of Rv2430c: Having shown that the PPE protein is expressed during infection, it was of imminent

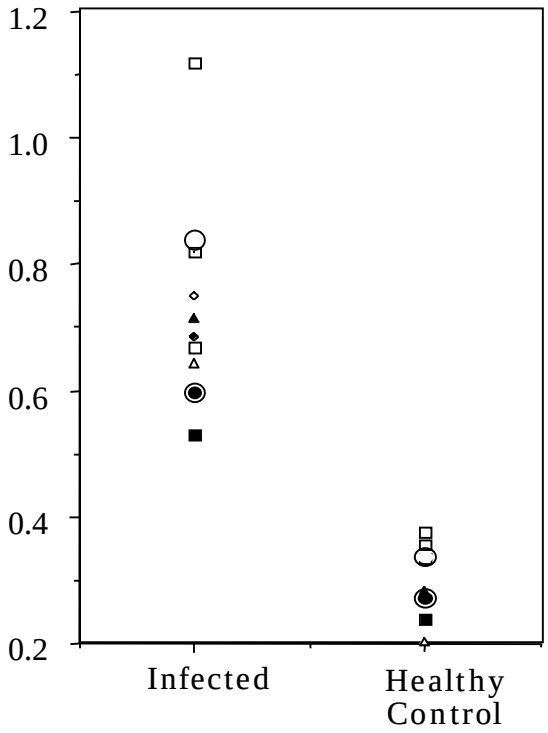
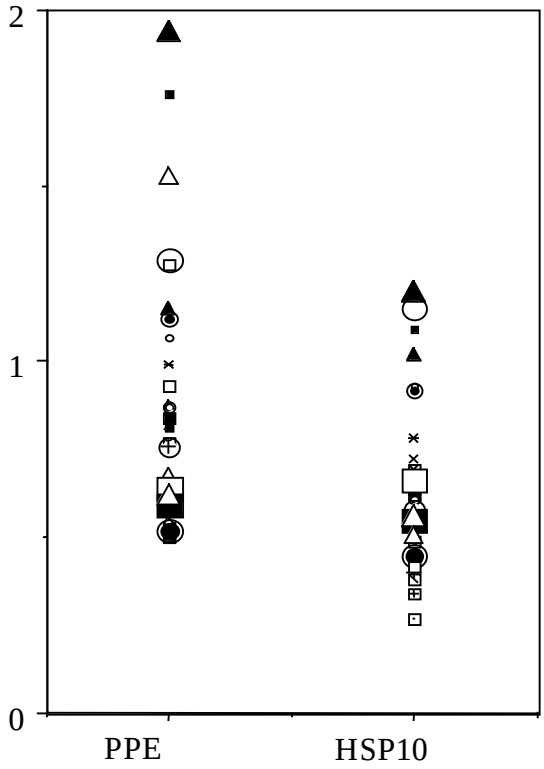


Fig. 9. The recombinant Rv2430c PPE protein elicits strong antibody response in *M.tuberculosis* infected patients as opposed to healthy controls. ELISA reactivity of IgG anti-Rv2430c antibodies were assayed in sera of either *M.tuberculosis* infected patients or healthy controls. The healthy controls were *M.bovis* BCG vaccinated whereas the infected subjects comprised patients from all three groups.

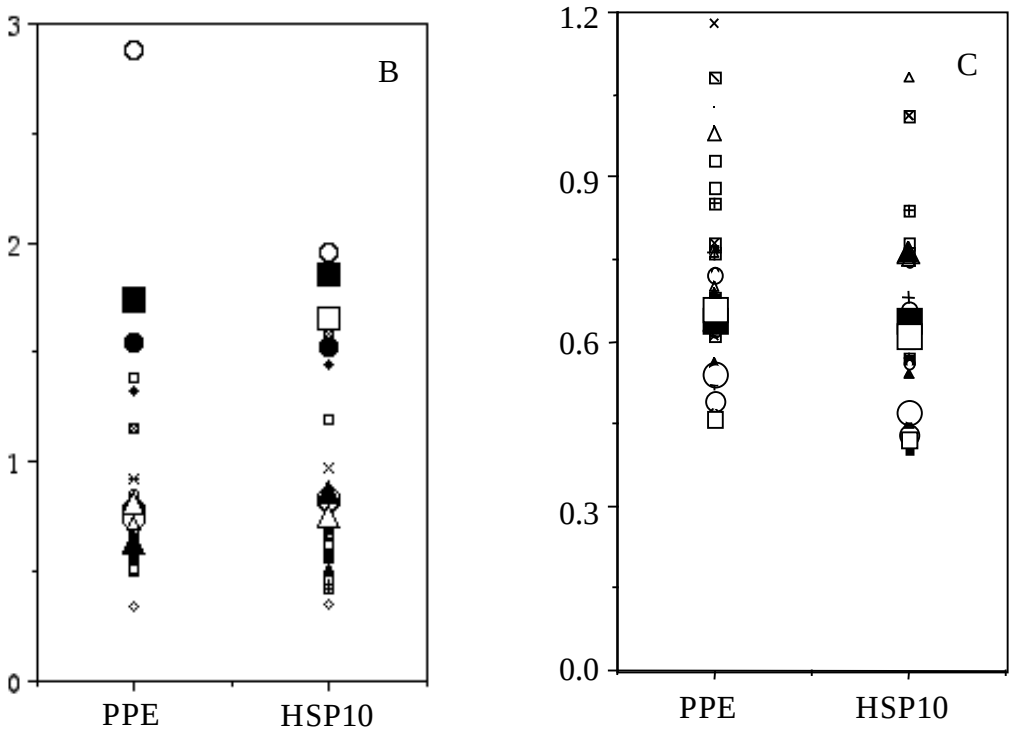


Category 1
(Fresh Infection)

Fig. 10A. The antibody response to Rv2430c is significantly higher than that elicited by Hsp10 in patients with fresh infection with *M.tuberculosis*. Reactivity to both recombinant Rv2430c and Hsp10 of *M.tuberculosis* in Category 1 patients was estimated by ELISA.

interest to compare the reactivity of Rv2430c protein in the four classified categories of patients, namely : 1) patients with fresh infection, 2) patients with relapsed tuberculosis, 3) extrapulmonary cases, and 4) Multi Drug resistant TB patients. Hence, recombinant PPE protein was used to screen sera obtained from all the four patient categories. In order to ascertain the immunodominant nature of Rv2430c, Hsp10, a well documented immunodominant antigen of *Mycobacterium tuberculosis*, was selected and the serological reactivities of both the antigens to the sera obtained from all the four clinical categories were compared. Results presented earlier (Fig. 9) showed that Rv2430c elicits a strong immune response in the sera obtained from a pool of TB infected patients. The antibody response to Rv2430c and Hsp10 was comparable in relapsed tuberculosis (Fig, 10B), extrapulmonary cases (Fig.10C) and Multi Drug Resistant TB patients (Fig. 10D). However in the case of fresh infection cases (Fig. 10A), the immune response to Rv2430c was significant as compared to Hsp10 ($P<0.003$). These results (Fig. 10A, B, C, D) convincingly demonstrate the immunodominant nature of Rv2430c.

3.4.3 Serological sensitivity of Rv2430c: In order to compare the serological sensitivity of Rv2430c over Hsp10, data obtained from the screening of categories 1, 2 and 3 were recalculated as percentage of



Category 2
(Relapsed)

Category 3
(Extra pulmonary)

Fig. 10B and 10C. Rv2430c and Hsp10 mount more or less similar immune response in patients with relapsed and extrapulmonary TB. Reactivity to both recombinant Rv2430c and Hsp10 of *M.tuberculosis* in Categories 2 and 3 was estimated by ELISA.

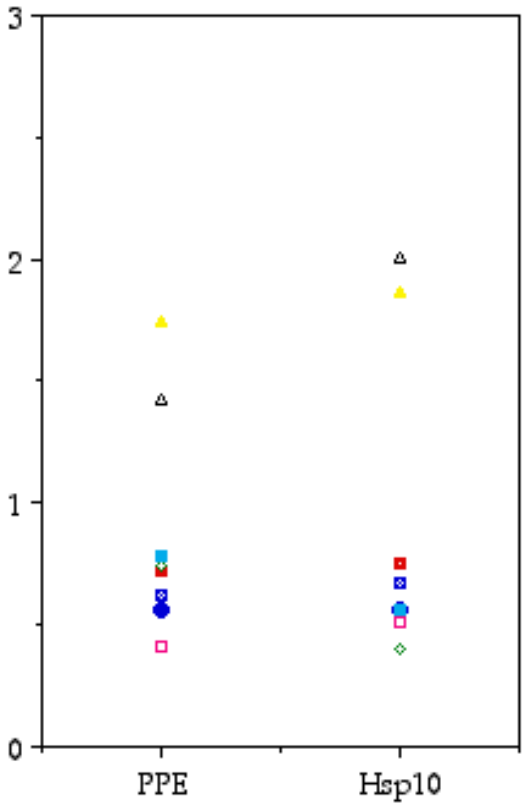


Fig. 10D. Patients belonging to the MDR category TB also displayed similar antibody responses to both the antigens.

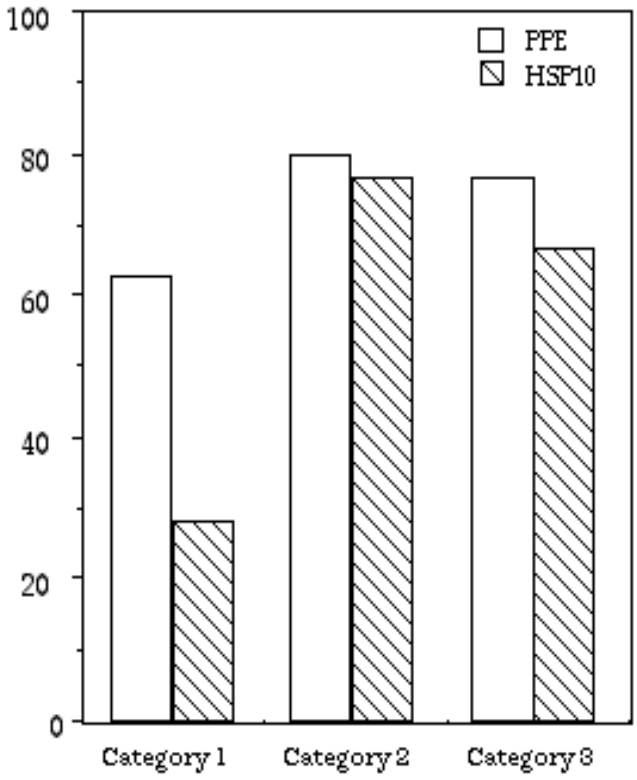


Fig. 11. A higher percentage of individuals belonging to Category 1, but not Category 2 and Category 3, show stronger reactivity to Rv2430c than for Hsp10. Immunological response to PPE Rv2430c and Hsp10 as a function of % of individuals was calculated. Results described in Figs. 8A, B and C were recalculated as percentage individuals showing absorbance value above 0.65 at 492 nm.

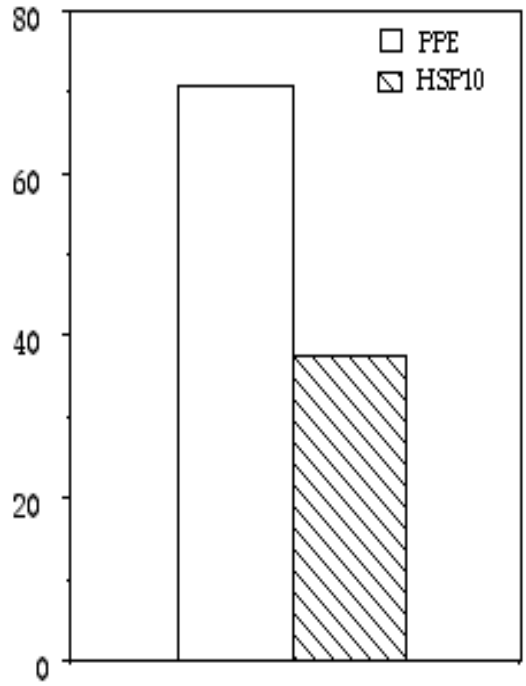


Fig. 12. A higher percentage of individuals show anti-IgM antibody against PPE Rv2430c antigen as compared to the Hsp10. Percentage individuals showing anti-IgM antibody above 0.5 absorbance at 492 nm in Category 1 is presented.

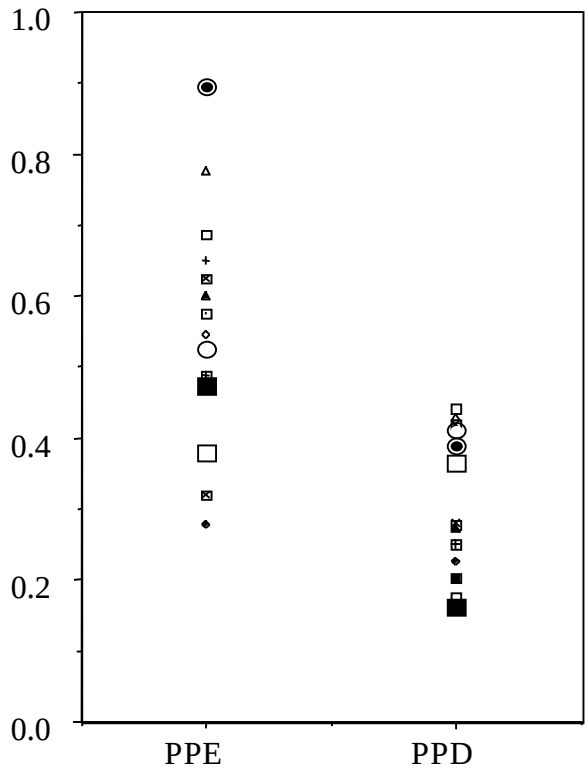


Fig. 13. Patients belonging to Category 1 show significantly higher antibody responses to the Rv2430c protein as compared to the PPD antigen. ELISA reactivity of sera of *M. tuberculosis* infected patients belonging to category 1 to Rv 2430c and PPD was scored.

