STUDIES ON BIOMEDICALLY IMPORTANT *Mycobacterium tuberculosis* GENES INVOLVED IN NUTRIENT ASSIMILATION AND IMMUNE RESPONSE

Thesis submitted to

Department of Biochemistry
School of Life Sciences
University of Hyderabad
HYDERABAD
INDIA

for the degree of
Doctor of Philosophy

by

PRACHEE

Laboratory of Molecular and Cellular Biology
Centre For DNA Fingerprinting and Diagnostics
HYDERABAD
INDIA

Registration Number: 99LBPH14

2004

DEPARTMENT OF BIOCHEMISTRY
SCHOOL OF LIFE SCIENCES
UNIVERSITY OF HYDERABAD
HYDERABAD
INDIA

CENTRE FOR DNA FINGERPRINTING
AND DIAGNOSTICS
HYDERABAD
INDIA

CERTIFICATE

This is to certify that this thesis entitled "Studies on biomedically important *Mycobacterium tuberculosis* genes involved in nutrient assimilation and immune response" comprises the work done by Ms. Prachee under my guidance at the Centre for DNA Fingerprinting and Diagnostics, Hyderabad, India. This work is original and has not been submitted in part or full for any degree or diploma of any University.

22/9

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

Dr. Seyed E Hasnain
Thesis Supervisor
CDFD, Hyderabad

22/09

Dr. C K Mitra
Head, Department of Biochemistry
University of Hyderabad

CENTRE FOR DNA FINGERPRINTING AND DIAGNOSTICS
HYDERABAD
INDIA

DECLARATION

I hereby declare that the work presented in this thesis entitled, "Studies on biomedically important *Mycobacterium tuberculosis* genes involved in nutrient assimilation and immune response" has been carried out by me under the supervision of Dr. Seyed E Hasnain at the Centre for DNA FIngerprinting and Diagnostics, Hyderabad, India. This work is original and has not been submitted in part or in full for any degree or diploma of any other University earlier.

Prachee

Prachee

Candidate

To my parents,

for being a constant source of inspiration. Your love and support motivated me to work hard and complete my research objectives.

Acknowledgements

The work presented in this thesis was accomplished with the help of many colleagues and associates. I would like to express my gratitude to each one of them.

I am deeply indebted to my research supervisor Dr Seyed E. Hasnain whose guidance, help, stimulating suggestions and encouragement enriched my graduate experience. I also thank him for his constructive criticism and excellent advice during the preparation of this thesis. Apart from his scientific skills, I have tried to imbibe at least a fraction of his excellent literary and oratory skills, an absolute requirement for sharing one's research with the scientific community. I also appreciate his generosity in sending students to various national and international conferences that are important eye openers in the field of research.

I express an extreme sense of gratitude to the earlier graduate students in the laboratory, Dr Noman Siddiqi, Dr Aruna Ramachandran, Dr Sudip Ghosh and Dr Priya Viswanathan, who introduced me to research in contemporary biology. They also offered an extremely conducive laboratory atmosphere I required to initiate my research work.

I extend my sincere thanks to Dr Bandi Aruna and Niteen Pathak for their contribution towards the completion of many experiments. I am grateful to Y. Sailu, whose computational skills generated several leads for my work. I thank Battu Aruna for her help with biophysical experiments as well as for being an admirable friend. Sincere thanks to Anil, Aisha and G. Savithri for extending their help in various ways. Thanks are due to Dr MD Bashyam whose knowledge and experience in the field of molecular biology was very useful in the formulation of my research objectives. I acknowledge Dr Sangita Mukhopadhyay for her help in the interpretation of the immunological data and Dr Shekhar C. Mande and Dr Akash Ranjan for many helpful discussions.

I am extremely grateful to Dr Abhijit A. Sardesai and Dr. J. Gowrishankar for sharing with me their excellent insights in the field of genetics. Their help was instrumental in the completion of related experiments.

I would like to acknowledge all other members of the Laboratory of Molecular and Cellular Biology for creating a highly inspiring and cheerful atmosphere. The scientific exposure I received at CDFD was a great learning experience.

I thank our research collaborators Dr KJR Murthy and Dr VM Katoch for helpful discussions. I also thank Dr Nasreen Z. Ehtesham for extending her lab facilities at the National Institute of Nutrition, Hyderabad and for providing useful suggestions.

I gratefully acknowledge the financial support of several institutions, the Council of Scientific and Industrial Research for providing a research fellowship and the Department of Biotechnology for funding the projects I worked for. I thank the "Global Alliance for TB Drug Development" for sponsoring my participation in the Gordon Research Conference on TB drug development at Oxford, United Kingdom. This Conference allowed me to interact with leaders in the field of tuberculosis research and gave me an invaluable opportunity to present my research before an international audience.

Heartfelt thanks are reserved for my fiance, Dr Rajat Prakash, who, despite his expertise in mathematical sciences, expressed profound interest in my work, took the effort to go through my thesis drafts and offered useful suggestions for improvement. I look forward to a great life ahead with Rajat.

Finally, I thank the almighty for his benevolence.

Prachee

Prachee

CONTENTS IN BRIEF

CHAPTER 3

THE HYPOTHETICAL ORF Rv1885c OF *Mycobacterium tuberculosis* ENCODES A MONOFUNCTIONAL AroQ CLASS OF PERIPLASMIC CHORISMATE MUTASE

List of Tables

CHAPTER – I

Introduction and Objectives of the Study

Contents

* * *

CHAPTER – I

Introduction and Objectives of the Study

1.0 Concept of Volatility

Volatility is a measure of risk. It can be used to measure the market risk of a single instrument or a portfolio of instruments. The volatility of an asset indicates the variability of its returns. Academic studies have defined volatility in terms of statistical measure of the variability of percentage price changes or rates of return. Statistical measure of volatility is the standard deviation of return. It is used as a measure of risk in theories of portfolio selection. Many investors consider volatility as episodes of extreme or rapid price movements within certain days. An important aspect of volatility is its emphasis on the variability, rather than the direction of price movements, which can be upward, downward or flat. Thus, volatility as a concept, can be treated as synonymous with variability in general or variance in particular.

In all financial models, variance is used as a proxy for variability. Any measure of variability that is based on past prices that conform to the present variability has been conceptualized as 'volatility' in the financial market. There are two ways of modelling stock market volatility *viz*, 'conditional volatility' and 'stochastic volatility'. Under the first approach, is possible to estimate the volatility that would prevail today given the set of information relating to the past changes in stock price. This approach is also known as the 'Deterministic Approach' or the 'Conditional Volatility' approach. Technically speaking, if the variance (σ_t^2) is pre-determined and conditional on a set of observable variables at time t-1 and the structure of the model is known, then volatility is said to follow an Autoregressive Conditional Heteroskedasticity (ARCH) process. Alternatively, when volatility is seen as a function of some random variables/events in the market which may be observable or latent, this approach is termed as 'Latent Approach' to the study volatility or

'Stochastic Volatility'. Here, the term 'stochastic' refers to the randomness that affects volatility. To be specific, if σ_t^2 is influenced either by a separate contemporaneous stochastic vector or by an unobserved stochastic vector, then σ_t^2 is not pre-determined in the above sense and is not identical to conditional volatility. Hence it is called stochastic volatility.

Considerable research effort has already gone into modeling time–varying conditional heteroskedastic asset returns. It is important because if both returns and volatility can be forecasted, then it is possible to construct dynamic asset allocation models that use time dependent mean – variance optimization over each period. Financial econometrics suggests the use of non- linear time series structures to model the attitude of investors toward risk and expected return. In this context Bera and Higgins (1993) remarked, "a major contribution of the ARCH literature is the finding that apparent changes in the volatility of the economic time series may be predictable, and result from a specific type of non-linear dependence rather than exogenous structural changes in variables." When the variance is not constant, it is more likely that there are more outliers than expected from the normal distribution, i.e., when a process is heteroskedastic it will follow heavy–tailed or outlier–prone probability distribution. According to Mc Nees (1979), "the inherent uncertainty or randomness associated with different forecast periods seem to vary over time, and large and small errors tend to cluster together". This suggests the application of ARCH formulation to model in financial returns data. In finance, portfolios of financial assets are held as function of expected mean and variance of rate of return. Since any shift in asset demand must be associated with changes in expected return and variance of rate of return. ARCH models are more suited to capture this aspect of assets. In most applications of the ARCH model, a relatively long lag in the conditional variance is often called for and this leads to the problem of negative variance and non-stationary. To avoid this problem, generally a fixed lag structure is typically imposed. Therefore, it is necessary to extend the ARCH models to a new class of models, allowing for both a long memory and much more flexible lag structure. Bollerslev (1986) introduced a

Generalized ARCH (GARCH) model, which allows long memory and flexible lag structure.

Every time the returns deviate from the mean, either in the positive or in the negative direction, they would tend to move in opposite direction, i.e., back towards the mean. This is synonymous to negative autocorrelation. Negative autocorrelation cancels out up- moves and down-moves over a period of time, yielding lower volatility over time. Positive autocorrelation, on the other hand, reinforces the moves in one direction and results in higher volatility over time. Positive autocorrelation in the returns, follows clusters of positive moves.

Stock returns are characterized by statistical distributions. Most of the high frequency financial time series are found deviating from normality. They exhibit characteristics such as skewness and excess kurtosis. This means that the size of the up-moves is not the same as the size of the down-moves, and more returns are to be found in the distribution's tails. High-frequency financial time series, such as stock return series, exhibit a tendency to revert to their long–term mean (Lo and Mackinlay 1988).

1.1 Some Stylized Facts about Volatility

The search for an appropriate model for explaining stock return behaviour has been always guided by empirical stylized facts. A model's ability to reproduce such stylized facts is a desirable feature. A failure. to do so is often a criterion to dismiss a specification, although a single model is not expected to fit or explain all possible empirical. Stylized facts about volatility have been well documented in the ARCH literature (Engle 1982, Bollerslev 1986, Nelson 1991).

Stock market volatility changes over time. There is a scatter of stylized facts which would explain the volatility changes. Some of the stylized facts are as follows:

1.1.1 Thick Tails:

Since the early sixties it was observed, notably by Mandelbort (1965) and Fama (1963, 1965), among others, that asset returns have leptokurtic distribution.

1.1.2 Volatility Clustering:

It refers to a situation wherein large changes tend to be followed by large changes, of either sign, and small changes tend to be followed by small changes [Mandelbort 1965]. A close look at returns data reveals bunching of high and low volatility episodes. Volatility clustering and thick tails of asset returns are intimately related. The ARCH models provide a formal link between dynamic (conditional) volatility behaviour and (unconditional) heavy tails [Diebold (1988), Drost and Nijman (1993)]. ARCH models introduced by Engle (1982) and numerous extensions thereafter, as well as stochastic volatility models are essentially built to mimic volatility clustering. It is also widely documented that ARCH effects disappear with temporal aggregation,

1.1.3 Leverage Effect:

A phenomenon coined by Black (1976), leverage effect suggests that stock price movements are negatively correlated with volatility. Because falling stock prices imply an increased leverage of firms, it is believed that this entails more uncertainty and hence volatility. Empirical evidence reported by Black (1976), Christie(1982) and Schwert (1989) suggests, however, that leverage alone is too small to explain the empirical asymmetries one observes in stock prices. This is closely related to the observation that prices fall more rapidly than they rise.

1.1.4 Long Memory and Persistence:

Generally speaking, volatility is highly persistent. Particularly for high frequency data, one finds evidence of near unit root behavior of the conditional variance process. In the ARCH literature, numerous estimates of GARCH models for stock returns are consistent with an Integrated Generalised Autoregressive Conditional Heteroscedasticity (IGARCH) specification. Likewise stochastic volatility models show similar patterns of persistence (Jacquire, *et.al* 1994). These findings have led to a debate on modelling persistence in the conditional variance process, either *via* a unit root or a long memory process.

1.1.5 Non-trading Periods:

Information that is accumulates until the time financial markets are closed, is reflected in prices after the markets reopen. If for example, information accumulates at a constant rate over calendar time, then the variance of returns over the period from Friday close to Monday close should be three times the variance from the Monday close to Tues day close.

1.1.6 Information Arrivals:

Researchers have long established an empirical relationship between trading volume and stock return volatility. Empirical work with intra-day transaction based data has shown systematic patterns in the relationship between stock return volatility and trading volume, the number of transactions, the bid–ask spread, or market liquidity. There is an association between stock return volatility and trading activity and vice versa.

1.1.7 Volatility Co-movement:

Increase in volatility in some markets, such as the U.S., has led to increased volatility in other world market, such as in Japan, Europe and Latin America, by one day or even up to one month. These dynamic patterns came to be known as "volatility spillovers". Larger volatility spillovers following large, negative returns compared to large positive returns were unable to find any fundamental or macroeconomic news that could explain volatility spillover. In addition to fundamental factors, spillover arises due to institutional factors such as the number of stocks that are cross-listed across major international markets, the scope of closed-end country funds, international portfolio flows and even margin regulations across market. East Asian crisis in 1997 and Russian crisis in 1998 prompted researchers to cite "contagion" effect as the only remaining explanation.

1.1.8 Implied Volatility Correlation:

Implied volatilities are model-based as they are calculated from a pricing equation of a specified model, namely the Black-Scholes model (1974). As they are computed on a daily basis, there appears an internal inconsistency since the model presumes constant volatility. However, deterministic patterns of time-varying volatility are consistent with a generalization of the Black–Scholes model. Most option prices are quoted through their implied volatilities, and it is the implied volatilities that practioners are concerned with. Often, one computes a composite measure since synchronous option prices with different strike price and maturities for the same underlying asset yield different implied volatilities. There has been ample research on the question of whether trading in options and futures causes increases in stock return volatility. Academic research by Edwards (1988) and Stoll and Whaley (1987) has demonstrated that derivatives do not exacerbate volatility.

1.1.9 The Term Structure of Implied Volatility:

The Black-Scholes (1974) model predicts a flat term structure of volatilities. In reality, the term structure of at-the-money implied volatilities is typically upward sloping when short-term volatilities are low and reverse when they are high. Stein and Stein (1989), Taylor and Xu (1994) found that the term structure of implied volatilities from foreign currency options reversed slope every few months. Stein and Stein (1989) also found the actual sensitivity of medium to short term implied volatilities to be greater than the estimated sensitivity from the forecast term structure, and concluded that medium term implied volatilities were greater than the estimated sensitivity from the forecast term structure and concluded that medium term implied volatilities overreacted to information

1.1.10 Smiles:

If option prices in the market conform to the Black-Scholes formula, all Black-Scholes implied volatilities corresponding to various options written on the same asset would coincide with the volatility parameter of the underlying asset. The so-called smile refers to the U–shaped pattern of implied volatilities across different strike prices. It is widely believed that volatility smiles have to be explained by a modelling of stochastic volatility.

1.2 Issues and Motivation of the Study

As a concept, volatility is simple and intuitive. It measures the variability or dispersion about a central tendency. However, there are some subtleties that make volatility challenging to analyze and implement. Since volatility is a standard measure of financial vulnerability, it plays a key role in assessing risk/return trade-offs. Volatility is central to many investment decisions in new product areas and is the critical variable in options. The volatility of an underlying asset dictates the extent and likelihood of the option's pay - out. Most option pricing models, including Black-

Scholes require a volatility input defined as standard deviation of log relative prices. Option traders typically make option quotes in terms of volatility rather than dollar or rupee value. Option related products such as interest rates caps and floors vary in price according to volatility. The pricing and investment attractiveness of mortgage – based securities, callable bonds and other fixed income instrument with embedded options depends on market volatility. The cost of many sophisticated hedging strategies, such as portfolio insurance, also depends on the level of market volatility. Understanding and appropriately modeling the statistical characteristics of asset returns volatility will aid individual investors in their quest for accurate and precise portfolio and risk management. In other words, measures of volatility assist the investor in the choice of individual assets in a portfolio. An investor could use the same methodology to analyze, model and measure the risk on the entire portfolio held.

Financial market and institutions play a key role in an economy by channeling funds from savers to investors. Some volatility in the process of financial assets is a normal part of the process of allocating investible funds among competing users. However, excessive or extreme volatility of stock returns may be detrimental, because such volatility may impair the smooth functioning of the financial system and adversely affect economic performance. Stock return volatility may also affect business investment spending. Investors may perceive a rise in stock market volatility as an increase in risk in the equity investments. If so, investors may shift their funds to less risky assets. This reaction would tend to raise the cost of funds to firms issuing stock. Moreover, small firms and new firms might bear the brunt of this effect as investors might gravitate towards the purchase of stock in larger, well-known firms.

1.3 The Present Study

Despite the developments in the area of stock market volatility, there has been very little empirical work relating to the Indian stock markets. Studies focusing on time varying second moment character of stock returns, and issues such as asymmetric

effects of volatility and days-of-the-week effect have been few and far between. To date, there is no empirical work in the Indian context examining these issues at the level of individual scripts. In view of the high degree of relevance of the issues mentioned above – volatility as captured by time varying second moments, asymmetric effects of volatility and days-of-the-week effect, it has been felt that a comprehensive study should be undertaken to fill the gaps at an empirical level. The present study is an attempt in this direction.

1.4. Objectives of the Study

In the light of the above, the major objectives of the study are formulated as below.

i) The first major objective of the thesis is to re-examine the issue of volatility of Indian stock returns using script level data as well as aggregate indices, using ARCH – GARCH framework.

ii) The second objective is to examine the asymmetric effects of volatility.

iii) The third objective is to confirm the presence of days-of-the-week effect.

1.5 Methodology

The returns are calculated as log price differences. The daily return data for the entire period of the study has been first tested for some statistical characteristics such as serial correlation in the returns, serial correlation in the squares of returns, asymmetry in the distribution of returns, before being used in the estimation of volatility models.

In ARCH model, the conditional variance of the error term is modeled as a linear combination of squared past errors of a specified lag. In the present study, stock market volatility is examined using the ARCH, ARCH in Mean and GARCH models to fulfill the first objective of the thesis. In ARCH-in-Mean process, the conditional

variance (or standard deviation) is allowed to influence stock returns. In GARCH model, the conditional variance is modeled as a linear combination of a specified lag of squared previous errors and conditional variances of a specified lag.

Though the GARCH models successfully capture thick tail returns and volatility clustering, they are not well suited to capture the "leverage effect" since the conditional variance is a function only of magnitude of the lagged residuals and not their signs. In the exponential GARCH (EGARCH) model of Nelson (1991) σ_t^2 is formulated to depend on both the size and sign of lagged residuals. The Glosten-Jagannathan-Runkle model is a simple extension of GARCH with an additional term added to account for possible asymmetries-distribution that explains the asymmetric and fat - tail characteristics of the return. In addition to these two, the student-t distribution test is also employed in this study to examine asymmetric volatility.

Three different specifications are employed for the return and volatility equations to explain the days - of - the - week effect. The first one investigates the effect only in return equation. The second and third models incorporate the effect into both the return and volatility equations. While the second model uses the GARCH (1, 1) specification for the volatility equation, the third model considers a group of conditional variance specifications.

1.6 Nature and Sources of Data

The study makes use of daily data on ninety five individual companies and eight aggregate indices. The study period spans over the period January 1, 1990 to November 30, 2004, thus involving around 3414 number of data points, which provides rich data set for the analysis. The indices chosen are BSE 100 Index, BSE 200 Index, BSE 500 Index, BSE –Sensitive Index, CNX Nifty Junior, S&P CNX Nifty, S&P CNX 500, S&P CNX Defty. The companies that are chosen such that they make a fair representation of the different sectors.The data is drawn from the

Center for Monitoring Indian Economy (CMIE) PROWESS data base. CMIE indices of stock markets are computed in two ways namely, those weighted by market capitalization and others using equal weighted.

The index weighted by market capitalization gives us return on a portfolio whose total value is distributed among scrips in proportion to the market capitalization of the scrips. The equally weighted index provides returns on a portfolio, in which equal investment is made in each of the constituent companies. Returns on a portfolio of companies on a day are the weighted average of the returns on each of the companies in the index on the day. In the case of value-weighted indices, the weights are the proportion of market capitalization of the individual companies in the total market capitalization of all companies on the preceding day. In the case of equally weighted indices, the weights are unity. Thus in the case of value weighted indices, the returns on the index is the sum of the product of the preceding days market capitalization of each constituent company and the returns of the current day, divided by the total market capitalization of all companies on the preceding day. Broadly, there are two groups of sub-indices. One set of sub-indices is computed with each company in an index being weighted by market capitalization and the other is a set of sub-indices where each scrip is given equal weight. The first one is called the market capitalization weighted indices and the other the equally weighted indices.

Each company reported in PROWESS database necessarily belongs to one and only one industry group and size group. One of the biggest advantages of having a stock market index, which is very broad based, is that one can compute industry indices, which are reasonably disaggregated and have a reasonably large representation of the industries. The CMIE family of stock market indices contains over 150 industry sub indices. For each of these we have a set of market capitalization weighted indices and a set of equally weighted indices. Usually, large indices are divided into large market capitalization, middle level market capitalization and small market capitalization companies respectively.

1.7 Organization of the Thesis

The thesis is organized into seven chapters. The first Chapter presents the concept of stock return volatility and stylized facts about volatility. The objective of the study and methodology adopted are also discussed in this chapter. In the second chapter, the behavioural characteristics of Indian stock market are introduced. A brief note on reforms in the Indian stock market is also presented. The price behaviour and movements in daily volatility is also presented in this chapter. In the third chapter, the conditional volatility framework is presented. This chapter basically deals with the methodology and present a review of the models employed in this study. The fourth chapter presents the empirical evidences of the ARCH-GARCH model. The issues involved in the study of asymmetric volatility – theory and methods, are discussed in the fifth chapter. The empirical analysis of volatility asymmetry is presented in chapter six. The last chapter deals with days-of-the-week effect. The theoretical discussions, review of previous work and empirical findings are presented there. Towards the end of the thesis, a brief summary of the thesis is presented.

CHAPTER – II

Stock Market in India: Reforms and Trends in Prices

Contents

* * *

CHAPTER – II

Stock Market in India: Reforms and Trends in Prices

2.0 Aim of the Chapter

India's stock market continues to show wild fluctuations, often unrelated to real economy. The market crisis is a phase of such wild behaviour. Such a phenomenon is neither new nor isolated. It may be noted that the amplitude of fluctuations in the state of Indian stock market from euphoria to despondency has been regularly high, specially since 1990. Such wild movements show the extent of irrationality of the Indian stock market's behaviour. An important task for the policy makers is to rectify such erraticism which has become a characteristics of the Indian stock market. The basic aim of this chapter is to present a review of the behavioural characteristic of Indian stock market and the price behaviour in the secondary securities market.

2.1 Stock Market and the Economy

The growth of an economy depends on the rate of savings and investments. The financial market provides a linkage between savings and the preferred investment across the entities, time and space. It mobilizes savings and channelises them through securities into preferred enterprises. The securities market fosters economic growth to the extent that it: (a) augments quantities of real savings and capital formation from any given level of national income, (b) increases net capital inflow from abroad, (c) raises productivity of investment by improving allocation of investible funds, and (d) reduces cost of capital. The securities market facilitates the internationalization of an economy by linking it with the rest of the world. This linkage assists through inflow of capital in the form of portfolio investment. This implies that the domestic economy is opened up to international competitive pressures, which helps to raise efficiency. It is also very likely that the existence of a domestic securities market will deter capital outflow and provide attractive investment opportunities within the domestic

economy. A developed securities market successfully monitors the efficiency with which the existing capital stock is deployed, and thereby significantly increases its average return. As much as the securities market enlarges the financial sector, promoting additional and more sophisticated financing, increases opportunities for specialization, division of labour and reductions in costs of financial activities.

There are also other developmental benefits associated with the existence of a securities market. First, the securities market provides a fast breeding ground for the skills and judgment needed for entrepreneurship, risk bearing, portfolio selection and management. Second, an active securities market serves as an 'engine' of general financial development, and may in particular, accelerate the integration of informal financial systems with the institutional sector. Securities directly displace traditional assets such as gold and stocks of produce, or may indirectly provide portfolio assets for unit trusts, pension funds and similar financial institutions that raise savings from the traditional sector. Third, the existence of securities market enhances the scope and provides institutional mechanisms for the operation of monetary and financial policies.

2.2 Some Behavioral Characteristics of Indian Stock Market

The character of a stock market and the type of market participants it attracts are inter-related. A volatile market attracts speculation and drives away genuine investors, making it even more volatile and more dominated by speculative interests. Stock market in India came into prominence in the 1980s.The 1980s experienced the emergence of a new source of demand for Indian equities, due to the gradual opening up of the Indian stock market to portfolio investment from overseas. In the context of economic liberalization, greater reliance has been placed on the market. The wholesale changes in the Indian macro economic architecture had been due to the onset of liberalization in the early 1990's. Different sectors have changed the way they operate and indeed in the way they are put together. Amongst all the sectors, the most visible changes have come about in the financial sector, and within it, the stock market has been a star performer. In a sense this was to be expected, since it is easier

to introduce legislative and institutional changes in the financial sector in general, and stock market in particular. Capital market, even though very important, has come under the purview of reforms only since 1991. The Narasimham Committee (1991) emphasized the need for strengthening the powers of SEBI, vesting in it Capital Issues Control powers and freeing of operations in the capital market with SEBI as the supervisory and regulatory authority. As per its recommendations, the SEBI Act was passed in March 1992, vesting legal powers in SEBI to act as a regulatory authority on the stock and capital markets in India. The twin objectives with which the SEBI has been created are: development of the capital market and protection of the interest of investors, and the regulation of the securities market.

The market's erratic functioning has been a matter of concern for policy-makers. The causes of such erratic behaviour may be traced to certain trading practices on Indian stock exchanges. These fluctuations expose genuine investors to grave settlement risks, delays, hassles and poor quality of market services. Abnormally high or abnormally low prices (i.e, prices which deviate far from true values) should get corrected automatically and quickly through market forces themselves in a market, which is rational. The behavioral characteristic (euphoria and despondency) of the stock market is going to be a serious problem for the economy in the coming years. This is because of the dependence of the country's future industrial growth on the stock market, which is going to increase in pursuance of liberalization policy. It is therefore high time that the systems and practices of the stock market are set right to ensure that it behaves more rationally.

The main factor underlying the Indian stock market's upward movement since the mid- 1980's was a policy-induced rise in the demand for equities among domestic investors, reinforced later by a gradual opening up of market to overseas investors. The market gathered momentum during the early 1990's, leading to an unprecedented boom. However, by the mid-1990's, the market had turned full circle and collapsed. The demand side momentum in favor of equities as regards the domestic investors had completely dissipated by the end of 1996. As a result, the Indian market

underwent a substantial downward revaluation, pushing the market Price/Earnings ratio even below the mid–1980's level. The unprecedented boom of the early 1990's was followed by an unprecedented crisis in 1995-97. The recovery from the crisis has been proving difficult.

The report of the Securities Exchange Board of India – National Council of Applied Economic Research (NCAER) working group (1999) on the development of the capital market, was concerned more with the market's development than its regulation. It is recognized that even though there was considerable expansion of the market in the last few years, "the annual net flow of investible funds from the rest of the economy to the private corporate sector has been modest", that there remained a need for "greater thrust towards equity culture" but the basic ground rules on the operation of the capital market in India are inadequate for the altered situation".

Gupta (1999) observed that the primary task of a capital market is to ensure efficient resource allocation. Investors are not willing to invest in Financial Institution shares at one–fourth of their unsecured bonds giving tax incentives. Since quantitative measures of investor's confidence are not available, the level of investor's confidence in the equity or debt markets cannot be found empirically. In fact, all the problems that would cause loss of investor's confidence are there in the debt market. The remedy for the market crisis requires that (i) the regime of investor protection should be much tightened to strengthen the investor's confidence; (ii) there should be much faster implementation of reforms of the settlement and share transfer system, and (iii) market regulation should be quickly extended to cover large parts of trading channels which still remain outside the regulatory framework, through which the bulk of retail investors have to trade.

There is a widely held belief among market community, economists and market authorities that higher liquidity of the shares of larger companies means substantially lower market volatility, has not been confirmed by the Indian market's behaviour. The higher liquidity enjoyed by shares of larger companies has not helped to reduce

market volatility at all, as measured by Gupta (1999). The reason for this may lie in the fact that highly liquid shares in India attract most of the speculative activity, which accentuates, rather than stabilizes, price fluctuations in the Indian share market. Besides, smaller company shares remain infrequently traded, sometimes not traded for a whole year.

2.3 Reforms in Indian Securities Market

Securities are the most insecure instruments. The only ingredient common to all types of securities is its associated insecurity. If there is a market for such insecure instruments, that market would collapse if somebody does not regulate away the insecurities. During the 1980s and 1990s, it was increasingly realized that an efficient and well-developed securities market is essential for sustained economic growth. In order to improve the quality of the market i.e., to improve market efficiency, enhance transparency, prevent unfair trade practices and bring Indian market up to the international standards, a package of reforms consisting of measures to liberalize, regulate and develop the securities market has been launched since the early 1990s. The practice of allocation of resources among different competing entities as well as its terms by a central authority was discontinued. The issuers complying with the eligibility criteria were allowed freedom to issue securities at the market-determined rates. The secondary market overcame the geographical barriers by moving to screen based trading. All kinds of securities – debt and equity, government and corporate – are traded on exchange side by side. The trading cycle shortened to a day and trades are settled within three working days, while all deferral products are banned. Physical security certificates have almost disappeared. It is clear that there is an apparent contradiction that the reforms aim at liberalization while regulations appear to restrict liberalization. Liberalization means replacement of one set by another set of more liberal code/statute, which allow full freedom to economic agents, but influence or prescribe the way they should carry out their activities, so that the liberalized markets operate in an efficient and fair manner while the risks of systematic failure are minimized. Some of the major reforms since the early 1990's as follows:

a) Control over Issue of Capital: A major initiative of the liberalization policy was the repeal of the Capital Issues (Control) Act, 1947, in May 1992. With this, Government's control over issue of capital, pricing of the issues, fixing of premium and rates of interest on debentures, etc., ceased and the market was allowed to allocate resources to competing uses. In the interest of investors, SEBI issued Disclosures and Investor Protection (DIP) guidelines. The guidelines allow issuers complying with the eligibility criteria to issues securities at the market-determined rates.

b) Establishment of Regulator: A major initiative of regulation was establishment of a statutory autonomous agency called SEBI to provide reassurance that it is safe to undertake transactions in securities. It was empowered adequately and assigned the responsibilities to: (a) protect the interests of investors in securities, (b) promote the development of the securities market, and (c) regulate the securities market. Its regulatory jurisdiction extends over corporate in the issuance of capital and transfer of securities, in addition to all intermediaries and persons associated with securities market. All market intermediaries are registered and regulated by the SEBI.

c) Screen Based Trading: A major developmental initiative was a nation-wide on-line fully-automated screen based trading system (SBTS), through which a member can punch into the computer the quantities of securities and the prices at which he/she would like to transact. And the transaction is executed as soon as it finds a matching sale or buy order from a counter party. SBTS electronically matches orders on a strict price/time priority and hence cuts down on time, cost and risk of error, as well as on fraud resulting in improved operational efficiency. It allows faster incorporation of price sensitive information into prevailing prices, thus increasing the informational efficiency of the markets. It enables market participants to see the full market on real-time, making the market transparent. It allows a large number of participants, irrespective of their geographical locations, to trade with one another simultaneously, improving the depth and liquidity of the market – over 10,000 terminals are creating waves by clicks from over 400 towns/cities in India. It provides full anonymity by

accepting orders, big or small, from members without revealing their identity, thus providing equal access to everybody. It also provided a perfect audit trail, which helps to resolve disputes by logging-in the trade execution process in entirety.

d) Risk Management: A number of measures have been introduced to manage the risks in the market, so that the participants are safe and the market integrity is protected. These include:

i) Trading Cycle: One of the far-reaching changes for the Indian Securities market in 2001-02 has been the introduction of compulsory T+5 rolling settlement for all scrips listed and traded in any Stock Exchange in India. Rolling settlement was introduced on voluntary T+5 basis in the Demat segment of the Stock Exchanges on January 15,1998, to expedite the trading and settlement process and to improve the efficiency of the securities market. Thus, the total number of scrips traded on rolling settlement was 414. By December 31,2001, rolling settlement had been extended to the remaining scrips listed on all exchanges. It was subsequently decided to further shorten the settlement cycle to T+3 for listed scrips from April 1, 2002.

ii) Dematerialization: Settlement system in Indian stock exchanges gave rise to settlement risk, due to the time elapse before trades were settled. Physical movement of paper settled the trades. This had two aspects: first, the settlement of trade in stock exchanges by delivery of shares by the seller and payment by the purchaser. The process of physically moving securities from the seller to the ultimate buyer through the seller's broker and buyer's broker took time with the risk of delay somewhere along the chain. The second aspect related to transfer of shares in favour of the purchaser by the company. The system of transfer of ownership was grossly inefficient, as every transfer involved physical movement of paper securities to the issuer for registration, with the change of ownership being evidenced by an endorsement on the security certificate. To remove these problems, the Depositories Act 1996 was passed to provide for the establishment of depositories in securities, with the objective of ensuring free transferability of securities with speed, accuracy and security by (a) making securities of public limited companies freely transferable,

subject to certain exceptions; (b) dematerializing the securities in the depository mode; and (c) providing for maintenance of ownership records in a book entry form. In order to streamline both the stages of settlement process, the Act envisages transfer of ownership of securities electronically by book entry without making the securities move from person to person.

iii) Derivatives: To assist market participants to manage risks better through hedging, speculation and arbitrage, Securities Control (Regulation) Act was amended in 1995 to lift the ban on options in securities. The SC(R)A was amended further in December 1999 to expand the definition of securities to include derivatives, so that the whole regulatory framework governing trading of securities could apply to trading of derivatives also. A three-decade old ban on forward trading, known as *BADLA*, which had lost its relevance and was hindering introduction of derivatives trading, was withdrawn. Derivative trading took off in June 2000 in two exchanges.

iv) **Settlement Guarantee**: Various measures were under taken to address the risk in the market. Clearing corporations emerged to assume counter party risk. Trade and settlement guarantee funds were set up to guarantee settlement of trades, irrespective of default by brokers. These funds provide full notation and work as central counter party. The Exchanges/clearing corporations monitor the positions of brokers on real time basis.

2.4 Significance of Liberalization in Indian Economy

As a result of these reforms, market design has changed drastically for better. This has boosted the confidence of international investors in India's, securities market. Indian market is getting integrated with the global market, though in a limited way, through euro issues. Since 1992, when they were permitted access, Indian companies have raised about Rs 37,000 crores through American Depository Resources / Global Depository Resources. More than 500 Financial Institutional Investors (FIIs) are currently registered with the SEBI. They have made net cumulative investments of about 16 billion U.S dollar by the end of February 2003. The extent of reforms in the

securities market has been visible from the transaction cost. The market impact cost of transactions in equity on stock exchanges reduced from 0.75 per cent in mid-1993 to 0.10 per cent in January 2003, while the brokerage reduced from 3 per cent to 0.15 per cent respectively. Low transaction costs hold growth attraction for investors in securities and issuers of securities to participate in the market on a larger scale.

Bajpai (2003) identified three main sets of entities, which depend on security market. While the corporate and government raise resources from the securities market to meet their obligations, the households invest their savings in the securities. He highlighted a few statistics, mostly taken from the Indian Securities Market Review (a publication of the National Stock Exchange), to indicate the level of significance. While corporate sector and government together raised a total of Rs. 226,911 crores from the securities market during 2001-2002, there are about 20 millions investors who have invested in the securities. The Indian economy witnessed a descent growth of 6 per cent per annum in the 1990s against the euphemistically described Hindu growth rate of 3.5 per cent over the preceding four decades. This was possible by the contributions mostly made by the organized secondary and tertiary sectors (industry and services). The securities market helped these organized sectors, corporate and government to raise resources to realize a growth rate of 6 per cent. Of late, the activity in the securities market has slowed down, so also the level of activity in the economy.

a) Corporate Sector: The 1990s witnessed the emergence of securities market as a major source of finance for trade and industry. A growing number of companies are accessing the securities market, rather than depending on loans from FIs/Banks. The corporate sector is increasingly depending on external sources for meeting its funding requirements. There appears to be a growing preference for direct financing (equity and debt) to indirect financing (bank loans) within the external sources. According to the *Center for Monitoring Indian Economy* data, the share of capital market based instruments in externally resource raised increased to 53 per cent in 1993-94, but declined thereafter to 31 per cent by 2000-01.

The average annual capital mobilization from primary market, which used to be about Rs.70 crore in the 1960s and about Rs.90 crore in the 1970s, increased mani-fold during the 1980s, with the amount raised in 1990-91 being Rs. 4,312 crore. It received a further boost during the 1990s, with the capital raised by non-government public companies rising sharply to Rs. 26,417 crore in 1994-95. However, the market appears to have dried up in the late 1990s due to the interplay of various factors. The corporates have shifted focus to other avenues for raising resources, like private placement where companies are much less. Available data, although scanty, indicate that private placement has become a preferred means of raising resources for the corporate sector. It contributed about Rs. 65,000 crore during 2001-02,while a total of Rs. 74,403 crore was raised from securities market by the sector during 2001-02.

b) **Governments:** Along with increased fiscal deficits of the government, its dependence on market borrowings to finance them has increased over the years. During the year 1990-91, the state and the central governments respectively financed nearly 14 per cent and 18 per cent of their fiscal deficits through market borrowings. In percentage terms, the dependence of state governments on market borrowings did not increase much during the decade 1991-2001. In the case of central government, it increased to 69.4 per cent by 2001-02.Together the central and state governments borrowed Rs. 110,510 crores from market during 2001-02, against Rs. 10,557 crores in 1990-91.

c) **Households:** According to *Reserve Bank of India* data, the household sector accounted for 89 per cent of the gross domestic savings during 2000-2001. They invested only 4 per cent of their savings in securities, including government securities and units of mutual funds, during 2000-2001. The share of financial savings of the household sector in securities [shares, debentures, public sector bonds and units of UTI and other mutual funds and government securities] is estimated to have gone down from 22.9 per cent in 1991-92 to 4.3 percent in 2000-01.

d) Investor population: The Society for Capital Market Research and Development carries out periodical survey of household investors to estimate their number. Their first survey (1990) recorded the total number of share owners at 90-100 lakhs, their second survey (1993) at 140-150 lakhs in mid 1993 and the third survey at around two crores during 1997 end. According to the SEBI – NCAER survey of Indian investors conducted in early 1999, an estimated 12.8 million (7.6 per cent), of all Indian households representing 19 million individuals had directly invested in equity shares and debentures by the end of the financial year 1998-99. An estimated 15 million (nearly 9 per cent), of all households, representing at least 23 million unit holders, had invested in units of mutual funds.

e) **Disinvestment Programme** The disinvestments programme in India would not have been successful without a well-developed securities market. So far, Government of India has been able to disinvest to the tune of over Rs 31,000 crore, including about Rs. 5,000 during the current fiscal.

2.5 Price Behavior in the Secondary Securities Market During 1996-2003

For most emerging markets, liberalization is an essential policy tool that attracts much needed foreign capital for investment purposes. In an era of increased financial integration, stock market liberalization could become a policy option that many, if not all, economies may well adopt in the near future. Some studies have examined the issue of liberalization and volatility in the stock market and empirically found that market opening does not always result in increased volatility but may even lead to decreased volatility. The effect of stock market liberalization on market volatility in particular is an important issue that the emerging market economies must consider before their decision to liberalize and perhaps even after. This is because volatility is an unattractive feature that has adverse implications for decision pertaining to the effective allocation of resources, and therefore for investment. Volatility makes

investors more averse to holding stocks due to uncertainty. A greater risk premium results in a higher cost of capital, which then leads to less private physical investment. In addition, greater volatility may increase the value of the 'option to wait' thereby delaying investment. Also, weaker regulatory systems in developing markets reduce the efficiency of market signals and processing of information, which further magnifies the problem of volatility.

The Indian stock markets has witnessed a gradual rise in stock prices during the first half of the year 1997-98, but thereafter, various macro economic factors both at the national and international levels, have brought down the valuations again. On a year to year basis, the 30 stock sensitive index (Sensex) of the Bombay Stock Exchange, Mumbai, increased from 3360.89 on March 31,1997 to 3892.75 on March 31,1998, an increase of 531.86 points. Nifty index comprising of fifty stocks also rose by 146.6 points to close the year at 1116.90. The broad based S&P CNX 500 index, which represents 500 scrips, increased from 631.42 to 715.54. The currency crises in the South-East Asian countries had also affected the Indian capital markets, but the impact was not severe during the year. The net FIIs flows were negative for three consecutive months in November, December 1997 and January 1998, for the first time since they were allowed to invest in the Indian securities markets. The lowest value of Sensex for the year 1997-98 was 3209.55, recorded in January 28,1998. The markets again showed a rising trend thereafter. The Sensex closed the financial year 31st March 1998 at 3892.75, which reflected a gain of 15.82 per cent over the last year.