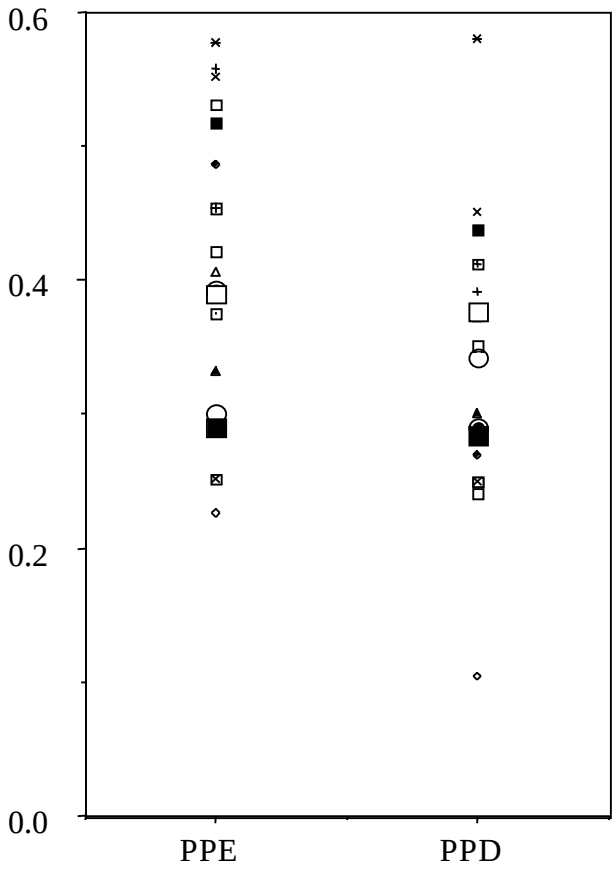


Fig. 14. Patients belonging to Category 3 show more or less identical antibody responses to both Rv2430c protein and the PPD antigen. ELISA reactivity of sera of *M. tuberculosis* infected patients belonging to Category 3 to Rv2430c and PPD was scored.

individuals showing absorbance above 0.65 at 492 nm. Though the reactivities to Categories 2, 3, 4 between Rv2430c and Hsp10 were comparable, a higher percentage of individuals belonging to Category 1 mounted antibody response of IgG type to Rv2430c as compared to Hsp10 [(62.5 v/s28.12,), (Fig. 11)]. Similar results were obtained when IgM antibodies were assayed against Rv2430c and Hsp10 in fresh infection cases (Fig. 12). From these results it is clear that Rv2430c mounts a stronger antibody response as compared to Hsp10, to patients in fresh infection cases as compared to relapsed, extrapulmonary and MDR TB cases.

3.4.4. Rv2430c is more immunodominant than PPD: The Purified Protein Derivative (PPD) antigen of *Mycobacterium tuberculosis* has often been used to diagnose TB. Therefore, the immunodominancy of Rv2430c *vis-a-vis* PPD was compared with respect to patients belonging to Categories 1 and 3. The results demonstrate that Rv2430c elicits strong and statistically significant immune response as compare to PPD in the case of fresh infection cases whereas a comparable immune response was observed in the case of extrapulmonary TB patients (Fig. 13 and 14) .

3.4.5. Recombinant Rv2430c can differentiate between TB and non TB patients: In order to asses the diagnostic utility of Rv2430c,

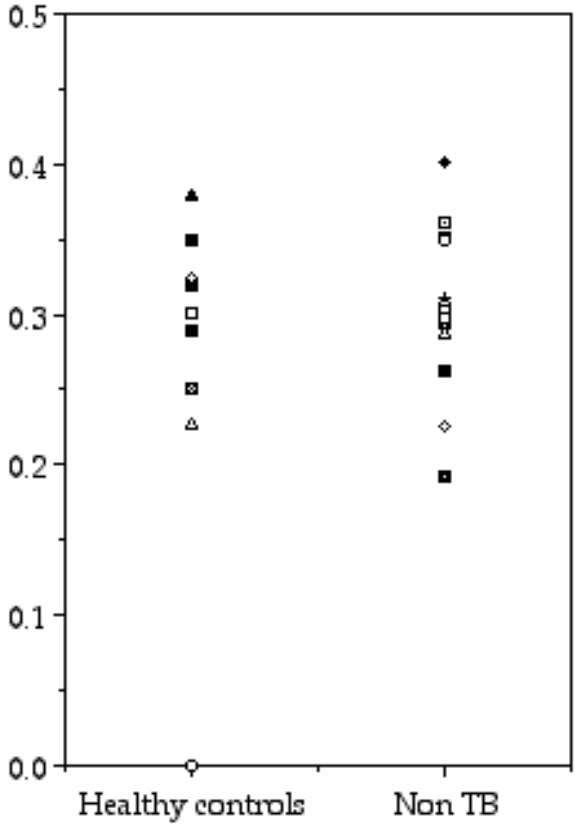


Fig. 15. The response to Rv2430c in sera from both healthy controls and Non TB infected group was statistically insignificant ($p<0.92$). Immunoreactivity of recombinant Rv2430 was scored against sera obtained from healthy controls and non TB infected patients.

the potential of recombinant Rv2430c to specifically distinguish between TB and non-TB patient sera (those essentially culture negative for *Mycobacterium tuberculosis* but harboring other pathogens) was tested by examining the cross-reactivity of the recombinant protein with non-TB patient sera and comparing the reactivity with *Mycobacterium bovis* vaccinated healthy controls and infected categories. The non-TB patient sera included patients which tested negative for tuberculosis by staining and culture based techniques but harboring infections other than tuberculosis. Fifteen non-TB patient sera were checked for their immunogenic response against recombinant Rv2430c using ELISA and the reactivity was compared with *M.bovis* BCG vaccinated healthy controls and infected patients. The reactivity of Rv2430c with non-TB patient sera as seen by ELISA absorbance values was found to be similar or almost equal to that of healthy controls [($p<0.92$) (Fig. 15)]. Since the non-TB patients (those harboring other pathogens but not *M. tuberculosis*) are also BCG vaccinated, the reactivity of the recombinant protein to non-TB patient sera mimicked that of BCG vaccinated healthy controls. However the recombinant Rv2430c showed significant reactivity with samples obtained from infected subjects. The immune response against the Rv2430c was statistically significant in the case of all the three infected categories as compared to the non-TB patients ($p<0.0001$ in all the three categories). These results

demonstrate the presence of negligible levels of anti-Rv2430c antibodies in the sera of non-TB patients as compared to the infected patients. Thus, our results illustrate that Rv2430c can distinguish between infected and non-TB patients and can be used as a diagnostic marker for TB.

3.4.6. Rv2430c is a cytosolic protein as evident from the coupled *in vitro* transcription and translation system: Results obtained from the serological characterization of Rv2430c using the sera obtained from infected patients clearly demonstrated the immunodominant nature of this protein. With a strong B cell response it is highly probable that Rv2430c is a secretory protein. Hence we employed a coupled *in vitro* transcription and translation system to ascertain the probable nature of Rv2430c with respect to its localization. *In vitro* translation of the ORF Rv2430c, using Rabbit reticulocytes, under the control of the T7 promoter yielded a protein of expected size, i.e ~23kDa (Fig. 16). Digestion with Trypsin resulted in the complete degradation of this protein. *In vitro* translation in the presence of Canine Pancreatic Microsomal membrane did not result in the reduction of the size of the protein, pointing to the unlikely possibility of the protein having a signal peptide. Furthermore, Trypsin digestion of Rv2430c translated *in vitro* in the presence of Canine Pancreatic Microsomes also resulted in complete digestion of

this protein indicating the absence of transmembrane regions in this protein as predicted by web based analyses. This therefore, demonstrates that Rv2430c is a cytosolic protein.

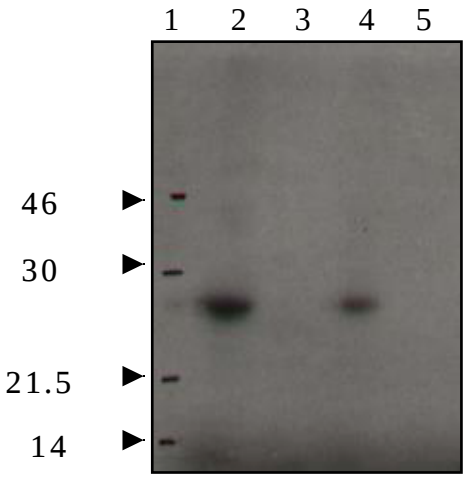


Fig. 16. Rv2430c appears to be a cytosolic protein. Coupled *in vitro* transcription-translation of Rv2430c was carried out to determine the localization of the recombinant protein. Rv2430c was translated *in vitro* in the presence and absence of Canine Pancreatic Microsomal membrane and the reactions were trypsinized in order to check the presence and absence of secretory and transmembrane signals. The different lanes are : Lane 1, Marker; Lane 2, *In vitro* translated protein; Lane 3, Trypsinized translated protein; Lane 4, *In vitro* translation of the protein in the presence of Canine Pancreatic Microsomal membrane; Lane 5, Trypsinization of the protein translated in the presence of Canine Pancreatic Microsomal membrane.

CHAPTER 4

DISCUSSION

DISCUSSION

One third, approximately 1.7 billion people, of the world's population is infected with tuberculosis. Tuberculosis is the number one cause of mortality from infectious agents world-wide, killing an estimated 3 million people each year, more than 10 times the number of deaths due to AIDS. Global tuberculosis control will unlikely come from development of new antituberculosis drugs because they will be too expensive to be used in developing countries. Control of tuberculosis depends on development of an effective antituberculosis vaccine. Development of a safer and more effective vaccine has been made feasible by two recent developments. First, it was found that immunization of guinea pigs with a crude mixture of proteins secreted by the tuberculosis bacterium protected against development of tuberculosis. Second, in June 1998, the genome of the tuberculosis bacterium, *Mycobacterium tuberculosis*, was sequenced, and the post genomic era may allow identification of new genes that may produce proteins that elicit protective immunity (Ahmed and Hasnain, 2004, Chakhaiyar and Hasnain, 2004). These advances will make it possible to make a "subunit" protein vaccine, which will be much safer than a live vaccine. However, it is essential to first identify the best vaccine candidate proteins, and second, to develop methods to produce large quantities of these proteins.

4.1. The PE and PPE gene families of *Mycobacterium tuberculosis*

Among the most interesting gene families found in the mycobacteria are the PE and PPE families. Ten percent of the *Mycobacterium tuberculosis* genome is devoted to these genes, encoding acidic glycine-rich proteins. The field is rife with speculation on the functions of these proteins, including their functions to provide antigenic variation and interfere with immune responses. The amino acid composition of these protein families differs radically from that of the bulk of the proteins of the Mycobacterial genome. Both the PE and PPE proteins are exceptionally glycine-rich while the PPE proteins also contain copious amounts of asparagine, an amino acid that is generally rare in the proteome. Curiously, asparagine is the preferred nitrogen source for *Mycobacterium tuberculosis*, and this raises the possibility that the PPE proteins may also serve as storage proteins. Prior to completion of the genome sequence, the existence of these protein families, whose genes occupy 10% of the total coding sequence, was unknown. It was clear, however, that the genome contained two dispersed simple sequence repeats referred to as PGRS (polymorphic G+C-rich sequence) and MPTR (major polymorphic tandem repeat) and these have since been shown to

correspond to part of the 3´ ends of the PE and PPE genes (Cole *et al*, 1998).

Review of literature reveals substantial information regarding the PE families of proteins. Banu *et al* showed 10 of these genes to be variable surface antigens (Banu *et al*, 2002). A significant humoral immune response was observed against recombinant Rv1818, a member of the PE-PGRS family (Brennan *et al*, 2001). Differential expression of PE and PE_PGRS genes in *Mycobacterium tuberculosis* strains has also been reported (Flores and Espitia, 2003). Singh *et al* showed that Rv3367, a member of the PE-PGRS family is recognized by pooled sera from TB patients and not from healthy controls, confirming its *in vivo* expression during active infection in humans (Singh *et al*, 2001). pH dependent expression of Rv0834c, a member of the PE-PGRS family of *Mycobacterium tuberculosis* has been reported (Saviola *et al*, 2003). However very few reports are available regarding the PPE family of *M.tuberculosis*. The primary objective of this study therefore, was to identify novel immunodominant antigen(s) within the PPE family of *M.tuberculosis*.

4.2. *In silico* and web based analyses

The 69 members of the PPE protein family have a conserved N-terminal domain that comprises ~180 amino acids followed by C-terminal segments that vary markedly in sequence and length. Based on our pattern search analysis of the Tuberculist database (http://genolist.pasteur.fr/TubercuList/) these proteins were categorised into three groups. Subgroup 1, represented by 20 members, constitutes the MPTR class characterized by the presence of multiple, tandem copies of the motif Asn-X-Gly-X-Gly-Asn-X-Gly. The second subgroup, comprising of 21 members, contains a characteristic well conserved motif Gly-X-X-Ser-Val-Pro-X-X-Trp around position 350, and the third subgroup proteins, with 28 members, are unrelated except for the presence of the common PPE domain. Proteins belonging to the third sub group was the focus of the present study as members of these sub-group display significant sequence and length variation in their C-terminal region and are unrelated except for the presence of the common PPE domain.

ORFs belonging to the third subgroup with coding capacity equal to or less than 200 amino acids were shortlisted. This shortlist was further narrowed down based on two very important criteria –

namely, antigenic profile and the association of the ORF with pathological conditions.

Antigenic profiles were generated using the Protein Analysis Software, Protean, DNASTAR. Two ORFs, Rv2430c and Rv3425 displayed major antigenic stretches. Furthermore, DNA Microarray results demonstrated that Rv2430c was one of the genes induced in IdeR mutant of *M. tuberculosis* (Rodriguez *et al*, 2002), pointing to its possible role in pathogenesis. Rv2430c was accordingly shortlisted for the present study and was evaluated for its role as an antigen in a clinical setting. To confirm whether Rv2430c indeed represented a functional gene, mRNA extracted from H37Rv, the virulent strain of *Mycobacterium tuberculosis* was used in an RT-PCR reaction. The RT-PCR reaction was found to be positive indicating the expression of Rv2430c in liquid cultures of *Mycobacterium tuberculosis*.

The ORF corresponding to Rv2430c was amplified from the genomic DNA of H37Rv and cloned in an *E.coli* expression vector. The recombinant protein was purified from *E.coli* under denaturing conditions and was used for further studies.