The equity price in general registered receding trend during the major part of the financial year 1998-99. On point-to-point basis, the Bombay Sensex decelerated from 3892.75 as on March 31,1998 to 3739.96 as on March 31, 1999, recording a decrease of 152.79 points or 3.92 per cent. The BSE 100 had a fall of 45.77 points (-2.69 percent) from 1697.14 to 1651.37, followed by S&P CNX Nifty registering a fall of 38.85 points or 3.47 per cent from 1116.92 to 1078.05. The S&P CNX 500 Index with 500 scrips recorded a rise of 42.71 points (5.96 per cent) from 715.54 to 758.25.

The rupee was volatile for sometime and the exchange rate fell sharply on a few occasions in 1998, in turn affecting the sentiments and inhibiting fresh FII inflows. The FIIs inflows were negative for many months during 1998-99 on account of the East Asian Crisis.

The year 1999-2000 was marked by a smart recovery and strong spurt in equity prices. During the largest part, the equity market registered upward movement. The BSE Sensex on point-to-point basis recorded a growth from 3739 as on March 31,1999 to 5001 as on March 31, 2000. The BSE 100 also shot up by 76 per cent, 1651 to 2902 during the same period. On the other hand, S&P CNX Nifty increased by 42 per cent. The performance of stocks prices was not linear during the year as shown by the movement of daily indices. In the beginning of the year, the BSE Sensex was in the grip of bearish phase following the presentation of the Budget 1999-2000 at the end of February 1999. (see table 2.1)

The stock prices in the secondary market remained subdued and were dominated by bearish sentiment during 2001-02. Almost all major stock market indicators like marketing capitalization turn over and P/E ratios showed a falling trend during 2001-02. However, the fall in major indices and other market indicators was substantially less than that noticed during 2000-01. The BSE Sensex on a point-to-point basis decreased from 3604 in March 30, 2001 to 3496 by March 28,2002, a fall of four per cent. The S&P CNX Nifty witnessed a decline of two per cent from 1148 in March 30, 2001 to 1129 by March 28, 2002. On the contrary, the other indices on a point-to-point basis, namely the BSE Natex, the S&P CNX 500 and the BSE Dollex registered a rise of two per cent, and three per cent respectively during the 2002 under review.

The stock prices however, suffered several reversals during the 2002 on account of several factors, such as investor perception of slowing down of global economies combined with reports on the irregularities in the trading of shares on the Indian stock markets. The BSE Sensex had lost 177 points by March 2, 2001 and further 625 points on March 30, 2001 over the closing of March 1, 2001. The other indices also

suffered similar losses. The bearish movement in stock prices was further precipitated by stock market crash in USA, following the terrorist attack on September 11, 2001 and massive sales of stocks by FIIs in September 2001. The BSE Sensex dropped to 2681 as on September 17, 2001 and further to 2600 by September 21, 2001 from 3184 in September 10, 2001, recording a loss of 503 points and 584 points respectively. The Sensex picked up slowly and reached 3287 by November 21, 2001 and further to 3443 by December 10, 2001, before coming down sharply to 3132 by December 27, 2001, following the terrorist attacks on the Indian Parliament. Despite such reversals, the stock price regained stability as the BSE Sensex recovered to 3402 as on January 7, 2002 and further to 3469 by March 28,2002. These trends were visible in other indices also. The S&P CNX Nifty, which had risen from 987 in November 7, 2001 to 1115 in December 10, 2001, declined to1020 as of December 27, 2001 shedding nearly 95 points. Thereafter, Nifty recovered to 1189 by February 27, 2002 before declining to 1130 by March 28, 2002 (see table 2.2)

Table 2.1 Movements of Various Share Price Indices During 1997-99

Index	31-Mar-97	31-Mar-98	31-Mar-99	1997-98*	1998-99*
BSE Sensex	3360.89	3892.75	3739.96	15.82	- 3.92
BSE Natex	1463.69	1697.14	1651.37	15.95	- 2.69
S&PCNXNifty	968.3	1169.9	1078.05	15.35	-3.47
S&P CNX 500	631.42	715.54	758.25	13.32	5.96
BSE Dollex	151.9	158.75	149.11	4.51	6.07

**percentage increase/decrease over the previous year*
Source : SEBI

Table 2.2 Movements of Various Share Price Indices During From 2000 - 02

Index	31-Mar-00	30-Mar-01	28-Mar-02	2000-01*	2001-02*
BSE Sensex	5001.28	3604.38	3469.35	-27.9	-3.7
BSE Natex	2902.2	1691.71	1716.28	-41.7	1.5
S&PCNXNifty	1528.45	1148.2	1129.55	-24.9	-1.6
S&P CNX 500	1322.9	754.18	775.5	-43	2.8
BSE Dollex	237.86	130.89	134.39	-45	2.7

**percentage increase/decrease over the previous year*
Source : SEBI

2.6 Movements in Daily Volatility

Daily volatility is calculated as a standard deviation of the natural log of daily returns on the indices for the respective months. Volatility in stock prices in India, by and large has exhibited a declining trend except over certain months. The movement in volatility of BSE stocks reflected a rising trend during the months of April to July 1998-99 and were higher than that observed during the same months in the previous year. The volatility during the following months registered a continuous decline to 1.83 per cent by March 1999 from 5.26 per cent in June 1998. Moreover, volatility was substantially lower during the second half of 1998-99 than during the same period in 1997-98. Further, the volatility of Natex also witnessed decline trend.

The BSE 100 volatility, which was 2.36 per cent in April 1998, increased to 5.92 per cent in June 1998 and thereafter gradually decreased to 1.67 by March 1999. S&P CNX Nifty and CNX Junior Nifty also reflected the same trend. The BSE Sensex which exhibited volatile movement in April 1999 with a volatility of 3 per cent improved its performance in March 2000 with the volatility registering a decline to 1.72 per cent. Natex more or less, reflected a downward trend, recording a decline in the volatility from 3.09 per cent in April 1999 to 2.23 per cent in March 2000. Thus, the sensex volatility recorded in 1999-2000 was lower than the one observed during 1998-99, whereas, Natex volatility showed a rising trend during 2000, which declined

in March 2000. Thus, the volatility movements of S&P CNX Nifty and CNX Junior indices exhibited a falling trend in their volatility from 3.35 per cent in April 1999 to 1.82 per cent in March 2000. Likewise CNX Junior Nifty also recorded a similar behavior. (see Table 2.3)

The BSE Sensex, which had a volatility of 2.4 per cent in April 2001, witnessed an improved performance with only 0.9 per cent volatility in May 2001, which further fell to 0.6 per cent in August 2001. However, September 2001 was marked by a pronounced volatility in the market, as reflected in BSE Sensex measuring volatility at 2.8 per cent,which receded to 1.3 per cent by March 2002 . Volatility of BSE 100 Index, which was 2.7 per cent in April 2001, declined to 0.7 per cent in August 2001 but rose back to 1.1 per cent in March 2002.(see table 2.4).

During 2001-02, the sensex daily volatility declined from 2.4 per cent in April 2001 to 0.7 per cent in August 2001. The volatility, however, sharply jumped up to 2.8 per cent in September 2001. This was the month when BSE 100 also showed high volatility following the September 11, 2001 crash of the Nasdaq market. The volatility at BSE Sensex , thereafter declined to 0.9 per cent in January 2002 before rising to 1.5 per cent in February and again declined to 1.3 percent in March 2002. The pattern of movement of BSE 100 Index volatility was similar to Sensex volatility. It showed high volatility in September 2001 at 2.9 per cent that declined to 1.1 per cent in the Month of March 2002.The indices of S&P CNX Nifty and CBX Nifty junior registered the same trend in their respective volatilities (see table 2.5).

Table 2.3: Daily Volatility: BSE Sensex and BSE 100 Index During 1997 - 2000
(in percentage)

Month	BSE Sensex			BSE 100 Index		
	1997-98	1998-99	1999-00	1997-98	1998-99	1999-00
April	3.57	2.17	3.17	3.06	2.36	3.09
May	0.85	3.24	2.17	0.81	2.93	2.14
June	3.24	5.26	1.43	3.29	5.92	1.27
July	1.74	3.45	1.60	1.46	3.25	1.61
August	4.72	2.66	1.33	4.56	2.26	1.40
September	2.15	3.21	1.16	2.21	2.91	1.20
October	2.51	1.93	2.11	2.30	1.89	2.30
November	3.49	2.60	1.52	3.37	2.48	1.45
December	3.27	2.85	1.45	3.10	2.93	1.30
January	4.83	2.87	1.90	4.74	2.84	1.92
February	2.62	1.70	2.51	2.70	1.87	2.78
March	2.30	1.83	1.72	2.60	1.67	2.23

Note: Daily volatility is calculated as standard deviation of natural log of daily returns on the indices for the respective months.
Source: SEBI

Table 2.4: Daily Volatility: BSE Sensex and BSE 100 Index During 2000-03
(in percentage)

Month	Sensex			BSE 100 Index		
	2000-01	2001-02	2002-03	2000-01	2001-02	2002-03
April	3.93	2.42	0.99	3.99	2.77	1.06
May	2.89	0.96	1.55	3.09	1.08	1.43
June	1.50	1.24	1.18	1.69	1.45	1.10
July	2.13	1.19	1.03	2.27	1.12	1.06
August	1.16	0.66	0.92	1.34	0.69	0.86
September	2.20	2.83	0.83	2.30	2.94	0.73
October	1.61	1.46	0.96	2.06	1.38	0.86
November	1.51	1.26	0.69	1.45	1.25	0.59
December	1.52	1.36	0.83	1.92	1.55	1.01
January	1.31	0.93	0.72	1.52	1.01	0.74
February	1.64	1.51	0.78	1.88	1.68	0.80
March	2.85	1.31	1.16	3.26	1.12	1.08

Note: Daily volatility is calculated as standard deviation of natural log of daily returns on the indices for the respective months.
Source: SEBI

A comparative analysis of volatility trends among the indices shows that market have, by and large, remained stable except for short periods of bursts in volatility. Volatility was the highest in case of BSE 100 Index during the year of 2001-02 followed by BSE Sensex, CNX Nifty Junior and Junior Nifty. It may be mentioned that BSE 100 Index registered the highest level of volatility in September 2001 at 2.9 per cent followed by a volatility of 2.8 per cent recorded by the BSE Sensex . Junior Nifty recorded the volatility in the same month at 2.6 per cent by CNX Nifty and 2.8 per cent. (see table 2.6)

Table 2.5: Daily Volatility of S&P CNX Nifty and Junior Nifty During 1998-2000

(in percent)

Month	S&P CNX Nifty		Junior Nifty	
	1998 - 99	1999 – 2000	1998 –99	1999 –2000
April	1.96	3.35	6.04	3.58
May	3.09	2.19	2.93	2.08
June	5.11	1.36	7.72	1.25
July	3.12	1.60	3.89	1.85
August	2.57	1.35	1.84	1.61
September	3.1	1.33	2.16	1.60
October	1.81	2.09	2.54	3.31
November	2.58	1.73	1.98	2.42
December	2.69	1.45	3.1	2.07
January	2.91	1.95	3.15	2.97
February	1.79	1.89	2.89	2.62
March	1.53	1.82	2.74	3.06

Note: Daily volatility is calculated as standard deviation of natural log of daily returns on the indices for the respective months.

Source: SEBI

Table 2.6: Daily Volatility of S&P CNX Nifty and Junior Nifty During 2000 - 03

(in percent)

Month	S&P CNX Nifty			Junior Nifty		
	2000-01	2001-02	2002-03	2000-01	2001-02	2002-03
April	3.60	2.23	1.11	4.59	2.63	0.87
May	2.43	0.89	1.35	4.04	1.21	1.78
June	1.39	1.25	1.14	2.19	1.34	1.35
July	1.76	1.03	0.99	2.00	1.18	1.47
August	1.08	0.54	0.86	1.88	0.59	1.06
September	2.00	2.62	0.73	2.79	2.80	1.40
October	1.61	1.26	0.85	2.35	0.91	1.04
November	1.43	1.20	0.69	1.86	0.98	0.62
December	1.45	1.24	0.92	2.44	1.85	0.98
January	1.20	1.00	0.80	1.68	0.90	1.10
February	1.56	1.48	0.87	2.42	0.21	1.20
March	2.91	1.10	1.08	3.40	1.17	1.18

Note: Daily volatility is calculated as standard deviation of natural log of daily returns on the indices for the respective months.

Source: SEBI

2.7 Concluding Remarks

In this chapter we have explained the behavioural characteristics of Indian stock market. It has been observed that Indian stock market continues to exhibit wild fluctuation often unrelated to real economy, as a result of which it has acquired the characteristics of a 'bubble' market and not a market governed by economic fundamentals. An important task for the policy makers and regulators is to rectify such erratic behaviour. There are some major reforms that have taken place since the 1990's. The significance of the reforms has been discussed in the Chapter. The price behaviour of Indian stock market shows a gradual rise during the first half of 1997-98, but thereafter various economic factors brought the valuation down. The equity price in general registered a rising trend during the major part of the financial year 1998-99. The year 1999-2000 was marked by a smart recovery and strong spurt in equity prices. The stock prices in the secondary market remained subdued and were dominated by bearish sentiments during 2001-02.

CHAPTER – III

Stock Market Volatility: A Review of Models

Contents

* * *

CHAPTER - III

Stock Market Volatility: A Review of Models

3.1 Introduction

In financial economics literature, there are several non-linear models for explaining the return behaviour. However, only a small number of non-linear models have been found to be useful for modelling financial data. The most popular non-linear financial models are the ARCH-GARCH class of models, which are used for modelling and forecasting volatility, and also the switching models, which allow the behaviour of a series to follow different processes at different points in time.

A non-linear model may be used where the relationship between variables as such involves complexities. But the choice between linear and non-linear models may also be made partly on statistical grounds, i.e, deciding whether a linear specification is sufficient to describe all of the most important features of the data at hand. The traditional tools of time series analysis (such as autocorrelation or partial autocorrelation function, or 'spectral analysis', are available to detect non-linear behaviour in financial time series. But they seem to be of little use. A number of tests for non-linear patterns in time series are now available to the researcher. These tests may be broadly divided into two types, *viz*, general tests and specific tests. General tests, also sometimes called 'portmanteau' tests, are usually designed to detect many departures from randomness in data. The implication is that, such tests will detect a variety of non-linear structures in data. One of the widely used tests is known as Brock test (BDS-1996) which is a pure hypothesis test. That is, it has its null hypothesis that the data are pure noise (completely random). It has been argued that this test has the power to detect a variety of departures from randomness like linear or non-linear stochastic processes, deterministic chaos, etc. The BDS test follows a standard normal distribution under the null hypothesis. Specific tests, on the other

hand, are usually designed to have the power to find specific types of non-linear structure.

The basic aim of the chapter is to present a quick review of various volatility models that fall mainly in the ARCH-GARCH framework. Such a review provides the background for empirical implementation of stock market volatility models in the Indian context which forms the core of the present thesis.

3.2 Volatility Models

Modelling and forecasting stock market volatility has been the subject of vast empirical and theoretical investigation over the past decade or so, among academics and practitioners alike. There are a number of motivations for this line of inquiry. Arguably, volatility is one of the most important concepts in the whole of finance literature. Volatility, as measured by the standard deviation or variance of returns, is often used as a crude measure of the total risk of financial assets. Many value-at-risk models for measuring market risk require the estimation or forecast of a volatility parameter. The volatility of stock market prices also directly enter into the Black–Scholes (1974) formula for deriving the prices of traded options.

Volatility of daily stock returns over a month, year or the entire period has been calculated as the standard deviation of daily returns of individual stocks or market indices. The returns are calculated as log price differences. The daily return data for the entire period of study is first tested for some statistical characteristics, such as serial correlation in the returns, serial correlation in the squares of returns, asymmetry in the distribution of returns. Later, the time series of daily stock returns has been estimated by using the Auto Regressive Conditional Heteroskedastic (ARCH) or Generalized ARCH (GARCH) class of models. Considerable research effort has gone into modelling time varying conditional heteroskedasticity in asset returns. It is important because if both returns and volatility of stock returns can be forecasted,

then it becomes possible to construct dynamic asset allocation models, which use time dependent mean and variance optimization over each period.

The prime motivation behind the development of conditional volatility models emanated from the fact that the then existing linear time series models were inappropriate, in the sense that they provided very poor forecast intervals and it was contended that like conditional mean, variance (volatility) could as well evolve over time. Engle proposed the ARCH model for the first time. Though the model was initially developed in the macroeconomic context, the same was found to be extremely useful in explaining the time varying second moment or volatility in other fields as well. Applied research in this direction thereafter surged ahead tremendously, leading to literally hundreds of papers have been published with various extensions, modifications and empirical applications. Some good surveys on the developments in this line may be found in literature (Bollerslev, *et. al.*, 1994, Nelson 1991, and Pagan 1984).

It is apparent from the literature on financial markets that stock market volatility is not a directly observable variable. Hence, persistent efforts have been made to develop a suitable model to measure and explain stock market volatility, and the quest is still on. In this regard, two distinct ways used in the literature to define volatility may be identified. The first approach measures volatility directly from the data, using a transformation of the return series, while the second approach considers volatility as an unobservable variable, using either a statistical or an economic model to obtain an estimate of the volatility process.

3.3 Direct Measurement

Volatility is measured directly from the returns series in two distinct ways. The first the "naïve" measurement, approach which involves constructing volatility series from simple transformations of residuals from a return generating model which is estimated using sample data. The second way is to measure volatility directly from

the data. This approach is motivated by the great wealth of high frequency data that has been made increasingly available to researchers in recent years. In this case, construction of a volatility series at a given frequency uses high frequency data. For example, standard deviation of the intra–daily observations on stock returns (say, at the five – minute frequency) is an estimate of that day's market volatility. This is the approach followed in some of the recent studies [Anderson and Bollerslev, 1997; Ebens, 1999; and Hsieh, 1989]. It has been contended that "naïve" volatility series are particular cases of aggregative series, where the number of observations used in the construction of the aggregative series is equal to one. This suggests that aggregative series are a better way of measuring volatility than the naïve approach, although the questions associated with infrequent trading may pose some difficulties in the construction of aggregative series, namely, the frequency of the intra–daily data to be used [Ebens (1999)].

Models of volatility be broadly classified into: historical volatility. implied volatility models, autoregressive volatility models, autoregressive conditionally heteroscedastic models.

3.4 Historical Volatility:

The simplest model for volatility is the historical estimate. Historical volatility simply involves calculating the variance (or standard deviation) of returns in the usual way over some historical period, and this then becomes the volatility forecast for all future periods. The historical average variance (or standard deviation) was traditionally used as volatility input in the options pricing models, although there is a growing body of evidence suggesting that the use of volatility predicted from more sophisticated time series models will lead to more accurate option valuations [Akgiray 1989]. Historical volatility is still useful as a benchmark for comparing the forecasting ability of more complex time models.

3.1 ABSTRACT

Chorismate mutase [CM] is a branchpoint enzyme of the aromatic amino acid biosynthesis pathway and catalyzes a committed step in the biosynthesis of tyrosine and phenylalanine. Naturally occurring variants of the enzyme chorismate mutase are known to exist that exhibit diversity in enzyme structure, regulatory properties, association with other proteins etc. These features have been the apparent reasons behind extensive studies that have been carried to date on the enzyme. Chorismate mutase was not annotated in the initial sequenced genome of *Mycobacterium tuberculosis* on account of low sequence similarity that exists between known CMs. In the present work, CM activity was assigned to the hypothetical protein corresponding to ORF Rv1885c of *M. tuberculosis*. Extensive biochemical and biophysical characterization of the recombinant protein suggests its semblance to the aroQ class of CMs, prototype examples of which include the *E. coli* and yeast CMs. These results also demonstrate that unlike the P and the T proteins of *E. coli*, *M. tuberculosis* CM does not have any associated prephenate dehydratase or dehydrogenase activity, indicating its monofunctional nature. Further, the ability of the 33 amino acids at the N-terminus of *M. tuberculosis* CM to act as a signal sequence to export *E. coli* alkaline phosphatase into the periplasmic space suggests that *M. tuberculosis* CM belongs to the periplasmic subclass of aroQ class of CMs. A redundancy of CM in *M. tuberculosis* genome was evident from the observation that the hypothetical protein coded by ORF Rv0948c also shows *in vitro* CM activity. The present study is the first report wherein biochemical evidence has been provided for the existence of two monofunctional CMs in any pathogenic bacteria. Considering the fact that the enzyme does not have a human homologue, understanding the mechanism of its action would open up avenues for the development of novel therapeutic interventions for tuberculosis.

3.2 INTRODUCTION

Mycobacterium tuberculosis has developed ingenious mechanisms to survive inside the hostile environment presented by the host and to acquire essential nutrients from this adverse environment [Wayne *et al.*, 1994; Betts *et al.*, 2002]. These attributes enable the bacterium to establish a successful infection inside the host. For the development of new therapeutic strategies to check tuberculosis, there is a need for the identification of novel targets that are not only unique to *M. tuberculosis* but blocking of which would either prove lethal to the bacterium or render it extremely susceptible to the host immune response. In this context, understanding the mechanism of action of the aromatic amino acid pathway enzymes of *M. tuberculosis* assumes utmost importance as most of the corresponding genes have been reported to be essential for the bacterium and there is no human or mammalian counterpart of the same. [Parish and Stoker, 2002; Sassetti *et al.*, 2003]. Moreover, amino acid auxotrophs of *M. tuberculosis* do not survive or multiply in macrophages [Bange *et al.*, 1996; Gordhan *et al.*, 2002], suggesting that these nutrients are not available within the compartment of the macrophage in which the bacteria reside.

Chorismic acid is the last common precursor in the aromatic amino acid biosynthesis pathway and is a substrate for multiple enzymes [Gibson and Jackman, 1963; Dosselaere *et al.*, 2001] [Figure 3.1]. Hence it has always been of interest to explore how a microbe partitions chorismate into diverse metabolic pathways. Various enzymes that utilize chorismate as a substrate include chorismate mutase [CM], anthranilate synthase [AS], isochorismate synthase [ICS] and p amino benzoate synthase [Haslam, 1993; Weiss and Edwards, 1980].

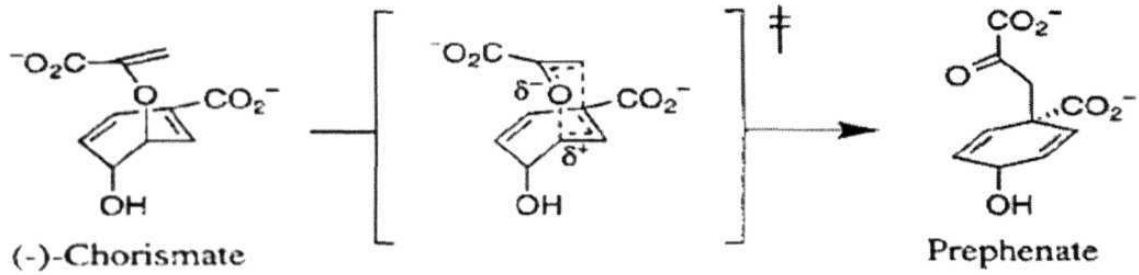
3.2.1 Chorismate mutase

Chorismate mutase [EC 5.4.99.5] is a branchpoint enzyme of the shikimate pathway and plays a crucial role in the biosynthesis of aromatic amino acids. The enzyme channels chorismate towards the biosynthesis of phenylalanine and tyrosine away from that of tryptophan. Chorismate mutase is thus important for regulating the balance of aromatic

amino acids in the cell. Since the shikimate pathway and hence chorismate mutase only exists in bacteria, fungi and higher plants, the enzyme may be extremely useful in developing not only herbicides but also anti- bacterial and anti-fungal therapeutics [Kishore and Shah, 1988].

Chorismate mutase [CM; EC 5.4.99.5] catalyzes the first committed step in the biosynthesis of the aromatic amino acids phenylalanine and tyrosine in bacteria, fungi and higher plants [Haslam, 1993]. CM catalyzes the conversion of chorismate to prephenate, which is then converted to the precursors of tyrosine or phenylalanine by the subsequent action of prephenate dehydrogenase or prephenate dehydratase respectively.

The conversion of chorismate to prephenate involves breaking an ether bond between the ring moiety and a carboxylic side-chain of the substrate, and reconnecting the two components at another spot by alkylation [Hur and Bruice, 2002]. It is generally agreed that the conversion proceeds *via* a pericyclic [Claisen type] rearrangement, involving a chair-like transition state as shown below.



This reaction is a rare example of an enzyme catalyzed pericyclic reaction and is one of the reasons behind extensive studies that have been carried to date on the enzyme [Guilford *et al.*, 1987; Kast *et al.*, 2000; Hur and Bruice, 2003; Guimaraes *et al.*, 2003; Kienhofer *et al.*, 2003; Heather *et al.*, 1996]. Despite extensive studies, however, the enzyme mechanism and in particular the 10^6 to 10^7 fold rate acceleration by chorismate

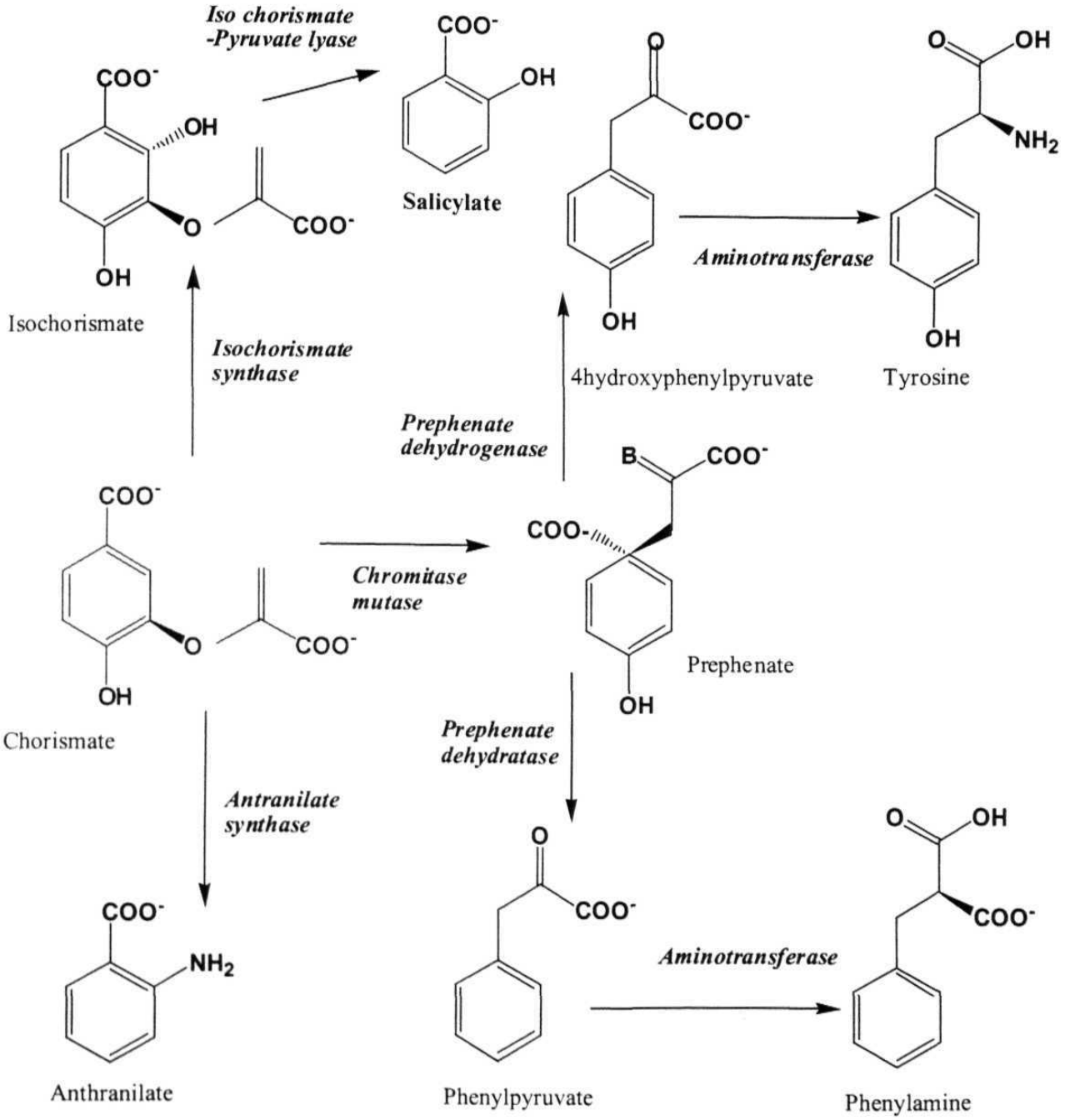
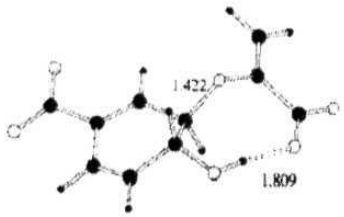


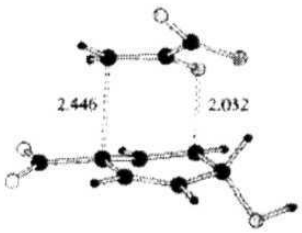
Figure 3.1: Chorismate is a substrate for multiple enzymes

mutase is poorly understood. The proposed mechanism by which CM catalyzes the conversion of chorismate to prephenate is by preferential binding of the enzyme to the less stable pseudo-diaxial form of the enolpyruvate side chain of chorismate [Sogo *et al.*, 1984]. It thus positions chorismate for a rearrangement in the chair-like transition state [Copley *et al.*, 1985; Guo *et al.*, 2001].The pseudo-diaxial and pseudo-diequitorial conformers of chorismate as proposed by Wiest and Houk [1994] are shown below.

pseudo- diequitorial



pseudo-diaxial



Transition state

Chorismate to prephenate reaction is also known to occur spontaneously. Knowles *et al.* have worked out the detailed mechanism of catalyzed and uncatalyzed conversion of chorismate to prephenate [Addadi *et al.*, 1983]. In the uncatalyzed reaction, the pseudodiequitorial form of chorismate is more populated as its stability is increased by the formation of an internal hydrogen bond between the hydroxyl group and the side-chain carboxylate. However, the pseudodiaxial conformer is the more reactive conformer that is capable of progressing towards the transition state [Carlson *et al.*, 1996]. As stated in the previous paragraph, chorismate mutases catalyzes the conversion of chorismate to prephenate by stabilizing the less stable but more reactive pseudo-diaxial form of chorismate.

Apart from a unique reaction mechanism, the CM enzyme has also gathered attention due to the presence of completely different protein fold and ligand binding pockets in different organisms, which suggests separate evolutionary origins of the enzyme. A brief description of the same is provided in the proceeding sections.

3.2.1.1 Classification of chorismate mutase

Chorismate mutase enzyme can be classified into three groups; AroQ class, prokaryotic type; AroQ class, eukaryotic type; and AroH class. These three classes fall into two structural folds [AroQ and AroH class], which are completely unrelated. Additionally, there are subsets of the AroQ class that are periplasmic or the ones that have large regulatory domains. The former are often referred to as periplasmic CMs and the latter are referred to as AroR CMs.

The two types of the AroQ structural class [the *Escherichia coli* CM dimer and the yeast CM monomer] can be structurally superimposed, and the topology of the four-helix bundle forming the active site is conserved. However, the AroH class of CMs have an entirely different structure. While the AroQ proteins are generally all-helix-bundle proteins [Lee *et al.*, 1995], while the AroH proteins possess a trimeric pseudo α/β barrel structure [Chook *et al.*, 1993]. A subset of the AroQ CMs, commonly referred to as AroR are structurally similar to AroQ with a four-helix-bundle structure around the active site. AroR CMs are seen only in eukaryotes and can be viewed as a version of AroQ that has inserted loops required for allosteric regulation [Strater *et al.*, 1997; Goers *et al.*, 1984; MacBeath *et al.*, 1998]. The periplasmic chorismate mutases have an N-terminal signal sequence, that defines their periplasmic location [Calhoun *et al.*, 2001; Xia *et al.*, 1993]. The extent of dissimilarity, yet a common reaction mechanism of the two classes of chorismate mutase enzymes suggests an independent evolutionary origin of analogous enzymes [Galperin *et al.*, 1998].

Chorismate mutases encoded by both *aroQ* and *aroH* are widely distributed among prokaryotes. AroH is the more rare group of CMs and most familiar examples of the same include those of *Bacillus* species [*B. subtilis, B. halodurans etc*], cyanobacteria [*Synechocystis* sp., *Synechococcus* sp], *Streptomyces coelicolor,* and *Thermus thermophilus* etc. In contrast, AroQ class of CMs are more widely distributed [*Escherichia coli, Saccharomyces cerevisisae*]. Even multiple paralogs of *aroQ* are commonly present in a single organism. In *Bacillus subtilis,* genes encoding both AroQ and AroH co-exist [Jensen and Nester, 1965].

Chorismate mutase frequently exists as a fusion protein and the most common associations are seen with other enzymes of the aromatic amino acid biosynthesis pathway like prephenate dehydrogenase, prephenate dehydratase, DAHP synthase etc. *E. coli* has two CMs, one in association with prephenate dehydrogenase [P protein] and the other with prephenate dehydratase [T protein]. In other organisms like *Bacillus* sp, *Methanococcus*, Yeast etc, CM exists in two copies, one as a monofunctional protein and the other as a fusion protein with other aromatic amino acid biosynthesis enzymes [Llewellyn *et al.*, 1980].

The crystal structures of CMs from *Escherichia coli* [EcCM] [Lee *et al.*, 1995], *Bacillus subtilis* [BsCM] [Chook *et al.*, 1994], and *Saccharomyces cerevisiae* [Xue *et al.*, 1994] have been solved and their significant features are described in the following sections.

3.2.1.1.1 Chorismate Mutase of *Escherichia coli* [AroQ; prokaryotic]

The Chorismate mutase of *E. coli* consists of residues I-I09 of the P-protein. *E. coli* CM is the the prototypical AroQ chorismate mutase and consists of homodimers of intimately entwined α-helices. The individual subunit has three helical segments [residues 6-42, 49-65, 70-I00]. Coiled-coil and helix-helix interactions between the two longest segments that create a catalytically functional, elongated homo-dimer with two equivalent, elbow-shaped active sites that are highly charged and completely enclosed [Lee *et al.*, 1995] [Figure 3.2].

AroQ CMs have composite active sites. Each of the two active sites of *Ec*CM dimer is formed by 3 helices from one monomer and a helix from the other subunit. It has been proposed that such domain-swapped dimers might have evolved from active, monomeric precursors [Bennet *et al.*, 1995]. By inserting a 180^0 turn into the middle of helix 1 of the thermostable *Methanococcus jannaschii*.CM, active site was generated with residues from a single polypeptide chain [MacBeath *et al.*, 1998].

Aromatic amino acid biosynthesis is generally very well regulated because of the high energy costs to synthesize these amino acids. In fact, amongst all amino acids, aromatic

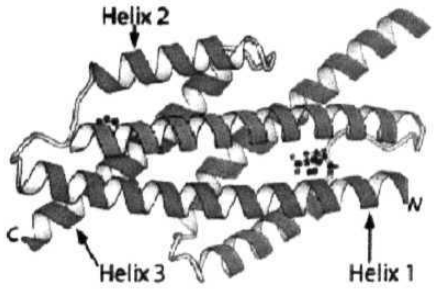


Figure 3.2: Structure of *Escherichia coli* chorismate mutase shown in a ribbon diagram representation. Each subunit has three helices. A transition state analog inhibitor is bound in each active site and is represented in ball-and-stick form [Lee *et al*, 1995].

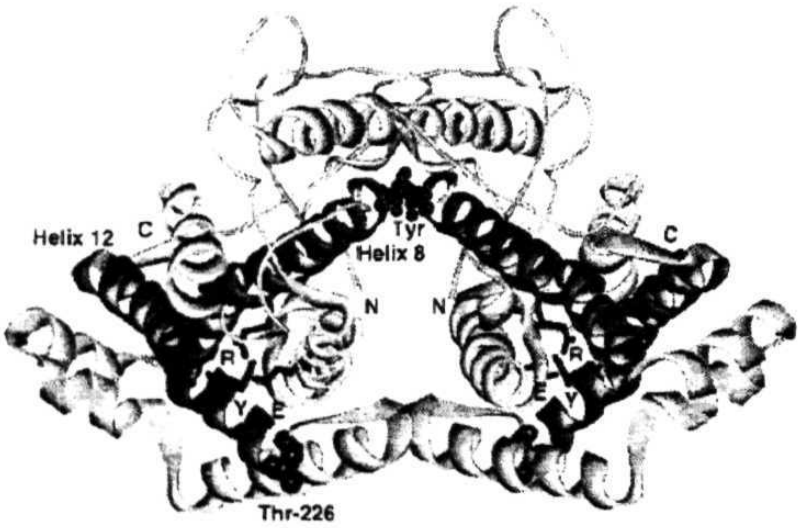


Figure 3.3: Ribbon diagram of the structure of the T state of yeast chorismate mutase. The N and C termini are labeled as N and respectively. Tyr represents the two tyrosine molecules located at the interface of dimer. Helix 8, which connects the allosteric active sites, and helix 12 are shown in dark color. The side chains Glu-23 [E] and Tyr-234 [Y], which are involved in the hydrogen bonding interaction with the active residue Arg-157 [R], are drawn stick model. Thr-226, which connects the 220s loop and helix 12 is indicated. [Strater *et al* 1997]

amino acids are the most expensive to synthesize [L-Phe: 65, L-Tyr: 62, and L-Trp: 78 ATP metabolic equivalents, respectively] [Walker-Simmons and Atkinson, 1977]. In *Escherichia coli*, the biosynthesis of phenylalanine and tyrosine is subjected to feedback regulation. The catalytic and regulatory domains of *E. coli* P-protein have been studied by Zhang *et al.* [1998] using genetically engineered proteins. The authors have shown that the mutase, dehydrogenase and regulatory activities of P-protein reside within specific domains. Residues 286-386 were found to be crucial for phenylalanine binding.

3.2.1.1.2 Chorismate mutase of *Saccharomyces cerevisiae* [Aro Q; eukaryotic]

The yeast CM is a homodimer consisting of two 30-kDa polypeptide chains coded by the ARO7 gene and shows homotropic activation by the substrate chorismate. The crystal structures of wild-type yeast CM co-crystallized with tryptophan and an *endo*-oxabicyclic transition state analogue inhibitor [Bartlett and Johnson, 1985] has been determined [Strater *et al.* 1997]. YCM has essentially an all-helix structure, which has similarities to the CM domain of the P protein from *E. coli.* Both proteins contain a four-helix bundle that surrounds the active site cavity. Using 94 residues of 22% sequence identity, these helix bundles can be superimposed with a root-mean square deviation [rmsd] of 1.06 Å [Xue and Lipscomb, 1995] [Figure 3.3].

Yeast CM shows approximately 10-fold allosteric activation by tryptophan and 10-fold inhibition by tyrosine [Schmidheini *et al.*, 1990]. This antagonism plays an important role in maintaining the correct balance of aromatic amino acids in the cell. The structural basis for these regulatory processes has also been described [Strater *et al.*, 1996, 1997].

Phenylalanine is not an inhibitor of yeast chorismate mutase. This is in contrast to the homologous chorismate mutase from *Arabidopsis thaliana*, which is inhibited by both tyrosine and phenylalanine [Eberhard *et al.*, 1993].

The kinetic properties of the enzyme can be viewed in terms of a transition from a conformation having a low affinity for chorismate, the T [tense] state, to a conformation with high affinity for the substrate, the R [relaxed] state [Monod *et al.*, 1965]. The influence of the allosteric regulators is best explained by local conformational changes induced by their competitive binding to the regulatory sites, thereby stabilizing either of the two states or inducing a conformational change in the overall enzyme structure.

3.2.1.1.3 Chorismate mutase of *Bacillus subtilis* [AroH class]

BsCM is a member of the AroH class of CMs. This class adopts a trimeric, pseudo-α/β barrel fold [Chook and Lipscomb, 1993]. Chook *et al.* [1993] were the first to solve the crystal structure of the monofunctional chorismate mutase [EC 5.4.99.5] of *Bacillus subtilis* at 2.2^{0}A resolution. The enzyme was found to be a homotrimer, with beta-sheets from each monomer packing to form the core of a pseudo-alpha beta-barrel with helices on the outside of the trimer [Figure 3.4]. In addition, the active sites were located by using data from the enzyme structure complex with an endo-oxabicyclic inhibitor that mimics the transition state of the reaction. The structure of this complex was refined to 2.2^{0}A with a current R value of 0.182 for a model that included 1388 residues, 12 inhibitor molecules, and 530 water molecules in the asymmetric unit. Three active sites were seen in each trimer at the interface of two adjacent subunits.

The structure was refined further when it was solved at 1.3^{0}A resolution. In this structure, the C-terminal tails of the subunits of the trimer were refined further and were shown to be hydrogen bonded to residues of the active site of neighboring trimers in the crystal. This bonding could therefore cross-link the molecules in the crystal lattice.

AroH and AroQ CMs share some common active site features. For example, in the crystal structures with an oxabicyclic transition state analog [TSA], both enzymes show multiple cationic groups [Arg and Lys] interacting with the carboxylates and the ether oxygen [Lee *et al.*, 1995; Chook *et al.*, 1993]. Additionally, both possess a Glu residue that hydrogen bonds to the hydroxyl group of the TSA. Despite their different folds, both enzymes are likely to utilize similar catalytic mechanisms.

Figure 3.4: Structure of *Bacillus subtilis* chorismate mutase. The monofunctional chorismate mutase from *B. subtilis* is a homotrimer and adopts a pseudo-α/β barrel fold. A transition state analog, shown in a ball-and-stick representation, is bound in each of the active sites, which are located at the trimer interfaces [Chook and Lipscomb, 1993].

It was earlier believed that there were apparently no catalytic residues for CM and catalysis was due to the geometry and polarity of the active site of CM which stabilizes the transition state [Chook *et al.* 1993]. Later however, a different mechanism was proposed for EcCM as well as BsCM, that involves a catalytic group [Xue and Lipscomb 1995 Strater *et al.* 1996].

Chorismate mutase was not annotated in the initial genome sequence *of M. tuberculosis*. However, the recent COG [Cluster of Orthologous groups, NCBI] annotation of the *M. tuberculosis* genome suggests that CM activity can be attributed to two ORFs, Rv1885c and Rv0948c [http://www.ncbi.nlm.nih.gov/COG/]. The ORF Rv1885c has been annotated in the TubercuList web-server [http://genolist.pasteur.fr/TubercuList/], Institute Pasteur, as a conserved hypothetical protein with some similarity to the monofunctional chorismate mutases of *Erwinia herbicola* [28.6% identity in a 133aa overlap]. Rv0948c has also been annotated as a conserved hypothetical protein, equivalent to a conserved hypothetical protein [105 aa] from *Mycobacterium leprae* [NP_301237.1|NC_002677] which is also similar to N-terminus of some chorismate mutase/prephenate dehydratase enzymes.

This chapter addresses the question whether these two hypothetical proteins indeed show *in vitro* chorismate mutase activity and if they do, what is their proximity or distance with respect to the known two classes of CMs [AroQ and AroH]. The approach involved expressing the *M. tuberculosis* ORFs, Rv1885c and Rv0948c in *E. coli* and determining the biochemical and biophysical parameters of the encoded proteins. While more extensive studies were carried out with the protein coded by ORF Rv1885c, it was also demonstrated that Rv0948c also possesses CM activity, though with a reduced turnover. Kinetic and regulatory studies of rRv1885c indicate several unique properties of the enzyme which include feedback regulation by pathway specific as well as cross pathway specific ligands in the same manner. A genetic approach was used for functional characterization of the predicted N-terminal signal sequence of *M. tuberculosis* CM. These studies provide sufficient evidence to conclusively place the protein coded by ORF Rv1885c of *M. tuberculosis* in the aroQ class of periplasmic CMs.

3.3 EXPERIMENTAL PROCEDURES

3.3.1 Bacterial strains and plasmids

All *E. coli* and *M. tuberculosis* strains used in this part of the work are briefly described in Table 3.1. The plasmid vectors [with their sources] and the recombinant plasmids constructed are also described in the same table. The details of primers used in this part of the work are given in Table 3.2

3.3.2 Media, chemicals, buffers, and enzymes

Luria Broth (LB) supplemented with 200 μg/ml of ampicillin was used for selection and amplification of plasmids in *E. coli* hosts. LB with 4% glycerol was used for growth of bacterial cells for protein purification purposes.
XP medium was used for direct AP [Alkaline phosphatase] assay on bacterial plates. The composition of the medium is as follows.

10X XP medium [1Litre]

15g of Agar-agar was dissolved in 900ml of water and autoclaved. To this, 100ml of 10X XP medium [sterile] was added. The XP medium consisted of

1M CaCl2	2ml
10mM FeCl3	2ml
10mM ZnCl2	2ml
20% Glucose	10ml
Vitamin B1	500μl
20mg/ml XP [fresh]	2ml

3.3.3 Cloning, overexpression and purification of recombinant proteins in *E. coli*

E. coli BL21DE3 cells transformed with the 6xHis tagged chimeric construct [Rv1885c / Rv0948c cloned in the *NdeI*/*XhoI* sites and *NdeI*/*HindIII* sites respectively of pET23a expression vector [Novagen]] were grown in 1L of LB medium supplemented with

Table 3.1: **Bacterial Strains and Plasmids**

Strains/plasmids	Description	Source / Reference
Mycobacterium tuberculosis		
H37Rv	Laboratory Strain [Virulent]	Cole ST *et al.*, 1998
E. coli		
DH5α		
BL21 DE3		
MG1655	Wild Type	
MG1655*phoA+proC-*	*E. coli* proline auxotroph	This work
MG1655*phoA-proC+*	*E. coli* Alkaline phosphatase negative prototrophic strain	This work
Plasmids		
pET23a	Expression vector	Novagen
pET23aCM1	pET23a derivative with *M. tuberculosis* CM1 [Rv1885c] cloned in *NdeI/XhoI* sites	This work
pET23aCM2	pET23a derivative with *M. tuberculosis* CM2 [Rv0948c] cloned in *NdeI/XhoI* sites	This work
pGEMTEasy	Cloning Vector	Promega
pGEMT AP	pGEMT Easy derivative with ss less *E.coli phoA* cloned in vector MCS	This work
pETCMssAP	pET23aCM1 derivative with ss less *E.coli phoA* cloned in *SalI/XhoI* sites downstream of *M. tuberculosis* CM ss	This work
pBAD18 MCS2	Expression vector with the arabinose inducible promoter, P_{BAD}	Guzman *et al.*
pBADCMssAP	pBAD MCS2 derivative with ss less *E. coli* phoA cloned in *XbaI/HindIII* sites	This work

ss: signal sequence

Table 3.2: Sequences of primers used to amplify *M. tuberculosis* chorismate mutase (Rv1885c and Rv0948c) and *E. coli* Alkaline phosphatase (AP)

ORF/ Gene	FORWARD PRIMER	REVERSE PRIMER
	NdeI	*XhoI*
Rv1885c	AT***CATATG***TTGCTTACCCGTCCACGTGA	AT***CTCGAG***GGCCGGCGGTAGG
	NdeI	*HindIII*
Rv0948c	AAT***CATATG***AGACCAGAACCCCCACATCACGA	ATA***AAGCTT***GTGACCGAGGCGGCCCCT
	SalI	*XhoI* *HindIII*
E coli AP	AT***GTCGAC***ACACCAGAAATGCCTGTTCTG	AT***CTCGAGAAGCTT***TTATTTCAGCCCCAGAGC

100μg/ml of ampicillin and 10% glycerol. IPTG was added to the mid log phase culture. The cells were kept in an incubator shaker for another five hours at 27^{0}C/150rpm to allow protein expression. A low temperature was used to allow the protein to enter the soluble phase. After induction, purification of the protein was carried out in a manner similar to that described for IdeR in Chapter 2 [Section 2.3.2]

3.3.4 Enzyme assays and kinetic studies

Chorismate mutase activity assays were carried out as described elsewhere with a few modifications [Davidson and Hudson, 1987]. This assay is based on enzymatic conversion of chorismate to prephenate, which is terminated by the addition of HCl. Prephenate can be further converted to phenylpyruvate under alkaline conditions which can be measured spectrophotometrically [Krappmann *et al.*, 1999]. Reaction volumes of 100 μl contained 10mM Tris HCl, pH7.5, chorismic acid 0.5 to 4mM, optionally tyrosine, phenylalanine or tryptophan [100 μM to 500 μM]. The reaction was started by the addition of 100 to 500 pico moles of the purified protein to the prewarmed chorismate solution [at 30^{0}C] and was stopped by the addition of 10μl of 1.6NHCl. After a second incubation of the mix at 37^{0}C for about 15 minutes, 890μl of 1.58N NaOH was added and the absorbance was recorded at 320nm. To exclude the absorbance caused by the uncatalyzed arrangement of chorismate, blanks of increasing chorismate concentrations without enzyme were prepared and their absorbance was also recorded. These readings were then subtracted from the absorbance measured for enzyme activities at different concentrations of the substrate.

For determination of the optimum pH for CM activity, buffers of different pKa [CPB, pH4, CPB pH4.5, MES pH6, HEPES pH7, Tris HCl pH7.5, Tris HCl pH8] were used CM activity was also assessed at different temperatures [15^{0}C to 80^{0}C]. One unit of enzyme was defined as the amount of enzyme required to form 1 μmol of product/min at 37 °C. Prephenate dehydratase [PDT] and prephenate dehydrogenase [PDH] assays were carried out as described elsewhere [Gething *et al.*, 1976; Davidson and Hudson, 1987].

The PDT assay is described in detail in Chapter 5, section 5.3.3.

Phe and Tyr Feedback inhibition assays

Allosteric regulation of CM activity by L-Phe and L-Tyr was measured at 100-1000μM concentrations of the effectors.

3.3.5 Isolation of the periplasmic fraction of *E. coli* BL21 cells

The periplasmic fraction of *E. coli* was isolated following an osmotic shock procedure [Qiagen, USA]. About 1l of induced culture was harvested by centrifugation at 4000g for 20 minutes and the pellet was resuspended in 30mM Tris HCl, pH 8.0 at 80ml per gram wet weight. The cells were kept on ice and to it 500mM EDTA was added dropwise to a final concentration of 1mM. This was followed by another incubation on ice for 5-10 minutes with gentle agitation. The cell suspension was then centrifuged at 8000g for 20 minutes at 4^0C. The supernatant contained the osmotic shock fluid consisting of periplasmic proteins. The periplasmic and the cytoplasmic fractions were quantified and used in western blots for a comparative estimation of protein expression.

3.3.6 Western Blot

Western blot with the cytosolic and periplasmic fractions of *E. coli* BL21 cells expressing *M. tuberculosis* CM [Rv1885c] was carried out with monoclonal anti Histidine antibody using the manufacturer's instructions [Santa Cruz Biotech, USA].

3.3.7 Limited Proteolysis

To study the effects of various ligands on enzyme stability, integrity and accessibility of active sites, limited proteolysis of *M. tuberculosis* CM was carried out. Trypsin was taken as the protease of choice. 10μg recombinant protein [1μg/μl concentration] was taken to which different concentrations [100μM to 5mM] of the ligand [Tyrosine, Phenylalanine, tryptophan, Salicylate] was added. The reaction buffer contained 50 mM Tris pH 7.5 and 100mM NaCl. Trypsin protein molar ratio was kept at 1:1000 and the reaction was

incubated at 25^{0}C for 30 minutes, following which trypsin was inactivated by the addition of 0.5mM PMSF. The reaction was mixed with equal volume of 2X Tris-Tricine loading dye and loaded on a 10% Tris Tricine gel. The gel was silver stained and scanned.

3.3.8 Silver staining of protein gels:

3.3.8.1 Gel Preparation

After the completion of electrophoresis, the gel was soaked in SOLUTION 1 [50ml MeOH + 10ml Acetic Acid + 40ml MQ-H_2O]. Following this, the gel was transferred to SOLUTION 2 [5ml MeOH + 7ml Acetic Acid + 88ml MQ-H_2O] for 30 minutes. Solutions 1 and 2 were prepared in advance and stored at room temperature. The gel was further transferred to SOLUTION 3 [20ml 25% glutaraldyhyde [Sigma G6257] + 30ml MQ-H2O] for 30 min. Solution 3 was made fresh. The gel was washed with MQ- H_2O either 4 times for 30 minutes [> 100 ml/gel/wash] or 1 time overnight [> 500 ml/gel/wash]. The gel was then soaked in DTT solution [0.5 ml 100x DTT + 50ml MQ-H_2O] [100X DTT = 0.5 g/ml, at -20 C]

3.3.8.2 Gel Staining

The gel was kept for 30 min in SILVER SOLUTION [0.05g Silver Nitrate + 50ml MQ-H_2O]. STOP solution [2.42 g Citric Acid monohydrate + 5 ml MQ-H2O] was prepared in advance. DEVELOPER solution [30 ml 37% Formaldehyde + 3g Sodium Carbonate (anhydrous) + 100 ml MQ-H_2O] was also prepared before hand.

The gel was dumped in Silver Solution and rinsed briefly with MQ-H_2O. Next, the gel was rinsed with ~20ml Developer solution for 5 min then the remaining developer solution was added. Development of bands was closely monitored. Just before bands were dark enough, the entire 5ml of STOP solution was added. When bubbling stopped [~5min], the gel was rinsed with MQ-H_2O. Gel was stored in MQ-H_2O until it was scanned.

3.3.9 CD Spectrometry

The CD spectrum of the recombinant protein [Rv1885c] was recorded in a wavelength range of 190-250nm in steps of 1nm with four accumulations at each step. The spectral baseline was corrected by subtracting the respective blanks. Ellipticity, represented in milli degrees was plotted as a function of wavelength [nm]. Percentage helicity for secondary structure determination was calculated using the online available K2D software [http://www.embl-heidelberg.de/~andrade/k2d/]

3.3.10 Analytical size exclusion chromatography

The oligomeric state of the recombinant protein corresponding to ORF Rv1885 of *M. tuberculosis* was determined using analytical size exclusion chromatography on a Superdex 200 [HP 10/30] FPLC column from Pharmacia. Chromatography was performed at room temperature with 10mM Tris and 100mM NaCl as the running buffer, the same buffer in which the recombinant protein was eluted. A standard curve was prepared according to the instruction manual using the LMW Gel Filtration calibration kit from Pharmacia Biotech. The void volume was determined using Blue dextran 2000. The elution parameter Kav was calculated as Kav= Ve-Vo/ Vt-Vo, where Ve= elution volume for the protein, Vo= column void volume, Vt= total bed volume. Kav was plotted against log MW. The protein sample was loaded on the gel filtration column at a concentration of 4mg/ml in the presence of DTT. The elution volume of the protein was used to calculate its molecular weight using the standard curve.

3.3.11 Construction of an *E. coli pho*A negative strain and N-terminal signal sequence characterization of *M. tuberculosis* chorismate mutase

Wild type *E. coli* MG1655 strain was transduced with P1 phage carrying a *pro*C mutation linked to a kanamycin resistance marker [15% linkage]. The resultant transductants were selected for proline auxotrophy. A second round of transduction was carried out in this strain using a Δ*pho*A P1 phage where *phoA*- character was linked to *pro*C+ character [90% linkage]. 90% of these transductants were able to grow on minimal medium but

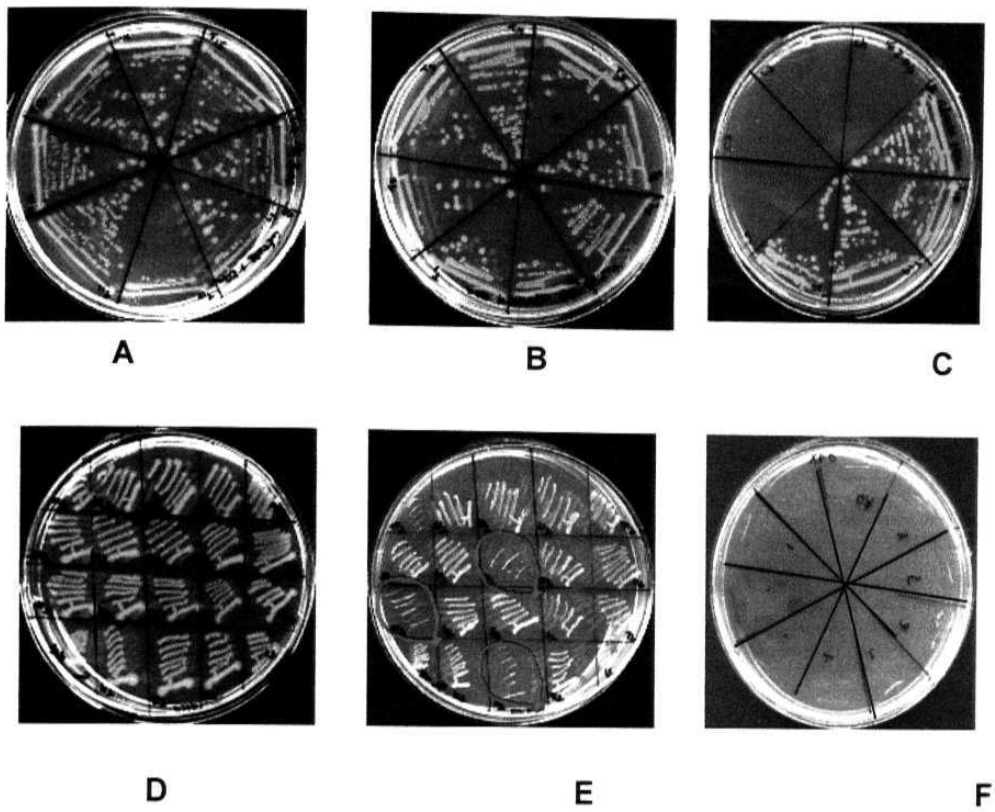


Figure 3.5: Construction of an *E. coli* strain lacking intrinsic alkaline phosphatase activity (*E. coli pho*A- strain)

E. coli strain MG1655 (wild type) was transduced with P1 phage carrying a mutation in *pro*C, where the *pro*C- character (for proline auxotrophy) was linked to a kanamycin resistance marker. The transductants were selected on LB kanamycin plates.

A,B,C: Purified kanamycin resistant colonies obtained after first transduction. Few colonies which grew in the control plates (without P1 transduction) did not grow upon restreaking (empty sectors in plate C)

D, E: The *pro*C- character of transductants was tested on LB (D) and Minimal medium (E). Colonies unable to grow on minimal medium (encircled red) were designated as MG1655 *pho*A+*pro*C- and selected for second round of transduction with P1 phage carrying the *pro*C+ *pho*A- segment.

F: After the second P1 transduction, the transductants were selected on Minimal XP plates. The sector with blue colonies is a positive control (*pho*A+) *E. coli* strain. 90% of the transductants were *pho*A negative (white on Minimal XP plates). This strain was referred to as a *pho*A negative *E. coli* strain (MG1655*pho*A-*pro*C+).

failed to develop blue colour on XP medium. This strain was designated as an *E. coli phoA* negative strain [Figure 3.5] and was used for functional characterization of N-terminal signal sequence of *M. tuberculosis* chorismate mutase. *E. coli pho*A [alkaline phosphatase] lacking the signal sequence was amplified from wild type *E. coli* genomic DNA and cloned in the MCS of pGEMTEasy vector followed by subcloning in *Sal*I/*Xho*I sites downstream to the 33amino acid signal sequence of the pET23aCM1 construct. This construct was named pETCMssAP. The *Xba*I/*Hind*III cassette from pETCMssAP alongwith the optimized RBS was transferred to pBAD18MCS2 vector [Guzman *et al.*, 1995] [Figure 3.6]. Expression from pBAD is driven by the arabinose inducible promoter P_{BAD}. Subsequently, the *pho*A negative *E. coli* strain was transformed with the pBADCMssAP chimeric construct and plated on LB-XP plates. As *E. coli* AP is known to be active only in the periplasmic space [Brickman and Beckwith, 1975], blue colour was expected only if *M. tuberculosis* CM signal sequence could export *E. coli* AP to the periplasmic space.

3.3.11.1 Preparation of P1 phage lysates

The donor strain was grown overnight in 0.3ml Z broth and mixed with 10^7 pfu [plaque forming unit] of P1 phage. Adsorption was allowed to occur at 37^0C for 20 minutes followed by the addition of 10ml of Z broth and incubation at 37^0C in a 100ml flask with slow shaking until growth. Lysis of the culture followed [after approximately 4 hours]. The lysate was treated with chloroform, centrifuged and the clear supernatant was stored at 4°C.

Z broth:

LB medium	100 ml
$CaCl_2$ [0.5 M]	0.5 ml

3.3.11.2 P1 Transduction

The recipient cells were grown to $1x10^9$ cfu/ml in rich broth. $CaCl_2$ was added to a concentration of 1mM. 0.1 to 0.2 ml cells were mixed with P1 lysate to give a MOI (multiplicity of infection) of 0.01 to 0.1 and allowed adsorption at 37°C for 15 min.. About 1/4 to 1/2 of the mixture were plated on selective plates containing 0.01M sodium citrate and appropriate antibiotic and incubated overnight at 37°C.

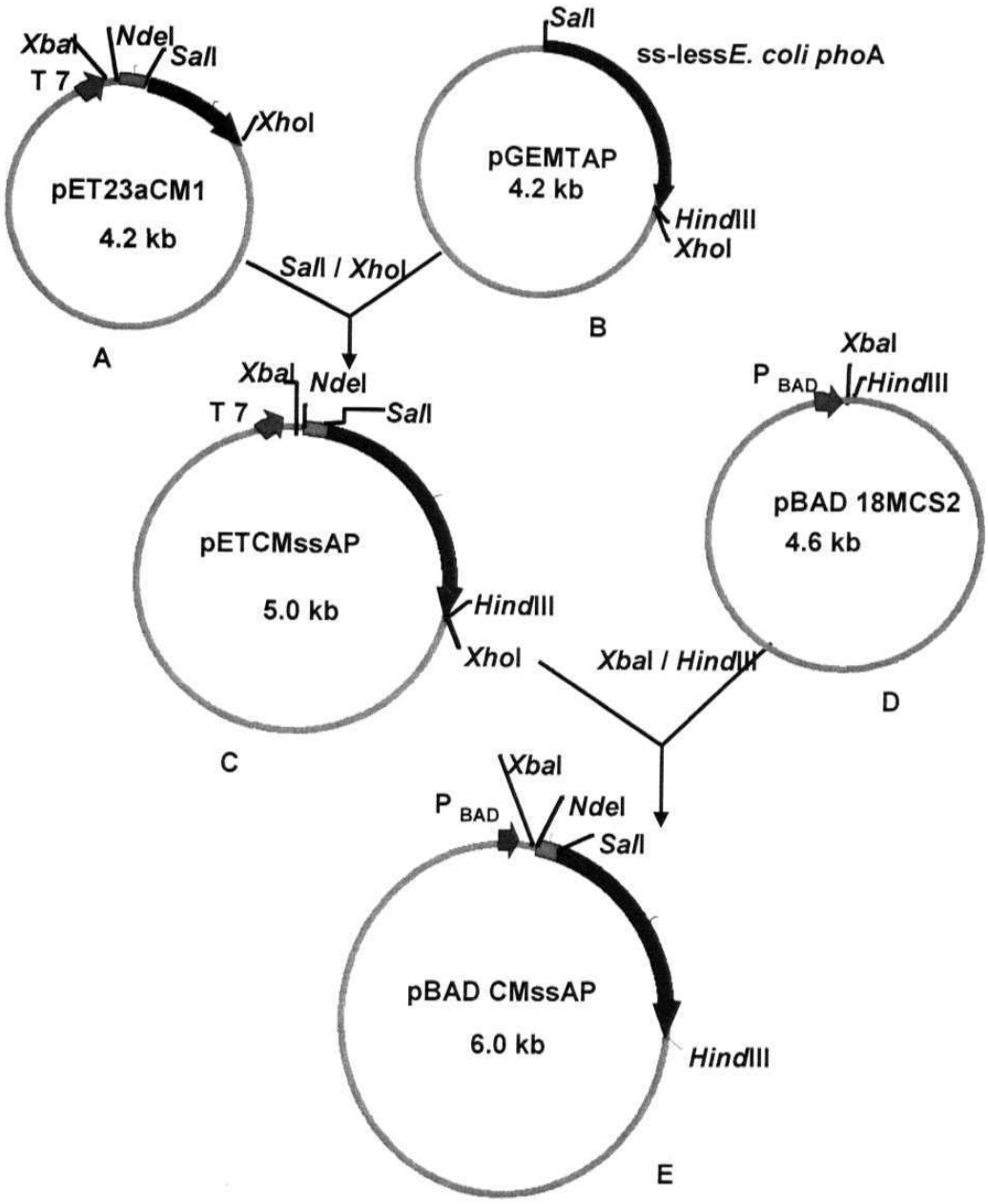


Figure 3.6: Construction of a recombinant plasmid with the signal sequence (ss) of *M. tuberculosis* chorismate mutase (CM) cloned upstream of ss less *E. coli pho*A. A: ORF Rv1885c of *M. tuberculosis* cloned in the *Nde*I/*Xho*I sites of pET23a vector. The region between *Nde*I and *Sal*I is the putative signal sequece of *M. tuberculosis* CM B: *E. coli pho*A (gene coding for alkaline phosphatase) lacking the ss cloned in the MCS of pGEMTEasy vector. C: A recombinant plasmid generated by transferring the *Sal*I/*Xho*I cassette from pGEMTAP to the *Sal*I/*Xho*I site of pET23aCM1 E: The *Xba*I/*Hind*III fragment from pETCMssAP was transferred to pBAD 18 (D) to generate pBADCMssAP (E). This was the final construct used for functional characterization of the of ss of *M. tuberculosis* CM.

3.4 RESULTS

3.4.1 The hypothetical ORF, Rv1885c of *M. tuberculosis* encodes a functional chorismate mutase enzyme

The gene corresponding to the ORF Rv1885c of *M. tuberculosis* was PCR amplified, cloned and expressed in *E. coli* BL21 cells and recombinant protein was purified to near homogeneity [Figure 3.7]. The kinetic properties of the purified protein were studied using CM activity assays. The substrate saturation curves were hyperbolic for the enzyme i.e. the enzyme followed Michealis Menton's kinetics [Figure 3.8]. There was no indication of positive homotropic co-operativity of *M. tuberculosis* CM towards chorismate. Kinetic parameters were subsequently obtained by fitting the initial rate data to the Michealis Menton's equation. Km for the enzyme was calculated as 1.2mM and Vmax as 71.42μmoles min-1 mg-1. The molar catalytic activity [Kcat] was 26.33 sec^{-1}. The effects of temperature and pH on enzyme kinetics were also studied. Since spontaneous arrangement of chorismate to prephenate is strongly dependent on temperature, blanks of the same reaction without the enzyme were also kept at different temperatures and the readings were subtracted from those obtained in presence of the enzyme. The optimum temperature for optimum activity of *M. tuberculosis* CM ranged from 37 to 50^{0}C [Figure 3.9 A]. While high temperature did allow a spontaneous conversion of chorismate to prephenate, there was a distinct enhancement of the reaction rate upon addition of the enzyme. pH also has a significant role in determining the activity of CM. *M. tuberculosis* CM was found to be active between pH 6 to 9 with a pH optimum of 7 [Figure 3.9 B]. The pH optimum is similar to *E. coli* CM, where again maximum activity is at pH 7.3. CM activity of rRv1885 was greatly inhibited at acidic pH. While fungal CMs are reported to be more active at acidic pH, bacterial CMs [*E. coli* P protein, *Salmonella typhimurium* CM] are reported to be active in alkaline pH [Schnappauf *et al.*,1997; Helmstaedt *et al.*, 2001]. *Bacillus* CM has maximum activity in the range of pH 5 to 9 [Gray *et al.*, 1990]. Hence the pH range at which *M. tuberculosis* CM works efficiently is similar to that of *Bacillus subtilis* with the optimum pH being closer to that of *E. coli* CM.

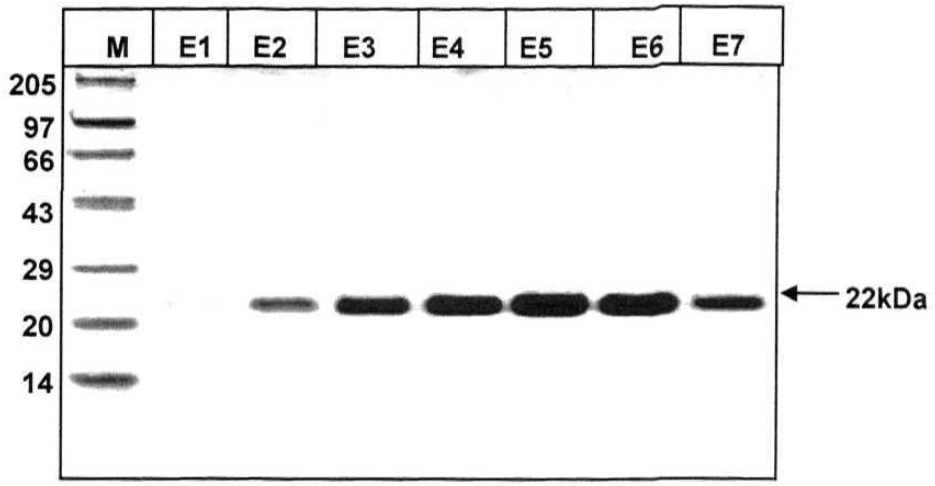


Figure 3.7: Purification of the chorismate mutase enzyme (Rv1885c) of *Mycobacterium tuberculosis* as a recombinant protein in *E. coli*. The ORF corresponding to Rv1885c was cloned in the *Nde*I and *Xho*I sites of pET23a vector with a C-terminal Histidine tag and expressed in *E. coli* BL21 cells. Affinity purification of recombinant protein was carried out using Talon resin (Clonetech). The purified protein was resolved on 10% Tris-Tricine gel. The gel was stained with Coomassie Brilliant blue dye. M represents the protein molecular size marker (Broad range, Genei), E1-E7 show the successive eluted fractions of the recombinant protein. Arrowhead indicates the position of the 22kDa protein, which is the predicted mass of *M. tuberculosis* chorismate mutase.

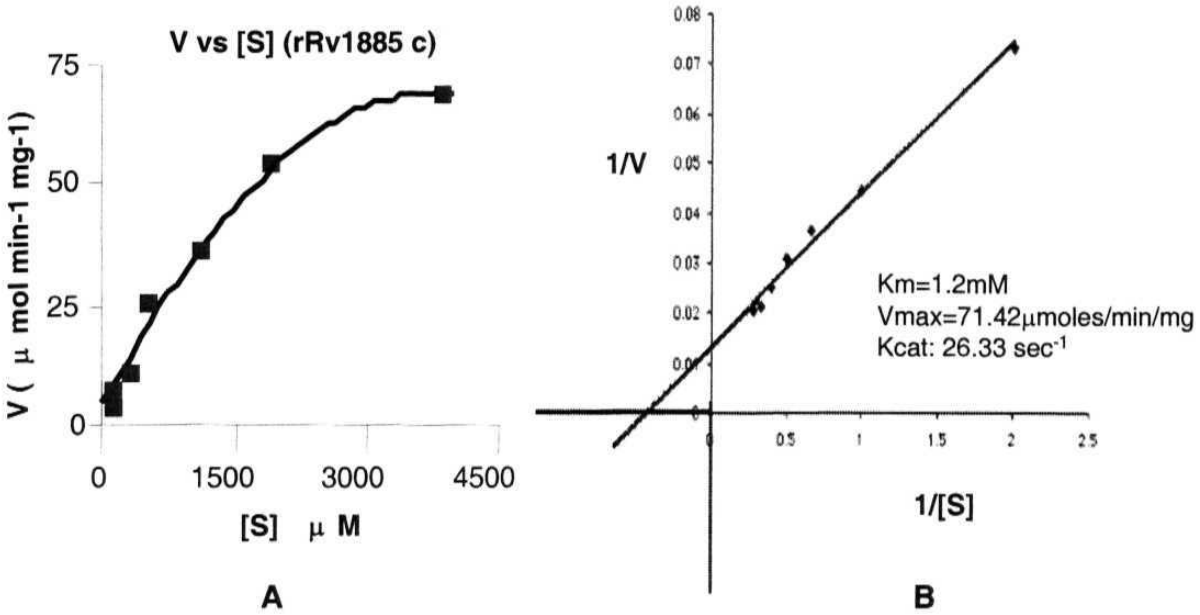


Figure 3.8 . A: Substrate saturation plot of *M. tuberculosis* chorismate mutase. Purified recombinant protein corresponding to ORF Rv1885c was assayed for chorismate mutase activity in the absence of any effector molecule. The data were fitted to functions describing Michaelis–Menton type saturation.B: Lineweaver-Burk plot of *M. tuberculosis* CM that was used to calculate Km and Vmax of the enzyme.

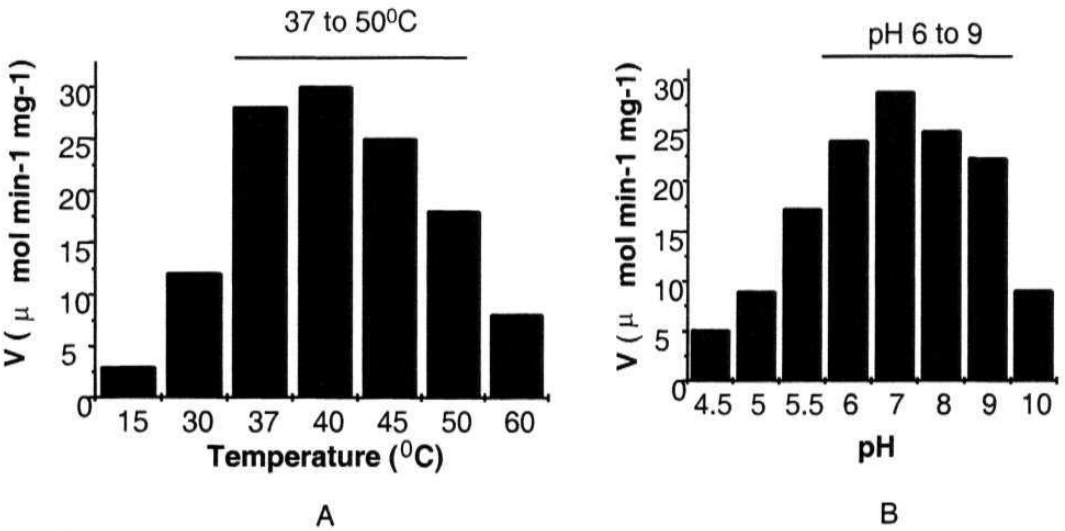


Figure 3.9 : Effect of pH and temperature on the activity of *M. tuberculosis* chorismate mutase. The temperature for optimum activity of *M. tuberculosis* CM ranged from 37 to 50°C (A). The enzyme was found to be active between pH 6 to 9 with a pH optimum of 7 (B).

3.4.2 *M. tuberculosis* chorismate mutase [Rv1885c] does not display prephenate dehydratase [PDT] or dehydrogenase [PDH] activity

Though PDT or PDH domains have not been predicted to be present in the ORF Rv1885c, on account of several examples of divergent evolution of enzymes involved in aromatic amino acid biosynthesis [Dosselaere and Vanderleyden, 2001], it was decided to determine whether *M. tuberculosis* CM displays PDT or PDH activities. PDT and PDH assays were accordingly carried out with the recombinant protein. Results conclusively demonstrate that *M. tuberculosis* CM does not possess any intrinsic PDT or PDH activity. This is unlike the CMs of many other enteric bacteria where CM and PDT/PDH are present as fusion proteins [Ahmad and Jensen, 1988].

3.4.3 *M. tuberculosis* chorismate mutase [Rv1885c] shows allosteric regulation by tyrosine, phenylalanine and tryptophan

Having shown that the hypothetical protein coded by ORF Rv1885c encodes a functionally active chorismate mutase enzyme with no associated PDT or PDH activity, it was tested whether L-Phenylalanine, L-tyrosine, L-tryptophan, could act as potential modifiers of the CM activity of rRv1885c. This experiment revealed an unusual property of the enzyme. It was observed that while these ligands moderately enhance CM activity at low concentrations, they completely inhibit the enzyme at higher concentrations [Figure 3.10]. All known forms of CM that show allosteric regulation are usually large fusion proteins [Helmstaedt *et al.*, 2001]. The monofunctional CM of *Bacillus subtilis* is entirely resistant to feedback inhibition by aromatic amino acids [Gray *et al.*, 1990]. *E. coli* CM exists as a fusion protein and shows allosteric inhibition by only pathway specific inhibitors [Helmstaedt *et al.*, 2001]. The allosteric regulation of *M. tuberculosis* CM is also unlike yeast CM where tryptophan causes activation while tyrosine and phenylalanine cause inhibition of enzyme activity [Helmstaedt *et al.*, 2002]. Fluorimetric assays were also carried out to assess if phenylalanine binding causes a conformational change in *M. tuberculosis* CM. The blue shift in the spectra and the quenching of net fluorescence indicates that tryptophan residues of the enzyme are buried inside as a consequence of phenylalanine binding [Figure 3.11].

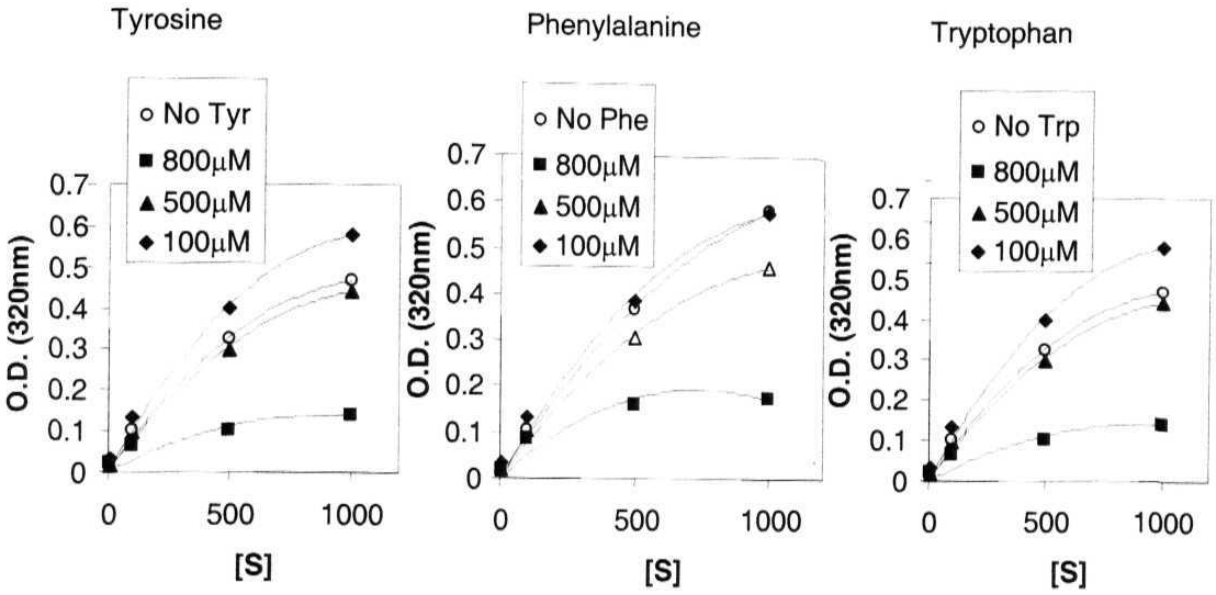


Figure 3.10: *M. tuberculosis* chorismate mutase (Rv1885c) is allosterically regulated by aromatic amino acids: The CM activty of Rv1885c was tested in the presence of various allosteric effectors (tyrosine, phenylalanine and tryptophan). Low concentration (up to 100 μM) of all the three aromatic amino acids led to a modest increase in enzyme activity. However, at higher concentrations of the effectors (800μM and above), CM activity was greatly inhibited.

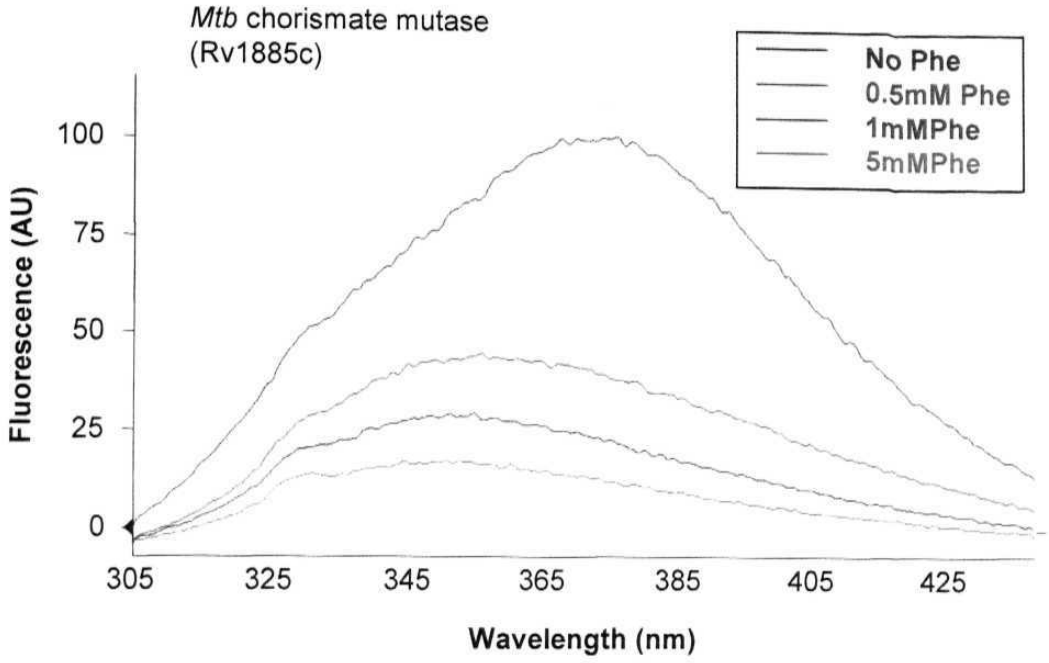


Figure 3.11:Phenylalanine binding leads to a change in the fluorescence spectrum of *M. tuberculosis* chorismate mutase (Rv1885c). Fluorescence emission spectra were recorded by exciting the protein at 280nm and recording the emission spectra in the range of 305nm-440nm. A blue shift was absorbed upon incubation of rRv1885c with phenylalanine indicating that the tryptophan residues of the protein are buried inside as a consequence of phenylalanine binding.