4.3. Biophysical characterization of the PPE ORF Rv2430c

Though there have been several reports about immunodominant antigens of *Mycobacterium tuberculosis*, information about the structure of antigens is very little. Ronning *et al* have described the crystal structure of the secreted form of antigen 85C (Ronning *et al*, 2000). The antigen 85 (ag85) complex, composed of three proteins (ag85A, B and C), is a major protein component of the *Mycobacterium tuberculosis* cell wall. Each protein possesses a mycolyltransferase activity required for the biogenesis of trehalose dimycolate (cord factor), a dominant structure necessary for maintaining cell wall integrity. The crystal structure of recombinant ag85C from *M. tuberculosis*, refined to a resolution of 1.5 A, reveals an alpha/beta-hydrolase polypeptide fold, and a catalytic triad formed by Ser 124, Glu 228 and His 260. ag85C complexed with a covalent inhibitor implicates residues Leu 40 and Met 125 as components of the oxyanion hole. Renshaw *et al* have determined the secondary structures of ESAT-6 and CFP-10, secretory proteins of *Mycobacterium tuberculosis*, which are potent B cell and T cell antigens (Renshaw *et al*, 2002). Using a combination of fluorescence, circular dichroism, and nuclear magnetic resonance spectroscopy they have shown that ESAT-6 contains up to 75% helical secondary structure, but little if any stable tertiary structure, and exists in a

molten globule-like state. In contrast, CFP-10 was found to form an unstructured, random coil polypeptide. An exciting discovery was that ESAT-6 and CFP-10 form a tight, 1:1 complex, in which both proteins adopt a fully folded structure, with about two-thirds of the backbone in a regular helical conformation. This clearly suggests that ESAT-6 and CFP-10 are active as a complex. Goulding *et al* have determined the crystal structure MPT63, a small, major secreted protein of unknown function from *Mycobacterium tuberculosis* at 1.5 A resolution that has been shown to have immunogenic properties and has been implicated in virulence (Goulding *et al,* 2002). The structure of MPT63 is an antiparallel beta-sandwich immunoglobulin-like fold, with the unusual feature of the first beta-strand of the protein forming a parallel addition to the small antiparallel beta-sheet. MPT63 has weak structural similarity to many proteins with immunoglobulin folds, in particular, *Homo sapiens* beta2-adaptin, bovine arrestin, and *Yersinia pseudotuberculosis* invasin. Although the structure of MPT63 gives no conclusive evidence to its function, structural similarity suggests that MPT63 could be involved in cell-host interactions to facilitate endocytosis/phagocytosis. In another study, the purification and characterization of three immunodominant antigens of *Mycobacterium tuberculosis*, namely the 38kDa, 30kDa and 16kDa antigens have been described. The 38-kDa form was purified by

preparative isoelectric focusing, followed by preparative electrophoresis (Devi *et al,* 2002). Gel-permeation chromatography was employed for the isolation of the 16-kDa form, from the cytosol fraction of *M. tuberculosis* H37Rv. The purified proteins were characterized by Circular Dichroism studies. Analysis of the CD data revealed that the 38kDa and the 16kDa proteins have a beta sheet like structure.

No information, however, is available regarding the biophysical or structural features of the PE and PPE family of proteins. Hence, the present study was also carried out to gain insights into the structural characteristics of the recombinant PPE protein. Secondary structure prediction employing Protean 4.0 revealed a very high content of α- helical structure in Rv2430c. This was evident from both the Garnier-Robson and Chou-Fasman methods of secondary structure prediction. Web based analysis of Rv2430c using PSIPRED and Predict protein also suggested a predominant α helical composition. To confirm these predictions, the recombinant PPE protein was extracted under denaturing conditions from *E.coli* as expression of the protein led to its localization in the inclusion bodies. Initial attempts to refold the protein using dialysis resulted in massive precipitation of the protein. Therefore, an on-column refolding strategy (Pullakhandham *et al,* 2004) was used in the

presence of L-arginine that is known to act as a chemical chaperone (Srinivas *et al*, 2003), and glutathione was included to provide reducing equivalents during folding. The on-column refolded protein was soluble and was found to be pure by SDS-PAGE analysis. The purified protein displayed a CD spectra characteristic of α helical proteins, confirming the *in silico* predictions of secondary structure. The possible environment of the aromatic amino acid residues in refolded protein was also checked. The refolded protein exhibited emission maximum at 340nm. Interestingly, incubation of protein with urea resulted in the significant red-shift (340 to 350nm) in the emission maxima and also an increase in the fluorescence intensity was observed, indicating that the aromatic amino acid residues are present in a hydrophobic environment. Thus, there appears to be a well-formed hydrophobic core in the protein, which becomes exposed under the influence of 8M urea. In the absence of any functional assay described for this protein, conformation of the protein analyzed by CD and fluorescence studies compared with *in silico* predictions of protein based on primary structure suggest that the protein is properly folded.

The current method of refolding can be employed for other members of the PE/PPE protein family to obtain large quantities of protein for crystallization purposes. Such on-column refolding strategy has been

successfully used for a protein which is known to generate oligomers (Pullakhandham *et al*, 2004). That such properly folded protein will likely display differential immunoreactivity to patient sera renders it possible to evaluate other members of this family for their likely biological roles.

4.4. Serological characterization of the PPE ORF Rv2430c

Having shown that the ORF Rv2430c was expressed at the mRNA level in liquid cultures of *M.tuberculosis*, experiments were designed to assess its role in eliciting a humoral immune response. The recombinant His-tagged protein was used to assess its immunological potential. Human sera obtained from a total of 101 patients were used in the present study. These patients belonged to Fresh infection cases (Category 1, n=32), Patients with relapsed tuberculosis (Category 2, n=30), Extrapulmonary cases (Category 3, n=30) and Multi Drug Resistant Cases (Category 4, n=9). Also included were 10 clinically healthy controls which were *M.bovis* BCG Vaccinated and sera from 15 non-TB patients (culture negative for *Mycobacterium tuberculosis*). Initial screening of the sera revealed that the recombinant protein coded for by Rv2430c was recognized by infected patients at a serum dilution of 1:200 where as a poor reactivity was observed in clinically healthy controls. This result

confirmed that Rv2430c is expressed in active infection with *M.tuberculosis* and is involved in disease manifestation and progression.

Having demonstrated that the PPE ORF Rv2430c is associated with disease manifestation and progression, it was of interest to dissect the immune response to Rv2430c. For this the recombinant protein was used to screen the panel of sera obtained from all the four clinical groups. We used Hsp10, a well documented immunodominant antigen of *M.tuberculosis* (Young and Garbe, 1991) as a reference to compare the immune response against Rv2430c in all the four clinical categories. Screening of the sera revealed that recombinant protein Rv2430c was well recognized by the sera of all the four clinical categories. However, as compared to Hsp10, patients belonging to Category 1 mounted a strong immune response to Rv2430c ($P<0.003$). The immunoreactivity to Rv2430c in the remaining three categories, however was comparable to Hsp10 with values of $P<0.8$, $P<0.06$ and $P<0.912$ for Categories 2, 3, and 4, respectively. Serological sensitivity of Rv2430c as compared to Hsp10 was demonstrated by the fact that a higher percentage of individuals belonging to Category 1 mounted a strong antibody response to Rv2430c as compared to Hsp10. Similar results were obtained when IgM antibodies were assayed.

The Purified Protein Derivative has often been used to diagnose infection with TB. This test has remained essentially unchanged since it was developed 100 years ago by Robert Koch, who also discovered *M. tuberculosis*. Intradermal inoculation of purified protein derivative (PPD, or tuberculin), a crude precipitate of *M. tuberculosis* culture supernatant containing over 200 antigens widely shared among mycobacteria other than *M. tuberculosis*, including *M. bovis* BCG, elicits a local cutaneous delayed type hypersensitivity response in sensitised individuals. This *in vivo* cellular immune response results in a bump in the skin. The size of the bump, or induration, is measured using a scale three to seven days after the test is placed. A large bump is interpreted as indicating *M. tuberculosis* infection, while the absence of a bump implies no infection. Intermediate sized bumps are difficult to interpret. Several reports have emphasized the utility of PPD in the serodiagnosis of tuberculosis. Zeiss *et al* obtained sera from patients with active tuberculosis and sera from appropriate control individuals and screened the same for immunoglobulin G antibody activity to purified protein derivative by a polystyrene tube radioimmunoassay and an enzyme-linked immunosorbent assay. Both assays showed a marked increase in immunoglobulin G antibody activity in patients with active tuberculosis. There was no overlap between the values for the patient group and the values for the purified protein derivative skin

test-positive control individuals. The replication of these assays was excellent, and both could provide quantitative measurements of immunoglobulin G antibody activity to purified protein derivative antigen within 24 h (Zeiss *et al*, 1982). Radin *et al* assayed the sera of patients with active tuberculosis (TB) and sera from control groups for IgG, IgA, secretory IgA, IgM and IgE antibody activity to purified protein derivative (PPD) using the enzyme-linked immunosorbent assay. Patients with active TB clearly had higher levels of IgG antibody activity to PPD antigen than did healthy patients who were skin test positive or negative. There was a clear separation between the diseased and healthy groups. Similar, but not as marked, increase was seen in IgA and secretory IgA antibody activity in diseased patients. No correlation between the presence of disease and antibody levels were found with IgM, and no IgE antibodies were found (Radin *et al*, 1983). Viljanen *et al* developed an Enzyme-linked immunosorbent assay (ELISA) for IgM, IgA and IgG antibodies against PPD. PPD antibodies of 44 patients with active pulmonary tuberculosis were measured at admission to hospital. The control material consisted of 35 healthy blood donors. The mean antibody levels in all three immunoglobulin classes correlated with the extent of the tuberculous infection, i.e. the severer the disease the higher the mean antibody level. The mean antibody levels also were significantly higher in the tuberculosis patients than in the controls,

the only exception was specific IgA antibodies in the patients with minimal disease. ELISA-positivity of the patients, i.e. positive result in at least one immunoglobulin class was as follows: 33% of the grade I, 64% of the grade II and 100% of the grade II patients (Viljanen *et al*, 1982). Kalish *et al* studied three patients with culture-proven *Mycobacterium tuberculosis* meningitis. Analysis of cerebrospinal fluid with an enzyme-linked immunosorbent assay (ELISA) method measuring IgG antibody to purified protein derivative rapidly yielded positive results, whereas results of acid-fast smears were negative and cultures took several weeks before growth appeared. They carried out serial studies of cerebrospinal fluid and sera from one patient. Initially, greater amounts of IgG antibody to purified protein derivative were present in the cerebrospinal fluid than in the serum. The antibody level in the cerebrospinal fluid paralleled the patient's clinical course, cerebrospinal fluid cell count, protein level, and glucose level. Cerebrospinal fluid samples from 33 hospitalized control patients were negative for antibody to purified protein derivative (Kalish *et al*, 1983a; Kalish *et al*, 1983b).

4.5. Diagnostic potential of PPE ORF Rv2430c

The diagnostic utility of PPD has become a question of debate primarily because of false positives it generates in the case of healthy

controls. The broad antigenic cross-reactivity of PPD is responsible for the poor specificity of the Tuberculin Skin Testing (TST); a positive reaction is consistent with BCG vaccination as well as *M. tuberculosis* infection. Since one third of the world's population is believed to be infected with *M. tuberculosis* and the majority have been BCG-vaccinated, accurate identification of *M. tuberculosis* infected people for targeted chemoprophylaxis is very difficult. There are also numerous operational drawbacks with the TST: administration and reading of the TST are highly operator dependent, standardization of PPD is problematic and a strongly positive TST can cause painful ulceration and scarring, Taking into account the above facts the need for identification of novel antigens which can also serve the purpose of serodiagnosis has become one of the primary goals in *M.tuberculosis* research. We compared the immunopotentiality of Rv2430c with PPD. Interestingly, patients belonging to Category 1 mounted statistically significant immune response to Rv2430c as compared to PPD.

In order to confirm the diagnostic ability of Rv2430c, we compared the reactivity of Rv2430c in sera obtained from clinically healthy donors and non-TB patients. The non-TB patients were confirmed TB negative both by acid fast staining and culture and therefore they could be harboring any pathogen other than *M.tuberculosis*. ELISA

results revealed that Rv2430c displayed similar reactivity to sera obtained from clinically healthy donors and non-TB patients. However the reactivity to the sera obtained from TB patients was statistically significant as compared to sera obtained from clinically healthy controls and non-TB patients indicating the presence of antibodies against Rv2430c in infected patients and their absence in non-TB patients. These results clearly indicate that Rv2430c does not cross react with the sera of non-TB patients and can differentiate between infected and non-TB patients.

Several reports have emphasized on the observation of lack of sufficient immune responses in TB patients against many promising serodiagnostic antigens of *M. tuberculosis*. The fact is more distressing in case of fresh infection or active infection where for majority of the cases the immune system is not sufficiently primed to elicit a strong antibody responses against most of the *M.tuberculosis* antigens. The recombinant Rv2430c protein was very strongly recognized by all the four categories of patients including the fresh infection or so called active infection group.

The development of a sensitive and rapid serodiagnostic test in tuberculosis (TB) would complement present methods of diagnosis including skin testing, DNA amplification, bacterial culture and

radiological imaging. The major focus of tuberculosis research also includes serodiagnosis of tuberculosis, since an early knowledge about the patient's disease would be an effective measurement for global control of tuberculosis. The ELISA based serodiagnosis appears to be rapid and inexpensive assay which could reduce the cost of diagnosis too.

It is pertinent to note that although several antigens have been tested for their use in serodiagnosis (Amara *et al,* 1998, Batoni *et al,* 2002, Devi *et al,* 2002, Dillon *et al,* 2000, Florio *et al,* 2002, Houghton *et al,* 2002, Laal *et al,* 1997, Ljungqvist *et al,* 1990, Lim *et al,* 2000, Lodes *et al,* 2001, Mustafa *et al,* 2002), no single antigen has proved to be able to achieve sensitivity and specificity in a study population suitably large and heterogenous. The factors responsible include a) the stage of the disease, b) the location of the infection, and c) the genetic background. Our results reveal that Rv2430c is expressed during active infection with TB and shows better reactivity to sera from fresh infection or the so called active infection cases when compared to Hsp10 and PPD. These results clearly demonstrate the immunodominant as well as immunodiagnostic nature of Rv2430c. It would be interesting to speculate on the use of Rv2430c alongwith other immunodominant antigens for vaccine development (Dhar *et al,* 2000). The PPE ORF Rv2430c used in our

study belongs to the third subgroup of the PPE class of proteins. Members of this subgroup are unrelated. Given the known antigenic variability between the different members of this family the antigenic epitopes will not be common. Therefore, the chances of cross-reactivity would be minimum, if at all.

Data obtained from microarray studies reveal that Rv2430c is one of the genes upregulated in Rel Mtb knock out of *Mycobacterium tuberculosis* (Dahl *et al*, 2003). Long-term survival of nonreplicating *Mycobacterium tuberculosis* is ensured by the coordinated shutdown of active metabolism through a broad transcriptional program called the stringent response. In *M.tuberculosis*, this response is initiated by the enzymatic action of RelMtb and deletion of RelMtb produces a strain (H37RvΔrelMtb) severely compromised in the maintenance of long-term viability. Microarray analysis revealed that H37RvΔrelMtb suffers from a generalized alteration of the transcriptional apparatus, as well as specific changes in the expression of virulence factors, cell-wall biosynthetic enzymes, heat shock proteins, and secreted antigens that may alter immune recognition of the recombinant organism. Hence, Rv2430c could be one of the genes responsible for the long term survival of non replicating *Mycobacterium tuberculosis*.