3.4.4 Pathway specific as well as cross pathway specific ligands protect *M. tuberculosis* CM from proteolytic cleavage

Having shown that tyrosine, phenylalanine and tryptophan are allosteric inhibitors of *M. tuberculosis* CM activity, experiments were designed to determine if binding of these ligands induce a conformational change in the enzyme that leads to inaccessibility of the active site. Limited proteolysis was employed as a structural probe to ascertain enzyme conformational change. Trypsin was taken as the protease of choice and limited proteolysis was carried out to study the amenability of the enzyme to proteolytic cleavage in the presence of various ligands. It was observed that in the presence of high concentrations [3mM and above] of effectors [Trp, Tyr and Phe], the enzyme was completely protected from proteolytic degradation [Figure 3.12 A-C]. Salicylic acid, another cross pathway secondary metabolite derived from chorismate did not confer any protection on the enzyme [Figure 3.12 D]. Higher concentrations (3mM or above) of all the three aromatic amino acids also protected the protein coded by ORF Rv0948c from trypsin cleavage [Figure 3.12, E-F].

3.4.5 *M. tuberculosis* chorismate mutase is a dimeric protein with a predominantly alpha helical structure

While catalytic and regulatory activity of *M. tuberculosis* CM point towards some novel properties of the enzyme, the study was continued to determine the biophysical parameters of the enzyme to define the actual class to which it belongs. Size exclusion chromatography was performed to determine the oligomeric state of the protein. The output was a single peak corresponding to the dimeric state of the recombinant protein [Figure 3.13]. In this context, *M. tuberculosis* CM is similar to the *E. coli* or Yeast CMs, which are also dimers of identical subunits [Lee *et al.*, 1995; Xue and Lipscomb, 1995]. To determine the secondary structure of *M. tuberculosis* CM, the CD spectrum was recorded on a spectropolarimeter [JASCO]. The data were analyzed using the online available K2D software. The results suggest a predominantly alpha helical structure for the enzyme [Figure 3.14]. This reminisces of the corresponding aroQ class of enzymes from yeast and *E. coli* [Xue Y, 1993; Strater *et al.*, 1996; Lee *et al.*, 1995] that are also

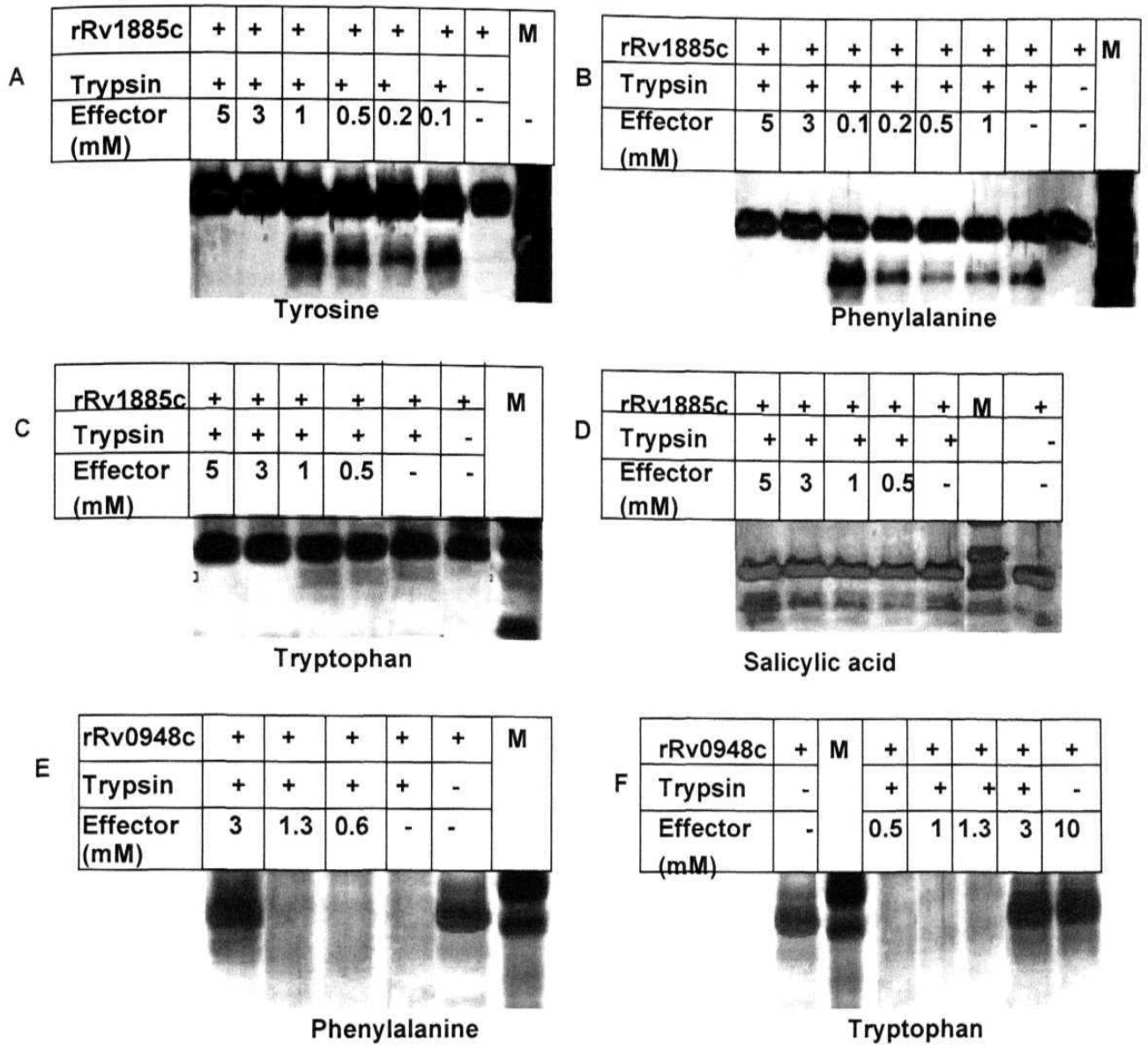


Figure 3.12 : Aromatic amino acids at high concentrations (3mM and above) provide protection to *M. tuberculosis* chorismate mutase from trypsin cleavage (Limited proteolysis studies). As evident from the above figure, tyrosine, phenylalanine and tryptophan (concentration 3mM and above) protect *M. tuberculosis* chorismate mutase enzymes (Rv1885c and Rv0948) from proteolytic cleavage (A,B,C,E,F). Salicylic acid, another cross pathway effector molecule does not protect *M. tuberculosis* CM (Rv1885c) from proteolytic cleavage (D).

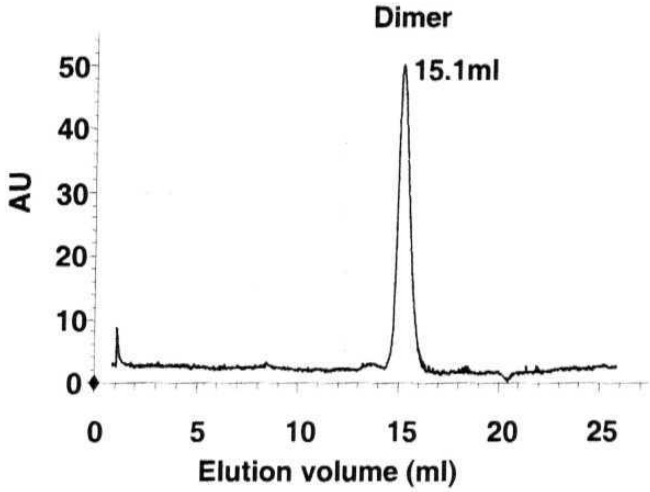


Figure 3.13: Size exclusion chromatography reveals that *M. tuberculosis* chorismate mutase (Rv1885c) is a dimeric protein (two subunits of individual MW 22kDa) rRv1885c was loaded on a Superdex 200 column (Pharmacia Biotech) and absorbance at 280nm (AU) was plotted as a function of elution volume. The elution parameter Kav vs log MW plot of protein standards was used to determine the actual MW of the recombinant protein.

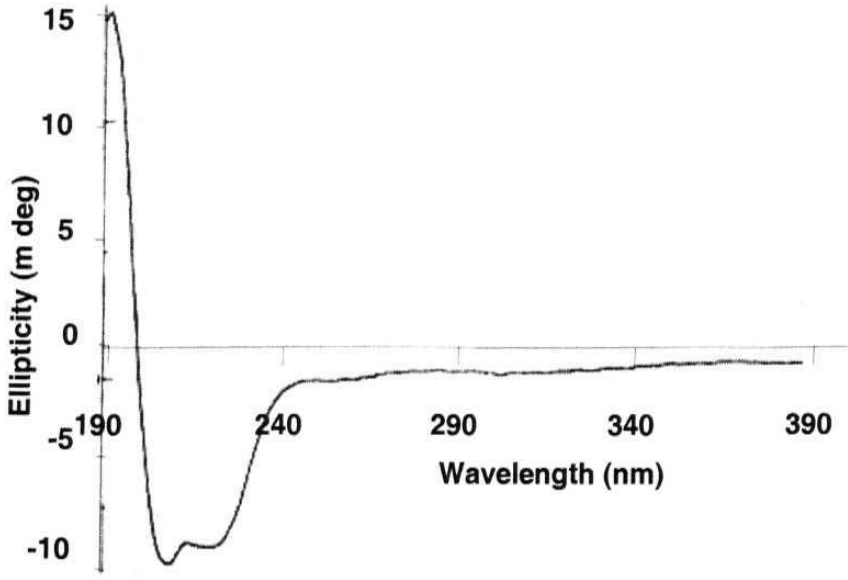


Figure 3.14: CD spectrum of *M. tuberculosis* chorismate mutase (Rv1885c) suggests a predominantly alpha helical secondary structure of the protein. The graph has been plotted as ellipticity (mdeg) as a function of wavelength (nm). Percentage helicity for secondary structure determination was calculated using the online available K2D software (http://www.embl-heidelberg.de/~andrade/k2d/) The spectrum confirms a highly helical structure of the recombinant protein with signature peaks at 208 and 222nm.

all helical proteins. The AroQ class of CMs consist of unregulated and regulated [AroQR] enzymes and are unusually divergent among closely related organisms [Hall *et al.*, 1982].

3.4.6 The N-terminal signal sequence of *M. tuberculosis* chorismate mutase can export *E. coli* alkaline phosphatase to the periplasmic space

While our experimental evidence in support of a dimeric alpha helical structure for *M. tuberculosis* CM classified the enzyme as an AroQ class of CM, the prediction of a signal peptide in the N-terminus of the protein was intriguing to determine if *M. tuberculosis* CM belongs to a periplasmic subclass of CMs. The CM of *Erwinia herbicola*, *Pseudomonas aeruginosa* and *Salmonella typhimurium* also possess an N-terminal signal sequence and differ in this respect from other members of the aroQ class of CMs [Xia *et al.*, 1993]. The study was therefore continued with the functional assessment of the N-terminal signal sequence of *M. tuberculosis* CM.

The approach used involved tagging *M. tuberculosis* CM signal sequence to PCR amplified *E. coli pho*A lacking its native signal sequence. This recombinant DNA molecule was cloned in pBAD18MCS2 vector [Guzman *et al.*, 1995] where expression is driven from an arabinose inducible promoter. The resultant construct was used to transform an *E. coli pho*A negative strain. In *E. coli*, Alkaline phosphatase [AP] is coded by *pho*A. AP is inactive in the cytosol due to its highly reducing environment. Wild type *E. coli pho*A has a signal sequence its N-terminus, which exports the protein to the peiplasmic space, wherein AP is functional [Brickman and Beckwith, 1975]. This property of *E. coli pho*A was used to test if the N- terminal probable signal sequence of *M. tuberculosis* CM could export *E. coli* AP [lacking the signal sequence] to the periplasmic space thereby making it active. AP activity can be detected on solid media containing X-P [a colorless compound like X-gal, that can be cleaved by AP to form a blue coloured product].

It was observed that the *E. coli pho*A negative strain, transformed with the above mentioned chimeric construct, could demonstrate *pho*A activity. Functional activity of AP

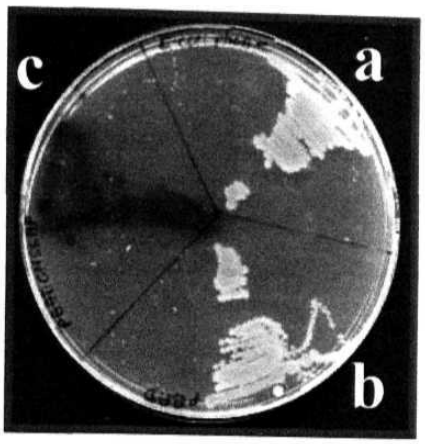
was seen in its ability to cleave the substrate XP, which could be monitored by the appearance of blue colour of the colony [Figure 3.15 A]. This indicated that *M. tuberculosis* CM ss could export AP to the periplasmic space. The *E. coli pho*A negative strain transformed with pBAD18MCS2 vector alone remained white. Localization of *Mtb* CM in the periplasmic and the cytoplasmic fractions of *E. coli* BL21 cells expressing rRv1885c were also assessed using western blot with anti-Histidine antibody. Almost the entire CM activity was confined to the periplasmic space [Fig 3.15 B].

3.4.7 Rv1885c is not the sole chorismate mutase enzyme of *M. tuberculosis*: Rv0948c also shows CM activity though with a reduced turnover

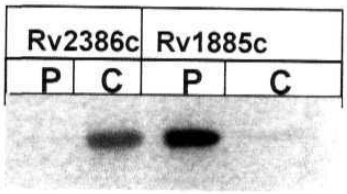
The COG [Cluster of Orthologous groups] annotation of the *M. tuberculosis* genome suggests that CM activity could be attributed to two ORFs of *M. tuberculosis*; Rv1885c and Rv0948c. Hence, *M. tuberculosis* ORF Rv0948c was also cloned and expressed in *E. coli* and the recombinant protein was assayed for CM activity. Results demonstrate that even rRv0948c shows a CM activity though its molar catalytic activity is much lower [50 fold lower than Rv1885c] [Figure 3.16]. The existence of another CM in *M. tuberculosis* genome is consistent with the hypothesis that a free monofunctional CM never exists in a single copy in any genome [Gu *et al.*, 1997]. However, it is inconsistent with the hypothesis that amongst two copies of CMs, at least one is a fusion protein. rRv0948c also showed no PDT or PDH activity indicating that it is also a monofunctional protein. rRv0948c was also found to be resistant to proteolytic digestion in the presence of phenylalanine and tryptophan [Figure 3.12 E, F] indicating that it is a regulated protein.

3.5 DISCUSSION

Chorismate mutase is known to be the sole example of an enzyme that catalyzes a pericyclic rearrangement reaction and has therefore drawn extensive attention of biologists and chemists. Earlier reports that have referred to differences in the structure and regulatory properties of the enzyme from different sources was the basis of this study of the respective properties of the equivalent protein from *M. tuberculosis*. This



A



B

Figure 3.15: The N-terminus of *M. tuberculosis* chorismate mutase (Rv1885c) encodes a functional signal sequence.

A: *E coli pho*A- strain (Sector a) was transformed with pBAD18MCS2 vector(Sector b) and pETCMssAP (Sector c) and streaked on LB-XP plates. Blue colour of colonies was observed only in Sector c, which indicates that signal sequence less *E. coli* alkaline phosphatase could be exported to the periplasmic space by the putative signal sequence in the N-terminus of *M. tuberculosis* CM.

B: Western blot demonstrating the presence of *M. tuberculosis* CM in the periplasmic space: 5 mg of protein from the cytosolic (C) and periplasmic (P) fractions of *E. coli* BL 21 cell expressing *M. tuberculosis* AS (Rv2386c) and CM (Rv1885c) were transferred to a nitrocellulose membrane and probed with monoclonal anti-Histidine antibody. As evident from the blot, only CM (Rv1885c) of *M. tuberculosis* shows a periplasmic localization.

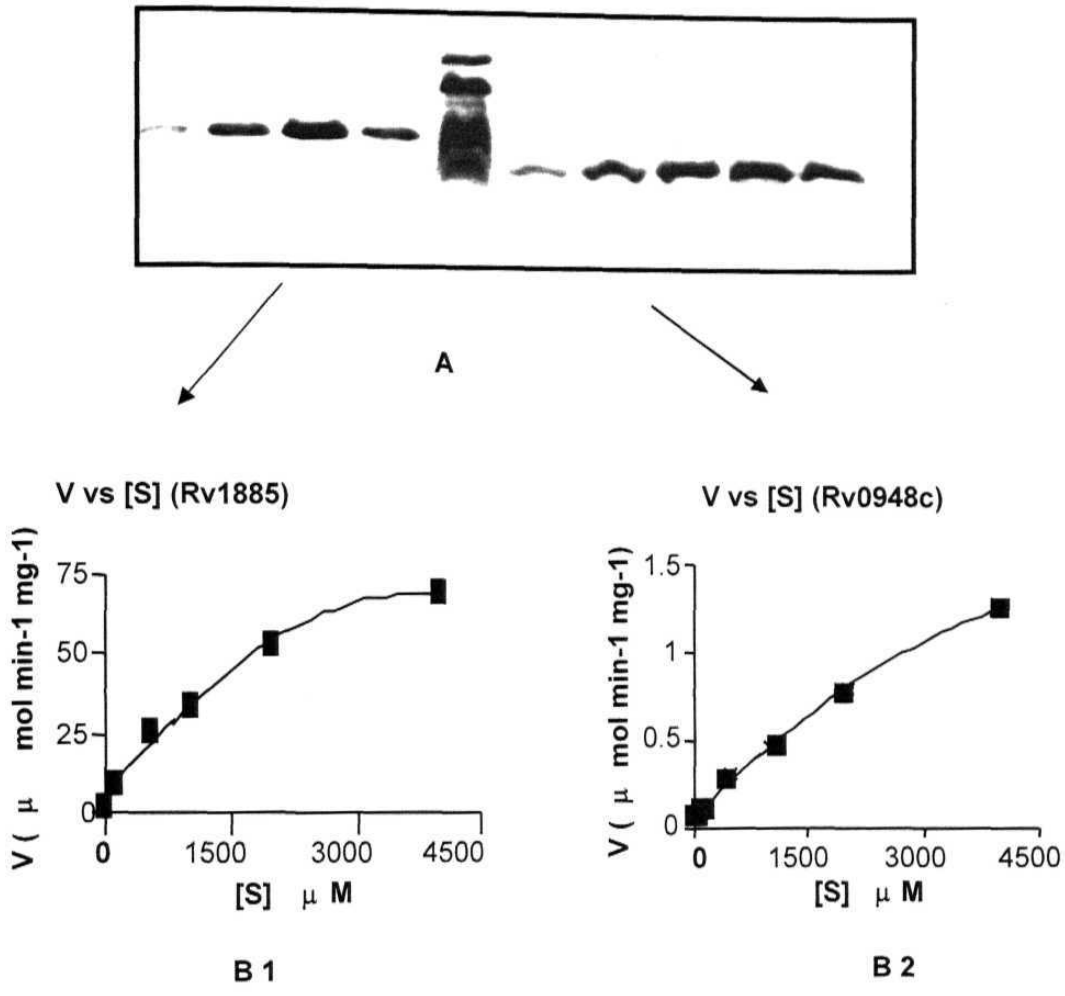


Figure 3.16: Comparative study of the chorismate mutase activity of the ORFs Rv1885c and Rv0948c. A: Purified fractions of both the proteins used for CM activity assay. B: V vs [S] plot of rRv1885c (B1) and rRv0948c (B2). As evident from the graph, Rv0948c does possess CM activity. However, the specific activity of rRv0948c was about 50 fold lower than the specific activity of Rv1885c.

study involved a detailed biochemical and biophysical characterization of recombinant *M. tuberculosis* CM, encoded by ORF Rv1885, and comparison of these properties with enzyme from other sources. The results presented in this chapter conclusively place *M. tuberculosis* CM, coded by ORF Rv1885c in the AroQ class of periplasmic chorismate mutases.

In the course of this study, several unique properties of *M. tuberculosis* CM became evident which are important in the context of differentiating the enzyme from the bacterial and fungal counterparts. The first and foremost observation with respect to the regulation of *M. tuberculosis* CM by aromatic amino acids surfaced an unusual property of the enzyme. While *M. tuberculosis* CM was found to be allosterically regulated by pathway specific [Typ and Phe] as well as cross pathway specific [Trp] ligands in the same manner, it showed a modest increase in enzyme activity at low concentrations and complete inhibition at higher concentrations of effectors. Though the reason behind this modest increase is still unclear, a logical explanation would be that, when intracellular amino acid concentrations are low, its cellular requirement is high and hence the enzyme is moderately activated. It is proposed that this unusual property of *M. tuberculosis* CM must be coordinately regulated with other enzymes of aromatic amino acid biosynthesis pathway. Here, it would be worthwhile to mention that *M. tuberculosis* PheA [Prephenate dehydratase], the enzyme that catalyzes the subsequent step in the biosynthesis of Phenylalanine, is regulated in exactly the opposite manner i.e. it is inhibited by low concentrations and activated by higher concentrations of aromatic amino acids [described in section 4.4.3, Chapter 4].

As binding of ligands causes inhibition of enzyme activity, further experiments were designed to study if ligand binding would also protect the enzyme from proteolytic cleavage. This experiment was expected to generate another line of evidence that ligand binding induces a conformational change in the enzyme. Limited proteolysis studies suggest that ligand binding imparts a much more rigid structure to the enzyme leading to enhanced stability and inaccessibility of the active site and protease cleaving sites. A cross-pathway specific secondary metabolite, salicylic acid could not protect *M. tuberculosis* CM from proteolytic cleavage. Salicylic acid is also derived from chorismate.

This suggests that only aromatic amino acids and not secondary metabolites have a binding pocket on the enzyme.

The chorismate to prephenate reaction is also known to occur spontaneously and an increase in product formation at higher temperatures was also observed. However, there was a distinct enhancement of reaction rate upon addition of the enzyme. It is possible that in spite of a spontaneous conversion of chorismate to prephenate, there is a need for enzyme catalyzed conversion as chorismate is a substrate for multiple enzymes and the fate of the compound is to be determined by the enzyme that acts upon it that in turn is dependent on the requirement of the organism.

It is also interesting to note that unlike most enteric bacteria where chorismate mutase exists in fusion with other enzymes of the aromatic amino acid biosynthesis pathway, *M. tuberculosis* CM is predicted to be a non-fusion protein [http://www.sanger.ac.uk/cgi-bin/Pfam/speciesview.pl?family=CM_2&acc=PF01817&ncbicode=83332&name=Mycobacterium%20tuberculosis]. Yet, on account of several reports of divergent evolution of aromatic amino acid biosynthesis enzymes, experiments were designed to test if the amino acid sequence of *M. tuberculosis* CM could also contribute to PDT or PDH activity. The results presented in this chapter confirm that *M. tuberculosis* CM is a monofunctional protein i.e it has no associated PDT or PDH activity. The CMs that show allosteric regulation are usually large fusion proteins. *M. tuberculosis* CM appears to be unique in the sense that inspite of being a small monofunctional protein, it shows a complex pattern of allosteric regulation. While *M. tuberculosis* does not possess a discrete ACT domain that is required for amino acid binding, it is possible that the complex regulatory activity of the enzyme rests in the uncharacterized C-terminus of the enzyme.

Among the monofunctional AroQ proteins, existence of a cleavable signal peptide has been reported in *Erwinia herbicola*. Later, AroQ CMs from *Salmonella typhimurium* and *Pseudomonas aeruginosa* were also shown to have a periplasmic location [Calhoun *et al.*, 2001]. Periplasmic CMs possess a signal peptide at the N-terminus and usually have a C-terminus of an unknown function. The C-terminus of *M. tuberculosis* CM is also

unique and its precise function is still unclear, though we suggest that it could have a regulatory role. Based upon the ability of *M. tuberculosis* CM signal sequence to export *E. coli* PhoA to the periplasmic space, it was demonstrated that *M. tuberculosis* CM also belongs to this unique periplasmic subclass of AroQ proteins. However, this study also necessitates the requirement for identification of other periplasmic enzymes of aromatic amino acid biosynthesis pathway of *M. tuberculosis*. This study would shed further light on whether there exists a periplasmic pathway of aromatic amino acid biosynthesis in *M. tuberculosis*. In the absence of existence of such a pathway, it is proposed that the periplasmic location of *M. tuberculosis* CM could either have a role to play in the modulation of the host immune response or the enzyme could have a chemotactic role, [Maddock and Shapiro, 1993]. It is significant to note that the ORF Rv1885c exists in the same operon as Antigen85A, a major secreted antigen of *M. tuberculosis*. It is also worthwhile to mention here that chorismate mutase is a virulence factor of nematode parasites that secrete CM from their oesophagal glands to alter the shikimate pathway of the plants which they parasitize [Lambert *et al.*, 1999; Bekal *et al.*, 2003]. While the shikimate pathway does not exist in humans, *M. tuberculosis* chorismate mutase might be released as a virulence factor to alter the shikimate pathway of other bacteria that would have invaded the same host cell. Such a mechanism might allow *M. tuberculosis* to be the sole invader of the host macrophages.

Finally, the evidence for the existence of another CM in *M. tuberculosis* genome is in accordance with the hypothesis that a free monofunctional CM never exists in a single copy in any genome [Gu *et al.*, 1997]. Interestingly however, this is the first report that demonstrates the existence of two monofunctional CMs in any bacterial system. In other bacteria, where CM is present in two copies, at least one of them is a fusion protein [Huang *et al.*, 1974]. Evolution of these fusion proteins was brought about with the purpose of coordination of gene expression as per the requirement of the cell. Hence, the issue of co-regulation of *M. tuberculosis* CMs and PDT/PDH still remains. While IdeR is known to govern the expression of *M. tuberculosis* prephenate dehydratase [Gold *et al.*, 2001], regulation of expression of *M. tuberculosis* CM is still not known. The apparent lack of synchrony in gene expression also suggests that *M. tuberculosis* CM

could be involved in some other pathway that is independent of the involvement of PDT/PDH.

While the data presented in this chapter support the existence of two monofunctional CMs in *M. tuberculosis* genome, yet the question remains, Why two CMs? The clue lies in the observation of a huge difference in turnover number of the two CMs, which suggests that one of these enzymes could have an additional function. Additionally, the existence of one of the CMs as a secreted protein also suggests an as yet undeciphered function of the enzyme. The second chorismate mutase of *M. tuberculosis* [Rv0948c] does not possess a signal sequence which suggests that there are two pathways for aromatic amino acid biosynthesis in *M. tuberculosis*. The role of the periplasmic pathway is still not clear. Our other novel observation that the periplasmic CM of *M. tuberculosis* shows a higher turnover number also demands elucidation of some other function of the enzyme. As chorismate utilizing enzymes are known to have divergently evolved, it is likely that these enzymes could have some other function as well. In *Pseudomonas aeruginosa*, the isochorismate pyruvate lyase [IPL] enzyme, PchB has been reported to show chorismate mutase activity [Gaille *et al.*, 2003].

In summary, these results suggest that the hypothetical proteins coded by the ORFs Rv1885c and Rv0948c of *Mycobacterium tuberculosis* represent the chorismate mutase enzymes of the organism and the kinetic properties of the enzyme are under the control of strict regulatory mechanisms. Regulation is brought about by all the three aromatic amino acids *viz* Tryptophan, tyrosine and phenylalanine. This is the first detailed report that provides evidence for the existence of a monofunctional AroQ class of periplasmic CM in *M. tuberculosis* genome. Additionally, a low level sequence similarity of the protein with other known chorismate mutases make this protein an attractive molecule for the design of novel and specific compounds that would check tuberculosis.

CHAPTER 4

---

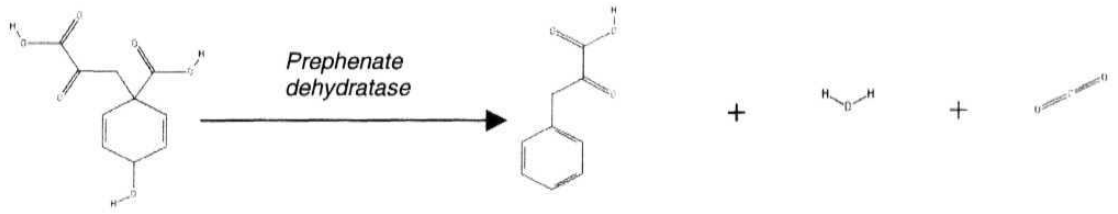
pheA, AN IdeR *REGULATED GENE OF Mycobacterium tuberculosis ENCODES A MONOFUNCTIONAL PREPHENATE DEHYDRATASE THAT REQUIRES BOTH CATALYTIC AND REGULATORY DOMAINS FOR OPTIMUM ACTIVITY*

4.1 ABSTRACT

Prephenate dehydratase [PDT] is a key regulatory enzyme in L-phenylalanine biosynthesis. In *M. tuberculosis, e*xpression of *phe*A, the gene encoding PDT has been earlier reported to be iron-dependent [Gold *et al.*, 2001; Rodriguez *et al.*, 2003]. The work described in this chapter shows that *M. tuberculosis phe*A is also regulated at the protein level by aromatic amino acids. All the three aromatic amino acids [Phenylalanine, Tyrosine and Tryptophan] are potent feedback activators of *M. tuberculosis* PDT. *In vitro* evidence that *M. tuberculosis* PDT does not show any chorismate mutase activity is also presented here which suggests that unlike many other enteric bacteria, where PDT exists as a fusion protein with chorismate mutase, *M. tuberculosis* PDT is a monofunctional and a non-fusion protein. Finally, the biochemical and biophysical properties of the catalytic and regulatory domains [ACT domain] of *M. tuberculosis* PDT were studied. It was observed that in the absence of the ACT domain the enzyme not only loses its regulatory activity but also its catalytic activity. This is the first report of a monofunctional prephenate dehydratase enzyme from a pathogenic bacteria that shows extensive feedback activation by aromatic amino acids and is absolutely dependent upon the presence of catalytic as well as the regulatory domains for optimum enzyme activity.

4.2 INTRODUCTION

*phe*A of *Mycobacterium tuberculosis* has been earlier reported to be an IdeR regulated gene [Gold *et al.*, 2001; Rodriguez *et al.*, 2003]. *phe*A codes for the enzyme prephenate dehydratase [PDT] that catalyzes the conversion of prephenate to phenylpyruvate with the elimination of water and carbon dioxide as shown below.



Prephenate Phenylpyruvate H_2O CO_2

Further, with the action of aromatic aminotransferase, phenylpyruvate is converted to phenylalanine [Figure 3.1]. Biochemical significance of the enzyme lies in its participation in the terminal steps of the biosynthesis of the aromatic amino acid, phenylalanine.

In enteric bacteria PDT usually exists as a fusion protein with chorismate mutase [Figure 4.1], thereby catalyzing the first two steps in the biosynthesis of phenylalanine. In *E. coli* as well, CM and PDT are fusion partners of the bifunctional P-protein, coded by *phe*A [Pittard, 1987]. In many other bacteria, PDT is a monofunctional protein that aligns well with the C-terminal part of P-proteins [Fischer and Jensen, 1987]. The N-terminal end of the bifunctional P-protein of *E. coli* specifies the chorismate mutase activity while the remainder of the sequence specifies the prephenate dehydratase enzymatic activity [Zhang *et al.*, 1998]. The native enzyme is a dimer of identical subunits, each containing a dehydratase active site, a mutase active site and a phenylalanine binding site. [Gething and Davidson, 1977]. PDT is reported as a highly regulated enzyme in bacterial systems [Pohnert *et al.*, 1999]. Protein engineering studies have established that the *E. coli* P-protein has three distinct domains and their functions are preserved when the individual domains are expressed separately as fragments [Zhang *et al.*, 1998]. While the CM domain rests in residues 1-109, the PDT domain includes residues 101-285. The

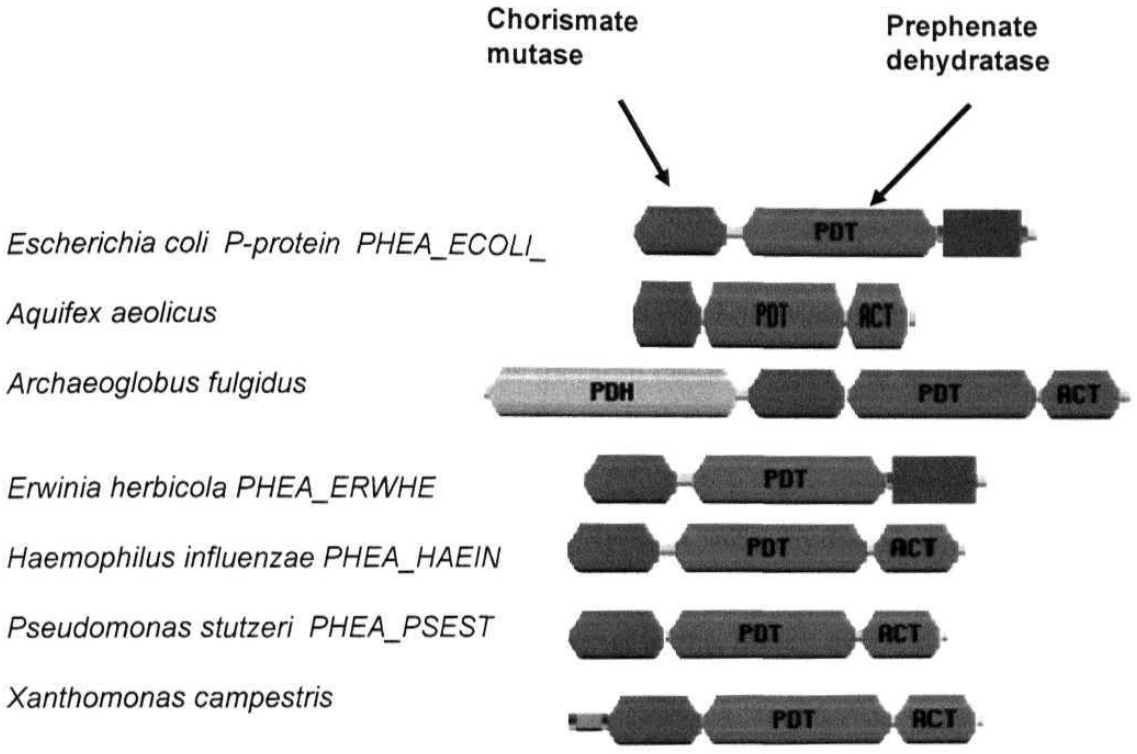


Figure 4.1: Prephenate dehydratase (PDT) exists as a fusion protein with chorismate mutase in most enteric bacteria
Source: Pfam

authors could also decipher a separate allosteric domain [residues 286–386], which was responsible for feedback inhibition by Phenylalanine. Further, the authors also showed that the Phe-binding domain of the P-protein might function as a 'modular' regulatory element when physically joined to an independent protein catalyst [Zhang *et al.*, 2003].

The chorismate mutase and prephenate dehydratase reactions in the P protein of *E. coli* occur at separate active sites. In contrast, the two active sites of the closely related enzyme chorismate mutase/prephenate dehydrogenase [*E. coli* T protein] are interacting sites [Hudson *et al.*, 1984]. Studies of the overall reaction using radioactive chorismate have shown that prephenate, which is formed from chorismate, dissociates from the mutase site and equilibrates with the bulk medium before combining at the dehydratase site. Also, the two activities are subject to differential inhibition. [Duggleby *et al.*, 1978]

In a few organisms like *Bacillus subtilis*, prephenate dehydratase is a monofunctional protein [Riepl *et al.*, 1979]. In *M. tuberculosis* as well, the ORF Rv3838c, annotated as a prephenate dehydratase appears monofunctional in the sense that it possesses only a prephenate dehydratase domain.

Regulatory domains or the ACT domains are present in many metabolic enzymes including enzymes of aromatic amino acid biosynthesis family [Chipman and Shaanan, 2001]. Amino acids that have a regulatory role in the activity of an enzyme usually bind to the ACT domain. The ACT domain is found in a variety of contexts and is proposed to be a conserved regulatory binding fold [Chipman and Shaanan, 2001]. ACT domains are linked to a wide range of metabolic enzymes that are regulated by amino acid concentration. The archetypical ACT domain is the C-terminal regulatory domain of 3-phosphoglycerate dehydrogenase [3PGDH], which folds with a ferredoxin-like topology [Schuller *et al.*, 1995]. The prephenate dehydratase of *M. tuberculosis* also possesses discrete ACT domain that is predicted to impart regulatory properties to the enzyme. In other bacteria as well, prephenate dehydratase is predicted to be a highly regulated enzyme.

PDT enzyme is known to exist in various oligomeric states in different bacterial systems. PDT of *Bacillus subtilis* has been reported to exist in three states of aggregation, an octamer, a dimer and a monomer (Riepl and Glover, 1979).

With the above background this part of the study was initiated to determine whether apart from genetic regulation, *M. tuberculosis pheA* is also regulated at the enzymatic level. The experimental approach involved expression of the *M. tuberculosis* ORF, Rv3838c in *E. coli* and determination of the biochemical parameters of the encoded protein. This chapter gives a brief description of the kinetic and regulatory properties of *M. tuberculosis* PDT and also details the use of engineered proteins to elucidate the role of the catalytic and regulatory domains of the recombinant protein. The significance of a differential feedback regulation of the terminal enzymes of the aromatic amino acid biosynthesis pathway of *M. tuberculosis* is also discussed.

4.3 EXPERIMENTAL PROCEDURES

4.3.1 Bacterial strains and plasmids:

E. coli XL1Blue and *E. coli* BL21 DE3 strains were used respectively as hosts for cloning and recombinant protein expression purposes. pET23a [Novagen] was taken as the expression vector.

4.3.2 Cloning, expression and purification of *M. tuberculosis* PDT and PDT-N and PDT-C

The ORF Rv3838c that corresponds to the prephenate dehydratase enzyme of the bacterium was amplified from *M. tuberculosis* H37Rv genomic DNA using primers carrying specific restriction enzyme sites [Table 4.1]. The amplicon were digested with *Nde*I/*Xho*I enzymes and cloned into the corresponding sites of pET23a expression

Table 4.1: Sequences of primers used for PCR amplification of *M. tuberculosis* prephenate dehydratase [coded by ORF Rv3838c] and its N and C termini [catalytic and regulatory domains]

Gene	Forward primer sequence	Reverse primer sequence
Rv3838c	*NdeI*	*XhoI*
	ATA***CATATG***GTGGTGCGTATCGCTTACCTCGGT	AT**CTCGAG**TGCTTGCGCCCCCTGGT
Rv3838N	*NdeI*	*XhoI*
	ATA***CATATG***GTGGTGCGTATCGCTTACCTCGGT	AT**CTCGAG**ATCGGCTCCGGTGCGCG
Rv3838C	*NdeI*	*XhoI*
	ATA***CATATG***CGCACGTCTGCAGTGCTGC	AT**CTCGAG**TGCTTGCGCCCCCTGGT

vector. The N and the C terminus of *M. tuberculosis* PDT corresponding to the catalytic and regulatory domains were cloned separately using other sets of primers described in Table 4.1. The resultant plasmids were labeled as pET3838, pET3838N and pET3838C.

The pET: 3838/3838-N/3838-C chimeric constructs were used to transform *E coli* BL21DE3 cells and protein expression and purification was carried out exactly as described in Chapter 2, section 2.3.2.

4.3.3 Enzyme assays and kinetic studies

4.3.3.1 Prephenate dehydratase activity assay

Prephenate dehydratase activity assays were carried out as described earlier with a few modifications [Gething *et al.*, 1977]. Essentially, the dehydratase activity of rRv3838c was assayed by measuring the rate of conversion of prephenate to phenylpyruvate. The reaction mixture contained 20mM Tris-HCl, pH 8.2, 1mM EDTA, 0.01% BSA, 1mM DTT and 0.5-2mM barium prephenate in a total volume of 400μl. The sample was pre-incubated at 37^0C followed by addition of 20-100 pico moles of purified recombinant *M. tuberculosis* PDT. After a second round of incubation at 37^0C for 5 minutes, the reaction was terminated by the addition of 800μl of 1.5M NaOH. Phenylpyruvate was measured spectrophotometrically at 320nm. Appropriate blanks without the enzyme were kept as controls. Allosteric regulation of CM activity by L-Phe, L-Tyr and L-Trp was measured at 100-500μM concentrations of the effectors.

4.3.3.2 Chorismate mutase activity assay

Chorismate mutase activity assays were performed exactly as described in Section 3.3.4, Chapter 3

4.3.4 Analytical size exclusion chromatography

Gel filtration or size exclusion chromatography was used to determine the oligomeric state of *M. tuberculosis* PDT as well as the individually expressed catalytic and regulatory domains. The chromatography experiment was performed on a Superdex 200 [HP 10/30] FPLC column from Pharmacia Biotech using 10mM Tris and 100mM NaCl as the running buffer. Void volume of the column was determined using Blue Dextran 200. Elution time of all the recombinant proteins was recorded and molecular weight was calculated by estimating the elution volumes of standards of known molecular weight. The procedure is described in detail in Section 3.3.10 of Chapter 3. The recombinant proteins were loaded on the gel filtration column at a concentration of 2mg/ml in the presence of 1mM DTT.