In order to ascertain the probable localization of Rv2430c, the ORF was cloned in pET23a expression vector and experiments were designed to ascertain its localization using a Rabbit Reticulocyte Lysate cell free coupled transcription and translation system. *In vivo* translation yielded a product of the expected size indicating lack of any post translational modification of the protein. Translation in the presence of Canine Pancreatic Microsomal membrane was set up to determine the localization of the protein. No reduction in size of the protein was observed in the above reaction pointing to the lack of a secretory signal. Digestion of the protein translated in the presence of Canine Pancreatic Microsomal membrane by Trypsin led to the complete digestion of this protein indicating the absence of any functional transmembrane domain in the protein. The Rv2430c, originally considered a hypothetical protein, but based on results presented in this thesis, could be ascribed a possible role in acting as an immunodominant antigen. On the question of localization of this protein, *in silico* analysis points to the unlikely possibility of Rv2430c being a transmembrane protein. With a strong B cell response it is highly probable that Rv2430c is a secretory protein.

SUMMARY

The PPE gene Rv2430c is expressed in liquid cultures of *Mycobacterium tuberculosis* suggesting that Rv2430c is not a hypothetical gene.

The gene product of Rv2430c was expressed in *E.coli* and the recombinant protein was purified in soluble form using an on-column folding strategy.

The gene product of the hypothetical PPE Rv2430c protein codes for a alpha helical protein as confirmed by the results of the Far UV CD spectroscopy.

The probable environment of the aromatic residues in the PPE Rv2430c protein is hydrophobic, which becomes exposed upon treatment with 8M urea as observed in the red shift.

Rv2430c elicits a strong immune response in all the four clinical categories of patients included in this study.

A strong immune response to Rv2430c was observed in the fresh infection cases as compared to Hsp10, whereas the reactivity to Hsp10 and Rv2430c in the remaining three categories of patients

was comparable, pointing to the immunodominant nature of this protein.

Rv2430c is more immunogenic than PPD in patients with fresh infection with TB whereas the immune response to Rv2430c and PPD was the same in the case of sera obtained from the extrapulmonary cases.

Rv2430c can be used as a possible diagnostic marker for active TB infection. The immune response to Rv2430c in the case of healthy controls and non-TB patients was comparable. Rv2430c, however mounted statistically significant immune response to sera obtained from TB patients including all the four infected categories, pointing to the ability of this ORF to differentiate between TB and non TB patients.

Rv2430c is a cytosolic protein as confirmed by Trypsin digestion of the product of coupled transcription and translation reaction in presence of Canine Pancreatic Microsomes.

In conclusion, data presented in this study describe for the first time, immunodominant as well as structural characteristics of a

hypothetical protein belonging to the PPE family of *Mycobacterium tuberculosis*.

REFERENCES

Abou-Zeid C, Garbe T, Lathigra R, Wiker HG, Harboe M, Rook GA and Young DB. **1991.** Genetic and immunological analysis of *Mycobacterium tuberculosis* fibronectin-binding proteins. Infect. Immun., 59:2712-2718.

Ahmed N and Hasnain SE. **2004**. Genomics of *Mycobacterium tuberculosis*: Old threats and new trends. Ind. Jour. Med. Res., 00: 000-000 (In press).

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.

Amara RR, Shanti S, and Satchidanandam V. **1998.** Characterization of novel immunodominant antigens of *Mycobacterium tuberculosis*. Microbiology. 144:1197-1203.

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

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

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

Batoni G, Bottai D, Esin S, Florio W, Pardini M, Maisetta G, Freer G, Senesi S and Campa M. **2002.** Purification, biochemical characterization and immunogenicity of SA5K, a secretion antigen of *Mycobacterium tuberculosis*. Scand. J. Immunol., 56:43-51.

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., 43:713-731.

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

Bothamley GH, Rudd R, Festenstein F and Ivanyi J. **1992.** Clinical value of the measurement of *Mycobacterium tuberculosis* specific antibody in pulmonary tuberculosis. Thorax, 47:270-275.

Bothamley G, Udani P, Rudd R, Festenstein F and Ivanyi J. **1988.** Humoral response to defined epitopes of tubercle bacilli in adult pulmonary and child tuberculosis. Eur. J. Clin. Microbiol. Infect. Dis., 7:639-645.

Brennan MJ, Delogu G, Chen Y, Bardarov S, Kriakov J, Alavi M, and Jacobs WR Jr. **2001.** Evidence that mycobacterial PE_PGRS proteins are cell surface constituents that influence interactions with other cells. Infect. Immun., 69:7326-7333.

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. U S A, 99:3684-3689.

Bujard H, Gentz R, Lanzer M, Stüber D, Müller M, Ibrahimi I, Häuptle MT and Dobberstein B. **1987.** A T5 promotor based transcription-translation system for the analysis of proteins in vivo and in vitro. Methods Enzymol., 155:416–433.

Coates AR, Nicolai H, Pallen MJ, Guy A, Chaparas SD and Mitchison DA. **1989.** The 45 kilodalton molecule of *Mycobacterium tuberculosis* identified by immunoblotting and monoclonal antibodies as antigenic in patients with tuberculosis. Br. J. Exp. Pathol., 70:215-225.

Chan J and Kaufmann SHE. **1994.** In Tuberculosis: Pathogenesis, Protection, and Control (ed. Bloom, B. R.). Am. Soc. Microbiol., Washington DC, 271–284.

Chakhaiyar P and Hasnain SE. **2004.** Defining the mandate of tuberculosis research in a post genome era. Medical Principles and Practice. (In press).

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 and Barrell BG. **1998**. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. Nature, 393:537-544. Erratum appeared in vol.396:190-198.

Dahl JL, Kraus CN, Boshoff HI, Doan B, Foley K, Avarbock D, Kaplan G, Mizrahi V, Rubin H and Barry CE 3rd. **2003.** The role of RelMtb-mediated adaptation to stationary phase in long-term persistence of *Mycobacterium tuberculosis* in mice. Proc. Natl. Acad. Sci. U S A, 100:10026-10031.

Damiani G, Biano A, Beltrame A, Vismara D, Mezzopreti MF, Colizzi V, Young DB and Bloom BR. **1988.** Generation and characterization of monoclonal antibodies to 28, 35, and 65 kDa proteins of *Mycobacterium tuberculosis*. Infect. Immun., 56:1281-1287.

Daniel TM, Ellner JJ, Todd LS, McCoy DW, Payne VD, Anderson PA and Bhe FT. **1979**. Immunobiology and species distribution of *Mycobacterium tuberculosis* antigen 5. Infect. Immun., 24 :77-82.

Daniel TM and Janicki BW. **1978.** Mycobacterial antigens: a review of their isolation, chemistry, and immunological properties. Microbiol. Rev., 42:84-113.

Daniel TM and Debanne SM. **1987.** The serodiagnosis of tuberculosis and other mycobacterial diseases by enzyme-linked immunosorbent assay. Am. Rev. Respir. Dis., 135:1137-1151.

Daniel TM. 1987. New approaches to the rapid diagnosis of tuberculous meningitis. J. Infect. Dis., 155:599-602.

Daniel TM. 1989. The chemical composition of immunoaffinity-purified *Mycobacterium tuberculosis* antigen 5. Am. Rev. Respir. Dis., 139:1566-1567.

Daniel TM, Ellner JJ, Todd LS, McCoy DW, Payne VD, Anderson PA and Bhe FT. **1979.** Immunobiology and species distribution of *Mycobacterium tuberculosis* antigen 5. Infect. Immun., 24:77-82.

Delogu G and Brennan MJ. **2001.** Comparative immune response to PE and PE_PGRS antigens of *Mycobacterium tuberculosis*. Infect. Immun., 69:5606-5611.

Devi KR, Kumar KS, Ramalingam B and Alamelu R. **2002.** Purification and characterization of three immunodominant proteins (38, 30, and 16 kDa) of *Mycobacterium tuberculosis*. Protein. Expr. Purif., 24:188-195.

Dhar N, Rao V, and Tyagi A.K. **2000**. Recombinant BCG approach for development of vaccines: cloning and expression of

immunodominant antigens of *Mycobacterium tuberculosis*. FEMS Microbiol. Lett., 190:309-316.

Dillon DC, Alderson MR, Day CH, Bement T, Campos-Neto A, Skeiky YA, Vedvick T, Badaro R, Reed SG, and R. L. Houghton. **2000.** Molecular and immunological characterization of *Mycobacterium tuberculosis* CFP-10, an immunodiagnostic antigen missing in *Mycobacterium bovis* BCG. J. Clin. Microbiol., 38:3285-3290.

Engers H, Bennedsen J and Buchanan T. **1986.** Results of a World Health Organization-sponsored workshop to characterize antigens recognized by mycobacterium specific monoclonal antibodies. Infect. Immun., 51:718-720.

Espitia C and Mancilla R. **1989.** Identification, isolation and partial characterization of *Mycobacterium tuberculosis* glycoprotein antigens. Clin. Exp. Immunol., 77:378-383.

Espitia C, Cervera I, Gonzalez R and Mancilla R. **1989.** A 38-kD *Mycobacterium tuberculosis* antigen associated with infection. Its isolation and serologic evaluation. Clin. Exp. Immunol., 77:373-377.

Fisher MA, Plikaytis BB and Shinnick TM. **2002.** Microarray analysis of the *Mycobacterium tuberculosis* transcriptional response to the acidic conditions found in phagosomes. J. Bacteriol., 184:4025-4032.

Flores J and Espitia C. **2003.** Differential expression of PE and PE_PGRS genes in *Mycobacterium tuberculosis* strains. Gene, 318:75-81.

Forte M, Maartens G, Rahelu M, Pasi J, Ellis C, Gaston H and Kumararatne D. **1992.** Cytolytic T-cell activity against mycobacterial antigens in HIV. AIDS, 4:407-411.

Florio W, Bottai D, Batoni G, Esin S, Pardini M, Maisetta 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 JL, Goldstein MM, Triebold KJ, Koller B and Bloom BR. **1992.** Major histocompatibility complex class I-restricted T cells are required for resistance to *Mycobacterium tuberculosis* infection. Proc. Natl. Acad. Sci. U S A, 89:12013-12017.

Flynn JL, Goldstein MM, Triebold KJ and Bloom BR. **1993.** Major histocompatibility complex class I-restricted T cells are necessary for protection against *Mycobacterium tuberculosis* in mice. Infect. Agents. Dis., 2:259-262.

Goulding CW, Parseghian A, Sawaya MR, Cascio D, Apostol MI, Gennaro ML and Eisenberg D. **2002.** Crystal structure of a major secreted protein of *Mycobacterium tuberculosis*-MPT63 at 1.5-A resolution. Protein. Sci., 11:2887-2893.

Harkin TJ and Harris HW. **1995**. Treatment of multidrug resistant tuberculosis, p. 843-850. *In* N. W. Rom and S. Garay (ed.), Tuberculosis, 1st ed. Little Brown and Company, Boston, USA.

Havlir DV, Ellner JJ, Chervenak KA and Boom WH. **1991.** Selective expansion of human gamma delta T cells by monocytes infected with live *Mycobacterium tuberculosis*. J. Clin. Invest., 87:729-733.

Havlir DV, Wallis RS, Boom WH, Daniel TM, Chervenak K and Ellner JJ. **1991.** Human immune response to *Mycobacterium tuberculosis* antigens. Infect. Immun., 59:665-670.

Hermans PW, van Soolingen D and van Embden JD. **1992.** Characterization of a major polymorphic tandem repeat in *Mycobacterium tuberculosis* and its potential use in the epidemiology of *Mycobacterium kansasii* and *Mycobacterium gordonae*. J. Bacteriol., 174:4157-4165.

Houghton RL, Lodes MJ, Dillon DC, Reynolds LD, Day CH, McNeill PD, Hendrickson RC, Skeiky YA, Sampaio DP, Badaro R, Lyashchenko KP and Reed SG. **2002.** Use of multiepitope polyproteins in serodiagnosis of active tuberculosis. Clin. Diagn. Lab. Immunol., 9:883-891.

Jackett PS, Bothamley GH, Batra HV, Mistry A, Young DB and Ivanyi J. **1988**. Specificity of antibodies to immunodominant mycobacterial antigens in pulmonary tuberculosis. J. Clin. Microbiol., 26:2313-2318.

Kalish SB, Radin RC, Levitz D, Zeiss CR and Phair JP. **1983a.** The enzyme-linked immunosorbent assay method for IgG antibody to purified protein derivative in cerebrospinal fluid of patients with tuberculous meningitis. Ann. Intern. Med., 99:630-633.

Kalish SB, Radin RC, Phair JP, Levitz D, Zeiss CR and Metzger E. **1983b.** Use of an enzyme-linked immunosorbent assay technique in the differential diagnosis of active pulmonary tuberculosis in humans. J. Infect. Dis., 147:523-530.

Kaufmann SHE. 1988. $CD8^+$ T lymphocytes in intracellular microbial infections. Immunol. Today, 9:168-174.

Khanolkar-Young S, Kolk AH, Andersen AB, Bennedsen J, Brennan PJ, Rivoire B, Kuijper S, McAdam KP, Abe C, Batra HV, et al. **1992.** Results of the third immunology of leprosy/immunology of tuberculosis antimycobacterial monoclonal antibody workshop. Infect. Immun., 60:3925-3927.

Kolk AH, Ho ML, Klatser PR, Eggelte TA, Kuijper S, de Jonge S, van Leeuwen J. **1984**. Production and characterization of monoclonal antibodies to *Mycobacterium tuberculosis*, *M. bovis* (BCG) and *M. leprae*. Clin. Exp. Immunol., 3:511-521.

Kumararatne DS, Pithie AS, Drysdale P, Gaston JS, Kiessling R, Iles PB, Ellis CJ, Innes J and Wise R. **1990**. Specific lysis of mycobacterial antigen bearing macrophages by class II MHC-

restricted polyclonal T cell lines in healthy donors or patients with tuberculosis. Clin. Exp. Immunol., 80: 314-323.

Laal S, Sharma YD, Prasad HK, Murtaza A, Singh S, Tangri S, Misra RS and Nath I. **1991.** Recombinant fusion protein identified by lepromatous sera mimics native *Mycobacterium leprae* in T-cell responses across the leprosy spectrum. Proc. Natl. Acad. Sci. U S A, 88:1054-1058.

Laal S, Samanich KM, Sonnenberg MG, Zolla-Pazner S, Phadtare JM and Belisle JT. **1997.** Human humoral responses to antigens of *Mycobacterium tuberculosis*: immunodominance of high-molecular-mass antigens. Clin. Diagn. Lab. Immunol., 4:49-56.

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. USA, 100:7213-7218.

Lewis K. **1994**. Multidrug resistance pumps in bacteria: variations on a theme. TIBS, 19:119-123.

Levitskaya J, Coram M, Levitsky V, Imreh S, Steigerwald-Mullen PM, Klein G, Kurilla MG and Masucci MG. **1995.** Inhibition of antigen processing by the internal repeat region of the Epstein-Barr virus nuclear antigen-1. Nature, 375:685-688.

Levitskaya J, Sharipo A, Leonchiks A, Ciechanover A and Masucci MG. **1997.** Inhibition of ubiquitin/proteasome-dependent protein degradation by the Gly-Ala repeat domain of the Epstein-Barr virus nuclear antigen 1. Proc. Natl. Acad. Sci. U S A, 94:12616-12621.

Levitsky V, Zhang QJ, Levitskaya J and Masucci MG. **1996.** The life span of major histocompatibility complex-peptide complexes influences the efficiency of presentation and immunogenicity of two class I-restricted cytotoxic T lymphocyte epitopes in the Epstein-Barr virus nuclear antigen. J. Exp. Med., 183:915-926.

Lim RL, Tan LK, Lau WF, Ming MC, Dunn R, Too HP and Chan L. **2000.** Cloning and expression of immunoreactive antigens from *Mycobacterium tuberculosis*. Clin. Diag. Lab. Immunol., 7:600-606.

Ljungqvist L, Andersen AB, Andersen P, Haslov K, Worsaae A, Bennedsen J and Heron I. **1990.** Affinity purification, biological

characterization and serological evaluation of defined antigens from *Mycobacterium tuberculosis*. Trop. Med. Parasitol., 41:333-335.

Ljungqvist L, Worsaae A and Heron I. **1988.** Antibody responses against *Mycobacterium tuberculosis* in 11 strains of inbred mice: novel monoclonal antibody specificities generated by fusions, using spleens from BALB.B10 and CBA/J mice. Infect. Immun., 56:1994-1998.

Lodes MJ, Dillon DC, Mohamath R, DayCH, Benson DR, Reynolds LD, McNeill P, Sampaio DP, Skeiky YA, Badaro R, Persing DH, Reed SG, and Houghton RL. **2001**. Serological expression cloning and immunological evaluation of MTB48, a novel *Mycobacterium tuberculosis* antigen. J. Clin. Microbiol., 39:2485-2493.

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, Feb 5 [Epub ahead of print].

Mauch H, Brehmer W, Sonneborn HH, Horn J and Wagner S. **1988.** Monoclonal antibodies selectively directed against the cell wall

surface of *Mycobacterium tuberculosis*. J. Clin. Microbiol., 26:1691-1694.

Meeker HC, Williams DL, Anderson DC, Gillis TP, Schuller-Levis G and Levis WR. **1989.** Analysis of human antibody epitopes on the 65-kilodalton protein of *Mycobacterium leprae* by using synthetic peptides. Infect. Immun., 57:3689-3694.

Munk ME, Gatrill AJ and Kaufmann SHE. **1990.** Target cell lysis and IL-2 secretion by gamma/delta T lymphocytes after activation with bacteria. J. Immunol., 145:2434-2439.

Mustafa AS, Cockle PJ, Shaban F, Hewinson RG and Vordermeier HM. **2002.** Immunogenicity of *Mycobacterium tuberculosis* RD1 region gene products in infected cattle. Clin. Exp. Immunol., 130:37-42.

Poulet S, and Cole ST. **1995a.** Repeated DNA sequences in mycobacteria. Arch. Microbiol, 163:79-86.

Poulet S and Cole ST. **1995b.** Characterization of the highly abundant polymorphic GC-rich-repetitive sequence (PGRS) present in *Mycobacterium tuberculosis*. Arch. Microbiol., 163:87-95.

Pullakhandham R, Ghosh S, Soundarya K, Haseeb A, Aruna B and Ehtesham NZ. **2004.** Dimerization of human recombinant resistin involves covalent and non-covalent interactions. Biochem. Biophys. Res. Commun., 313:652-656.

Radin RC, Zeiss CR and Phair JP. **1983.** Antibodies to purified protein derivative in different immunoglobulin classes in the diagnosis of tuberculosis in man. Int. Arch. Allergy. Appl. Immunol., 70:25-29.

Ramaswamy S and Musser JM. **1998**. Molecular genetic bases of antimicrobial agent resistance in *Mycobacterium tuberculosis*. Tubercle and Lung Disease, 79:3-29.

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.

Rindi L, Lari N and Garzelli C. **1999.** Search for genes potentially involved in *Mycobacterium tuberculosis* virulence by mRNA differential display. Biochem. Biophys. Res. Commun., 258:94-101.

Robertson BD and Meyer TF. **1992.** Genetic variation in pathogenic bacteria. Trends. Genet., 8:422–427.

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.

Ronning DR, Klabunde T, Besra GS, Vissa VD, Belisle JT and Sacchettini JC. **2000.** Crystal structure of the secreted form of antigen 85C reveals potential targets for mycobacterial drugs and vaccines. Nat. Struct. Biol., 7:141-146.

Salo WL, Aufderheide AC, Buikstra J and Holcomb TA. **1994**. Identification of *Mycobaterium tuberculosis* DNA in a pre-Colombian Peruvian mummy. Proc. Natl. Acad. Sci. USA, 19:2091-2094.

Sampson SL, Lukey P, Warren RM, van Helden PD, Richardson M and Everett MJ. **2001.** Expression, characterization and subcellular

localization of the *Mycobacterium tuberculosis* PPE gene Rv1917c. Tuberculosis (Edinb), 81:305-317.

Saviola B, Woolwine SC and Bishai WR. **2003.** Isolation of acid-inducible genes of *Mycobacterium tuberculosis* with the use of recombinase-based in vivo expression technology. Infect. Immun., 71:1379-1388.

Sela S, Thole JE, Ottenhoff TH and Clark-Curtiss JE. **1991.** Identification of *Mycobacterium leprae* antigens from a cosmid library: characterization of a 15-kilodalton antigen that is recognized by both the humoral and cellular immune systems in leprosy patients. Infect. Immun., 59:4117-4124.

Siddiqi NA, Mohammed S, 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. **2002.** Molecular characterization of multi-drug resistant isolates of *Mycobacterium tuberculosis* from patients in North-India. Antimicrobial Agents and Chemotherapy, 46:443-450.

Siddiqi N, Das R, Pathak N, Banerjee S, Ahmed N, Katoch V M, and Hasnain SE. **2004.** *Mycobacterium tuberculosis* isolate with a distinct

genomic identity overexpresses a TAP like efflux pump. Infection, 000:00-00 (In press).

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

Singh KK, Zhang X, Patibandla AS, Chien P Jr, and Laal S. **2001.** Antigens of *Mycobacterium tuberculosis* expressed during preclinical tuberculosis: serological immunodominance of proteins with repetitive amino acid sequences. Infect. Immun., 69:4185-4191.

Skeiky YA, Ovendale PJ, Jen S, Alderson MR, Dillon DC, Smith S, Wilson CB, Orme IM, Reed SG and Campos-Neto A. **2000.** T cell expression cloning of a *Mycobacterium tuberculosis* gene encoding a protective antigen associated with the early control of infection. J. Immunol., 165:7140-7149.

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.

Stüber D, Matile H, and Garotta G. **1990.** System for high-level production in *Escherichia coli* and rapid purification of recombinant proteins: application to epitope mapping, preparation of antibodies, and structure-function analysis. In: Immunological Methods, Lefkovits, I. and Pernis, B., eds., vol. IV, Academic Press, New York, pp. 121–152.

Srinivas V, Raman B, Rao KS, Ramakrishna T and Rao Ch M. **2003.** Structural perturbation and enhancement of the chaperone-like activity of alpha-crystallin by arginine hydrochloride. Protein. Sci., 12:1262-1270.

Van Vooren JP, Drowart A, De Bruyn J, Launois P, Millan J, Delaporte E, Develoux M, Yernault JC and Huygen K. **1992.** Humoral responses against the 85A and 85B antigens of *Mycobacterium bovis* BCG in patients with leprosy and tuberculosis. J. Clin. Microbiol., 30:1608-1610.

Van Vooren JP, Drowart A, De Cock M, Van Onckelen A, D'Hoop MH, Yernault JC, Valcke C and Huygen K. **1991.** Humoral immune response of tuberculous patients against the three components of

the *Mycobacterium bovis* BCG 85 complex separated by isoelectric focusing. J. Clin. Microbiol., 29:2348-2350.

Verbon A, Kuijper S, Jansen HM, Speelman P and Kolk AH. **1990.** Antigens in culture supernatant of *Mycobacterium tuberculosis*: epitopes defined by monoclonal and human antibodies. J. Gen. Microbiol.,136:955-964.

Verbon A, Kuijper S, Jansen HM, Speelman P and Kolk AH. **1992.** Antibodies against secreted and non-secreted antigens in mice after infection with live *Mycobacterium tuberculosis*. Scand. J. Immunol., 36:371-384.

Verbon A, Hartskeerl RA, Schuitema A, Kolk AH, Young DB and Lathigra R. **1992.** The 14,000-molecular-weight antigen of *Mycobacterium tuberculosis* is related to the alpha-crystallin family of low-molecular-weight heat shock proteins. J. Bacteriol., 174:1352-1359.

Viljanen MK, Eskola J and Tala E. **1982.** Enzyme-linked immunosorbent assay (ELISA) for antibodies to purified protein derivative of tuberculin (PPD). IgM-,IgA- and IgG- anti-PPD

antibodies in active pulmonary tuberculosis. Eur. J. Respir. Dis., 63:257-262.

Wallis RS, Paranjape R and Phillips M. **1993.** Identification by two-dimensional gel electrophoresis of a 58-kilodalton tumor necrosis factor-inducing protein of *Mycobacterium tuberculosis*. Infect. Immun., 61:627-632.

Wiker HG, Harboe M, Nagai S and Bennedsen J. **1990.** Quantitative and qualitative studies on the major extracellular antigen of *Mycobacterium tuberculosis* H37Rv and *Mycobacterium bovis* BCG. Am. Rev. Respir. Dis., 141:830-838.

Wiker HG and Harboe M. **1992.** The antigen 85 complex: a major secretion product of *Mycobacterium tuberculosis*. Microbiol. Rev., 56:648-61.

World Health Organization, Global Programme on AIDS. **1994.** The current global situation of the HIV/AIDS pandemic. Wkly. Epidemiol. Rev., 69:189-196

Young DB, and Garbe TR. **1991.** Heat shock proteins and antigens of *Mycobacterium tuberculosis*. Infect. Immun., 59:3086-3093.

Zeiss CR, Radin RC, Williams JE, Levitz D and Phair JP. **1982.** Detection of immunoglobulin G antibody to purified protein derivative in patients with tuberculosis by radioimmunoassay and enzyme-linked immunosorbent assay. J. Clin. Microbiol., 15:93-96.

PUBLICATIONS

INFECTION AND IMMUNITY, Nov. 2003, p. 6338–6343 Vol. 71, No. 11
0019-9567/03/$08.00+0 DOI: 10.1128/IAI.71.11.6338–6343.2003

PPE Antigen Rv2430c of *Mycobacterium tuberculosis* Induces a Strong B-Cell Response

Rakesh Kumar Choudhary,[1] Sangita Mukhopadhyay,[1] Prachee Chakhaiyar,[1] Naresh Sharma,[1] K. J. R. Murthy,[2] V. M. Katoch,[3] and Seyed E. Hasnain[1,4]*

Centre for DNA Fingerprinting and Diagnostics, Hyderabad 500-076,[1] Immunology Unit, Mahavir Hospital and Research Centre, A. C. Guards, Hyderabad 500-004,[2] Central Jalma Institute of Leprosy, Tajganj, Agra 282-001,[3] and Jawaharlal Nehru Centre for Advanced Scientific Research, Jakkur, Bangalore 560-064,[4] India

Received 8 April 2003/Returned for modification 20 May 2003/Accepted 8 August 2003

The variation in sequence and length in the C-terminal region among members of the unique PE (Pro-Glu) and PPE (Pro-Pro-Glu) protein families of *Mycobacterium tuberculosis* is a likely source of antigenic variation, giving rise to the speculation that these protein families could be immunologically important. Based on in silico analysis, we selected a hypothetical open reading frame (ORF) encoding a protein belonging to the PPE family and having epitopes with predictably higher antigenic indexes. Reverse transcriptase PCR using total RNA extracted from in vitro-cultured *M. tuberculosis* H37Rv generated an mRNA product corresponding to this gene, indicating the expression of this ORF (Rv2430c) at the mRNA level. Recombinant protein expressed in *Escherichia coli* was used to screen the sera of *M. tuberculosis*-infected patients, as well as those of clinically healthy controls ($n = 10$), by enzyme-linked immunosorbent assay. The panel of patient sera comprised sera from fresh infection cases (category 1; $n = 32$), patients with relapsed tuberculosis (category 2; $n = 30$), and extrapulmonary cases (category 3; $n = 30$). Category 2 and 3 sera had strong antibody responses to the PPE antigen, equal to or higher than those to other well-known antigens, such as Hsp10 or purified protein derivative (PPD). However, a higher percentage of patients belonging to category 1, as opposed to clinically healthy controls, showed stronger antibody response against the PPE protein when probed with anti-immunoglobulin M (IgM) (71 versus 37.5%) or anti-IgG (62.5 versus 28.12%). Our results reveal that this PPE ORF induces a strong B-cell response compared to that generated by *M. tuberculosis* Hsp10 or PPD, pointing to the immunodominant nature of the protein.

About 10% of the genome of *Mycobacterium tuberculosis* codes for the PE and PPE families of proteins (7), which are glycine rich and are exclusive to *M. tuberculosis*. The 69 members of the PPE protein family have a conserved N-terminal domain that comprises ~180 amino acids followed by C-terminal segments that vary markedly in sequence and length. These proteins fall into three groups, one of which constitutes the MPTR class characterized by the presence of multiple tandem copies of the motif Asn-X-Gly-X-Gly-Asn-X-Gly. The second subgroup contains a characteristic well-conserved motif, Gly-X-X-Ser-Val-Pro-X-X-Trp, around position 350. The proteins in the third group are unrelated except for the presence of the common PPE domain. The subcellular locations of a few PPE proteins are known (6, 25), and in only one case (7), that of a lipase (Rv3097), has a function been suggested. There are few studies supporting the notion that PE and PPE proteins could be of functional importance (7, 23). It is widely speculated that they could be responsible for generating antigenic variation (1, 4, 6, 8, 12, 27). However, the effects the PPE family proteins, unique in their protein sequences and possible structure, may have on the immune system have not been well documented. Furthermore, a qualitative and quantitative immune response to PPE proteins in a clinical setting has not been shown. Since 180 amino acid residues in the N-terminal regions of PPE proteins are conserved, it is interesting to speculate that the variation in sequence and length in the C-terminal region could represent a source of antigenic variability.

Based on in silico analysis and DNA microarray expression data (24), we selected an ORF, Rv2430c, whose product displays a high antigenic index and evaluated its importance in eliciting immune responses in a panel of human sera obtained from three well-classified categories of patients: (i) those reporting TB for the first time, (ii) those presenting with a relapse of TB, and (iii) those with extrapulmonary cases. Clinically healthy human sera were used as controls to compare the immunological responses to the protein. Enzyme-linked immunosorbent assay (ELISA) using the recombinant protein showed good specificity and sensitivity, suggesting that the product of this PPE family ORF, Rv2430c, induces a strong B-cell response in infected subjects.