4.3.5 Phenylalanine binding assays / Fluorimetric procedures

Binding of phenylalanine to the recombinant proteins was monitored using fluorescence spectroscopy. Fluorescence spectra of individual proteins were recorded in the presence and absence of phenylalanine on a Perkin-Elmer LS-3B spectrofluorimeter. Briefly, the recombinant proteins were excited using an excitation wavelength of 295nm and the tryptophan emission spectra were recorded from 305-440nm at 37^{0}C. Fluorescence slit width of excitation and emission were kept at 4nm and the scan speed was 50nm/sec. Protein concentrations were kept at 2.5μM in 10mM Tris, pH8 and the buffer signal was subtracted from the spectrum of recombinant proteins.

4.4 RESULTS

4.4.1 *pheA* (ORF Rv3838c) of *Mycobacterium tuberculosis* encodes a functional prephenate dehydratase enzyme

The recombinant protein corresponding to ORF Rv3838c, that is predicted to be the prephenate dehydratase [PDT] enzyme of the bacterium could be purified using affinity chromatography procedures [Figure 4.2]. Purified protein was quantified using Bradford assay and used for prephenate dehydratase activity assays. As evident from the substrate saturation plot, *M. tuberculosis* PDT followed Michealis Menton's kinetics [Figure 4.3A]. The double reciprocal plot of the enzyme was linear and Km for the enzyme was calculated as 1.38mM and Vmax as 270 μmoles/min/mg [Figure 4.3B]. Specific activity of the enzyme was determined as 270.27 units/mg pure protein. The effect of temperature and pH on enzyme activity was also studied. It was observed that the enzyme is maximally active between temperatures 37-42^{0}C and pH 6-7 [Figure 4.4: A,B]. Unlike the chorismate mutase of *M. tuberculosis*, *M. tuberculosis* prephenate dehydratase has a narrow range of tolerance of temperature and pH. Also, while other bacterial PDTs are active at an acidic pH [Friedrich *et al.*, 1976], *M. tuberculosis* PDT was found to be active from mildly acidic to near neutral pH.

4.4.2 Ionic interactions are required for optimum PDT activity

To determine if ionic interactions are important for the activity of *M. tuberculosis* PDT, the enzyme was assayed in the presence of various concentrations of NaCl. It was observed that the enzyme activity was completely abolished in the presence of 200mM and higher concentrations of NaCl [Figure 4.5]. This suggests that enzyme substrate interactions are brought about by ionic interactions and disruption of the same leads to inhibition of enzyme activity.

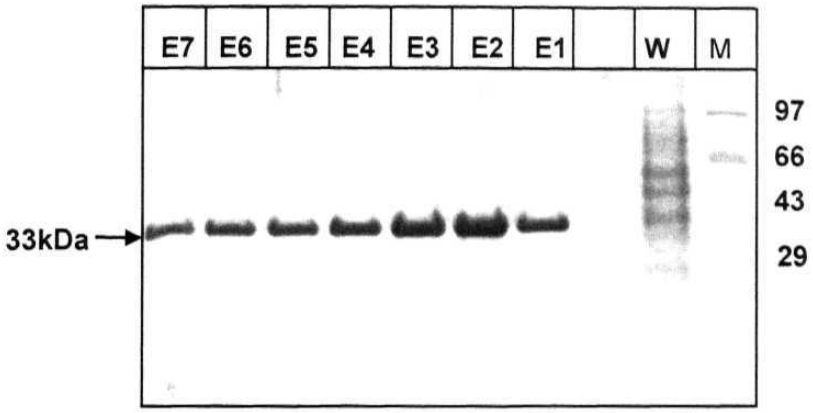


Figure 4.2: Purification of *M. tuberculosis* prephenate dehydratase (PDT) as a recombinant protein in *E. coli*. *Mtb phe*A corresponding to ORF Rv3838c was cloned in the *Nde*I and *Xho*I sites of pET23a vector with a C-terminal Histidine tag and expressed in *E. coli* BL21 cells. Affinity purification of recombinant protein was carried out using Talon resin (Clonetech). Purified protein resolved on a 10% Tris-Tricine gel and stained with Coomassie Brilliant Blue dye. M represents the protein molecular size marker (Genei, India), W represents the wash fraction during affinity purification of recombinant protein and E1-E7 show the successive eluted fractions of the 33kDa recombinant protein.

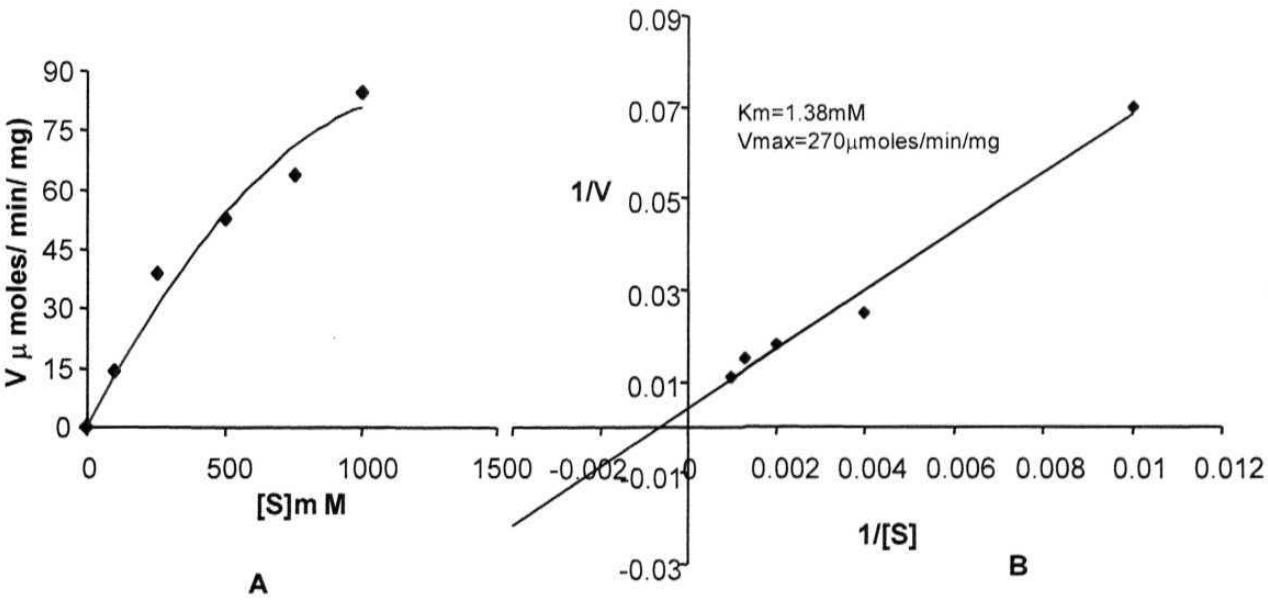


Figure 4.3: A: Substrate saturation plot of *Mtb* Prephenate dehydratase. The enzyme follows Michealis Menton's kinetics B: Double reciprocal plot of *Mtb* PDT enzyme activity that was used to calculate Km and Vmax. The specific activity of recombinant *Mtb* PDT was calculated as 270 units/mg pure protein.

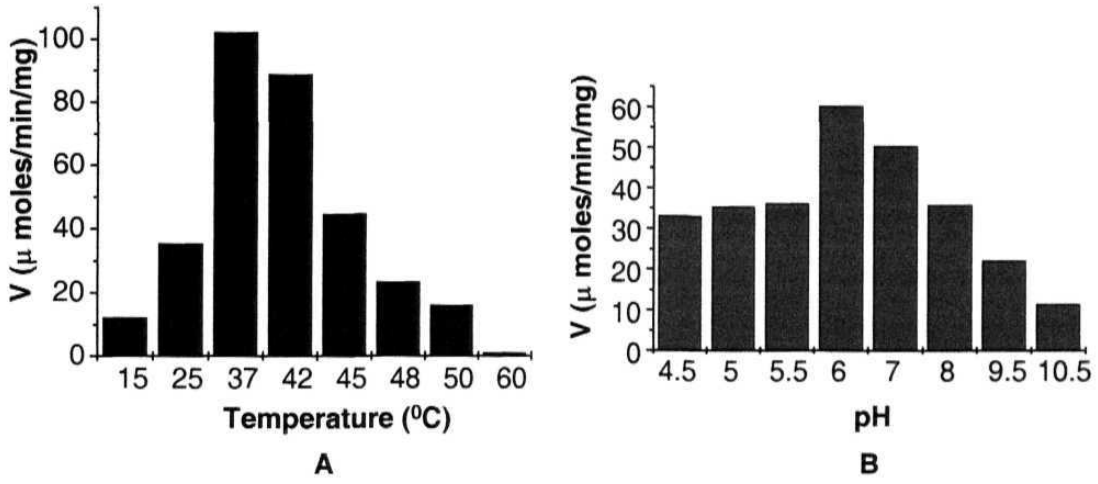


Figure 4.4: Effect of temperature and pH on the prephenate dehydratase (PDT) activity of rRv3838. Optimum pH for *M. tuberculosis* PDT activity is between pH 6-7. Optimum temperature for *M. tuberculosis* PDT activity is 37°C. Temperatures above 50°C completely inactivate the enzyme

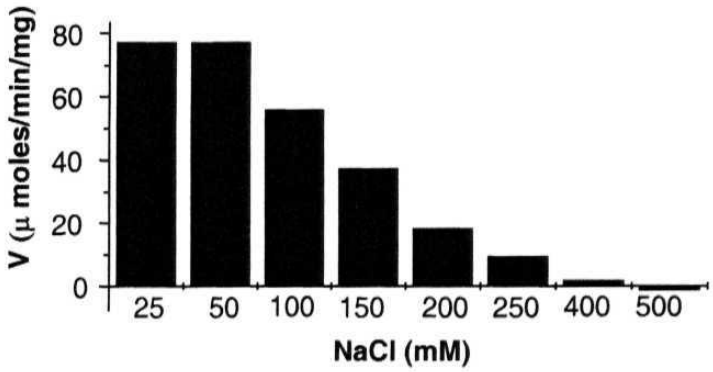


Figure 4.5: Ionic interactions are required for optimum catalytic activity of *M. tuberculosis* prephenate dehydratase (PDT) . Higher concentrations of NaCl (200mM and above completely abolish the PDT activity of *M. tuberculosis* ORF, Rv3838c.

4.4.3 Aromatic amino acids are potent feedback activators of *M. tuberculosis* prephenate dehydratase

Control of the terminal enzymes in any biosynthetic pathway is crucial for maintaining the correct balance of the end product in accordance with the requirement of the organism. This holds true for the aromatic amino acid biosynthesis as well. PDT being a terminal enzyme in phenylalanine biosynthesis, it was decided to determine the effect of phenylalanine as well as other cross pathway specific aromatic amino acids [tyrosine and tryptophan] on enzyme activity. It was observed that PDT activity is allosterically regulated in a manner exactly opposite to that of *M. tuberculosis* CM [refer section 3.4.3, chapter 3]. More specifically, *M. tuberculosis* PDT was inhibited by lower concentrations of aromatic amino acids [up to ~100μM] and highly activated at higher concentrations [Figure 4.6 A,B and C]

4.4.4 *M. tuberculosis* prephenate dehydratase does not display any chorismate mutase activity

While the chorismate mutase enzyme of *M. tuberculosis* did not show any prephenate dehydratase activity [Section 3.4.2], it was determined if the reverse also holds true i.e. whether the PDT of *M. tuberculosis* displays any CM activity. Though *M. tuberculosis* PDT does not have a predicted CM domain, on account of several examples of convergent evolution of enzyme reaction mechanisms, it was decided to determine whether PDT also has any associated CM activity. Recombinant *M. tuberculosis* PDT was therefore used in a chorismate mutase activity assay. It was observed that *M. tuberculosis* PDT is completely devoid of CM activity. This experiment confirmed that like *M. tuberculosis* CM, *M. tuberculosis* PDT is also a monofunctional protein.

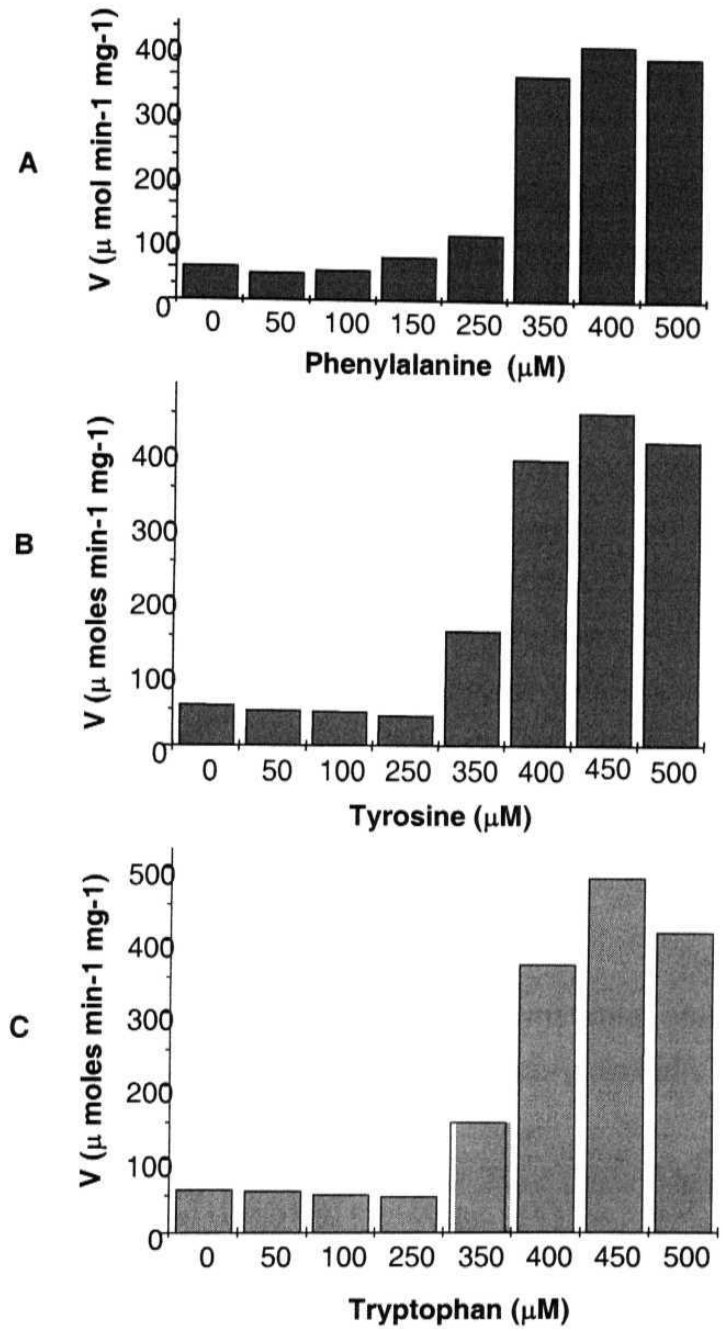


Figure 4.6: Aromatic amino acids at higher concentrations are potent feedback activators of *M. tuberculosis* PDT. Up to 100μM concentrations of aromatic amino acids negatively regulate *M. tuberculosis* PDT. However, a positive regulation was observed at higher concentrations (300μM and above).

4.4.5 The individually cloned, expressed and purified catalytic and regulatory domains of *M. tuberculosis* prephenate dehydratase are inactive in catalyzing the conversion of prephenate to phenylpyruvate

M. tuberculosis PDT has a distinct predicted catalytic and regulatory domain [Figure 4.7A]. Hence it was decided to test whether the catalytic activity of *M. tuberculosis* PDT is independent of the regulatory domain or it requires both the domains for optimum activity. For this purpose, the catalytic and regulatory domains of *M. tuberculosis* PDT were individually cloned and expressed in the *NdeI*/*XhoI* sites of pET23a vector and purified as recombinant proteins [Figure 4.7 B,C]. These proteins were labeled as PDT-N and PDT-C and were used in prephenate dehydratase activity assay. While the regulatory domain was completely inactive, it was surprising to note that even the individually cloned catalytic domain of *M. tuberculosis* PDT did not show any PDT activity [Figure 4.8]. Reconstitution of the enzyme was attempted by taking equimolar ratios of the catalytic and regulatory domains and using it in PDT activity assays. However, the activity assay demonstrated that the catalytic and regulatory domains can not be reconstituted *in vitro* to generate an active enzyme. These results are summarized in a tabular form [Table 4.2].

4.4.6 Phenylalanine binding leads to a conformational change in *M. tuberculosis* prephenate dehydratase enzyme and its regulatory domain

Fluorescence spectroscopy was used to determine if there is any conformational change in *M. tuberculosis* PDT upon ligand binding. This study was carried out with full length *M. tuberculosis* PDT as well as individual catalytic and regulatory domains. Tryptophan emission spectra for all the three recombinant proteins were monitored in the presence and absence of phenylalanine. *M. tuberculosis* PDT possesses four tryptophan residues and the individual catalytic and regulatory domains have 2 tryptophan residues each. As evident from the graph, the addition of phenylalanine leads to a change in the fluorescence emission spectrum of only the full length *M. tuberculosis* PDT and PDT-C [Figure 4.9: A, B].

Table 4.2: Activity profile of the engineered *M. tuberculosis* PDT proteins

Clones	Protein	Residues	Purified protein activity	
			PDT	CM
pET3838	PDT	1-321	+	_
pET3838N	PDT-N	4-192	_	NT
pET3838C	PDT-C	202-279	_	NT
	PDT-C + PDT-N		_	NT

NT: Not Tested;
+:active;
-: inactive

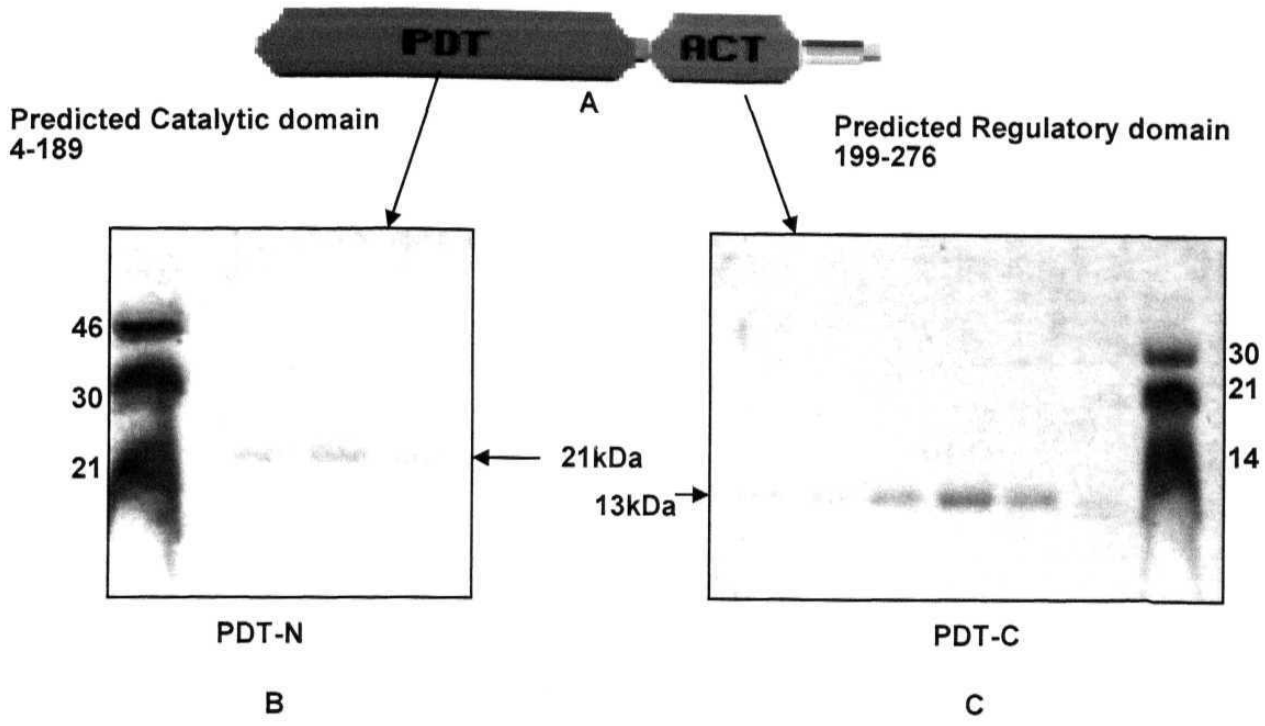


Figure 4.7: Expression and purification of the catalytic (PDT-N) and regulatory domains (PDT-C) of *M. tuberculosis* prephenate dehydratase. The individual domains were cloned in the *NdeI*/*XhoI* sites of pET23a expression vector and purified as recombinant proteins using affinity chromatography procedures. A: Schematic representation of the catalytic and regulatory domains B: Coomassie Blue stained polyacrylamide gel showing the eluted fractions of the pure recombinant proteins corresponding to catalytic and regulatory domains.

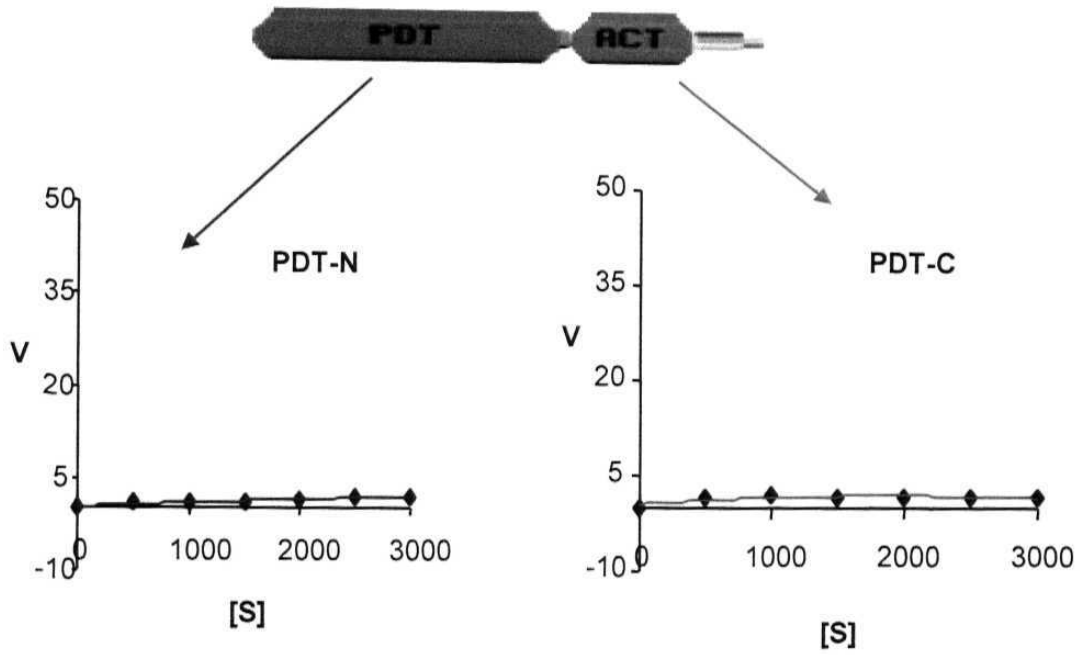
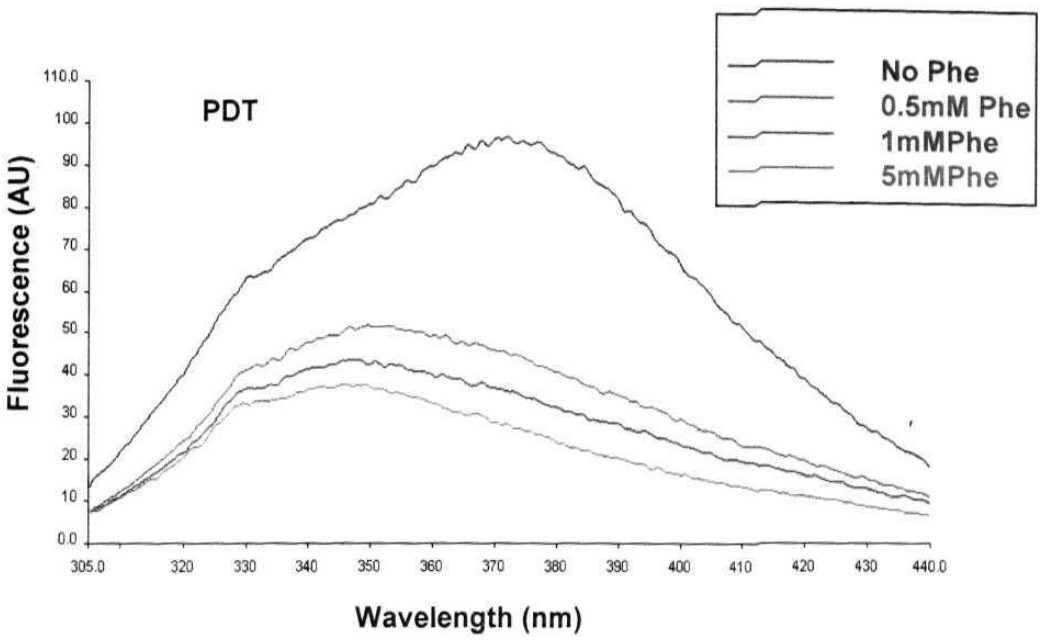
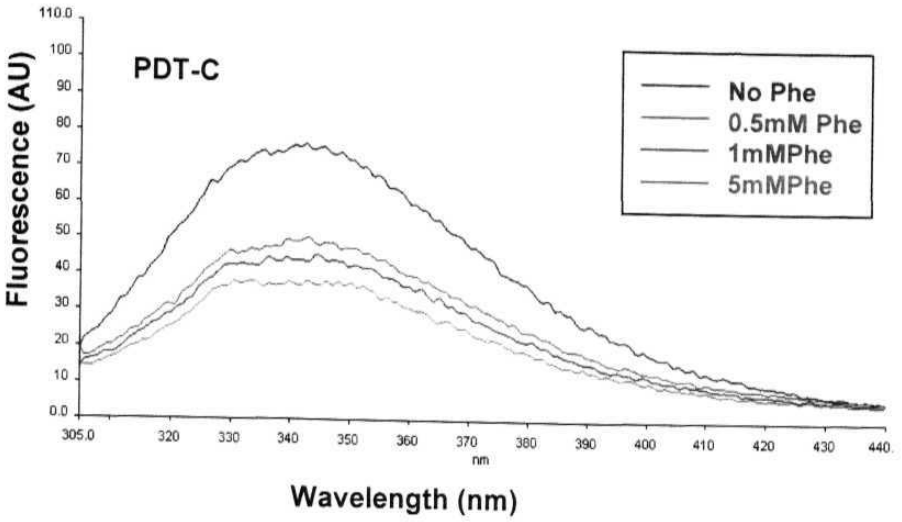


Figure 4.8: The catalytic (PDT-N) as well as the regulatory domains (PDT-C) of *M. tuberculosis* prephenate dehydratase in isolation can not catalyze the conversion of prephenate to phenylpyruvate. Lack of PDT activity of the individually purified catalytic and regulatory domains as evident from the above V vs. [S] plots suggest that interactions between both the domains are essential for enzyme activity.

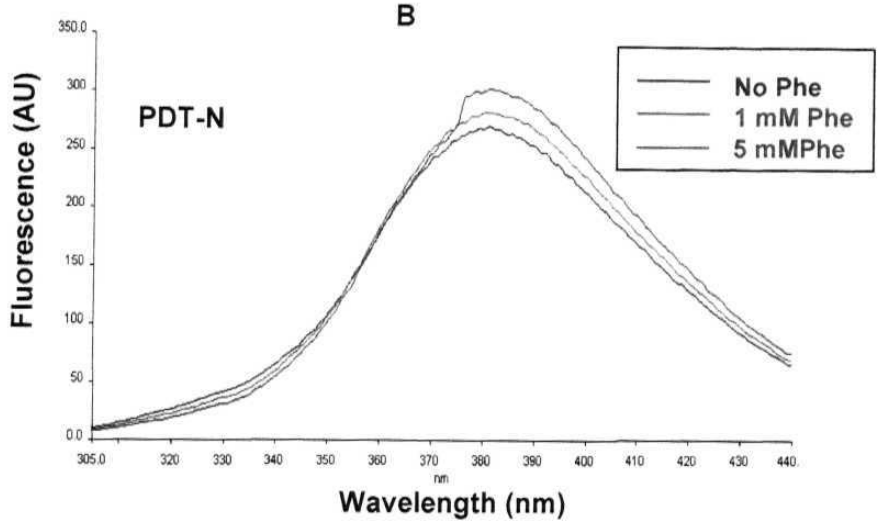


A

Figure 4.9: A: Fluorescence spectrum of *M. tuberculosis* prephenate dehydratase changes significantly in the presence of phenylalanine. The intrinsic fluorescence intensity decreased and the emission maxima shifted from 380nm to 350nm. This suggests that binding of phenylalanine renders a conformational change upon the enzyme. A blue shift in the spectra also indicates that tryptophan residues of the protein are buried inside upon phenylalanine binding.



B



C

Figure 4.9: B: Fluorescence emission spectrum of the PDT-C (Regulatory domain of *M. tuberculosis* prephenate dehydratase) is altered in the presence of phenylalanine. The intrinsic fluorescence intensity decreased though the emission maxima remained the same. This suggests that PDT-C undergoes a conformational change as a consequence of phenylalanine binding. C: Fluorescence emission spectrum of PDT-N (catalytic domain) remains unchanged upon the addition of phenylalanine. These experiments suggest that the phenylalanine regulatory domain resides only in the C-terminus of *Mtb* PDT.

The fluorescence emission spectrum of the catalytic domain or PDT-N remained unchanged even in the presence of phenylalanine [Figure 4.9C]. In case of full length *M. tuberculosis* PDT, a blue shift in the fluorescence spectrum along with quenching of fluorescence was seen [Figure 4.9A], while in case of PDT-C only quenching of fluorescence was seen [Figure 4.9B]. A blue shift in fluorescence spectrum and quenching indicates that the tryptophan residues of the native protein are buried inside as a consequence of the conformational change that results from the binding of phenylalanine.

4.4.7 Size exclusion chromatography reveals that *M. tuberculosis* PDT is an oligomer and PDT-N and PDT-C are monomers

Size exclusion chromatography was performed to determine the oligomeric state of *M. tuberculosis* PDT and its individual catalytic and regulatory domains. As the individual domains were catalytically inactive, this experiment was supposed to give insights into the possible reasons behind the catalytic insufficiency of the individual domain. Size exclusion was performed on a Superdex 200 column and a standard curve was drawn with standard protein MW markers. All the three proteins eluted as a single peak [Figure 4.10]. The molecular weight of all the proteins were calculated using a standard curve. The results showed that *M. tuberculosis* PDT is a hexameric protein and its individual catalytic and regulatory domains elute as monomers.

4.5 DISCUSSION

PDT exists both as a monofunctional protein and as a fusion protein with chorismate mutase. In *E. coli*, both CM and PDT activities reside in a single, bifunctional protein known as the P-protein, encoded by *phe*A [Pittard, 1987]. Amongst the monofunctional PDTs, almost all possess an ACT or amino acid binding domain that brings about feedback regulation activity of the enzyme.

In this work, *M. tuberculosis* PDT, an apparently monofunctional protein, was proved to

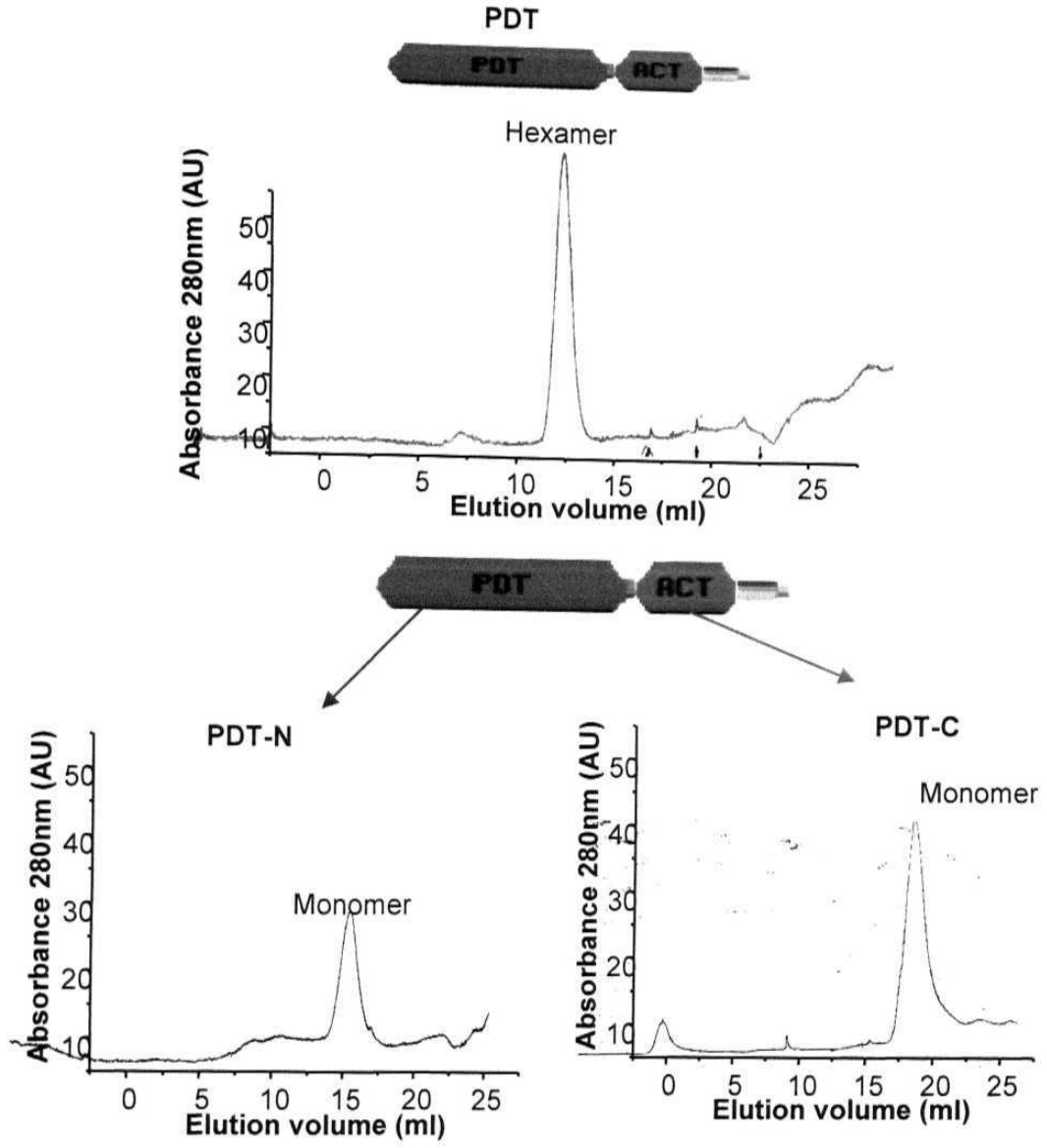


Figure 4.10: Size exclusion chromatography reveals that *M. tuberculosis* prephenate dehydratase (PDT) is a hexameric protein and its individually cloned catalytic (PDT-N) and regulatory domains (PDT-C) exist as monomers. The recombinant proteins were individually loaded on a Superdex 200 column (Pharmacia Biotech) and absorbance at 280nm (AU) was plotted as a function of elution volume. The elution parameter, Kav vs log MW plot of protein standards was used to calculate the actual MW of these proteins.

be so using a biochemical approach. The specific activity of full length recombinant PDT was very high [270 units/mg pure protein] and the enzyme was extensively feedback activated by all the three aromatic amino acids [Phe, Tyr and Trp]. Earlier reports have shown that *M. tuberculosis phe*A is an IdeR regulated gene [Gold *et al.*, 2001; Rodriguez *et al.*, 2003]. The present work demonstrates that *M. tuberculosis phe*A is also regulated at the protein level by aromatic amino acids. The high energy cost for the biosynthesis of aromatic amino acids [Atkinson, 1977] could be responsible for dual regulation [gene level as well as protein level] of *M. tuberculosis* PDT. However, dual regulation [specifically by IdeR] also suggests the involvement of the aromatic amino acid biosynthesis pathway in the biosynthesis of iron acquisition systems. Though phenylalanine as a precursor of salicylate *via* the phenyl ammonia lyase pathway is well documented in plants, this route of salicylate biosynthesis is not at all reported in bacteria [Wildermuth *et al.*, 2001]. It is therefore possible that *M. tuberculosis* could have an additional mechanism of salicylate biosynthesis that involves PDT. The dual regulation of *M. tuberculosis* PDT and specifically feedback activation by aromatic amino acids differentiates it form all other bacteria.

When the regulatory properties of *M. tuberculosis* CM and *M. tuberculosis* PDT were compared, the first observation was that both the enzymes display an exactly opposite pattern of regulation [Chapter 3, section 3.4.3]. While *M. tuberculosis* CM shows moderate activation by low concentrations of aromatic amino acids, activity was greatly inhibited at higher concentrations. The reverse holds true for *M. tuberculosis* PDT i.e. the enzyme is inhibited by low concentrations of aromatic amino acids and highly activated by higher concentrations. These results demonstrate for the first time that despite the occurrence of CM and PDT as monofunctional proteins in *M. tuberculosis*, correct balance of aromatic amino acids is brought about by opposite regulation of these two enzymes at the protein level.

Another important goal of the present study was to ascertain if the predicted catalytic domain of *M. tuberculosis* PDT, that lies in the N-terminus is independent of the feedback regulatory domain that is part of the C-terminus. Results presented in this chapter indicate that a discrete regulatory [ACT] domain along with a catalytic domain is

required for an optimally active *M. tuberculosis* PDT. These observations were found to be unlike the studies carried out on *E. coli* P protein where discrete domains of the complex P-protein retain their original activity [Zhang *et al.*, 1998]. However, in case of *E. coli* T protein wherein though CM and PDH domains can be expressed independently as functional proteins, the efficiency of the enzymes in isolation is much reduced as compared to the entire fusion protein [Chen *et al.*, 2003]. In neither case however, is the activity completely abolished, when the PDH or PDT domains are taken separately.

Various possible explanations can be considered for the complete loss of PDT activity upon removal of the regulatory domain. The dissociation of the oligomeric state of the enzyme upon removal of the C- terminus could be one possible reason. It is also possible that in the individually expressed and purified N-terminus of *M. tuberculosis* PDT, the substrate binding site is mechanistically altered leading to a loss of activity.

While the catalytic domain of *M. tuberculosis* PDT in isolation was functionally inactive, at least the individually cloned C-terminus of *M. tuberculosis* PDT retained its regulatory properties. Proof of the same was obtained using fluorimetric assays. These experiments generate sufficient line of evidence to conclude that there is a definite conformational change in the regulatory domain of *M. tuberculosis* PDT in the presence of phenylalanine. On the other hand, fluorescence emission spectra of the catalytic domain of *M. tuberculosis* PDT remain unchanged even in the presence of phenylalanine. It is possible that the C-terminus of *M. tuberculosis* PDT may serve as a modular regulatory domain that can be used to impart regulatory properties to other proteins of interest. These proteins can in turn be used as biosensors.

In summary, this study is the first report of a prephenate dehydratase enzyme that shows feedback activation by aromatic amino acids and is absolutely dependent upon the catalytic as well as the regulatory domains for optimum enzyme activity. The various levels at which the regulation of aromatic amino acid biosynthesis is brought about in *Mycobacterium tuberculosis* suggest that expression of the corresponding genes are strictly dependent upon the

requirement of the bacterium. Additionally, considering the absence of a human homologue of PDT, the enzyme might serve as a novel target for the design of chemotherapeutic compounds.

CHAPTER 5

MOLECULAR DISSECTION OF THE FUNCTIONS OF trpE AND trpE2 OF Mycobacterium tuberculosis

5.1 ABSTRACT

Anthranilate synthase [AS] and Isochorismate synthase [ICS] are other two chorismate utilizing enzymes that share extensive sequence similarity [Serino *et al.*, 1995]. In the *M. tuberculosis* genome, a tryptophan operon has been annotated with a probable anthranilate synthase or *trp*E [Rv1609]. However, in the initial genome sequence of *M. tuberculosis*, another ORF [Rv2386c], a part of the *mbt* operon [mycobactin biosynthesis operon] was reported as the second *trp*E [Cole *et al.*, 1998] and annotated as *trp*E2. Later, due to its unique location and regulation by IdeR, Rv2386c was reannotated as an isochosmate synthase [*mbt*I] [Quadri *et al.*,1998; Dussurget *et al.*, 1996]. This part of the study was carried out with the objective of actual assessment of the functions of the *trp*E and *trp*E2/*mbt*I genes of *M. tuberculosis*. The results presented in this chapter provide evidence that TrpE of *M. tuberculosis* [Rv1609] is an ammonia dependent anthranilate synthase and TrpE2/MbtI [Rv2386c] is a chorismate utilizing enzyme that catalyzes the conversion of chorismate to isochorismate. Furthermore, *trp*E as well as *trp*E2 could complement a *trp*E mutant *E. coli* strain indicating that Rv2386c possesses AS as well as ICS activities. These results have a bearing not only on the understanding of the role of these chorismate utilizing enzymes in the pathogenesis of *M. tuberculosis* but also on the behaviour and activity profile of divergently as well as convergently evolved group of enzymes.