MATERIALS AND METHODS

In silico analysis of Rv2430c. In silico pattern search analysis of the PPE family was carried out to classify the products of the various ORFs into three subgroups. Products of ORFs with ≤200 amino acids and belonging to the third subgroup of the PPE family were further analyzed using protein analysis software (Protean version 4.0, Lasergene Navigator; DNASTAR Inc., Madison, Wis.) to calculate their antigenic indexes.

RNA extraction and reverse transcriptase (RT) PCR. RNA was extracted from 10^9 $H_{37}Rv$ cells, cultured in vitro in Middlebrook 7H9 medium supplemented with albumin-dextrose complex according to the Qiaquick total RNA extraction kit (Qiagen Inc.) instructions, dissolved in 50 μl of nuclease-free water, and stored at −70°C until further use. First-strand synthesis was carried

* Corresponding author. Mailing address: C.D.F.D., ECIL Rd., Nacharam, Hyderabad 500-076, India. Fax: 91-040 27155610 and 91-040 27155479. Phone: 91-040 27155604 and 91-040 27155605. E-mail: hasnain@cdfd.org.in.

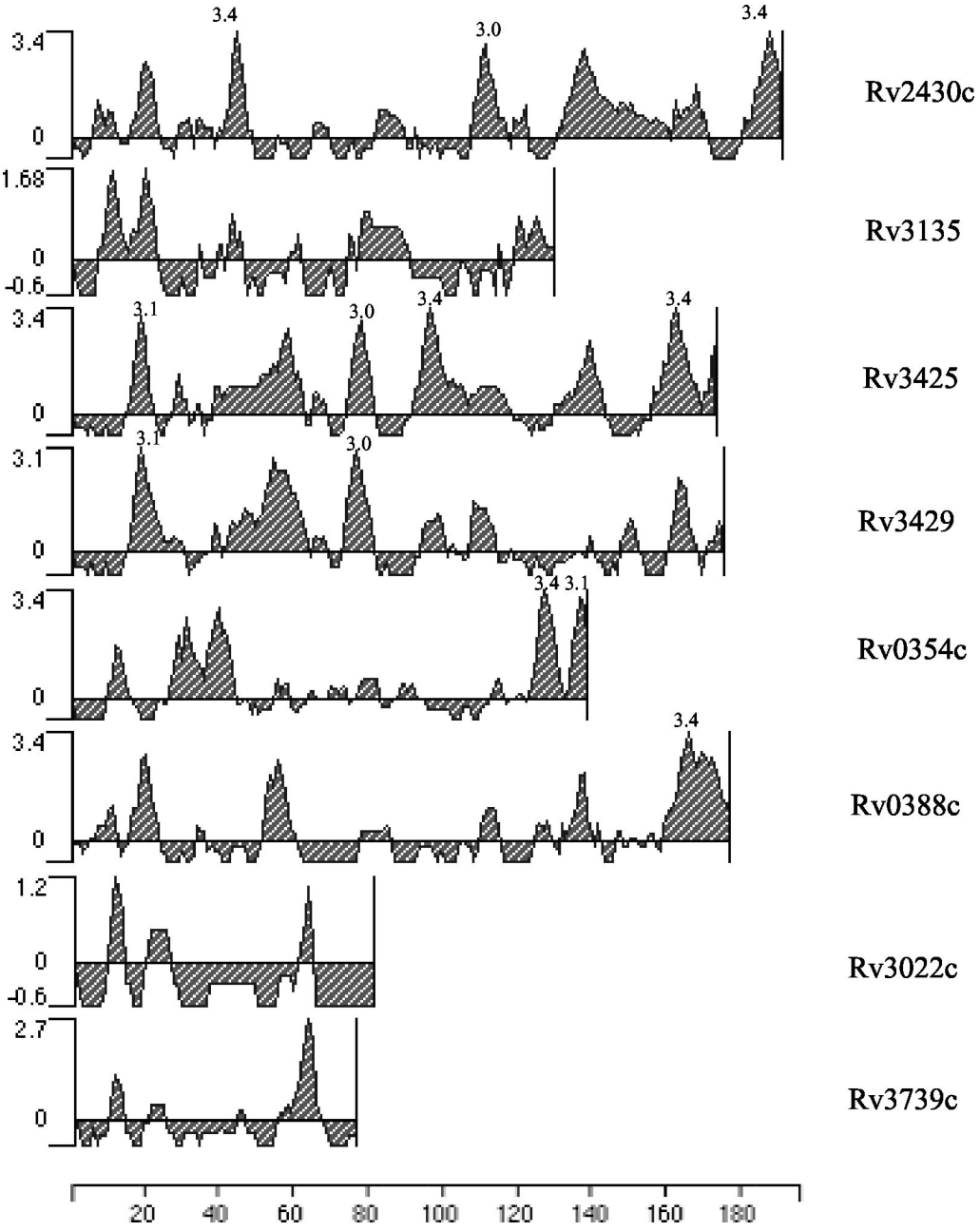


FIG. 1. In silico analysis of PPE ORFs. Only the ORFs belonging to subgroup 3 and encoding products ≤200 amino acids in length were subjected to analysis. The presence of stretches of high antigenic index are shown. The *x* axis indicates the ORF product length in amino acids, while the *y* axis represents the antigenic index.

out using avian myeloblastosis virus reverse transcriptase. This was followed by heat denaturation to inactivate the enzyme. Subsequent second-strand synthesis was performed using *Tfl* polymerase. The PCR product was visualized by electrophoresis in a 1% agarose gel.

Expression and purification of the recombinant protein encoded by Rv2430c. Genomic DNA of H37Rv was extracted using the Genome Extraction kit provided by Epicentre Technologies, as described earlier (26). The Rv2430c gene was PCR amplified from the genomic DNA of H_{37}Rv using upstream (5′-GGA TCCATGCATTTCGAAGCGTAC-3′) and downstream (5′-AAGCTTCTAAG TGTCTGTACGCGATGA-3′) primers. *Bam*HI and *Hin*dIII sites were incorporated in the 5′ and 3′ ends of the primers, respectively. The purified fragment was ligated into the pGEMT-Easy vector (Promega Inc.), and the recombinant clone carrying the Rv2430c insert was confirmed by DNA sequencing (ABI Prism 377 DNA Sequencer; PE Biosystems). The insert was then subcloned as a *Bam*HI and *Hin*dIII fragment into the PQE30 expression vector (Qiagen Inc.) to generate the plasmid construct PQERv2430c carrying Rv2430c as an N-terminal histidine-tagged fusion. PQERv2430c was transformed into the host M15pREP4 strain of *Escherichia coli*. A single colony of *E. coli* strain M15pREP4 harboring PQERv2430c was inoculated into 5 ml of Luria-Bertani broth with the appropriate antibiotics (100 μg of ampicillin/ml and 25 μg of kanamycin/ml) and grown overnight at 37°C with constant agitation; 100 μl of this overnight culture was inoculated into 5 ml of Luria-Bertani broth with the appropriate antibiotics and grown to an optical density at 590 nm of 0.6, at which time expression was induced with 1 mM IPTG. A separate aliquot of uninduced culture was kept as a control. Cells were harvested 3 h postinduction, suspended in 1× SDS sample buffer, and denatured by being heated at 100°C for 10 min. The recombinant

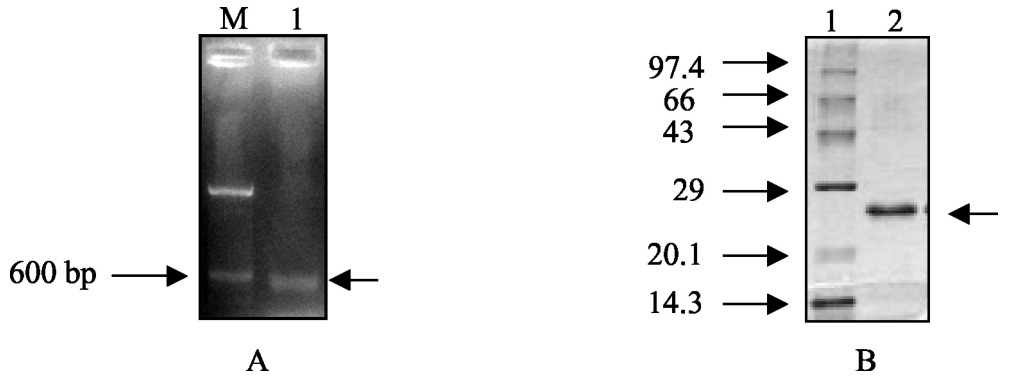


FIG. 2. (A) Transcription of hypothetical PPE ORF Rv2430c. RNA extracted from the virulent H37Rv laboratory strain of *M. tuberculosis* was used in an RT-PCR. A 597-bp RT-PCR product was observed after electrophoresis in a 1% agarose gel (lane 1). Lane M is the 100-bp DNA molecular size marker run alongside. The arrow on the right indicates the position of the 597-bp PCR product. (B) Expression and purification of the recombinant PPE protein. The recombinant protein was expressed in strain M15pREP4 of *E. coli* and was purified to homogeneity using the NiNTA protein purification kit. Lane 1, marker; lane 2, purified protein. The arrow on the right indicates the position of the 23-kDa protein.

protein was purified to homogeneity with the QIA Expressionist kit (Qiagen Inc.).

Serological characterization of the recombinant protein. ELISAs were performed in 96-well microtiter plates (Corning Costar) coated with the recombinant Rv2430c protein. After overnight incubation at 4°C, the plates were washed three times with phosphate-buffered saline (PBS) buffer and blocked with 200 μl of blocking buffer (PBS containing 1% bovine serum albumin) for 1 h at 37°C. The plates were then washed three times with PBS-Tween wash buffer (0.05% Tween 20 in 1× PBS, pH 8.0) and incubated for 1 h at 37°C with human sera (1:200 dilution in blocking buffer). The plates were washed with PBS-Tween and further incubated with either anti-human immunoglobulin G (IgG)-horseradish peroxidase (HRP) or anti-human IgM-HRP (Sigma). HRP activity was detected using a chromogenic substance, *o*-phenylenediamine tetrahydrochloride (Sigma), in citrate-phosphate buffer (pH 5.4) and 1 μl of H_2O_2 (Qualigens, India)/ml. The reactions were terminated using 1 N H_2SO_4, and the absorbance values were measured at 492 nm in an ELISA reader (Bio-Rad).

Study population. Serum samples were obtained from 92 TB patients who had reported to the Mahavir Hospital and Research Centre, Hyderabad, India, and 10 clinically healthy donors. The 92 patients belonged to three well-defined categories. Category 1 ($n = 32$) comprised patients who had contracted the pathogen for the first time and had no history of TB treatment. Category 2 ($n = 30$) comprised patients with relapsed TB, i.e., who were treated earlier for TB but the symptoms resurfaced after completion of the treatment. Category 3 ($n = 30$) comprised patients with extrapulmonary TB in which the disease was confirmed by tissue biopsy. In the cases of category 1 and category 2 patients, diagnosis was confirmed by examination of the sputum (acid-fast bacillus smear positive). Clinically healthy donors were vaccinated with *Mycobacterium bovis* BCG. The study was carried out after approval from the Institute Bioethics Committee.

Statistical analysis. Student's *t* test was used for analysis of statistical significance (*P* value). Graphpad Quickcalcs (online *t* test calculator [http://www.graphpad.com/quickcalcs/ttest1.cfm]) was used for this purpose.

RESULTS

The product of the hypothetical PPE ORF Rv2430c has a high antigenicity profile score and is expressed at the mRNA level. In silico analyses of the PPE ORF products, ≤200 amino acids in length, belonging to subgroup 3 of the PPE family (7) were carried out. Rv2430c and Rv3425 displayed major antigenic stretches (Fig. 1) with peak values of ≥3.0. However, an analysis of the microarray expression data (24) identified Rv2430c as one of the overexpressed genes in an IdeR mutant of *M. tuberculosis*, and it was thereby implicated in pathogenesis. Rv2430c was thus selected for our study. Rv2430c was further subjected to detailed analysis to predict its likely structure through various algorithms. The Predict Protein server (http://www.embl-heidelberg.de/predictprotein/) gave a very low score for Rv2430c, pointing to the unlikely possibility that this ORF encoded a transmembrane protein (data not shown). To check whether the hypothetical Rv2430c indeed represented a functional gene, the mRNA extracted from in vitro-cultured H_{37}Rv cells was used as a template for reverse transcription followed by PCR. The RT-PCR product was fractionated on a 1% agarose gel. A 597-bp band was observed upon staining with ethidium bromide, indicating that the ORF was expressed at the mRNA level in the liquid cultures of *M. tuberculosis* (Fig. 2A). The ORF Rv2430c was expressed as a His-tagged fusion protein in *E. coli* (Fig. 2B) and used for immunological studies.

The recombinant PPE protein Rv2430c displays strong B-cell responses during infection with TB. Having shown that the ORF encoding RV2430c of *M. tuberculosis* was expressed at the mRNA level, experiments were designed to evaluate the immune response of TB patients to recombinant Rv2430c PPE protein. For this, the recombinant Rv2430c protein was used to screen the TB patient sera by ELISA, using anti-human IgG-HRP and anti-human IgM-HRP as conjugates. The humoral immune responses directed against the recombinant protein by patients with TB and healthy controls were compared. The data (Fig. 3) reveal that the sera of all the infected patients mounted significantly higher antibody responses against Rv2430c than those of the healthy controls ($P < 0.0001$). Since negligible antibody responses were obtained in the healthy-control group (Fig. 3), it is likely that this protein is expressed during the course of *M. tuberculosis* infection, and it may be associated with disease manifestation and progression.

Immunodominant nature of Rv2430c. Since all the patients infected with TB revealed strong humoral responses against Rv2430c compared to the healthy controls, it was of interest to compare the antibody responses in various clinical categories and also to evaluate whether Rv2430c is immunodominant. For this, we chose recombinant Hsp10 of *M. tuberculosis*, which is a well-known immunodominant antigen of *M. tuberculosis* (28). Figure 4 clearly shows that strong antibody responses were elicited against Rv2430c in all three study groups (category 1, category 2, and category 3). Rv2430c elicited statisti-

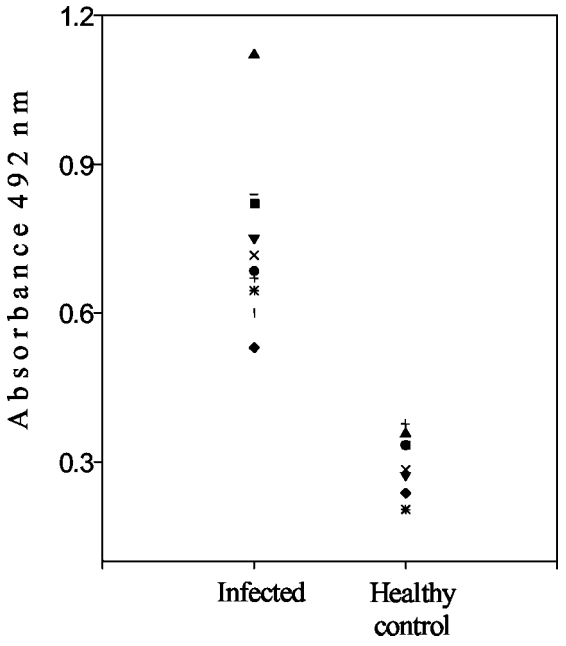


FIG. 3. The recombinant Rv2430c PPE protein elicits strong antibody responses in *M. tuberculosis*-infected patients as opposed to healthy controls. ELISA reactivities of IgG anti-Rv2430c antibodies were assayed in sera of either *M. tuberculosis*-infected patients or healthy controls ($P < 0.0001$). Symbols, patient population.

cally significant immune responses compared to Hsp10 ($P < 0.003$) in patients with fresh infections (Fig. 4A). Patients belonging to categories 2 and 3 exhibited Rv2430c-specific antibody equivalent to that for Hsp10 (Fig. 4B and C). Since the PPD antigen of *M. tuberculosis* is also used to diagnose TB infection (2, 15, 29), we compared the immunopotentiality of Rv2430c over PPD only in cases of fresh infection (Fig. 4A). The results clearly indicated that the PPE protein Rv2430c is far more immunogenic and could elicit a stronger B-cell response than PPD. Sera from the fresh-infection category responded better against the Rv2430c antigen than against PPD ($P < 0.0001$) (Fig. 4A).

In order to compare the serological sensitivities of Rv2430c and Hsp10, the data presented in Fig. 4 were recalculated as percentages of individuals showing absorbances above 0.65 at 492 nm. Although only 28.12% of individuals with fresh infections mounted strong antibody responses of the IgG type to Hsp10, a very high percentage of individuals (62.5%) recognized Rv2430c (Fig. 5A). Similar conclusions could be drawn when IgM antibodies were assayed (Rv2430c versus Hsp10, 71 versus 37.5%) (Fig. 5B). From these results, it is apparent that Rv2430c shows better reactivity vis-a-vis Hsp10 to sera from patients with fresh infection than to those of relapsed or extrapulmonary-TB patients. Our results, therefore, convincingly demonstrate the immunodominant nature of the product of the hypothetical PPE ORF Rv2430c.

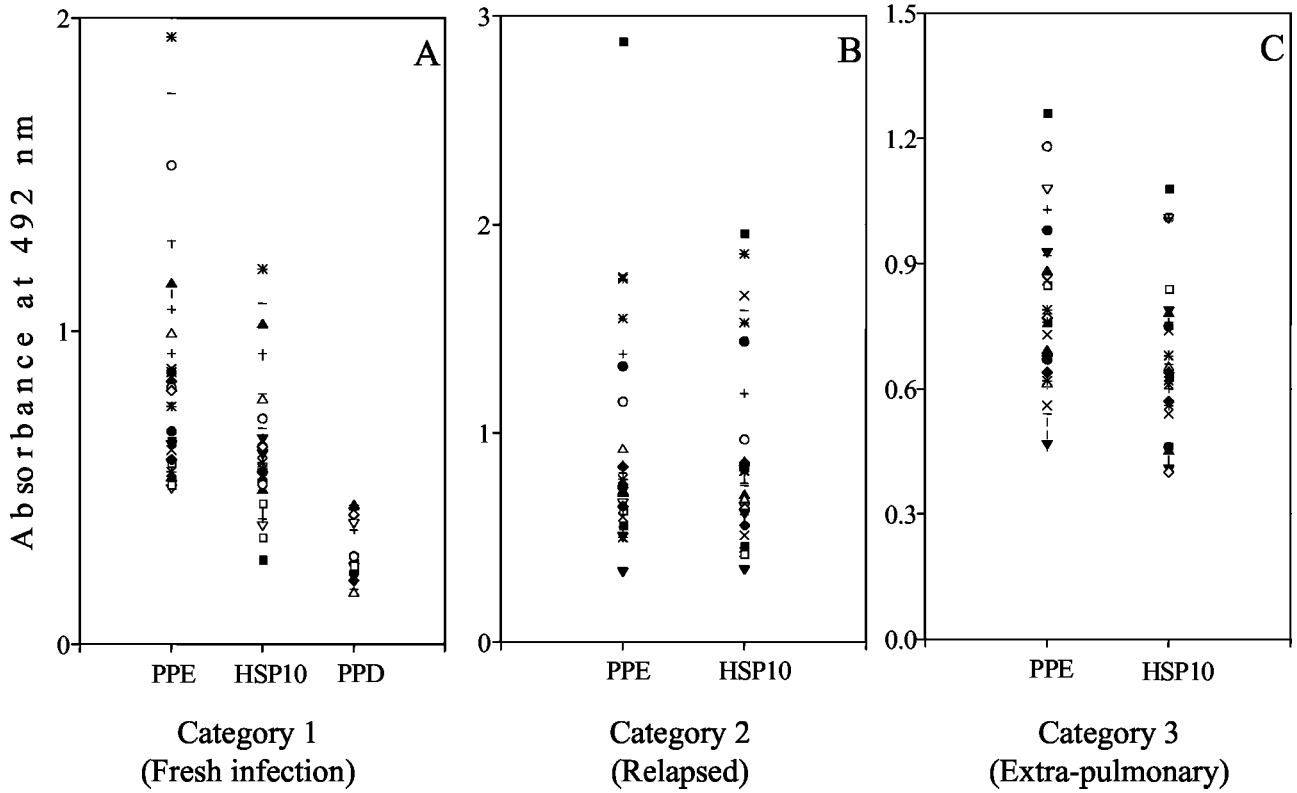


FIG. 4. PPE Rv2430c protein shows strong reactivities to sera from all three patient categories. Reactivities to both recombinant Rv2430c and Hsp10 of *M. tuberculosis* in the three categories of patients were estimated by ELISA. The patients belonging to categories 2 (B) and 3 (C) displayed similar antibody responses to both antigens. However, the antibody responses of category 1 patients (A) to Rv2430c were higher than those to Hsp10 ($P < 0.003$) or PPD ($P < 0.0001$).

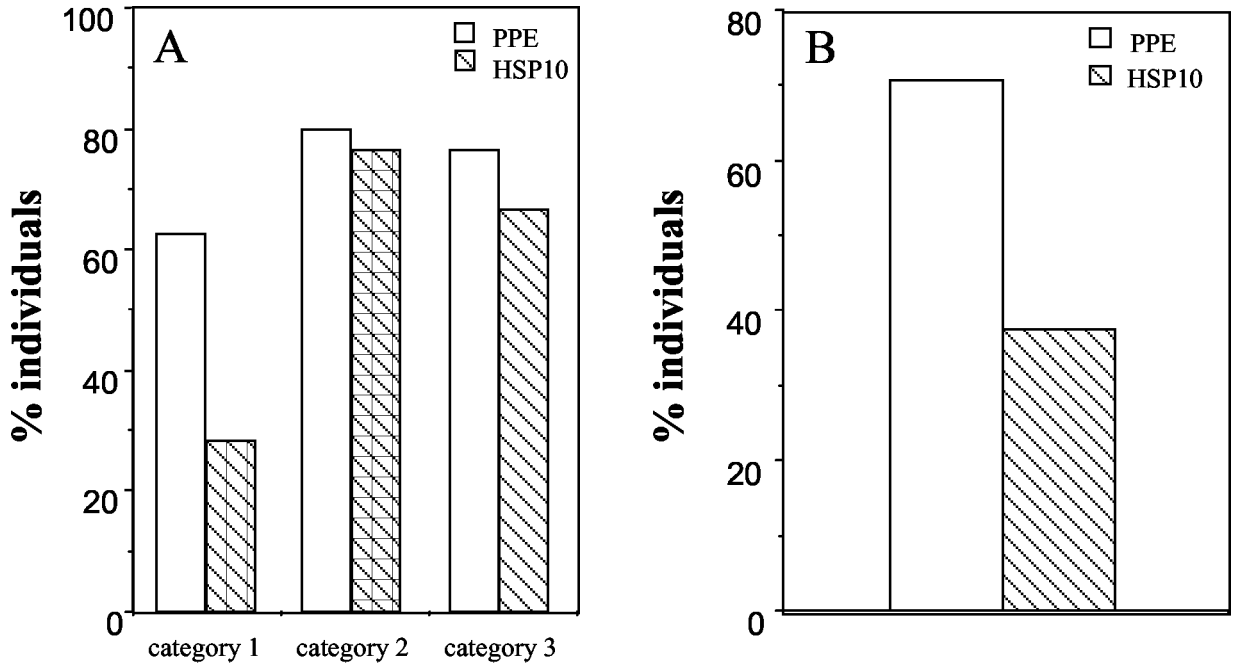


FIG. 5. Serological sensitivities of PPE Rv2430c and Hsp10 as a function of percentages of individuals. (A) The results shown in Fig. 4 were recalculated as percentages of individuals showing A_{492} values of >0.65. The anti-IgG responses against Rv2430c or Hsp10 were compared for all three categories of patients studied. A higher percentage of individuals belonging to category 1 showed stronger reactivity to Rv2430c than to Hsp10, but for the other two categories, the values were comparable. (B) Percentage of individuals in category 1 showing anti-IgM antibody A_{492} of >0.5. A higher percentage of individuals showed anti-IgM antibody A_{492} of >0.5 against Rv2430c than against Hsp10.

DISCUSSION

The 69 members of the PPE protein family have a conserved N-terminal domain that comprises ~180 amino acids followed by C-terminal segments that vary markedly in sequence and length (7). Based on our pattern search analysis of the TubercuList database (http://genolist.pasteur.fr/TubercuList/), these proteins were categorized into three groups. Subgroup 1, represented by 20 members, constitutes the MPTR class characterized by the presence of multiple tandem copies of the motif Asn-X-Gly-X-Gly-Asn-X-Gly. The second subgroup, comprising 21 members, contains a characteristic well-conserved motif, Gly-X-X-Ser-Val-Pro-X-X-Trp, around position 350. The proteins of the third subgroup, with 28 members, are unrelated except for the presence of the common PPE domain. That ORF products across the subgroup do not react immunologically was evident from our studies with Rv2608 (unpublished data). Rv2430c belongs to the third subgroup. ORFs belonging to the third subgroup with coding capacities of ≤200 amino acids were shortlisted. This shortlist was further narrowed down based on two very important criteria, namely, the antigenic profile and the association of the ORF with pathological conditions as evident from DNA microarray expression data (24). Rv2430c and Rv3425 were the ORFs with the highest antigenicity indexes. DNA microarray results demonstrated that of these two ORFs, Rv2430c was one of the genes induced in an IdeR mutant of *M. tuberculosis* (24), pointing to its possible role in pathogenesis. Rv2430c was accordingly shortlisted and evaluated for the role of its product as an antigen in a clinical setting.

The ORF was shown to be RT-PCR positive, pointing to the likelihood that it is expressed during infection. The Rv2430c ORF was expressed in *E. coli*, and the recombinant protein was purified and tested for its ability to recognize IgG antibody in the sera of TB patients and healthy individuals.

The TB patients used in our study represent a heterogeneous population, including fresh infection cases characterized by patients who contracted the pathogen for the first time (category 1); relapsed cases, in which the disease resurfaced after the completion of treatment (category 2); and extrapulmonary cases, which are mostly sputum negative (category 3). The immune response profiles of Rv2430c in different clinical categories were studied. The PPE protein Rv2430c was found to be recognized by antibodies in the sera of infected patients in ELISA with a serum dilution of 1:200, whereas poorer ELISA reactivity was observed for the sera of all of the healthy individuals. The presence of antibodies to Rv2430c in sera from TB patients (Fig. 3) and their absence in sera from healthy individuals suggests that the protein is expressed in vivo during active infection with *M. tuberculosis* and that the native molecule is immunogenic.

Several reports have emphasized the observation of a lack of sufficient immune responses in TB patients against many promising serodiagnostic antigens of *M. tuberculosis*. This fact is more distressing in cases of fresh infection, where in a majority of the cases the immune system is not sufficiently primed to elicit a strong antibody response against most of the *M. tuberculosis* antigens. The recombinant Rv2430c protein was very strongly recognized by all three categories of patients, including the fresh-infection group (category 1). The members

of the heat shock protein family, including Hsp70 (20, 28) and Hsp10 (28), have been known to elicit a strong B-cell response. Surprisingly, in our study, the immunodominant antigen Hsp10, though recognized by category 2 and 3 patient sera, was not sensitive enough to detect the patients with fresh infections (category 1). The picture remained unaltered when we used PPD in place of Hsp10. It is pertinent to note that although several antigens have been tested for use in serodiagnosis, no test with a single antigen has proved able to achieve sensitivity and specificity in a study population that is suitably large and heterogeneous (3, 5, 9, 11, 13, 14, 16, 17, 18, 19, 21, 22). The factors responsible for this include (i) the stage of the disease, (ii) the location of the infection, and (iii) the genetic background. Our results show that the PPE protein Rv2430c, which lacks a transmembrane domain and is therefore likely to be cytosolic or secretory in localization, is an immunodominant B-cell target antigen with apparent diagnostic potential. It is also interesting to speculate on the use of Rv2430c, along with other immunodominant antigens (10), for vaccine development.

ACKNOWLEDGMENTS

Recombinant Hsp10 was a kind gift from S. C Mande, Structural Biology Laboratory, CDFD. We thank the National Mycobacterial Repository, Central Jalma Institute for Leprosy, Agra, India, for providing us with liquid and Lowenstein-Jensen slant cultures of $H_{37}Rv$.

This study was supported by research grants from the Department of Biotechnology, Government of India, to S.E.H. R.K.C and P.C. thank the Council of Scientific and Industrial Research, Government of India, for the award of Senior Research Fellowships.