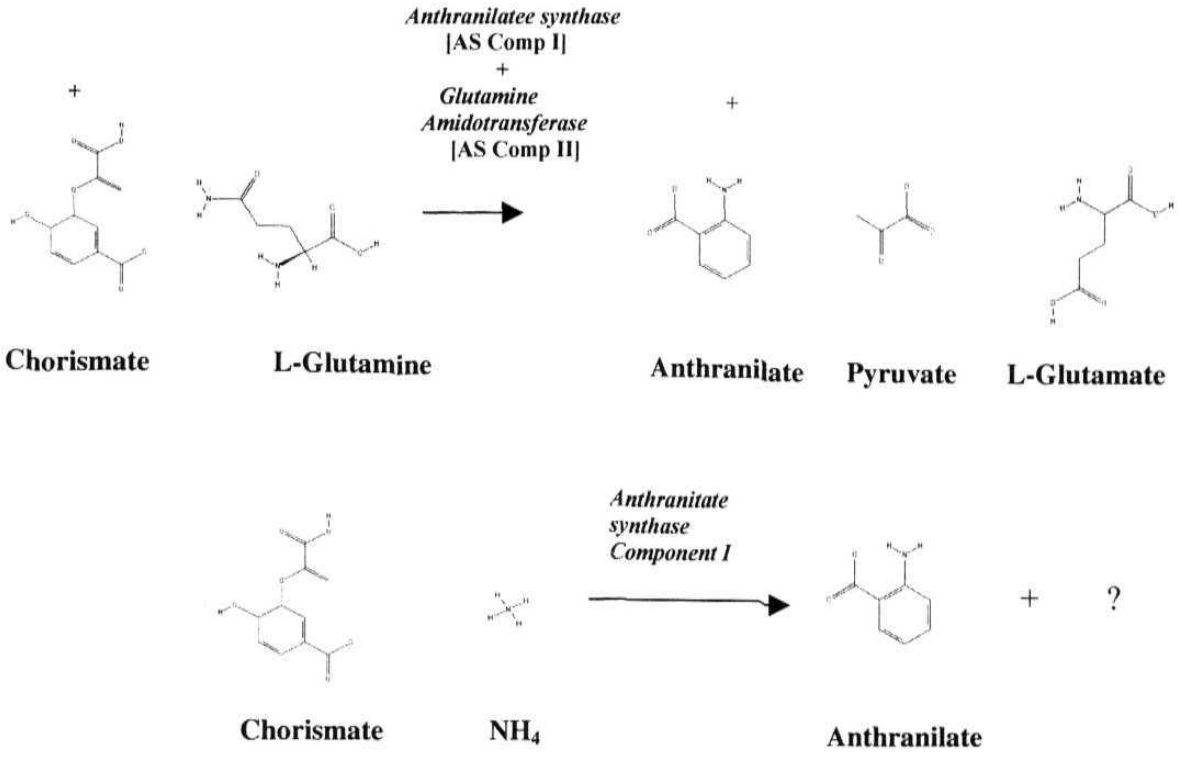
5.2 INTRODUCTION

In a continued effort to gain an insight into the physiological roles of the aromatic amino acid biosynthesis enzymes of *M. tuberculosis*, the study was extended to functional characterization of additional two chorismate utilizing enzymes, anthranilate synthase and isochorismate synthase of the bacterium. While the ORF Rv1609 is annotated as an anthranilate synthase [*trp*E], Rv2386c was earlier annotated as an anthranilate synthase [*trp*E2] and later reannotated as an isochorismate synthase [*mbt*I]. *trp*E2/*mbt*I has also been reported to be IdeR regulated [Gold *et al.*, 2001; Rodriguez *et al.*, 2003]. The proceeding sections provide a brief description of the reaction catalyzed by anthranilate synthase and isochorismate synthase.

5.2.1 Anthranilate synthase

Anthranilate synthase [AS] [EC 4.1.3.27] is a branch-point enzyme in the pathway for biosynthesis of aromatic amino acids that catalyzes the amination of chorismate to anthranilate. AS from various organisms has been purified and its catalytic mechanism and feedback regulation has been worked out [Caligiuri and Bauerle, 1991; Graf *et al.*, 1993].

The AS holoenzyme usually consists of two non-identical proteins and can use either glutamine or ammonia as the source of the amino group. The individual subunits are commonly referred to as Anthranilate synthase component I [α subunit] and Anthranilate Synthase Component II [β subunit]. Both the components are required for the Gln-dependent formation of anthranilate. Component II catalyzes the transfer of the amine group from Gln to component I, which in turn catalyzes the amination of chorismate [Ito and Yanofsky, 1966, 1969]. Component I alone can catalyze the conversion of chorismate to anthranilate only in the presence of very high concentrations of ammonium ions. Ammonium as well as Gln-dependent anthranilate synthase reactions require magnesium as a co-factor. The respective reactions are shown below.



In some organisms, Anthranilate synthase [TrpE] enzyme is part of a multifunctional protein, while in others only component I [TrpE] is present.

The X-ray crystal structure of the cooperative anthranilate synthase heterotetramer from *Salmonella typhimurium* has been determined at 1.9 Å resolution with the allosteric inhibitor L-tryptophan bound to the regulatory site in the TrpE subunit [Morollo *et al.*, 2001]. This structure was found to be different from the unliganded, noncooperative anthranilate synthase heterotetramer from *Sulfolobus solfataricus* [Knochel *et al.*, 1999]. These authors observed that even though the functional dimer pairs are structurally similar, they have completely different quarternary structures. Tryptophan binding to the TrpE subunit of *Salmonella typhimurium* was found to order a loop that in turn stabilizes the inactive T state of the enzyme by restricting closure of the active site cleft. This observation was consistent with differences in the cooperative behavior of the enzymes. The structural model is important for the study of the structural differences between cooperative and noncooperative anthranilate synthase homologues.

5.2.2 Isochorismate synthase [ICS]

ICS is also a chorismate utilizing enzyme that catalyzes the conversion of chorismate to isochorismate. Isochorismate is the precursor for salicylate. The enzyme is also required for enterochelin and menaquinone [Vitamin K] biosynthesis in *E. coli* [Liu *et al.*, 1990; Kaiser *et al.*, 1990; Dahm *et al.*, 1998].

Salicylate is the precursor of the siderophore pyochelin in *Pseudomonas aeruginosa*, and salicylate as well as pyochelin display siderophore activity [Ankenbauer and Quan, 1994; Serino *et al.*, 1997]. Similarly, salicylate is a precursor of the mycobacterial siderophore, mycobactin [Quadri *et al.*, 1998]. Salicylate biosynthesis in *P. aeruginosa* has been well studied after the characterization of the genes *pch*A [isochorismate synthase] and *pch*B [isochorismate pyruvate lyase], the two consecutive genes required for the conversion of chorismate to salicylate [Serino *et al.*, 1995; Gaille *et al.*, 2003]. While Rv2386c of *M. tuberculosis* appears to be a *pch*A equivalent, a gene homologous to *pch*B has not been reported in *M. tuberculosis* that can catalyze the conversion of isochorismate to salicylate. The *pch*A gene of P. *aeruginosa* encodes a protein of 52 kDa with extensive similarity to the chorismate-utilizing enzymes isochorismate synthase, anthranilate synthase [component I] and p-aminobenzoate synthase [component I]. Conversely, the 11kDa protein encoded by *pch*B does not show significant similarity with other proteins. The *pch*B stop codon overlaps the presumed *pch*A start codon [Serino *et al.*, 1995; Serino *et al.*, 1997]. Expression of the *pch*A gene in *P. aeruginosa* appears to depend on the transcription and translation of the upstream *pch*B gene. The *pch*BA genes of *P. aeruginosa* were in fact the first salicylate biosynthesis genes to be reported [Gaille *et al.*, 2002, 2003]. When *pch*BA genes were overexpressed in an *Escherichia coli ent*C mutant [lacking isochorismate synthase activity], salicylate formation was observed [Gaille *et al.*, 2003]. However, when *pch*B was expressed alone, there was no evidence of salicylate biosynthesis. In contrast, an *ent*B mutant of *E. coli* blocked in the conversion of isochorismate to 2,3-dihydroxybenzoate formed salicylate when transformed with a *pch*B expression construct. In the same study, salicylate formation was demonstrated *in vitro* when

chorismate was incubated with a crude extract of *P. aeruginosa* containing overproduced PchA and PchB proteins. Salicylate and pyruvate were formed in equimolar amounts. Furthermore, salicylate-forming activity could be detected in extracts of a *P. aeruginosa* siderophore-negative mutant grown under iron limitation, but not with iron excess. These results were able to establish a pathway leading from chorismate to isochorismate and then to salicylate catalyzed.

The structure of *Sulfolobus* anthranilate synthase has been solved and has been referred to as a prototype for the structure for other chorismate binding enzymes including isochorismate synthase.

5.2.3 The tryptophan operon of *M. tuberculosis*

While the genes involved in tryptophan biosynthesis are organized as an operon in *E. coli*, a scattered distribution is observed in the case of *M. tuberculosis* [Figure 5.1]. However, studies done so far suggest that *M. tuberculosis* possesses a homologue of all the genes that participate in tryptophan biosynthesis, suggesting the essentiality of the pathway in survival of the bacterium inside the host.

While *trp*E, B, C, A are clustered in the *M. tuberculosis* genome, *trp*G, *trp*D and *trp*F are quite distantly located. In fact, the occurrence of *trp*F was not known in *M. tuberculosis* genome until a recent report by Barona-Gomez and Hodgson [2003], which demonstrated that *his*A [Rv1603] of *M. tuberculosis*, a part of the histidine biosynthesis operon of *M. tuberculosis*, is a bifunctional protein and can complement a *trp*F mutant of *E. coli*. However, the reverse did not hold true i.e. the *E coli his*A could satisfy only histidine auxotrophy and *trp*F could satisfy only tryptophan auxotrophy in the respective *S. coelicolor* mutant strain. *M. tuberculosis his*A is therefore unique in the sense that it is a bifunctional protein with a broad substrate specificity.

In the absence of a glutamine amidotransferase [TrpG] enzyme in *M. tuberculosis trp* operon, *M. tuberculosis* TrpE is predicted to be an ammonia dependent anthranilate synthase. However, the pH requirement of this reaction is very high and far surpasses

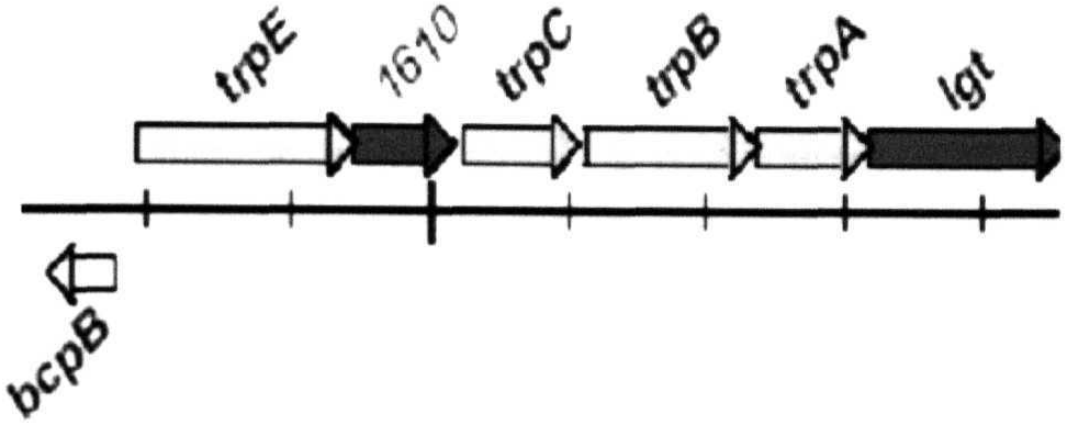


Figure 5.1: Organization of the tryptophan operon of *M. tuberculosis*. Members of this operon have been reported to be essential genes of the bacterium [*Sasseti et al, 2002*] *trp*D and *trp*G are not a part of this operon and are located elsewhere in the genome.

the physiological pH requirements. A distant location of TrpG emphasizes the need to explore the mechanism of co-regulation of these genes.

5.2.4 The *mbt* operon of *M. tuberculosis*

Iron assimilation is an important virulence attribute of *M. tuberculosis* and hence siderophore production is assumed to be associated with the virulence of the bacterium. Thus, it was obvious that the identification of genes involved in the iron sequestration pathway would lead to the identification of certain genes essential for survival of the bacterium inside the hostile environment of the host. The completion of the genome sequence of *Mycobacterium tuberculosis* has led to the identification of a gene cluster with homology to proteins of the non ribosomal peptide synthases and polyketide synthases. These genes were annotated as those possibly involved in siderophore biosynthesis [Cole *et al.*, 1998; Quadri *et al.*,1998]. This operon consists of ten genes [*mbt*A to *mbt*J], each of which have been assigned various functions. The individual roles of all these genes in the biosynthesis of siderophores has been briefly reviewed earlier [Crosa and Walsh, 2002; De Voss *et al.* 1999]. Essentially, there are seven genes [*mbt*A to *mbt*G] that appear to encode the desired modules required for the biosynthesis of the core of mycobactins. MbtA [Rv2384] activates salicylate as its acyl adenylate. The acyl adenylated salicylate is then covalently attached to a phosphopantetheine prosthetic group of MbtB. MbtB [Rv2383c] is an NRPS [non ribosomal peptide synthase] that is believed to activate serine and condense it with salicylate moiety and cyclize this product to a hydroxyphenyloxazoline. There are two other NRPSs encoded by *mbt*E [Rv2380c] and *mbt*F[Rv2379c] that have the appropriate activation, condensation and peptide carrier domains for donation of the two lysine-derived moieties of mycobactin. MbtF has a terminal domain that either functions as an epimerisation domain or a thioesterase responsible for releasing the mycobactin from the enzyme by lactamization of the terminal hydroxy lysine residue. Apart from these enzymes, the *mbt* gene cluster contains *mbt*C and *mbt*D genes, which encode proteins homologous to polyketide synthases. There are other genes in the operon [*mbt*H, *mbt*J] for which exact functions have not yet been determined.

Of all these proposed genes, only *mbt*A and *mbt*B have been experimentally shown to perform the function of acyl adenylation of salicylate and to encode the phosphopantetheinyl carrier domain, respectively [Quadri *et al.*, 1998]. Other genes require expression and characterization to solve the entire pathway of siderophore biosynthesis in *M. tuberculosis*.

Interestingly, the *mbt* operon is absent in *M. leprae*, which is a major setback to the pathogenicity of the bacterium [Cole *et al.*, 2001]. The importance of the salicylate derived siderophores in determining the pathogenicity of the bacterium was determined by generating *mbt*B knockouts of *M. tuberculosis* [De Voss *et al.* 2000]. The *mbt*B gene was replaced by a hygromycin resistant cassette through homologous recombination and the resulting bacterium was tested for production of siderophores. The assays showed that the delta *mbt*B:hyg strain of *M. tuberculosis* did not produce siderophores. The mutant bacterium was also tested for growth in macrophage like THP-1 cells under different concentrations of iron. It was observed that there was a rational decline in the growth of the mutant bacterium under iron limiting conditions. This laid the foundation for the belief that under iron limiting conditions, siderophores are essential for the survival of the bacterium.

Since salicylate derived siderophores have been shown to be required for the virulence of *M. tuberculosis*, the pathway for salicylate biosynthesis itself is vital for the bacterium.. Cole *et al.* [1998] have listed *ent*C and *ent*D genes as being possibly involved in the production of salicylate. These genes have similarity to enzymes involved in enterochelin biosynthesis pathway of *Pseudomonas aeruginosa*. However, Quadri *et al.* [1998] speculate that *mbt*I, which has similarity to anthranilate synthase, is responsible for the production of salicylate. Moreover, *ent*C and *ent*D do not contain any IdeR [Iron-dependent Regulator] binding site in their upstream sequences. Hence it is likely that MbtI could be involved indirectly in salicylate biosynthesis. In one study where *M. smegmatis* auxotroph for salicylic acid was assessed [Adilakshmi *et al.*, 2000] it was found that salicylate may even function as a signal molecule for recognition of cellular iron stress.

This Chapter gives a brief account of the studies carried on *trp*E and *trp*E2/*mbt*I of *Mycobacterium tuberculosis*. A biochemical as well as a genetic approach was employed to elucidate the function of the two. An overlap of function was observed which suggests convergent evolution of AS and ICS enzyme activities. The implications of the same are discussed.

5.3 EXPERIMENTAL PROCEDURES

5.3.1 Bacterial strains and plasmids

Table 5.1 gives a brief description of all *E. coli* and *M. tuberculosis* strains used in this study. The recombinant plasmids constructed along with their sources are also detailed.

5.3.2 Media, chemicals, buffers, and Enzymes

Luria Broth growth media supplemented with 200 μg/ml of ampicillin was used for selection and amplification of plasmids in *E. coli* hosts.

The composition of minimal media used for screening *E. coli* amino acid auxotrophs is described as follows

Minimal A medium:

K_2HPO_4	-	10.5 gm
KH_2PO_4	-	4.5 gm
$[NH4]_2SO_4$	-	1.0 gm
Sodium Citrate $2H_2O$	-	0.5 gm

Make up to 1 litre with water

After autoclaving, 1 ml of 1 M $MgSO_4$ and 0.1 ml of 1% Thiamine were added. Glucose at 0.2% was added separately. Amino acids, when required, were added to a final concentration of 40 μg/ml. Minimal A agar contained 2% bactoagar in minimal A medium, made from a separately autoclaved 4% bactoagar stock in water.

Unless otherwise specified, all chemicals and solvents, of reagent or HPLC grade, were procured from Sigma Chemicals, USA.

5.3.3 Cloning, over expression and purification of recombinant proteins

The ORFs Rv1609 and Rv2386c were PCR amplified from *M. tuberculosis* H37Rv genomic DNA using specific primers to enable directional cloning in pET23a expression vector. A brief description of the primers used is given in Table 5.2. The resultant constructs were used to transform *E. coli* BL21 DE3 competent cells and expression and purification of the recombinant protein was carried out by a method recently described [Ghosh *et al*, 2004]. The method employs the use of the dipeptide glycylglycine to enhance the solubility of a recombinant protein. In this method, the dipeptide is added to the culture media and all subsequent procedures are as described in chapter 2, section 2.3.2.

5.3.4 Anthranilate synthase activity assay [Rv1609 and Rv2386c]

The ammonia dependent AS activity of ORFs, Rv1609 and Rv2386c was studied using fluorimetric assays for anthranilate formation [Bohlmann J, 1996]. The reaction mixture contained 12.5mM Tris, 1.25% [v/v] glycerol, 0.25mM DTT, 10mM MgCl2, 1mM chorismate, 100mM NH4Cl, pH8.5. Reaction was initiated by addition of recombinant enzyme [rRv1609/ rRv2386c] to the assay mix followed by incubation at 32^0C for 30 minutes. The reaction was terminated with 10µl of 5N H_3PO_4. Anthranilate was extracted into 3ml of ethyl acetate and estimated using a fluorimeter at an excitation wavelength of 338nm and an emission wavelength of 440nm.

Table 5.1: Bacterial Strains and Plasmids

Strains/plasmids	Description	Source / Reference
M. tuberculosis		Cole *et al.*, 1998
H37Rv	Laboratory Strain [Virulent]	
E. coli		
DH5α		
BL21 DE3		
E. coli BL21 *trp- tet*	A tryptophan auxotrophic derivative of *E. coli*	
E. coli BL21 *trp- rho- kan tet*	BL21 strain	This work
	An anthranilate auxotrophic derivative of *E. coli* BL21 strain	This work
Plasmids		
pET23a	Expression vector	Novagen
pET23a1609	pET23a derivative with *M. tuberculosis trp*E [Rv1609c] cloned in *NdeI*/*XhoI* sites	This work
pET23a2386	pET23a derivative with *M. tuberculosis trp*E2/*mbt*I [Rv2386c] cloned in *NdeI*/*XhoI* sites	This work

Table 5.2: Sequences of primers used to amplify *M. tuberculosis trp*E [Rv1609] and *trp*E2 [Rv2386c]

ORF/ Gene	FORWARD PRIMER	REVERSE PRIMER
Rv1609	***NdeI***	***XhoI***
	AT***CATATG*** GTGCACGCCGACCTCG	AT***CTCGAG***GCAGCCACTGCGGTTC
Rv2386c	***NdeI***	***XhoI***
	AT***CATATG*** GTGTCCGAGCTCAGCGT	AT***CTCGAG***CTGGCGTGCAACCAG

5.3.5 Isochorismate synthase activity assay [Rv2386c]

ICS [EC 5.4.99.6] activity was determined according to the method of Poulsen *et al.* [1991] with slight modifications. The incubation mixture [250 μl] contained 0.1 M Tris-HCl, pH 7.5, 2 mM chorismate, 10 mM $MgCl_2$, and 100 pico moles of pure recombinant enzyme. The reaction mix was incubated for 60 min at 30°C followed by termination with 62.5 μl of methanol:sec-butanol [1:1, v/v]. The samples were centrifuged and loaded on a Nucleosil C-18 column. Waters 717 plus HPLC system equipped with a photodiode array detector was used to monitor isochorismate formation. Assay for direct salicylate biosynthesis was carried out using fluorimetric procedures.

5.3.6 Analytical size exclusion chromatography

The oligomeric states of the recombinant proteins were determined using size exclusion chromatographic procedures. The procedure is briefed in section 3.3.10 of Chapter 3

5.3.7 Limited proteolysis

Limited proteolysis was carried out to ascertain the existence of ligand binding pockets on *M. tuberculosis* TrpE and TrpE2. Anthranilate, salicylate and tryptophan were taken as the ligands of choice. The procedure for limited proteolysis is described in Chapter 3, section 3.3.7. The gels were viewed by silver staining.

5.3.8 Silver staining of protein gels

Silver staining was carried out as per method described in section 3.3.8 of chapter 3.

5.3.9 Construction of an *E. coli* BL21 *trp*E mutant strain

E. coli BL 21 was taken as the parent strain for introduction of an amber mutation in *trp*E. This strain was transduced with a P1 phage carrying *trp*E amber mutation linked to a tetracycline resistant marker [70% linkage]. The resultant strain was a tryptophan auxotroph [designated *E. coli* BL21 *trp- tet*]. A second round of transduction was carried

out in this strain with another P1 phage carrying a mutation in *rho*, linked to kanamycin resistance marker [60% linkage]. The second transduction was carried out to relieve polarity introduced by an amber mutation in *trp*E and expression of the downstream genes [Harinarayan and Gowrishankar, 2003]. The resultant strain was an anthranilate auxotroph [designated *E. coli* BL21 *trp- rho-tet kan*].

5.3.9.1 Preparation of P1 phage lysates/ P1 transduction

These procedures are detailed in section 3.3.11.1 and 3.3.11.2 of chapter 3.

5.3.9.2 Screening for auxotrophic mutants of *E. coli*

The minimal media were supplemented with appropriate amino acids and compounds [tryptophan, anthranilate] as per the requirements of the experiment.

5.4 RESULTS

5.4.1 *M. tuberculosis* TrpE [Rv1609] as well as TrpE2/MbtI [Rv2386c] show *in vitro* ammonia-dependent anthranilate synthase activity

*trp*E as well as *trp*E2 of *Mycobacterium tuberculosis* expressed in *E. coli* were purified as recombinant proteins [Figure 5.2 A and B]. The ammonia dependent anthranilate synthase activity of recombinant proteins were assessed using fluorimetric assays. It was observed that *M. tuberculosis* TrpE as well as TrpE2 display an ammonia dependent anthranilate synthase activity. However, the AS activity of TrpE2 was found to be 4 fold lower than that of TrpE [Figure 5.3]. The glutamine dependent AS activity of *M. tuberculosis* TrpE/TrpE2 was also studied. However, no evidence for the same was observed.

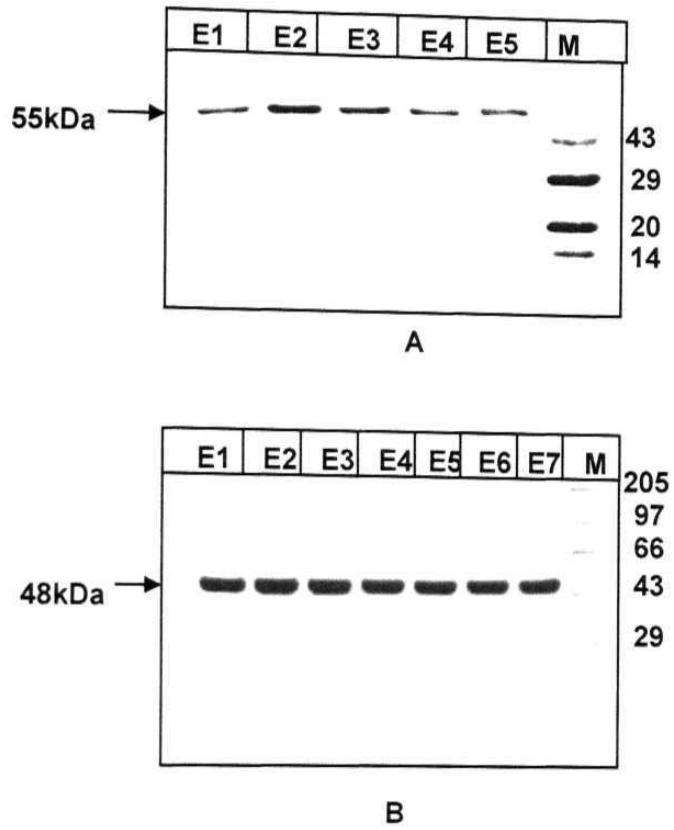


Figure 5.2: Purification of recombinant proteins corresponding to the ORFs Rv1609 (TrpE) (A) and Rv2386c (MbtI/TrpE2) (B) of *Mycobacterium tuberculosis*. Rv1609/Rv2386 were cloned in the *Nde*I and *Xho*I sites of pET23a vector with a C-terminal Histidine tag and expressed in *E. coli* BL21 cells. Affinity purification of recombinant protein was carried out using Talon resin (Clonetech, USA). The purified protein resolved on a 10% Tris-Tricine gel is shown. M represents the protein molecular size marker (Genei, India) and the E series represent the successive eluted fractions of the recombinant protein. Arrowheads indicate the position of the purified proteins.

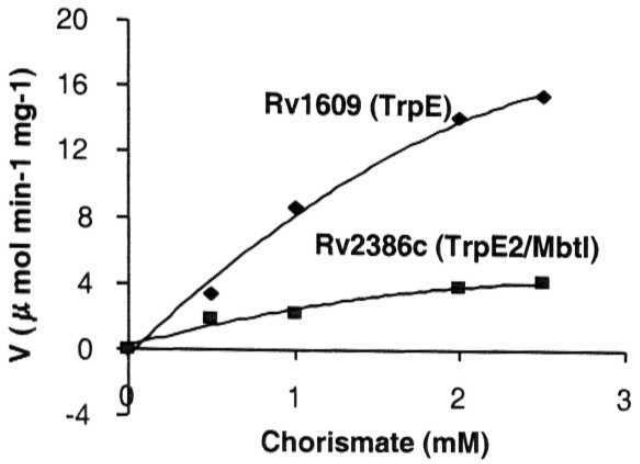


Figure 5.3: *In vitro* Anthranilate synthase (AS) activity of *M. tuberculosis* TrpE2/MbtI is 4-fold lower than that of TrpE. Ammonia dependent anthranilate synthase activity of *M. tuberculosis* TrpE (Rv1609) and TrpE2/MbtI (Rv2386c) were determined using fluorimetric assays. The above figure shows the substrate saturation plot of the two enzymes indicating that TrpE as well as TrpE2/MbtI possess AS activity. However, the AS activity of TrpE2/MbtI is ~4 fold lower than TrpE.

5.4.2 *M. tuberculosis* TrpE and TrpE2 are regulated proteins

Limited proteolysis was carried out to determine whether *M. tuberculosis* TrpE/TrpE2 are regulated by their end products. Tryptophan and anthranilate were used as the ligands of choice for TrpE and tryptophan and salicylate were used as ligands for TrpE2/MbtI [considering the possible involvement of TrpE2 in salicylate biosynthesis]. Tryptophan was found to protect TrpE [Figure 5.4B] as well as TrpE2 from trypsin cleavage [Figure 5.4 C]. Anthranilic acid also protected *M. tuberculosis* TrpE [Figure 5.4A] from proteolytic cleavage but salicylate could not do the same for *M. tuberculosis* TrpE2/ MbtI [Figure 5.4 D]. Phenylalanine binding to TrpE/TrpE2 were studied using fluorimetry. As there was no change in the fluorescence spectra of the proteins, it was concluded that these enzymes do not have phenylalanine binding pockets and are completely insensitive to the influence of cross pathway effector molecules [Figure 5.5: A, B].

5.4.3 *M. tuberculosis* ORFs Rv1609 [*trp*E] as well as Rv2386c [*mbtI*/*trp*E2] can complement an *E. coli trp*E mutant strain

Introduction of an amber mutation in *trp*E introduces polarity in the *trp* operon. Hence, as described in section 5.3.9, mutation in *rho* was introduced to relieve polarity. This strain was used as host for complementation tests with pET23a chimeric constructs, where *M. tuberculosis trp*E and *trp*E2 expression were driven from T7 promoter. It was observed that upon IPTG induction, both the constructs could enable *E. coli trp*E mutant strain to grow on minimal media [Figure 5.6]. This genetic complementation evidence demonstrated that Rv2386c has AS activity as well.

5.4.4 MbtI/TrpE2 [Rv2386c] shows *in vitro* isochorismate synthase activity

Taking into account the unique location of *M. tuberculosis* t*rp*E2/*mbt*I as part of the siderophore biosynthesis operon, and considering earlier reports by Quadri *et al.* [1998], it was decided to determine whether *M. tuberculosis* TrpE2 also shows isochorismate synthase activity. If indeed so, this would define the first step towards the biosynthesis of

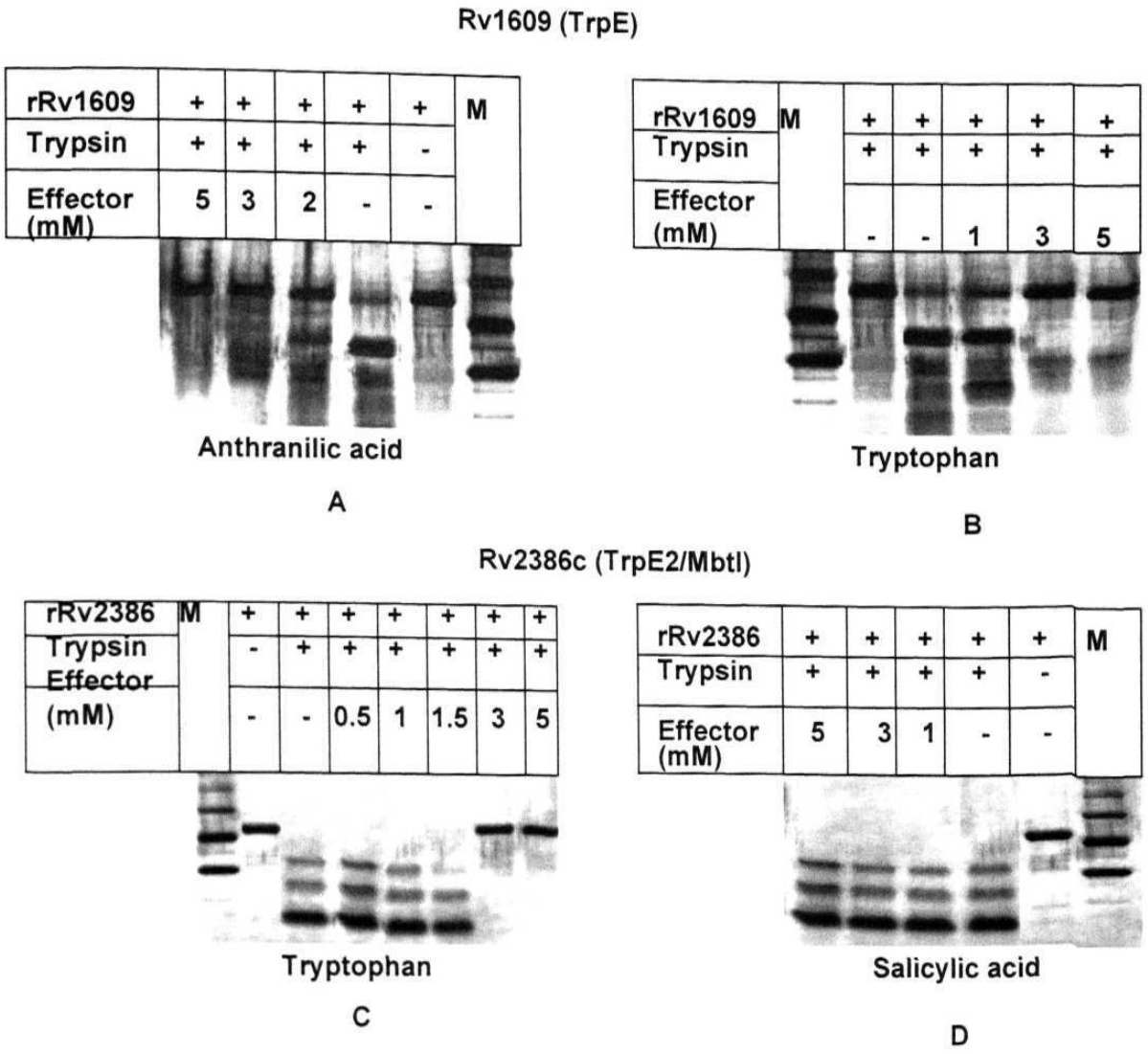


Figure 5.4: *M. tuberculosis* TrpE and TrpE2/MbtI are regulated at the protein level by effectors. The pathway specific effectors, Anthranilic acid as well as Tryptophan protect *M. tuberculosis* Anthranilate Synthase (TrpE/Rv1609) from proteolysis (A,B). *Mtb* TrpE2/MbtI is protected from proteolytic cleavage by tryptophan (C). Salicylic acid does not protect the protein from tryptic cleavage (D). M represents protein molecular size marker.

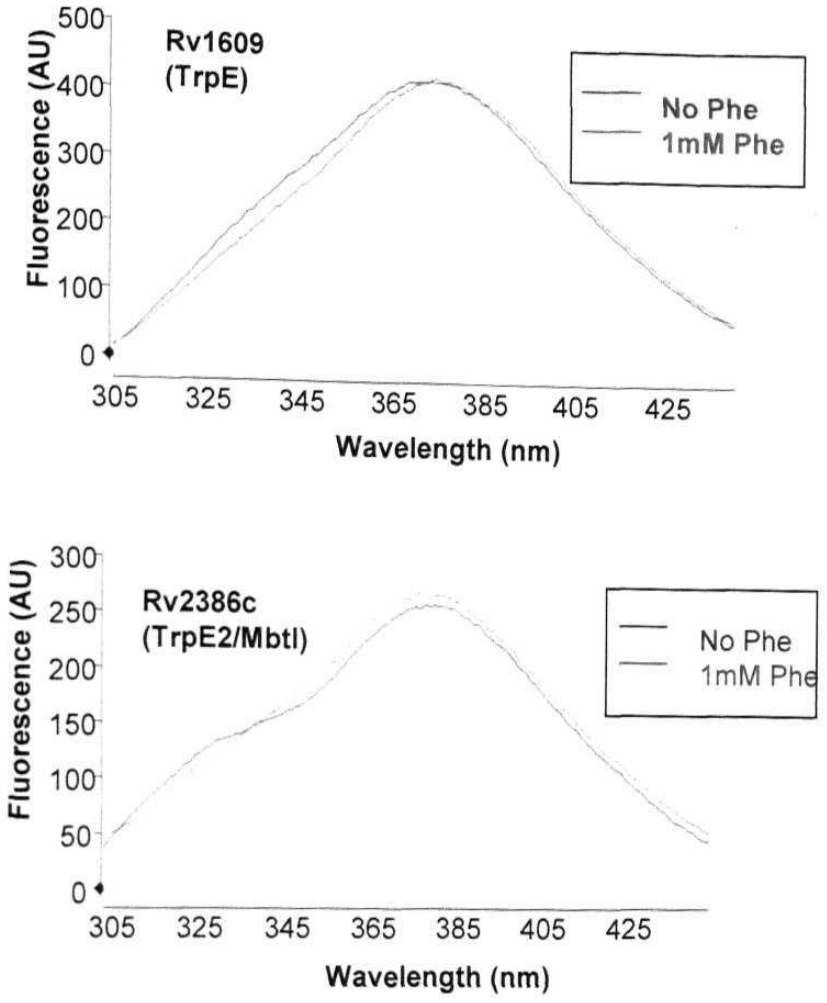


Figure 5.5: The cross pathway regulator phenylalanine does not change the fluorescence spectrum of *Mtb* TrpE (A) and TrpE2/MbtI (B). Phenylalanine does not have any modulatory effect on the activity of the enzyme.

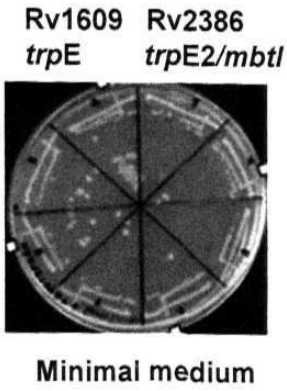


Figure 5.6: *M. tuberculosis trp*E (Rv1609) and *trp*E2/*mbt*I (Rv2386c) can complement an *E. coli* BL21 *trp*E mutant strain:

E. coli BL21 *trp*E mutant strain was transformed with pET23a constructs carrying *M. tuberculosis* genes Rv1609c (*trp*E) and Rv2386c (*trp*E2/*mbt*I) and its ability to grow on minimal mediium containing IPTG was tested. As evident from the figure above, Rv1609 (*trp*E) as well as Rv2386c (*trp*E2) could allow the *trp*E mutant *E. coli* BL21 strain (*E. coli trp- rho- kan tet)* to grow on minimal media.

salicylate, the precursor for salicylate. For this purpose, an HPLC assay was carried out to determine whether chorismate could be converted to isochorismate in the presence of rRv2386c. Results based on the retention time of isochorismate suggest that *M. tuberculosis* TrpE2 can convert chorismate to isochorismate [Figure 5.7]. The oligomeric state of *M. tuberculosis* TrpE2 was also determined using size exclusion chromatography. The native protein was found to be a dimer [Figure 5.8]. This observation was similar to other ASs which are also reported to be dimeric proteins. However, this is unlike the equivalent ICS enzyme PchA of *Pseudomonas aeruginosa* that exists as a monomer.

5.4.5 Coupled assays indicate that Rv2386c and Rv1885c are involved in the conversion of chorismate to salicylate

Earlier reports have documented the ability of isochorismate pyruvate lyase [IPL] enzyme of *P. aeruginosa* to catalyze a chorismate mutase [CM] reaction as well. It was accordingly decided to check if the CM of *M. tuberculosis* could execute an IPL reaction. Hence, both the enzymes were used in a coupled assay to monitor salicylate formation. While salicylate could not be directly formed by TrpE2/MbtI, detectable levels of salicylate was formed in a coupled assay involving recombinant TrpE2/MbtI and CM proteins [Figure 5.9]

5.4.6 TrpE2/MbtI [Rv2386c] does not show any evidence for direct salicylate biosynthesis

Recombinant TrpE2 was used in a fluorimetry based biochemical assay to assess if the enzyme could directly convert chorismate to salicylate. This experiment was done taking into account the absence of an isochorismate pyruvate lyase enzyme in *M. tuberculosis* genome, which is required for conversion of isochorismate to salicylate. It was observed that there is no evidence for direct salicylate biosynthesis by TrpE2 alone [Figure 5.9].