REFERENCES

1. **Abou-Zeid, C., T. Garbe, R. Lathigra, H. G. Wiker, M. Harboe, G. A. Rook, and D. B. Young.** 1991. Genetic and immunological analysis of *Mycobacterium tuberculosis* fibronectin-binding proteins. Infect. Immun. **59:**2712–2718.
2. **Almeida, L. M., M. A. Barbieri, A. C. Da Paixao, and L. E. Cuevas.** 2001. Use of purified protein derivative to assess the risk of infection in children in close contact with adults with tuberculosis in a population with high Calmette-Guerin bacillus coverage. Pediatr. Infect. Dis. J. **20**:1061–1065.
3. **Amara, R. R., S. Shanti, and V. Satchidanandam.** 1998. Characterization of novel immunodominant antigens of *Mycobacterium tuberculosis*. Microbiology **144:**1197–1203.
4. **Banu, S., N. Honore, B. Saint-Joanis, D. Philpott, M. C. Prevost, and S. T. Cole.** 2002. Are the PE-PGRS proteins of *Mycobacterium tuberculosis* variable surface antigens? Mol. Microbiol. **44:**9–19.
5. **Batoni, G., D. Bottai, S. Esin, W. Florio, M. Pardini, G. Maisetta, G. Freer, S. Senesi, and M. Campa.** 2002. Purification, biochemical characterization and immunogenicity of SA5K, a secretion antigen of *Mycobacterium tuberculosis*. Scand. J. Immunol. **56:**43–51.
6. **Brennan, M. J., G. Delogu, Y. Chen, S. Bardarov, J. Kriakov, M. Alavi, and W. R. Jacobs, Jr.** 2001. Evidence that mycobacterial PE_PGRS proteins are cell surface constituents that influence interactions with other cells. Infect. Immun. **69**:7326–7333.
7. **Cole, S. T., et al.** 1998. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. Nature **393:**537–544.
8. **Delogu, G., and M. J. Brennan.** 2001. Comparative immune response to PE and PE_PGRS antigens of *Mycobacterium tuberculosis*. Infect. Immun. **69:** 5606–5611.
9. **Devi, K. R., K. S. Kumar, B. Ramalingam, and R. Alamelu.** 2002. Purification and characterization of three immunodominant proteins (38, 30, and 16 kDa) of *Mycobacterium tuberculosis*. Protein Expr. Purif. **24:**188–195.
10. **Dhar, N., V. Rao, and A. K. Tyagi.** 2000. Recombinant BCG approach for development of vaccines: cloning and expression of immunodominant antigens of *Mycobacterium tuberculosis*. FEMS Microbiol. Lett. **190:**309–316.
11. **Dillon, D. C., M. R. Alderson, C. H. Day, T. Bement, A. Campos-Neto, Y. A. Skeiky, T. Vedvick, R. Badaro, S. G. Reed, and R. L. Houghton.** 2000. Molecular and immunological characterization of *Mycobacterium tuberculosis* CFP-10, an immunodiagnostic antigen missing in *Mycobacterium bovis* BCG. J. Clin. Microbiol. **38:**3285–3290.
12. **Espitia, C., J. P. Laclette, M. Mondragon-Palomino, A. Amador, J. Campuzano, A. Martens, M. Singh, R. Cicero, Y. Zhang, and C. Moreno.** 1999. The PE-PGRS glycine-rich proteins of *Mycobacterium tuberculosis*: a new family of fibronectin-binding proteins? Microbiology **145:**3487–3495.
13. **Florio, W., D. Bottai, G. Batoni, S. Esin, M. Pardini, G. Maisetta, and M. Campa.** 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.
14. **Houghton, R. L., M. J. Lodes, D. C. Dillon, L. D. Reynolds, C. H. Day, P. D. McNeill, R. C. Hendrickson, Y. A. Skeiky, D. P. Sampaio, R. Badaro, K. P. Lyashchenko, and S. G. Reed.** 2002. Use of multiepitope polyproteins in serodiagnosis of active tuberculosis. Clin. Diagn. Lab. Immunol. **9:**883–891.
15. **Kalish, S. B., R. C. Radin, J. P. Phair, D. Levitz, C. R. Zeiss, and E. Metzger.** 1983. Use of an enzyme-linked immunosorbent assay technique in the differential diagnosis of active pulmonary tuberculosis in humans. J. Infect. Dis. **147:**523–530.
16. **Laal, S., K. M. Samanich, M. G. Sonnenberg, S. Zolla-Pazner, J. M. Phadtare, and J. T. Belisle.** 1997. Human humoral responses to antigens of *Mycobacterium tuberculosis*: immunodominance of high-molecular-mass antigens. Clin. Diagn. Lab. Immunol. **4:**49–56.
17. **Lim, R. L., L. K. Tan, W. F. Lau, M. C. Ming, R. Dunn, H. P. Too, and L. Chan.** 2000. Cloning and expression of immunoreactive antigens from *Mycobacterium tuberculosis*. Clin. Diagn. Lab. Immunol. **7:**600–606.
18. **Ljungqvist, L., A. B. Andersen, P. Andersen, K. Haslov, A. Worsaae, J. Bennedsen, and I. Heron.** 1990. Affinity purification, biological characterization and serological evaluation of defined antigens from *Mycobacterium tuberculosis*. Trop. Med. Parasitol. **41:**333–335.
19. **Lodes, M. J., D. C. Dillon, R. Mohamath, C. H. Day, D. R. Benson, L. D. Reynolds, P. McNeill, D. P. Sampaio, Y. A. Skeiky, R. Badaro, D. H. Persing, S. G. Reed, and R. L. Houghton.** 2001. Serological expression cloning and immunological evaluation of MTB48, a novel *Mycobacterium tuberculosis* antigen. J. Clin. Microbiol. **39:**2485–2493.
20. **Mehlert, A., and D. B. Young.** 1989. Biochemical and antigenic characterization of the *Mycobacterium tuberculosis* 71kD antigen, a member of the 70kD heat-shock protein family. Mol. Microbiol. **3:**125–130.
21. **Mustafa, A. S., P. J. Cockle, F. Shaban, R. G. Hewinson, and H. M. Vordermeier.** 2002. Immunogenicity of *Mycobacterium tuberculosis* RD1 region gene products in infected cattle. Clin. Exp. Immunol. **130:**37–42.
22. **Mustafa, A. S.** 2002. Development of new vaccines and diagnostic reagents against tuberculosis. Mol. Immunol. **39**:113–119.
23. **Ramakrishnan, L., N. A. Federspiel, and S. Falkow.** 2000. Granuloma-specific expression of *Mycobacterium* virulence proteins from the glycine-rich PE-PGRS family. Science **288**:1436–1439.
24. **Rodriguez, G. M., M. I. Voskuil, B. Gold, G. K. Schoolnik, and I. Smith.** 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.
25. **Sampson, S. L., P. Lukey, R. M. Warren, P. D. van Helden, M. Richardson, and M. J. Everett.** 2001. Expression, characterization and subcellular localization of the *Mycobacterium tuberculosis* PPE gene Rv1917c. Tuberculosis **81:**305–317.
26. **Siddiqi, N., M. Shamim, S. Hussain, R. K. Choudhary, N. Ahmed, Prachee, S. Banerjee, G. R. Savithri, M. Alam, N. Pathak, A. Amin, M. Hanief, V. M. Katoch, S. K. Sharma, and S. E. Hasnain.** 2002. Molecular characterization of multidrug-resistant isolates of *Mycobacterium tuberculosis* from patients in North India. Antimicrob. Agents Chemother. **46:**443–450.
27. **Singh, K. K., X. Zhang, A. S. Patibandla, P. Chien, Jr., and S. Laal.** 2001. Antigens of *Mycobacterium tuberculosis* expressed during preclinical tuberculosis: serological immunodominance of proteins with repetitive amino acid sequences. Infect. Immun. **69:**4185–4191.
28. **Young, D. B., and T. R. Garbe.** 1991. Heat shock proteins and antigens of *Mycobacterium tuberculosis*. Infect. Immun. **59:**3086–3093.
29. **Zeiss, C. R., S. B. Kalish, K. S. Erlich, D. Levitz, E. Metzger, R. Radin, and J. P. Phair.** 1984. IgG antibody to purified protein derivative by enzyme-linked immunosorbent assay in the diagnosis of pulmonary tuberculosis. Am. Rev. Respir. Dis. **130**:845–848.

Editor: J. M. Mansfield

Dear Author,

Congratulations on the acceptance of your manuscript. A separate e-mail will follow from our office, requesting any additional materials if required. Thank you for your contribution to *Protein Expression and Purification*.

Sincerely,

Sarah Schnack
Protein Expression and Purification, Editorial Office
pep@elsevier.com

Protein Expression and Purification (In press).

[REVISED]

Title : Expression and characterization of Rv2430c, a novel immunodominant antigen of *Mycobacterium tuberculosis*

Rakesh K. Choudhary[1], Raghu Pullakhandam[2], Nasreen Z. Ehtesham[2], and Seyed E. Hasnain[1,3*]

[1]Centre for DNA Fingerprinting and Diagnostics, Hyderabad, INDIA;

[2]Molecular Biology Unit, National Institute of Nutrition, Hyderabad, INDIA;

[3]Jawaharlal Nehru Center for Advanced Scientific Research, Jakkur, Bangalore, INDIA.

[*]Corresponding author

Address for Correspondence: Seyed E. Hasnain, C.D.F.D., ECIL Road,
Hyderabad-500 076, INDIA
Fax no. 91-040 27155604, 91-040 27155479
Telephone no. 91-040 27155605
E.mail: ehtesham@cdfd.org.in

ANTIMICROBIAL AGENTS AND CHEMOTHERAPY, Feb. 2002, p. 443–450 Vol. 46, No. 2
0066-4804/02/$04.00+0 DOI: 10.1128/AAC.46.2.443–450.2002

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 *inh*A ribosome binding site; Δ, deletion at the indicated nucleotide position.

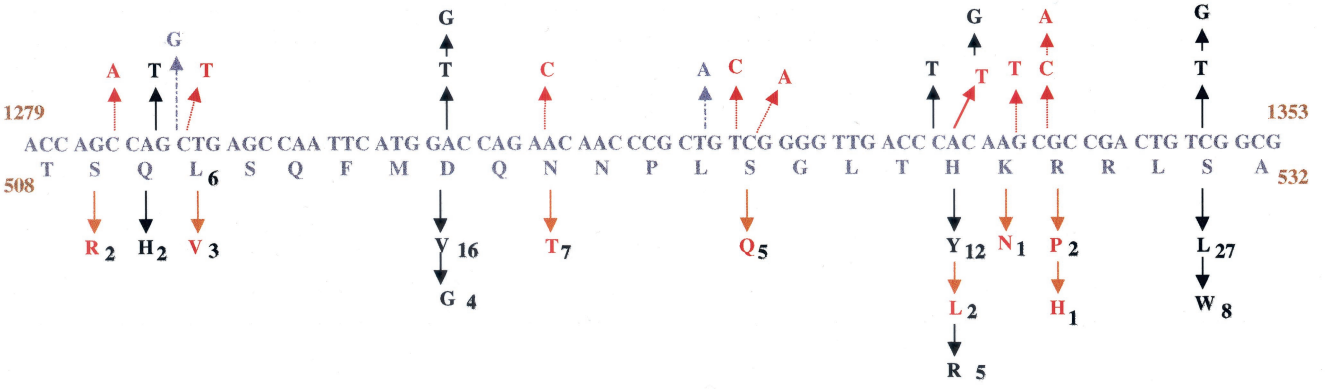


FIG. 1. Summary of mutations at codons 508 to 532 in the *rpoB* gene. The wild-type sequence and amino acids are shown in the middle frame. Nucleotide changes are marked with arrows in the top frame, and the corresponding amino acid changes are denoted in the bottom frame. The amino acids are subscripted 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).

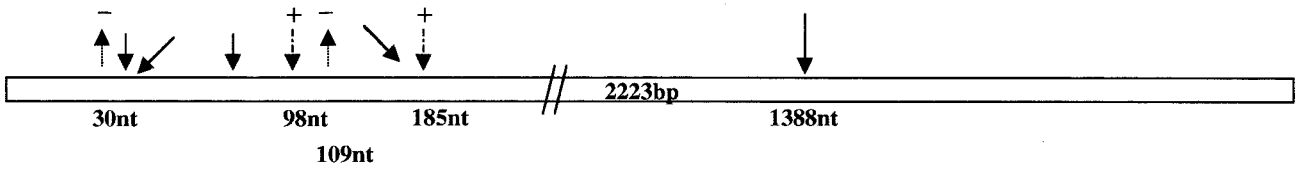


FIG. 2. Summary of mutations in the *katG* gene. Deletions are indicated by lines with a minus sign, while insertions are depicted by dashed lines with a plus sign. Solid lines show the substitutions. Codon 463 exhibited the highest degree of polymorphism, followed by the deletion at nucleotide 30.

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-

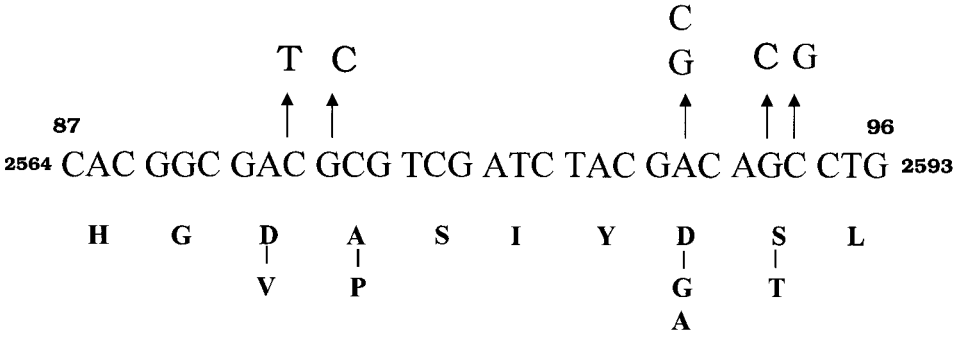


FIG. 3. Summary of missense mutations in the *gyrA* locus. Nucleotide changes are indicated on top of the wild-type sequence, and the corresponding amino acid changes are shown at the bottom. The most common mutation in this locus is Ser95Thr.

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 Harayana, 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.

BIO-DATA

RAKESH KUMAR CHOUDHARY

Laboratory of Molecular and Cellular Biology
Centre for DNA Fingerprinting and Diagnostics,
ECIL Road, Nacharam
Hyderbad 500 076
India

Date of Birth: November 23, 1972

Educational Qualifications:

Ph.D (Submitted)	: Centre for DNA Fingerprinting and Diagnostics (University of Hyderabad) Title of thesis " Cloning and Characterization of a PPE antigen of *Mycobacterium tuberculosis*".
M.Sc	: Biotechnology, 1997, Pondicherry University Pondicherry, INDIA Division - 1st.
B.Sc	: Botany (Hons.), Chemistry, Zoology, English, Hindi. 1994, Patna University, Bihar, INDIA Division - 1st .

List of Publications:

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

Choudhary RK, Pullakhandham R**,** Ehtesham NZ**,** Hasnain SE. Expression and characterization of Rv2430c, a novel immunodominant antigen of *Mycobacterium tuberculosis*. ***Protein Expression and Purification,*** (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, Hasnain SE (2001) Molecular characterization of multidrug-resistant isolates of *Mycobacterium tuberculosis* from patients in North India. ***Antimicrobial Agents and Chemotherapy,*** 46:443-450.