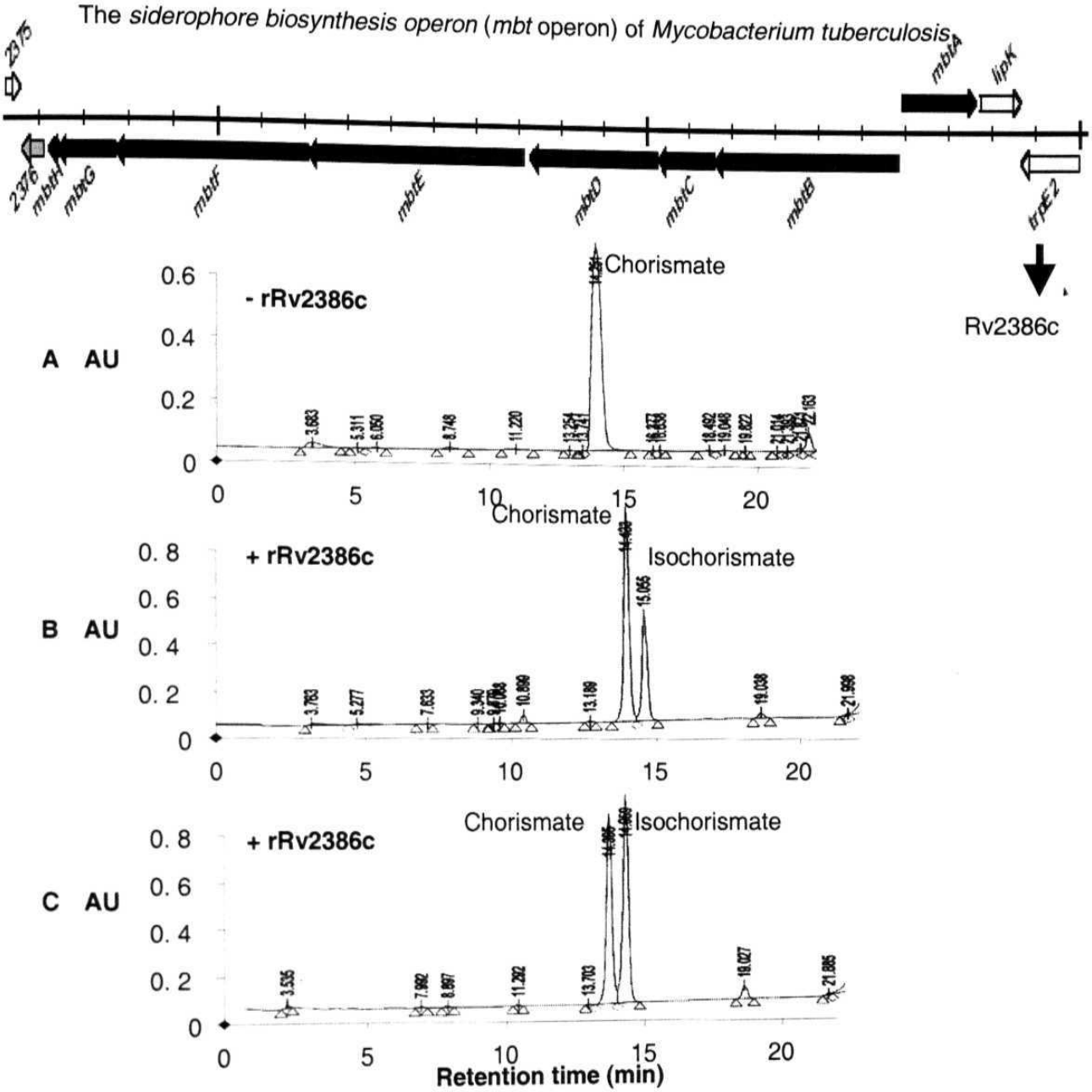


Figure 5.7: Recombinant Rv2386c (MbtI/TrpE2) can catalyze the conversion of chorismate to isochorismate. The reaction mixture containing 0.1M Tris-HCl, pH7.5, 2mM chorismate, 10mM $MgCl_2$ and rRv2386c was incubated at 30^{0}C for 60 minutes, terminated with the addition of 62.5µl pf MeOH: sec-BuOH (1:1 v/v), centrifuged and analysed by HPLC. As evident from the above figure, in the absence of rRv2386c, a single peak of chorismate was observed (A) and in the presence of rRv2386c, two peaks corresponding to chorismate as well as isochorismate could be seen (B and C, with increasing concentrations of chorismate). These results demonstrate the isochorismate synthase activity of *M. tuberculosis* MbtI/TrpE2

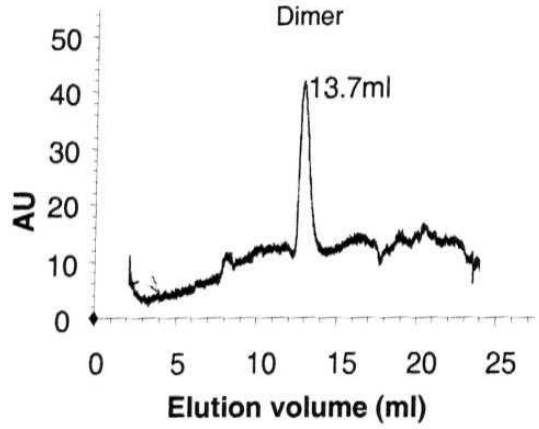


Figure 5.8: *M. tuberculosis* TrpE2/MbtI (Rv2386c) is a dimeric protein (two subunits of individual MW of 50kDa). rRv2386c was loaded on a Superdex 200 column (Pharmacia Biotech). Absorbance at 280nm (AU) was plotted as a function of elution volume. The elution parameter Kav vs log MW plot of standard proteins was used to determine the actual MW of TrpE2/MbtI .

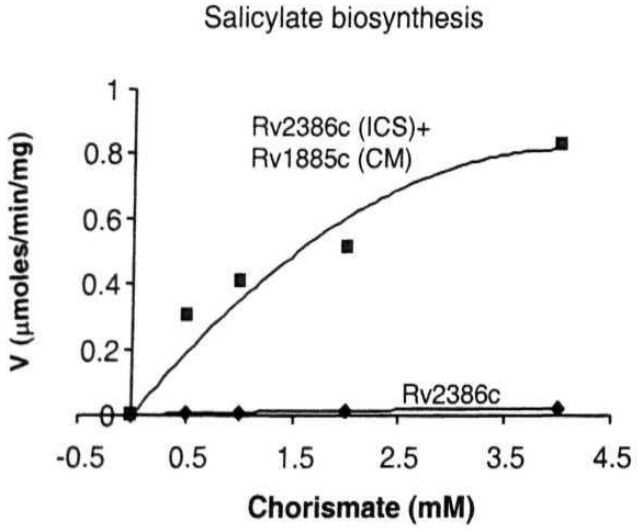


Figure 5.9: *M. tuberculosis* Chorismate mutase (CM) can substitute for an isochorismate pyruvate lyase (IPL) enzyme. Fluorimetric estimation of salicylate formation was carried out with *M. tuberculosis* MbtI/TrpE2 alone and in MbtI/TrpE2 and CM coupled reactions. While salicylate could not be formed by MbtI/TrpE2 (rRv2386c) alone, detectable amounts of salicylate were formed in a coupled assay involving both MbtI/TrpE2 and CM.

5.5 DISCUSSION

It has earlier been suggested that Anthranilate synthase [AS] and Isochorismate synthase [ICS] share a common reaction mechanism [He, 2004]. Extensive sequence similarity between AS and ICS enzymes of *M. tuberculosis* also suggest a commonality in function. *M. tuberculosis* ORFs 1609 and Rv2386c share about 40% aa sequence similarity amongst themselves as well as with the *E. coli* TrpE. Genetic and biochemical studies carried out on *M. tuberculosis* ORFs, Rv1609 and Rv2386 point to a convergent evolution of AS and ICS enzyme activities. A high sequence similarity of the two groups of enzymes is perhaps responsible for an overlap of function.

A glance at the biochemical profiles of TrpE and TrpE2 reveals that the AS activity of TrpE2 is 4 fold lower than TrpE. However, genetic evidence suggests that in spite of a weaker AS activity, even *trp*E2 could complement an *E. coli trp*E mutant strain. This observation could perhaps be attributed to the use of a strong promoter [T7], which in this case was sufficient for complementation purposes.

Chorismate can be converted to salicylate by the sequential action of the enzymes isochorismate synthase [ICS] and isochorismate pyruvate lyase [IPL] [Young *et al.*, 1969]. *M. tuberculosis* genome does not have an annotated IPL enzyme, and hence it has been suggested that Rv2386c [an ICS] could directly convert chorismate to salicylate [Gaille *et al.*, 2003]. However, results presented in this chapter indicate lack of evidence for direct salicylate biosynthesis by rRv2386c. It has been reported earlier that the isochorismate pyruvate lyase and chorismate mutase enzymes of *P. aeruginosa* are structurally and even functionally similar [Gaille *et al.*, 2002, 2003] as they possess a common chorismate binding domain. As *P. aruginosa* ICS also shows CM activity [Gaille C, 2002], the same analogy was used in *M. tuberculosis* to observe that chorismate mutase of *M. tuberculosis* could be a partner in salicylate biosynthesis. However, in *in vitro* coupled assays more than 2-3 μM salicylate could never be produced.

Considering the weak activity of *M. tuberculosis* CM as an IPL, it is still unclear whether *M. tuberculosis* CM participates in the conversion of chorismate to salicylate. The weak

IPL activity could be a mere relic of the broad substrate specificity of *M. tuberculosis* CM. However, in the absence of any *in vitro* evidence of the direct conversion of chorismate to salicylate by Rv2386c, it is still open for discussion whether *M. tuberculosis* depends upon the spontaneous conversion of isochorismate to salicylate or it depends on the weak IPL activity of CM [Young *et al.*,1969; Marshall and Ratledge, 1971, 1972]. As Rv2386 [mbtI/*trpE2*] even complements an *E. coli trp*E mutant strain, it also seems likely that *mbt*I could generate anthranilate, an amino analogue of salicylate which could get converted to salicylate by deamination.

In summary, the results presented in this chapter suggest that TrpE and TrpE2/MbtI *of M. tuberculosis* are interchangeable for their anthranilate synthase activity. Additionally TrpE2 can also execute the function of an isochorismate synthase. The evolutionary insight that could be drawn from the present findings is that despite the high sequence similarity at the amino acid level, these two chorismate utilizing enzymes can perform different functions. These results are consistent with the hypothesis that most ancestral enzymes have a broad substrate affinity which gradually get refined during the course of evolution to form enzymes with higher specific activities [Barona-Gomez and Hodgson, 2003, Xie *et al.*, 2003]. Additionally, the absence of a human homologue makes this protein an attractive drug target for the design of novel compounds against TB.

CHAPTER 6

Mycobacterium tuberculosis PPE/MPTR ORF, Rv2608 SHOWS A DIFFERENTIAL B CELL RESPONSE AND A LOW T CELL RESPONSE IN VARIOUS CATEGORIES OF TB PATIENTS

6.1 ABSTRACT

About 10% of the coding capacity of the *Mycobacterium tuberculosis* genome is contributed by two large, unrelated and functionally uncharacterized families of acidic, glycine rich proteins, the PE and the PPE families. The existence of polymorphic repetitive sequences, PGRS [Polymorphic GC Rich Sequences] and MPTRs [Major Polymorphic Tandem Repeats] in the C-terminal region of most of these genes suggest that these genes could be important in imparting antigenic variability to the bacterium. Earlier report from the present lab described a PPE family member Rv2430c as an immunodominant antigen [Choudhary *et al.*, 2003; 2004]. The present work describes the evaluation of a representative PPE_MPTR gene, Rv2608 for its ability to elicit humoral as well as well as T cell responses. Rv2608 was also found to be polymorphic in different clinical isolates as determined by PCR-RFLP analysis. 51 clinically confirmed TB patients belonging to fresh infection [n=22], relapsed infection [n=21], extrapulmonary infection [n=8], and 10 healthy controls were included in the study. The recombinant Rv2608 protein showed reactivity with sera of different category TB patients. Synthetic peptides corresponding to predicted regions of high antigenicity from the MPTR region of Rv2608 protein were used to assess cellular versus humoral immune response elicited by the protein. Interestingly, while the full length Rv2608 failed to show a differential humoral response as a function of patient infection category, the synthetic peptides elicited a predominantly humoral response in category II patients [relapsed TB]. The T cell response to the peptides as determined by T cell proliferation assays was low in all patient categories. The predominance of antibodies to this protein in relapsed infection cases and a relatively low cellular immune response indicate that this PPE member could be playing an important role in evading host immune responses by directing it towards a more humoral type that is not important to contain the infection. Additionally, for the first time the role of Gly-X-Gly-Asn-X-Gly repeat motifs present in a PPE family protein in eliciting a humoral immune response has been demonstrated.

6.2 INTRODUCTION

In the quest for identification of other attributes that would enable *M. tuberculosis* to establish a successful infection, the study was continued with the identification of novel antigens that would help the bacterium to evade host immune responses. In this context, it is significant to note that the only source of variation in the otherwise conserved genome of *M. tuberculosis* resides in the PE and the PPE gene families [Cole *et al.*, 1998]. The existence of PE/PPE gene families was evident even before the *Mycobacterium tuberculosis* genome was sequenced with occasional reports of occurrence of glycine and alanine rich multiple repetitive sequences in the genome [Poulet and Cole, 1995] or the identification of a few fibronectin binding proteins [Abou-Zeid *et al.*, 1991]. Sequencing categorized the PE/PPE gene families as two large unrelated families of highly acidic glycine rich proteins that constitute about 10% of the coding capacity of the genome [Cole *et al.*, 1998]. Comparative genome sequencing in various mycobacterial species revealed that by and large PE and PPE gene families are unique to *Mycobacterium tuberculosis* with few homologues in *M. leprae, M. bovis, M marinum etc* [Cole, 2002]. Amongst the *M. leprae* homologues, a major serine rich antigen is expressed in leprosy patients [Vega-Lopez *et al.*, 1993].

It is generally believed that the PE and PPE genes could be a source of antigenic variability. A recombinant PE_PGRS [Rv1759c] protein was shown to possess fibronectin binding properties and was also recognized by patient sera [Espitia *et al.*, 1999]. The same group also reported immense intra-strain variability in the PGRS domain with the N-terminal region staying constant. Transposon insertion in the PE_PGRS gene [Rv1818c] was shown to reduce macrophage infection ability of *M. tuberculosis* [Brennan *et al.*, 2001]. Surface localization of a PPE protein [Rv1917c] and many other PE_PGRS proteins has been reported [Sampson *et al.*, 2001; Banu *et al.*, 2002]. Few PE_PGRS genes have also been shown to be expressed during preclinical infection [Singh *et al.*, 2001]. Dissection of the PE_PGRS genes into PE and the PGRS domains to study their specific immunological response during mice infection revealed that the PE region can elicit an effective cellular immune response and the humoral response is largely directed against the Gly-Ala rich PGRS domain [Delogu *et al.*, 2001].

The involvement of PE/PPE genes in the virulence of the pathogen has also been reported [Ramakrishnan *et al.*, 2000]. Work done from the present laboratory has also shown that a PPE gene family member, Rv2430c is an immunodominant antigen of *M. tuberculosis* [Choudhary *et al.*, 2003, 2004].

The results presented in this chapter describe an *in-silico* approach to identify probable antigens from the PPE_MPTR [Major Polymorphic Tandem Repeat] subfamily. The humoral and cellular immune response to the same was studied using well characterized patient samples. Synthetic peptides corresponding to regions of high antigenic index of the protein were used to map the antigenic domains and assess the antigenic potential of the Gly-X-Gly-Asn-X-Gly repeat motif in eliciting a differential immune response. These results suggest that the PPE_MPTR ORF Rv2608 could be involved in directing the host towards development of a more humoral type of immune response.

6.3 EXPERIMENTAL PROCEDURES

6.3.1 PCR-RFLP analysis of the PPE ORF, Rv2608

PCR-RFLP was carried out to examine if Rv2608 exhibited polymorphism in different clinical isolates of *M. tubercuolsis*. Briefly, Rv2608 was PCR amplified from about 30 different clinical isolates and the amplified product was digested with *Sau*3AI enzyme. The digested product was separated on a 10% polyacrylamide gel and visualized under UV after ethidium bromide staining.

6.3.2 Cloning, overexpression and purification of Rv2608, a PPE MPTR subfamily member of *M. tuberculosis*

The PPE ORF, Rv2608 with predicted *in-silico* regions of high antigenic index was selected for cloning, expression and immunological characterization of the recombinant protein. The ORF was amplified from H37Rv genomic DNA using primers having specific restriction enzyme sites to enable directional cloning [Forward primer: AATGGATCCATGAATTTCGCCGTTTTG with a *Bam*HI overhang and Reverse primer:

AATAAGCTTGAAAAGTCGGGGTAGCGC with a *Hind*III overhang]. The amplified gene was first cloned in pGEMT easy vector followed by subcloning in pRSETa expression vector. The positive clones were used to transform *E. coli* BL21 cells. Protein expression was studied by initially growing the BL21 cells, transformed with the plasmid construct, overnight with appropriate antibiotics. 1% inoculum was taken from this primary culture from which a 5ml secondary culture was set up. When the absorbance value reached about 0.4, cells were induced to produce the recombinant protein by the addition of 1mM IPTG. 100µl of the culture was taken after every one hour to study the expression kinetics of the protein. A separate aliquot of uninduced culture was kept as a control. The cells were suspended in 1X SDS PAGE sample buffer and analyzed on a 12% SDS polyacrylamide gel to check for the expression of the 59.6kDa recombinant protein. The recombinant protein was then purified to homogeneity using the QIAExpressionist kit (Qiagen, USA). Cells harvested from 50 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]. 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 60kDa protein band was observed upon staining with Coomassie Brilliant Blue dye. The recombinant protein was tested for reactivity with patient sera after dialysis. Recombinant Hsp10 [Heat shock protein], a major secreted antigen of *M. tuberculosis* was used as control antigen.

6.3.3 Design of synthetic peptides

The PPE ORF, Rv2608 was scanned to identify regions of high antigenic index using the Protean software of Lasergene Navigator™ [DNA STAR]. Ten synthetic peptides of varying lengths corresponding to *in-silico* predicted regions of high antigenic index were commercially obtained as lyophilized powders. Peptide stocks of concentration 0.1 mg/ml were prepared in carbonate bicarbonate buffer and stored in aliquots at -70^0C.

6.3.4 Human patient sera

Fifty one TB patients confirmed by tuberculin skin test, radiographic examination and observation of Acid Fast Bacilli [AFB] in sputum for pulmonary TB and at the site of presumed TB in case of extrapulmonary infection were selected for this study. These patients were reporting to the Out Patient Department of the Mahavir Hospital and Research Centre Hyderabad, India. All the patients with confirmed diagnosis of TB were culture positive as well. The patients were categorized as follows: Category I: Individuals [n=22] diagnosed for TB for the first time; Category II: Individuals [n=21] with a relapsed TB and Category III: Extrapulmonary TB patients [n=8]. Sera were collected from all the subjects during early stage of infection when chemotherapy had just started. Healthy control [n=10] sera were taken from the laboratory staff of CDFD. These were individuals who did not have a prolonged direct contact with a TB patient. As this study was carried on a PPE gene family member of *M. tuberculosis*, members of which are unique to mycobacteria [3], cross reactivity to this protein would not be expected and therefore control subjects with other bacterial infection were not considered necessary for inclusion in our study.

6.3.5 Enzyme Linked Immunosorbent Assay [ELISA]

Serum samples from infected as well as control subjects were used to screen the recombinant protein and synthetic peptides using indirect ELISA technique. 96 well polystyrene microtiter plates [Costar] were incubated overnight with 2μg/ml of the synthetic peptides and 5 tuberculin units/ml of *M. tuberculosis* PPD. Carbonate bicarbonate buffer was used as the antigen coating buffer and each peptide was coated in triplicate. The plates were washed with PBS Tween [0.05% Tween 20 in PBS] and blocked by incubating with 250 μl of 2% BSA in PBS tween for one hour at 37^0C. Diluted Serum [1:200 in blocking buffer] was added to the antigen coated plate and the plates were incubated at 37^0C for another 60 minutes. After incubation, the plates were washed extensively with PBS tween followed by the addition of 100ul of secondary antibody [anti-Human IgG linked to the horse radish peroxidase enzyme]. The plates were

incubated again at 37^0C for one hour. After washing 6 to 8 times, the peroxidase activity was assessed by adding 100 µl of OPD [ortho phenylene diamine], the specific substrate of the enzyme dissolved in citrate phosphate buffer. OPD was added in dark and the plates were incubated at 37^0C for 30 minutes to allow colour formation. The reaction was stopped by the addition of 1N H_2SO_4. Absorbance was taken at 405nm.

6.3.6 Lymphocyte proliferation assay

Lymphocyte proliferation assays were carried out essentially as per method described earlier with a few modifications [Van de Loosdrecht *et al.*, 1994]. Heparinised blood was drawn and diluted with equal volume of RPMI1640 medium without serum. Diluted blood was layered on Ficoll gradient in 1:3 proportion. After a low speed [800g] centrifugation for 30 minutes, the peripheral blood mononuclear cells [PBMCs] were isolated and washed twice for 10 minutes at 800g to remove debris and platelets. Cell concentration was adjusted to 10^6/ml. Viability of the cells was checked using Trypan Blue. To each well of the microtiter plates, 0.1ml of cell suspension and 0.1 ml of antigen [2µg/ml] was added. ConA [Concanavalin A] was used as a positive control antigen. Control and experimental cultures were run in triplicate. The plate was incubated at 37^0C with 5% CO_2 for a period of 72 hours. At the end of the 3^{rd} day, 15µl of the tetrazolium salt MTT [2mg/ml] was added and incubated for another 4 hours. The culture was terminated and the MTT crystals were dissolved in 100 µl of acidified isopropanol. After one hour, the optical density was recorded using ELISA plate reader using a dual wavelength of 570 nm and 620 nm reference filter. Data were expressed as Stimulation Index [S.I.] i.e. ratio of the mean O.D. of experimental cultures [with test antigen] to the mean O.D. of control cultures [without antigen]. S.I. greater than or equal to 2 was considered as positive stimulation index.

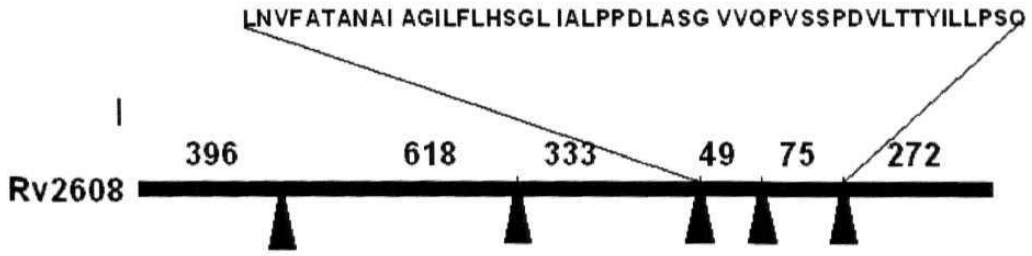
6.3.7 Statistical methods

Analysis of variance [ANOVA] as a test of statistical significance was performed using an online software [http://www.physics.csbsju.edu/stats/ANOVA.html] to calculate the p values and determine if there was any difference between different patient categories with respect to each antigen tested. The 95% confidence intervals for means were also determined for each set of data. Differences between groups were considered statistically significant if the 95% confidence interval limits did not overlap. To ascertain the results obtained by ANOVA, Kruskal Wallis non-parametric test [http://department.obg.cuhk.edu.hk/ResearchSupport/KruskallWallis.ASP] was also carried out. t tests were performed for paired comparison of means. $p<0.05$ was considered statistically significant.

6.4 RESULTS

6.4.1 Genetic variation in the PPE ORF, Rv2608

PE/PPE genes are predicted to be a source of antigenic variability of *M. tuberculosis* and polymorphism in a few of them based upon variation in the number of repeat sequences has already been reported [Espitia *et al.*, 1999]. A member of the PPE gene family, belonging to the major polymorphic tandem repeat [MPTR] subclass was analyzed by PCR-RFLP to score for the presence of genetic variation in different clinical isolates. The 1.7kb amplicon [Figure 6.1A] was digested with *Sau*3AI and the digest was electrophoresed on a 10% polyacrylamide gel. 16% of the clinical isolates showed a deviation from the normal band pattern. Figure 6.1 B gives the complete summary of the polymorphism obtained in 30 different clinical isolates. The disappearance of restriction fragments was restricted to the C terminus of the ORF, which is the predicted variable region of the PPE ORFs. It was therefore important to further evaluate Rv2608 in terms of its ability to elicit B and T cell response so as to study its role as a possible antigen for immune surveillance.



A

Location of *Sau*3AI sites	Expected bands (bp)	Additional bands (bp)	25 isolates	2 isolates	3 isolates
396	396		+	+	+
1014	618		+	+	+
1347	333		+	+	+
1396	49		+	+	-
1471	75		+	-	-
1753	272		+	+	+
		~120	-	+	+
		~100	-	+	+

B

Figure 6.1:A: *Sau*3AI Restriction map of PPE ORF, Rv2608. Arrowheads point to the *Sau*3A sites in the 1743bp Rv2608 ORF. Numbers above the line indicate the size of the restriction fragments (in base pairs) generated after *Sau*3AI digestion. B: Summary of *Sau*3AI PCR-RFLP pattern of 30 different clinical isolates of *Mycobacterium tuberculosis*.

6.4.2 Expression of Rv2608 in *E. coli* and purification of the recombinant protein

To evaluate the antigenic ability of Rv2608, the corresponding gene was expressed in *E. coli* BL21 cells and purified as a 6X His-tag fusion protein. Purified recombinant Rv2608 was fractionated by electrophoresis on a 12% polyacrylamide gel. A single band corresponding to 59.6kDa protein was observed upon staining the gel with Coomassie Brilliant Blue dye [Figure 6.2]. The expression of the gene was confirmed by probing the membrane containing the total cellular protein of *E. coli* BL21 cells harboring the Rv2608 construct with anti-Histidine antibody. There was no leaky expression of the protein in uninduced cells. The recombinant protein was largely present in the insoluble fraction and was therefore purified in the presence of 8M urea [Figure 6.2, LaneE]. The yield of the protein was 6mg/litre of culture. The recombinant protein was dialyzed overnight and used for immunoreactivity analysis.

6.4.3 Design of synthetic peptides based on antigenicity prediction of Rv2608

In-silico analysis of Rv2608 revealed the presence of two regions of high antigenicity: Two amino acid stretches [37 amino acids and 25 amino acids] corresponding to important antigenic epitopes within Rv2608 were selected for peptide synthesis [Figure 6.3 a]. Additional eight overlapping regions [Figure 6.3 b] which were essentially the subsets of the two main peptides were also selected for peptide synthesis. These peptides were used to map the antigenic domains of the protein. Table 6.1 shows the amino acid sequences of all the 10 synthetic peptides used in the present study. The peptides were part of the C terminal region of Rv2608 and apart from the high antigenic index also possessed the repeat motif Gly-X-Gly-Asn-X-Gly, characteristic of the PPE_MPTR gene family.

Table 6.1: Amino acid sequence of the synthetic peptides

PEPTIDE ID	AMINO ACID SEQUENCE*
P1	D**NIGN**A**NIGFGN**R**G**DA**NIGIGNIG**DR**NLGIGN**TG**N**WK [37]
P2	RP**G**LDELSFTLT**GNPN**RPD**GG**ILTK [25]
P1a	D**NIGN**A**NIGFGN**K [13]
P1b	**NIGFGN**R**G**DA**N**IK [13]
P1c	R**G**DA**NIGIGNIG**K [13]
P1d	**GIGNIG**DR**NLG**IK[13]
P1e	DR**NLGIGNTGN**WK [13]
P2a	RP**G**LDELSFTLTK [13]
P2b	LSFTLT**GNPN**RPK [13]
P2c	**GNPN**RPD**GG**ILTK [13]

*Residues in bold represent the Glycine-Asparagine repeat motifs

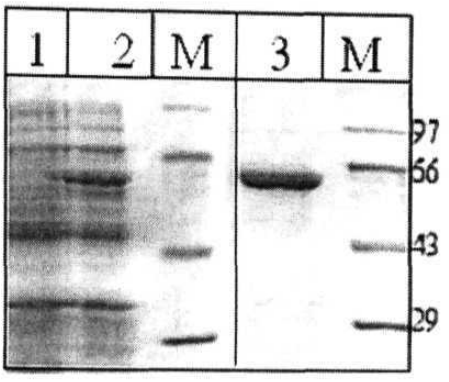


Figure 6.2: Expression and purification of *M. tuberculosis* protein corresponding to the PPE_MPTR ORF Rv2608. The left panel shows the uninduced and induced cell lysates and protein molecular size marker (Lanes 1, 2, M). The right panel shows the purified recombinant protein (Lane 3) and the protein molecular size marker (M).

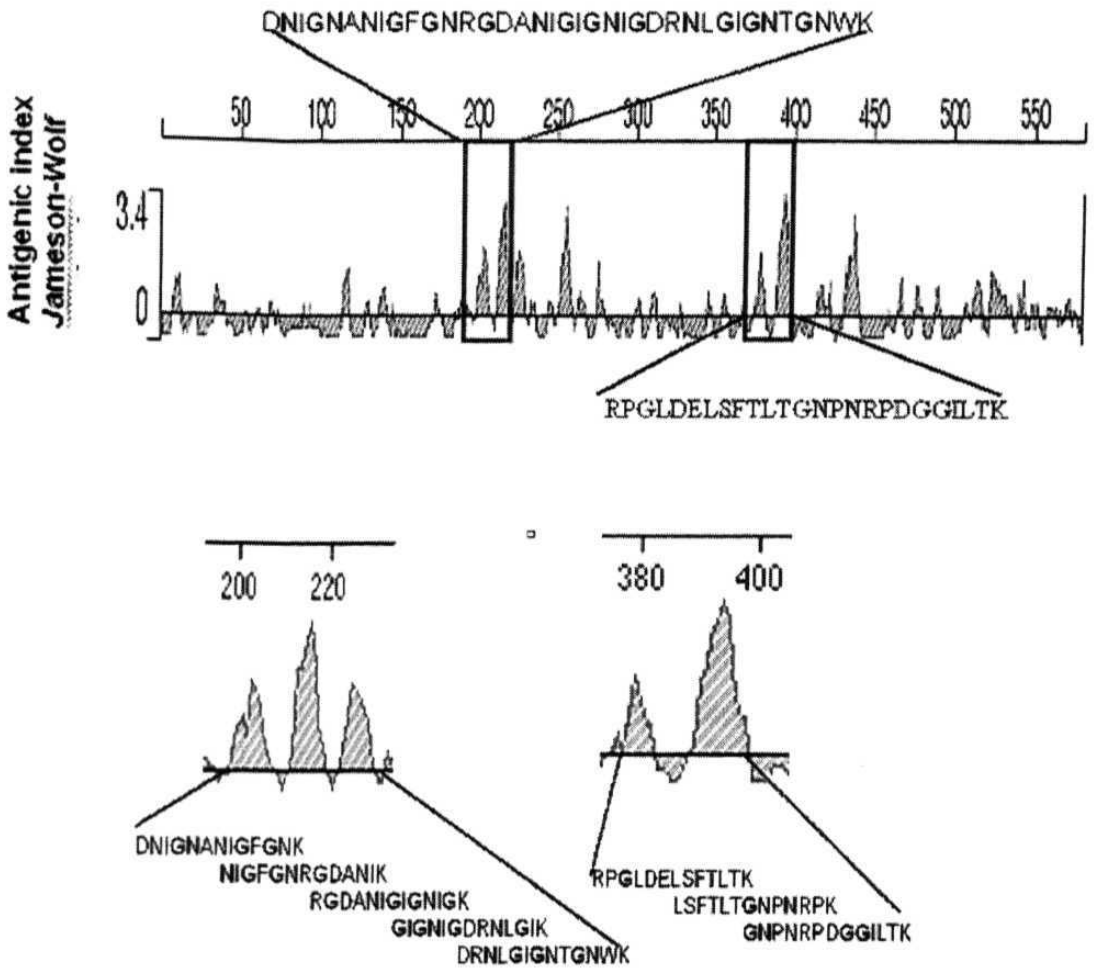


Figure 6.3:*In-silico* analysis of Rv2608 reveals regions of high antigenic index (potential antigenic determinants). Overall antigenic index of the protein was calculated using the Jameson Wolf method of the Protean software of Lasergene Navigator™. The boxed areas indicate the regions selected for designing synthetic peptides to map the region that was actually eliciting a variable immune response. As can be seen, one of the peptides (37mer) is largely composed of Gly-Asn repeats which is lesser in number in the other peptide (23mer). B: Stretches of overlapping peptides used for ELISA and T cell proliferation assay. These peptides were used to further map the region that was antigenic.

6.4.4 rRv2608 shows positive reactivity to sera from different categories of TB patients

The humoral response to the recombinant PPE protein was characterized by measuring serum IgG antibodies to the protein using ELISA. Antibody response was analyzed as a function of mean absorbance at 492nm. Recombinant Hsp10, a major secreted antigen of *M. tuberculosis* was used for comparison of the response to the rPPE protein. It was observed that for all the patient categories, serum reactivity to rRv2608 was equal to or higher than the response to Hsp10 [$p>0.05$, indicating no difference between the response to HSP10 and Rv2608] [Figure 6.4]. Healthy controls also showed some reactivity to the recombinant protein, however the response was significantly less when compared to that of patients [$p=0.0002$ using student's t test as a test of statistical significance for paired comparison of means between the patients and healthy controls].

6.4.5 Synthetic peptides corresponding to regions of high antigenicity within Rv2608 elicit strong humoral immune response in patients with relapsed TB infection:

Having shown that the recombinant protein coded by Rv2608 elicited an antibody response which was equal to or higher than that elicited by Hsp10 antigen in all the categories of TB patients selected for the study, an attempt was made to dissect differential responses if any as a function of patient category. For this, synthetic peptides spanning the two major antigenic regions within Rv2608 [P1 and P2] were used in ELISA [Table 6.1]. The results suggest that these peptides strongly react with patient sera [Figure 6.5] and hence the protein must be generating a strong humoral response in the host. Since a positive response was obtained with the peptides 1 and 2, patient sera were also tested for reactivity against the short overlapping peptide sequences 1a, 1b, 1c, 1d, 1e which were all components of peptide 1 and 2a, 2b and 2c which were a part of peptide 2. The results obtained indicate that even these overlapping peptide stretches react equally well with patient sera. Exact mapping of the antigenic region was not possible as most of the peptides showed a similar response. This was a reflection of

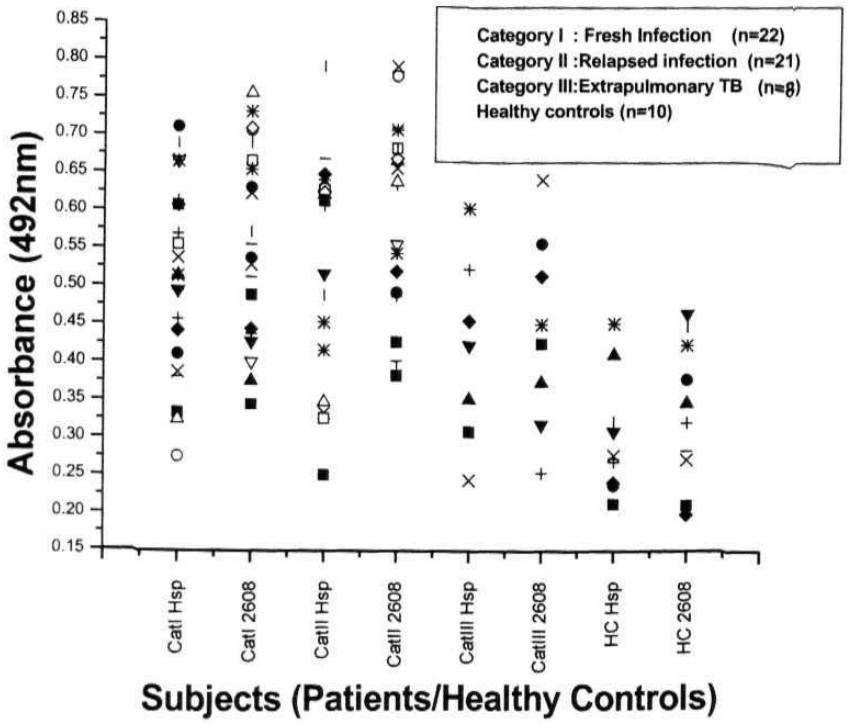


Figure 6.4:Antibody response of different categories of TB patients to rRv2608 is equivalent to the response to rHsp10, a well documented antigen of *M. tuberculosis*. Serum reactivity was measured by ELISA and the graph was plotted as patient response (O.D. at 492nm) to rHsp10 and rRv2608. The difference between patient response to Hsp10 and rRv2608 was not significant for all patient categories ($p>0.05$ using paired t tests). However, the response of healthy controls was lower and differed significantly from the patients ($p=0.0002$ using paired t test). (HC= Healthy Controls, Cat= Category)

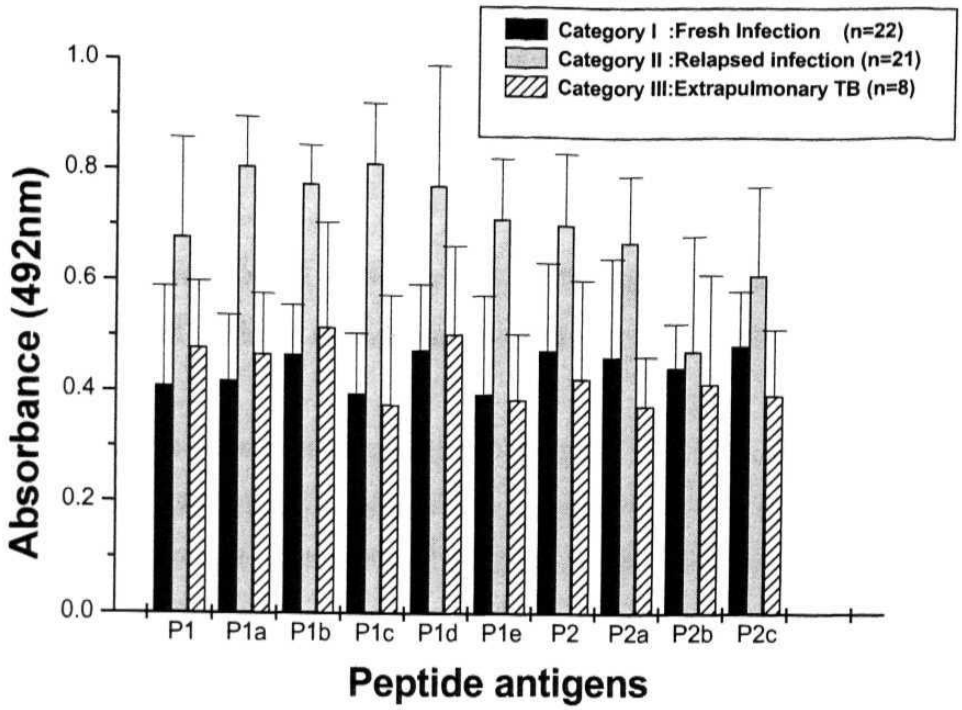


Figure 6.5:Antibody response of different categories of TB patients to different synthetic peptides (regions of high antigenicity within Rv2608) as determined by ELISA. Response to all the peptides was plotted as absorbance at 492nm (mean ± SD). The response of Category II patients was significantly higher than Category I or III ($p < 0.001$ for both, using paired t tests) with respect to each peptide antigen.

the fact that the Gly-X-Gly-Asn-X-Gly repeat motifs were present in all the peptides. Very interestingly, there was a significantly varied response to the peptides in different category of TB patients which was not so when the complete recombinant Rv2608 protein was used. The peptides could clearly distinguish between the patient categories [$p<0.001$ using ANOVA for each peptide antigen] [Table 6.2]. While humoral response observed in case of fresh infection cases [Category I] was similar to that of extrapulmonary TB patients [Category III], category II or the relapsed cases showed an unusually high antibody response to all the peptides. The response of Category II patients was significantly higher than Category I or III [$p< 0.001$ for both, using t test as a test of statistical significance].

6.4.6 The T cell response of TB patients to Rv2608 peptide antigens was low and the differences between various categories of patients were not evident

T cell proliferation assays were carried out to evaluate the response to different synthetic peptides. The overall T cell response of patients to these peptides was very low [S.I.<2] and the response could not distinguish between patient categories [$p>0.05$, using ANOVA and Kruskal Wallis test] at least for peptide 1 and its derivatives [Figure 6.6]. Peptide 2 and its derivatives exhibited a higher response in fresh infection cases as against relapsed and extrapulmonary cases [$p<0.05$ for both, using t test for paired comparison of means]. As can be noted from the amino acid sequence of the peptides [Table 6.1], peptide 2 has lesser number of glycine asparagine repeats and shows a higher T cell proliferative response in fresh infection cases.

6.5 DISCUSSION:

The ORF Rv2608 selected for the present study is a member of the PPE_MPTR class which is characterized by the presence of a conserved N-terminal region and a C-terminal domain with major polymorphic tandem repeats [MPTR] of Gly-X-Gly-Asn-X-Gly

Table 6.2: Summary of the results of statistical analyzes to estimate differences in humoral immune response to different peptide antigens

Peptide Antigen		Mean [O.D. at 492nm]	95% confidence interval of Mean	Degree of freedom	F value	P value	Difference between categories [Significant [S] / Not Significant [NS]
P1	I	0.412	0.338 to 0.486	2	12.69	<0.0001	S
	II	0.675	0.598 to 0.750				
	III	0.483	0.359 to 0.606				
P1a	I	0.426	0.360 to 0.492	2	29.69	<0.0001	S
	II	0.770	0.704 to 0.836				
	III	0.473	0.363 to 0.582				
P1b	I	0.469	0.416 to 0.521	2	35.47	<0.0001	S
	II	0.775	0.720 to 0.828				
	III	0.520	0.432 to 0.607				
P1c	I	0.416	0.364 to 0.467	2	70.48	<0.0001	S
	II	0.810	0.756 to 0.862				
	III	0.380	0.293 to 0.464				
P1d	I	0.482	0.427 to 0.535	2	33.47	<0.0001	S
	II	0.787	0.731 to 0.842				
	III	0.527	0.437 to 0.617				
P1e	I	0.407	0.348 to 0.464	2	31.26	<0.0001	S
	II	0.711	0.651 to.0.770				
	III	0.405	0.308 to 0.501				
P2	I	0.480	0.413 to 0.546	2	14.51	<0.0001	S
	II	0.706	0.637 to 0.774				
	III	0.441	0.331 to 0.551				
P2a	I	0.473	0.412 to 0.533	2	18.63	<0.0001	S
	II	0.689	0.627 to 0.750				
	III	0.392	0.291 to 0.491				
P2b	I	0.450	0.379 to 0.521	2	0.7688	0.4	NS
	II	0.498	0.425 to 0.570				
	III	0.422	0.304 to 0.540				
P2c	I	0.491	0.433 to 0.548	2	10.05	0.0002	S
	II	0.632	0.572 to 0.690				
	III	0.410	0.314 to 0.505				

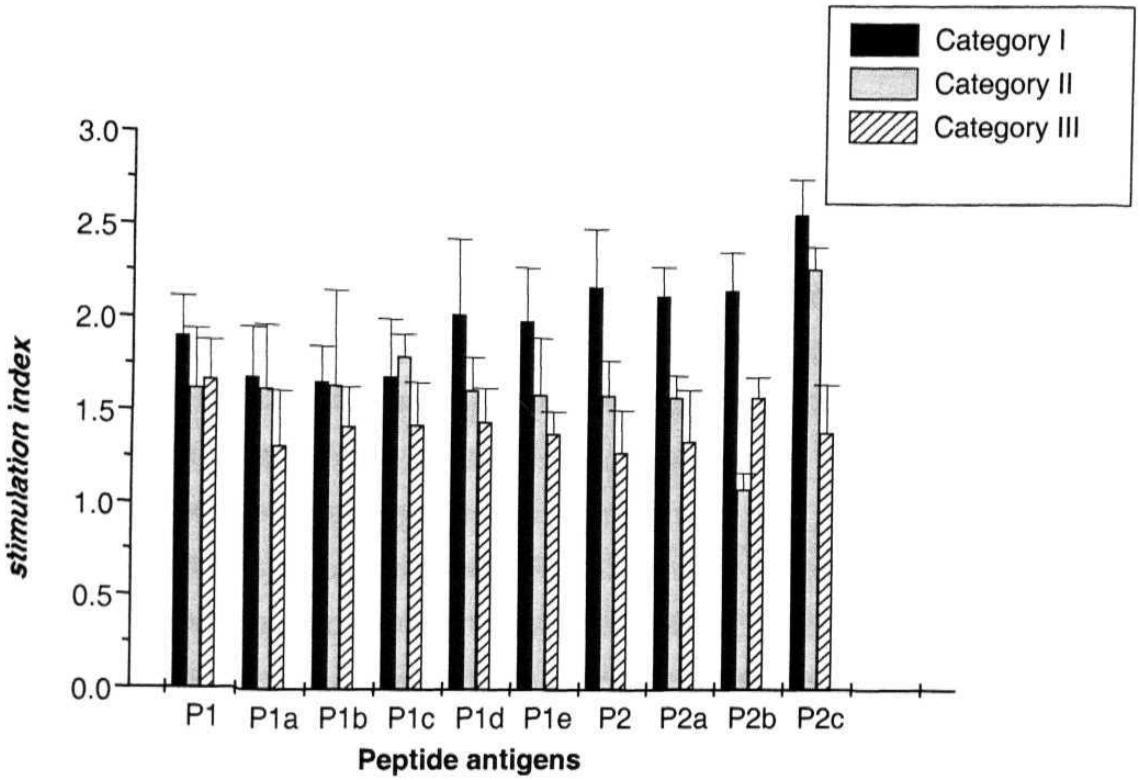


Figure6:The T cell proliferative response of different categories of TB patients as a result of stimulation by various synthetic peptides (regions of high antigenicity within Rv2608). Stimulation index(S.I.) = O.D. in test well (with antigen) /O.D. in control well (without antigen)> 2 was taken as a positive response, where test antigens are the peptides. The overall T cell response of TB patients to the peptide antigens was low. Only peptide 2 and its subsets show a positive S.I. Maximum S.I. is seen in Category I patients which progressively decreases in category II and extrapulmonary cases.

residues. Apart from this, the ORF also possesses regions of high antigenic index, which is a measure of overall hydrophilicity and surface probability. To test if polymorphism of the C-terminal region of this ORF exists in different clinical isolates of *M. tuberculosis*, PCR amplified Rv2608 was subjected to PCR-RFLP analysis. The observed variation in the band pattern lends weight to the hypothesis that PE/PPE genes, notably Rv2608 are perhaps a source of antigenic variability in the otherwise conserved genome of *M. tuberculosis*.

The rRv2608 protein was used in ELISAs to determine its reactivity to patient sera. The primordial observation that the recombinant protein reacted with patient sera indicates that this protein is definitely expressed during infection. Serum response of patients as well as healthy controls to rRv2608 was equivalent to or greater than the response to Hsp10, a well documented antigen of *M. tuberculosis* [Young and Garbe, 1991]. While category wise differentiation of serum reactivity towards the full length recombinant protein was not very apparent, it was significant to note that the extrapulmonary TB patients showed less reactivity with rRv2608 protein as compared to Category I or II [p=0.048]. It will be worthwhile to explore whether Rv2608 represents a protein[s] required by the bacterium to establish a pulmonary infection.

Since the serum response to the recombinant PPE protein was equal to or greater than Hsp10, it was decided to possibly map the antigenic domains of the probable PPE antigen using a synthetic peptide approach [Miguez *et al.*, 1998; Liljeqvist *et al.*, 1998; Benitez *et al.*, 1998]. Peptides corresponding to regions of high antigenic index were accordingly designed. The analyzes of the comparative humoral immune responses indicate that the serum response of patients to all the ten peptides is similar. This could be explained by the fact that all the peptides have a common repeat motif thereby eliciting similar response. While this negated the efforts to map the immunodominant epitope required for eliciting a strong humoral immune, a difference in the response of patients categorized according to different states of infection was surprisingly evident. Category II patients [relapsed infection cases] demonstrate the highest B cell response to the peptides followed by extrapulmonary TB cases.

The synthetic peptides were also used for T cell proliferation assays with the peripheral blood mononuclear cells of all category patients. It has been earlier shown that in about 90% of patients with active TB, there is a significant antibody response and/or T cell proliferative response to peptide specific single antigens of *M. tuberculosis* [Wilkinson *et al.*, 1998]. The 38kDa antigenic protein of *M tuberculosis* is a potent stimulus for both T cell and B cell responses in humans [Young *et al.*, 1986; Anderson *et al.*, 1989]. The T cell proliferative response to the synthetic peptides was of the order of fresh infection cases> relapsed TB > extrapulmonary TB cases at least for peptide 2 and its derivatives. However, the observed Stimulation Index [SI] with all the peptides was very low in all categories of TB patients [S.I.<2]. A high humoral response and a low cellular immune response to the peptides in category II patients points to an important possible function of the PE/PPE gene families. It is likely that these antigens play a role in evading host immune response and prevent the establishment of an effective cellular response, which is required to contain the disease. The positive T cell response in some cases could be explained by the fact that IgG antibody responses again require the involvement of helper T cells.

Antibody levels usually decrease in cured TB cases but dramatically increase in patients showing poor compliance [Bothamley *et al.*, 1993]. High antibody response to the peptides and a low T cell response hence explain the relapse of infection in category II patients. *In vivo*, it is possible that the responsive T cells are not able to expand as the glycine, asparagine repeat motifs somehow prevent antigen processing. The situation can be equated with the Epstein Barr Virus Nuclear antigen, where again the Gly-Ala repeat regions play an important role in preventing antigen processing [Levitskaya *et al.*, 1995]. Peptides 2 and 2c, which have lesser number of Gly-Asn repeats show a comparatively higher T cell response.

In conclusion, these results establish a relationship between immune responses to the PPE antigen and the status of the disease [fresh or relapsed TB]. The present study is the first report wherein, in a clinical setting, it has been demonstrated that the repeat sequences present within Rv2608 elicit a high humoral immune response and a low T cell response. Since PPE_MPTR is a gene family of *M. tuberculosis* of which Rv2608 is

a member sharing the MPTR motif, it is likely that other members of the same family may also serve the same function in the bacterium. These data contribute towards a better understanding of humoral as well as cellular immune responses elicited by PPE antigens. The practical utility of using these peptides for differentiating fresh infection from relapsed or reactivation cases is another interesting proposition and deserves large scale validation.

SUMMARY

SUMMARY

The present study represents the first systematic effort to characterize the aromatic amino acid biosynthesis pathway enzymes of *Mycobacterium tuberculosis* and establish its relationship with iron metabolism of the bacterium. While earlier reports have demonstrated the essentiality of the pathway [Parish and Stoker, 2002; Sassetti *et al.*, 2003], this study has been able to surface the intricacies of the catalytic and regulatory mechanisms employed by few member enzymes of the pathway.

The work was initiated with the identification of IdeR regulated genes of *M. tuberculosis*, which resulted in the identification of previously known as well as unknown targets of IdeR. The as yet unknown targets of IdeR identified in this study include *fec*B, a member of the ferric dicitrate uptake system; Rv1404, a MarR equivalent transcription regulator and the urease operon. Regulation of the urease operon by IdeR was demonstrated by *in vitro* binding assays. The known targets of IdeR that drew attention include the aromatic amino acid biosynthesis enzymes. Subsequently, the study was continued with the functional characterization of a few aromatic amino acid biosynthesis enzymes that might help the bacterium to survive inside the nutrient deficient environment of the host and possibly play a role in the pathogenicity of the disease.

Chorismate mutase [CM], prephenate dehydratase [PDT], anthranilate synthase [AS] and isochorismate synthase [ICS] enzymes of *M. tuberculosis* were selected for a detailed characterization. The corresponding genes were individually cloned and expressed in *E. coli* and the recombinant proteins were purified using metal affinity chromatography. Chorismate mutase activity was assigned to a hypothetical protein corresponding to ORF Rv1885c of *M. tuberculosis*. *In vitro* assays showed that the enzyme followed Michealis Menton's kinetics and the activity was temperature and pH dependent. Other biochemical assays demonstrated that *M. tuberculosis* CM does not have any associated PDT [prephenate dehydratase] or PDH [prephenate dehydrogenase] activity, indicating that it is a monofunctional protein. It was also demonstrated that *M. tuberculosis* CM, in spite of being a small protein is a highly regulated enzyme that shows feedback inhibition by pathway specific [tyrosine and phenylalanine] as well as cross pathway specific [tryptophan] aromatic amino acids.

Additionally, it was also shown that *M. tuberculosis* CM is a dimeric, alpha helical protein and possesses a functional signal sequence at its N-terminus. These studies could generate sufficient line of evidence to place the enzyme coded by *M. tuberculosis* ORF Rv1885c in the aroQ class of periplasmic chorismate mutases.

M. tuberculosis genome was scanned for the presence of another copy of chorismate mutase, as redundancy of the enzyme has been reported in many bacterial systems. According to the COG [Cluster of Orthologous Groups] database, Rv0948c could be the second chorismate mutase enzyme of *M. tuberculosis*. Hence, the gene corresponding to Rv0948c was expressed in *E. coli* and the recombinant protein was purified. It was observed that rRv0948c also possessed chorismate mutase activity though the specific activity was much lower than that of Rv1885c. This appears to be the first report of the existence of two chorismate mutase enzymes in a bacterium, both of which are apparently non-fusion proteins, having differential enzymatic activity.

Prephenate, the end product of the reaction catalyzed by CM is the substrate for the enzyme prephenate dehydratase [PDT]. Studies carried out on *M.tuberculosis* PDT indicate regulation of the enzyme in a manner exactly opposite to that of *M. tuberculosis* CM. While aromatic amino acids could bring about inhibition of activity of *M. tuberculosis* CM, they caused extensive activation of *M. tuberculosis* PDT. The other unique property of *M. tuberculosis* PDT that this study could decipher was the absolute requirement of the catalytic as well as the regulatory domains of the protein for optimum enzyme activity. Fluorescence spectroscopy could demonstrate that phenylalanine binding induces a conformational change only upon the full-length protein and the individually cloned and purified regulatory domain. The fluorescence spectrum of the catalytic domain was insensitive to the addition of phenylalanine. It was also proved that *M. tuberculosis* PDT does not possess any associated CM activity indicating that like CM, PDT is also a monofunctional protein. These studies on PDT and CM have therefore provided a definite mechanism for the control of terminal branches of aromatic amino acid biosynthesis in *Mycobacterium tuberculosis*.

Anthranilate synthase and isochorismate synthase are other two aromatic amino acid biosynthesis enzymes that were characterized in the present study. While the ORF

Rv1609 of *M. tuberculosis* is annotated as an anthranilate synthase [*trp*E] [Cole *et al.*, 1998], Rv2386c was earlier annotated as ananthranilate synthase [*trp*E2] [Cole *et al*, 1998] and later re-annotated as an isochorismate synthase [*mbt*I] [Quadri *et al.*, 1998]. These two ORFs were expressed in *E. coli* BL21 cells and the corresponding recombinant proteins were purified. Functions of these genes were studied using biochemical assays as well as using a genetic approach. A tryptophan auxotrophic *E. coli* BL21 strain was constructed that specifically lacked AS activity. This strain could not grow on minimal media but could grow on minimal media supplemented with anthranilate. The anthranilate auxotrophy of this strain could be satisfied with pET23a chimeric constructs carrying *M. tuberculosis* ORFs Rv1609 as well as Rv2386c. This experiment suggested that while Rv1609 [*trp*E] does code for an anthranilate synthase, Rv2386c also possesses AS activity. The ICS [Isochorismate synthase] activity of rRv2386c [*trp*E2/*mbt*I] was also tested using HPLC equipped with a diode array detector. It was observed that rRv2386c could convert chorismate to isochorismate, the precursor for salicylate. A direct proof for the involvement of *M. tuberculosis* TrpE2/MbtI [Rv2386c] and CM [Rv1885c] in salicylate biosynthesis was also provided.

In vitro ammonia-dependent AS activity assays for the protein coded by Rv1609c and Rv2386c were carried out using fluorescence spectroscopy. AS activity of Rv2386c was found to be four fold lower than Rv1609. It was also observed that Rv1609 [TrpE] as well as Rv2386c [TrpE2] could be protected from tryptic cleavage in the presence of tryptophan. Genetic and biochemical studies carried out on *M. tuberculosis* ORFs, Rv1609 and Rv2386c point to a convergent evolution of AS and ICS enzyme activities. The high sequence similarity of the two groups of enzymes is perhaps responsible for an overlap of function.

These studies suggests that the flux of chorismate into the biosynthesis of amino acids and secondary metabolites is regulated at both genetic and enzymatic levels in *M. tuberculosis*. While Rv2386c [*trp*E2/*mbt*I] and Rv3838c [*phe*A] have been reported to be under the regulatory control of IdeR [Gold B *et al.*, 2001], a 3 fold induction of *M. tuberculosis* ORF, Rv1885c has been reported in a *sig*E mutant *M. tuberculosis* strain [Manganelli *et al.*, 2001]. According to the present study, regulation of all these enzymes is also brought about by intracellular concentration of various ligands.

While assimilation of nutrients from the hostile environment of the host is an important virulence attribute of pathogenic microbes, evasion of the host immune response is another factor that determines the outcome of an infection. Off late, a lot of evidence has come up in support of the PE/PPE family proteins of *M. tuberculosis* as dominant antigens [Choudhary *et al.*, 2003]. The hypothesis whether the PPE family proteins of *M. tuberculosis* play a role in eliciting a variable host immune response was tested. A representative PPE_MPTR family ORF, Rv2608 was selected that possessed regions of high antigenic index. The corresponding recombinant protein showed positive reactivity to patient sera. Further, a synthetic peptide approach was used to map the antigenic domains of the protein. These peptides were used in ELISAs as well as lymphocyte proliferation assays. This study could identify the PPE_MPTR motif Gly-X-Gly-Asn-X-Gly as being responsible for eliciting a predominantly B cell response in patients suffering from relapsed tuberculosis [Chakhaiyar *et al.*, 2004]. This finding is important in the context of identification of novel antigens of *M. tuberculosis* that can be used as subunit vaccines.

In summary, the work presented in this thesis describes for the first time, the detailed characterization of few enzymes [Chorismate mutase, prephenate dehydratase, anthranilate synthase and Isochorismate synthase] of *M. tuberculosis* that participate in the biosynthesis of aromatic amino acids and/or iron acquisition systems. In light of essentiality of these genes and the absence of a human homologue, these enzymes could serve as novel and specific drug targets to check the growth of the tubercle bacillus. Additionally, the study also improves the current understanding of the antigenic proteins of *M. tuberculosis* that are used by the bacterium to establish a successful infection inside the host.

REFERENCES

REFERENCES

Abascal F, Valencia A. 2002. Clustering of proximal sequence space for the identification of protein families. *Bioinformatics*. **18**:908-21.

Abou-Zeid C, Garbe T, Lathigra R, Wiker HG, Harboe M, Rook GA, Young DB. 1991. Genetic and immunological analysis of *Mycobacterium tuberculosis* fibronectin-binding proteins. *Infect Immun*. **59**:2712-8.

Addadi L, Jaffe EK, and Knowles JR. 1983. Secondary Tritium Isotope Effects as Probes of the Enzymic and Non-enzymic Conversion of Chorismate to Prephenate, *Biochemistry*. **22**:4494-4501

Adilakshmi T, Ayling PD, Ratledge C. 2000. Mutational analysis of a role for salicylic acid in iron metabolism of *Mycobacterium smegmatis*. *J Bacteriol*. **182**:264-71.

Ahmad S, Jensen RA. 1988. The phylogenetic origin of the bifunctional tyrosine-pathway protein in the enteric lineage of bacteria. *Mol Biol Evol*. **5**:282-97.

Ahmed N, Alam M, Rao KR, Kauser F, Kumar NA, Qazi NN, Sangal V, Sharma VD, Das R, Katoch VM, Murthy KJ, Suneetha S, Sharma SK, Sechi LA, Gilman RH, Hasnain SE. 2004. Molecular genotyping of a large, multicentric collection of tubercle bacilli indicates geographical partitioning of strain variation and has implications for global epidemiology of *Mycobacterium tuberculosis*. *J Clin Microbiol*. **42**:3240-7

Ahmed N, Caviedes L, Alam M, Rao KR, Sangal V, Sheen P, Gilman RH, 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**:1712-6.

Ahmed N, Alam M, Abdul Majeed A, Asad Rahman S, Cataldi A, Cousins D, Hasnain SE. 2003. Genome sequence based, comparative analysis of the fluorescent amplified fragment length polymorphisms (FAFLP) of tubercle bacilli from seals provides molecular evidence for a new species within the *Mycobacterium tuberculosis* complex. *Infect Genet Evol*. **2**:193-9.

Alland D, Whittam TS, Murray MB, Cave MD, Hazbon MH, Dix K, Kokoris M, Duesterhoeft A, Eisen JA, Fraser CM, Fleischmann RD. 2003. Modeling bacterial evolution with comparative-genome-based marker systems: application to *Mycobacterium tuberculosis* evolution and pathogenesis. *J Bacteriol*. **185**:3392-9.

Andersen AB, Hansen EB.1989. Structure and mapping of antigenic domains of protein antigen b, a 38,000-molecular-weight protein of *Mycobacterium tuberculosis*. *Infect Immun*. **57**:2481-8.

Ankenbauer RG, Quan HN. 1994. FptA, the Fe(III)-pyochelin receptor of *Pseudomonas aeruginosa*: a phenolate siderophore receptor homologous to hydroxamate siderophore receptors. *J Bacteriol*. **176**:307-19

Atkinson, D. E. 1977. Cellular energy and metabolism and its regulation. *Academic Press, New York.*

Bøgg A, Neilands JB. 1987. Ferric uptake regulation protein acts as a repressor, employing iron (II) as a cofactor to bind the operator of an iron transport operon in *Escherichia coli*. *Biochemistry*. **26**:5471-7

Baker HM, Anderson BF, Baker EN. 2003. Dealing with iron: common structural principles in proteins that transport iron and heme. *Proc Natl Acad Sci USA*. **100**:3579-83.

Banerjee S, Nandyala A, Podili R, Katoch VM, Murthy KJR and Hasnain SE. 2004. *M.tb* isocitrate dehydrogenases display strong B-cell response and differentiate BCG-vaccinated healthy controls from TB patients. *Proc Natl Acad Sci USA*. **101**: 12652-7.

Bange FC, Brown AM, Jacobs WR Jr. 1996. Leucine auxotrophy restricts growth of *Mycobacterium bovis* BCG in macrophages. *Infect Immun*. **64**:1794-9.

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

Bao L, Chen W, Zhang H, Wang X. 2003. Virulence, immunogenicity, and protective efficacy of two recombinant *Mycobacterium bovis* bacillus Calmette-Guerin strains expressing the antigen ESAT-6 from *Mycobacterium tuberculosis*. *Infect Immun.* **71**:1656-61.

Barona-Gomez F, Hodgson DA. 2003. Occurrence of a putative ancient-like isomerase involved in histidine and tryptophan biosynthesis. *EMBO Rep.* **4**:296-300.

Bartlett, P.A. and Johnson, C.R. 1985. An Inhibitor of Chorismate Mutase Resembling the Transition State Conformation, *J. Am. Chem. Soc.* **107**:7792-7793.

Bateman,A., Birney,E., Durbin,R., Eddy,S.R., Howe,K.L. and Sonnhammer,E.L.L. 2000. The Pfam protein families database. *Nucleic Acids Res.* **28**: 263–266.

Bateman A, Coin L, Durbin R, Finn RD, Hollich V, Griffiths-Jones S, Khanna A, Marshall M, Moxon S, Sonnhammer EL, Studholme DJ, Yeats C, Eddy SR. 2004. The Pfam protein families database. *Nucleic Acids Res.* **1**:32. Database issue:D138-41

Bekal S, Niblack TL, Lambert KN. 2003. A chorismate mutase from the soybean cyst nematode Heterodera glycines shows polymorphisms that correlate with virulence. *Mol Plant Microbe Interact.* **16**:439-46.

Benitez J, Palenzuela D, Rivero J, Gavilondo JV. 1998. A recombinant protein based immunoassay for the combined detection of antibodies to HIV-1, HIV-2 and HTLV-I. *J Virol Methods.* **70**:85-9.

Bennett MJ, Schlunegger MP, And Eisenberg D. 1995. 3D Domain swapping: A mechanism for oligomer assembly. *Protein Sci* **4**: 2455-2468.

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

Betts JC, McLaren A, Lennon MG, Kelly FM, Lukey PT, Blakemore SJ, Duncan K. 2003. Signature gene expression profiles discriminate between isoniazid-, thiolactomycin-, and triclosan-treated *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother.* **47**:2903-13.

Bijlsma JJ, Waidner B, Vliet AH, Hughes NJ, Hag S, Bereswill S, Kelly DJ, Vandenbroucke-Grauls CM, Kist M, Kusters JG. 2002.The *Helicobacter pylori* homologue of the ferric uptake regulator is involved in acid resistance.*Infect Immun.* **70:**606-11.

Bohlmann J, Lins T, Martin W, Eilert U. 1996. Anthranilate synthase from Ruta graveolens. Duplicated AS alpha genes encode tryptophan-sensitive and tryptophan-insensitive isoenzymes specific to amino acid and alkaloid biosynthesis. *Plant Physiol.* **111**:507-14.

Bothamley GH, Schreuder GM, de Vries RR, Ivanyi J. 1993. Association of antibody responses to the 19-kDa antigen of *Mycobacterium tuberculosis* and the HLA-DQ locus. *J Infect Dis* 167:992-3.

Boyd J, Oza MN, Murphy JR 1990. Molecular cloning and DNA sequence analysis of a diphtheria tox iron-dependent regulatory element (dtxR) from *Corynebacterium diphtheriae. Proc Natl Acad Sci USA.* **87**:5968-72.

Brennan MJ, Delogu G, Chen Y, Bardarov S, Kriakov J, Alavi M, Jacobs WR Jr. 2001. Evidence that mycobacterial PE_PGRS proteins are cell surface constituents that influence interactions with other cells. *Infect Immun.* **69**:7326-33.

Brickman E, Beckwith J. 1975. Analysis of the regulation of *Escherichia coli* alkaline phosphatase synthesis using deletions and transducing phages. *J Mol Biol* . **96**:307-316.

Brosch R, Gordon SV, Billault A, Garnier T, Eiglmeier K, Soravito C, Barrell BG, Cole ST. 1998. Use of a *Mycobacterium tuberculosis* H37Rv bacterial artificial chromosome library for genome mapping, sequencing, and comparative genomics. *Infect Immun.* **66**:2221-9.

Brosch R, Gordon SV, Buchrieser C, Pym AS, Garnier T, Cole ST. 2000. Comparative genomics uncovers large tandem chromosomal duplications in *Mycobacterium bovis* BCG Pasteur. *Yeast.* **17**:111-23.

Byrd TF, Horwitz MA. 1993. Regulation of transferrin receptor expression and ferritin content in human mononuclear phagocytes. Coordinate upregulation by iron transferrin and downregulation by interferon gamma. *J Clin Invest.* **91**:969-76.

Calderwood SB, Mekalanos JJ. 1987.Iron regulation of Shiga-like toxin expression in *Escherichia coli* is mediated by the fur locus. *J Bacteriol.* **169**:4759-64

Calhoun DH, Bonner CA, Gu W, Xie G, Jensen RA. 2001.The emerging periplasm-localized subclass of AroQ chorismate mutases, exemplified by those from *Salmonella typhimurium* and *Pseudomonas aeruginosa*. *Genome Biol.* 2:RESEARCH0030. Epub 2001 Jul 27

Caligiuri MG, Bauerle R. 1991. Subunit communication in the anthranilate synthase complex from *Salmonella typhimurium*. *Science.* **252**:1845-8.

Caligiuri MG, Bauerle R. 1991. Identification of amino acid residues involved in feedback regulation of the anthranilate synthase complex from *Salmonella typhimurium*. Evidence for an amino-terminal regulatory site. *J Biol Chem.* **266**:8328-35

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

Camus JC, Pryor MJ, Medigue C, Cole ST. 2002. Re-annotation of the genome sequence of *Mycobacterium tuberculosis* H37Rv.*Microbiology.* **148**: 2967-73.

Carlson Heather A. and Jorgensen William L. 1996. Monte Carlo Investigations of Solvent Effects on the Chorismate to Prephenate Rearrangement *J. Am. Chem. Soc.* **118**:8475-8484.

Chakhaiyar P, Nagalakshmi Y, Aruna B, Murthy KJ, Katoch VM, Hasnain SE. 2004. Regions of High Antigenicity within the Hypothetical PPE Major Polymorphic Tandem Repeat Open-Reading Frame, Rv2608, Show a Differential Humoral Response and a Low T Cell Response in Various Categories of Patients with Tuberculosis. *J Infect Dis.* **190**:1237-44.

Chakhaiyar P and Hasnain SE. 2004. Defining the Mandate of Tuberculosis Research in a Postgenomic Era. *Medical Principles and Practice* **13**:177-184.

Chen S, Vincent S, Wilson DB, Ganem B. 2003. Mapping of chorismate mutase and prephenate dehydrogenase domains in the *Escherichia coli* T-protein. *Eur J Biochem.* **27**:757-63.

Chipman DM, Shaanan B. 2001. The ACT domain family. *Curr Opin Struct Biol* **11**: 694700

Chook YM, Ke H, Lipscomb WN. 1993. Crystal structures of the monofunctional chorismate mutase from *Bacillus subtilis* and its complex with a transition state analog. *Proc Natl Acad Sci USA.* **90**:8600-3.

Chook YM, Gray JV, Ke H, Lipscomb WN. 1994.The monofunctional chorismate mutase from *Bacillus subtilis*. Structure determination of chorismate mutase and its complexes with a transition state analog and prephenate, and implications for the mechanism of the enzymatic reaction. *J Mol Biol.* **240** :476-500.

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

Choudhary RK, Mukhopadhyay S, Chakhaiyar P, Sharma N, Murthy KJR, Katoch VM, and Hasnain SE. 2003. PPE antigen Rv2430c of *Mycobacterium tuberculosis* induces a strong B-cell response. *Infect Immun.* **71**:6338-43.

Clarke TE, Tari LW, Vogel HJ. Structural biology of bacterial iron uptake systems. 2001. *Curr Top Med Chem* ;**1**:7-30.

Clemens DL, Lee BY, Horwitz MA.1995. Purification, characterization, and genetic analysis of *Mycobacterium tuberculosis* urease, a potentially critical determinant of host-pathogen interaction. *J Bacteriol.* **177**:5644-52.

Cole ST *et al.* 1998. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature.* **393**:537-44

Cole ST *et al.* 2001. Massive gene decay in the leprosy bacillus. *Nature.* **409:**1007-11.

Cole ST. 2002. Comparative and functional genomics of the *Mycobacterium tuberculosis* complex. *Microbiology.* **148**:2919-28.

Copley, S.D. & Knowles, J.R. 1985. The uncatalyzed Claisen rearrangement of chorismate to prephenate prefers a transition state of chairlike geometry. *J. Am. Chem. Soc.* **107**: 5306-5308.

Covert BA, Spencer JS, Orme IM, Belisle JT. 2001. The application of proteomics in defining the T cell antigens of *Mycobacterium tuberculosis. Proteomics.* **1**:574-86.

Crosa JH, Walsh CT. 2002. Genetics and assembly line enzymology of siderophore biosynthesis in bacteria. *Microbiol Mol Biol Rev.* **66**:223-49.

Dahm C, Muller R, Schulte G, Schmidt K, Leistner E. 1998. The role of isochorismate hydroxymutase genes entC and menF in enterobactin and menaquinone biosynthesis in *Escherichia coli. Biochem Biophys Acta.* **1425**:377-86.

Davidson BE, Hudson GS.1987. Chorismate mutase-prephenate dehydrogenase from *Escherichia coli. Methods Enzymol.* **142**:440-50.

Delogu G, Brennan MJ. 2001. Comparative immune response to PE and PE_PGRS antigens of *Mycobacterium tuberculosis. Infect Immun.* **69**:5606-11.

De Voss JJ, Kerry Rutter, Benjamin G. Schroeder, and Clifton E. Barry, III. 1999. Iron Acquisition and Metabolism by Mycobacteria. *J. Bacteriol.* **181**: 4443-4451.

De Voss JJ, Rutter K, Schroeder BG, Su H, Zhu Y, Barry CE 3rd. 2000. The salicylate-derived mycobactin siderophores of *Mycobacterium tuberculosis* are essential for growth in macrophages. *Proc Natl Acad Sci USA.* **97**:1252-7.

Dosselaere F & Vanderleyden J. 2001 A metabolic node in action: chorismate-utilizing enzymes in microorganisms. *Crit Rev Microbiol.* **27**:75-131.

Duggleby RG, Sneddon MK, Morrison JF. 1978. Chorismate mutase-prephenate dehydratase from *Escherichia coli*: active sites of a bifunctional enzyme. *Biochemistry.* **17**:1548-54

Dussurget O, Rodriguez M, Smith I. 1996. An ideR mutant of Mycobacterium smegmatis has derepressed siderophore production and an altered oxidative-stress response. *Mol Microbiol.* **22**:535-44.

Dussurget O, Timm J, Gomez M, Gold B, Yu S, Sabol SZ, Holmes RK, Jacobs WR Jr, Smith I. 1999. Transcriptional control of the iron-responsive *fxb*A gene by the mycobacterial regulator IdeR. *J Bacteriol.* **181**:3402-8.

Eberhard J, Raesecke HR, Schmid J, Amrhein N. Cloning and expression in yeast of a higher plant chorismate mutase. 1993. Molecular cloning, sequencing of the cDNA and characterization of the *Arabidopsis thaliana* enzyme expressed in yeast. *FEBS Lett.*. **334**:233-6

Espitia C, Laclette JP, Mondragon-Palomino M, Amador A, Campuzano J, Martens A, Singh M, Cicero R, Zhang Y, Moreno C. 1999. The PE-PGRS glycine-rich proteins of *Mycobacterium tuberculosis*: a new family of fibronectin-binding proteins? *Microbiology*. **145**:3487-95.

Feese MD, Ingason BP, Goranson-Siekierke J, Holmes RK, Hol WG. 2001. Crystal structure of the iron-dependent regulator from *Mycobacterium tuberculosis* at 2.0-A resolution reveals the Src homology domain 3-like fold and metal binding function of the third domain. *J Biol Chem.* **276**:5959-66.

Fernandes ND, Wu QL, Kong D, Puyang X, Garg S, Husson RN. 1999. A mycobacterial extracytoplasmic sigma factor involved in survival following heat shock and oxidative stress. *J Bacteriol*. **181**:4266-74.

Finkelstein RA, Sciortino CV, McIntosh MA. 1983. Role of iron in microbe-host interactions. *Rev Infect Dis.* Suppl 4:S759-77. Review.

Fischer R, Jensen R. 1987. Prephenate dehydratase (monofunctional). *Methods Enzymol.***142**:507-12.

Fisher MA, Plikaytis BB, Shinnick TM. 2002. Microarray analysis of the *Mycobacterium tuberculosis* transcriptional response to the acidic conditions found in phagosomes *J Bacteriol.* **184**:4025-32.

Frank DW, Storey DG, Hindahl MS, Iglewski BH. 1989. Differential regulation by iron of regA and toxA transcript accumulation in *Pseudomonas aeruginosa. J Bacteriol.* **171**:5304-13.

Frees D, Qazi SN, Hill PJ, Ingmer H. 2003. Alternative roles of ClpX and ClpP in *Staphylococcus aureus* stress tolerance and virulence. *Mol Microbiol.* **48**:1565-78.

Friedrich B, Friedrich CG, Schlegel HG. 1976. Purification and properties of chorismate mutase-prephenate dehydratase and prephenate dehydrogenase from *Alcaligenes eutrophus. J Bacteriol.* **126**:712-22

Gaille C, Kast P, Haas D. 2002. Salicylate biosynthesis in *Pseudomonas aeruginosa*. Purification and characterization of PchB, a novel bifunctional enzyme displaying isochorismate pyruvate-lyase and chorismate mutase activities. *J Biol Chem.* **277**:21768-75.

Gaille C, Reimmann C, Haas D. 2003. Isochorismate synthase (PchA), the first and rate-limiting enzyme in salicylate biosynthesis of *Pseudomonas aeruginosa. J Biol Chem.*. **278**:16893-8.

Galperin MY, Walker DR, Koonin EV. 1998. Analogous enzymes: independent inventions in enzyme evolution.*Genome Res.***8** :779-90

Garnier T, Eiglmeier K, Camus JC, Medina N, Mansoor H, Pryor M, Duthoy S, Grondin S, Lacroix C, Monsempe C, Simon S, Harris B, Atkin R, Doggett J, Mayes R, Keating L, Wheeler PR, Parkhill J, Barrell BG, Cole ST, Gordon SV, Hewinson RG. 2003.The complete genome sequence of *Mycobacterium bovis. Proc Natl Acad Sci USA.* **100**:7877-82.

Gennaro ML. 2000. Immunologic diagnosis of tuberculosis. *Clin Infect Dis.* **30** Suppl 3:S243-6.

Gething MJ, Davidson BE. 1977. Chorismate mutase/prephenate dehydratase from *Escherichia coli* K12. Effects of chemical modification on the enzymic activities and allosteric inhibition. *Eur J Biochem.* **78**:111-7.

Ghosh S, Rasheedi S, Rahim SS, Banerjee S, Choudhary RK, Chakhaiyar Prachee, Ehtesham NZ, Mukhopadhyay S and Hasnain SE. 2004. A novel method for enhancing solubility of the expressed recombinant proteins in *E. coli. BioTechniques.* 00:000-000 [In press].

Gibson, F. and Jackman, L.M. 1963. Structure of chorismic acid, a new intermediate in aromatic biosynthesis. *Nature.* **198**, 388-389.

Gleissberg V. 1999. The threat of multidrug resistance: is tuberculosis ever untreatable or uncontrollable? *Lancet.* **353**:998-9.

Goers SK, Jensen RA: 1984. The differential allosteric regulation of two chorismate-mutase isoenzymes of *Nicotiana silvestris*. *Planta* **162**:117-124.

Gold B, Rodriguez GM, Marras SA, Pentecost M, Smith I. 2001. The *Mycobacterium tuberculosis* IdeR is a dual functional regulator that controls transcription of genes involved in iron acquisition, iron storage and survival in macrophages. *Mol Microbiol.* **42**:851-65.

Gomez M, Johnson S, Gennaro ML. 2000. Identification of secreted proteins of *Mycobacterium tuberculosis* by a bioinformatic approach. *Infect Immun.* **68**: 2323-7.

Gordhan BG, Smith DA, Alderton H, McAdam RA, Bancroft GJ, Mizrahi V. 2002. Construction and phenotypic characterization of an auxotrophic mutant of *Mycobacterium tuberculosis* defective in L-arginine biosynthesis. *Infect Immun.* **7**:3080-4

Graf R, Mehmann B, Braus GH. 1993. Analysis of feedback-resistant anthranilate synthases from *Saccharomyces cerevisiae*. *J Bacteriol* **175**:1061-8.

Gray JV *et al.* 1990. Monofunctional chorismate mutase from *Bacillus subtilis*: purification of the protein, molecular cloning of the gene, and overexpression of the gene product in *Escherichia coli*. *Biochemistry.* **29**:376-83.

Gu W *et al.* 1997. The aroQ and pheA domains of the bifunctional P-protein from *Xanthomonas campestris* in a context of genomic comparison. *Microb Comp Genomics.* **2**:141-58.

Guilford, W.J., Copley, S.D. & Knowles, J.R. 1987. On the mechanism of the chorismate mutase reaction. *J. Am. Chem. Soc.* **109**: 5013-5019

Guimaraes CR, Repasky MP, Chandrasekhar J, Tirado-Rives J, Jorgensen WL. 2003. Contributions of conformational compression and preferential transition state stabilization to the rate enhancement by chorismate mutase. *J Am Chem Soc*:**125**:6892-9

Guo H, Cui Q, Lipscomb WN, Karplus M. 2001. Substrate conformational transitions in the active site of chorismate mutase: their role in the catalytic mechanism. *Proc Natl Acad Sci USA.*; **98**: 9032-7.

Gupta S, Pandit SB, Srinivasan N, Chatterji D. 2002. Proteomics analysis of carbon-starved *Mycobacterium smegmatis*: induction of Dps-like protein. *Protein Eng.* **15**: 503-12.

Guzman LM, Belin D, Carson MJ, Beckwith J. 1995. Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J Bacteriol.***177**:4121-30.

Hall GC, Flick MB, Gherna RL, Jensen RA.1982. Biochemical diversity for biosynthesis of aromatic amino acids among the cyanobacteria. *J Bacteriol.***149**:65-78.

Harinarayanan R, Gowrishankar J. 2003. Host factor titration by chromosomal R-loops as a mechanism for runaway plasmid replication in transcription termination-defective mutants of *Escherichia coli*. *J Mol Biol.* **332** :31-46

Haslam E, 1993. Shikimic Acid Metabolism and Metabolites, John Wiley & Sons, New York.

Hasnain SE and Ahmed N. 2004. Leptospirosis: An epidemiologic, genomic and evolutionary perspective. *Lancet Infectious Diseases.***4**:543-544..

Hasnain SE. 2003. Molecular Epidemiology of infectious diseases: a case for increased surveillance. *Bulletin of the World Health Organization.* **81**: 474.

Hassan HM, Sun HC. 1992. Regulatory roles of Fnr, Fur, and Arc in expression of manganese-containing superoxide dismutase in *Escherichia coli*. *Proc Natl Acad Sci USA*. **89**:3217-21.

He Z, Stigers Lavoie KD, Bartlett PA, Toney MD. 2004. Conservation of mechanism in three chorismate-utilizing enzymes. *J Am Chem Soc*. **126**:2378-85.

Heather A. Carlson and William L. Jorgensen. 1996. Monte Carlo investigations of solvent effects on the chorismate to prephenate rearrangement. *J. Am. Chem. Soc*. **118,** 8475-8484.

Heimer SR, Welch RA, Perna NT, Posfai G, Evans PS, Kaper JB, Blattner FR, Mobley HL. 2002.Urease of enterohemorrhagic *Escherichia coli*: evidence for regulation by fur and a trans-acting factor. *Infect Immun*.**70**:1027-31.

Helmstaedt K *et. al.* 2001. Allosteric regulation of catalytic activity: *Escherichia coli* aspartate transcarbamoylase versus yeast chorismate mutase. *Microbiol Mol Biol Rev*. **65**:404-21

Hondalus MK, Bardarov S, Russell R, Chan J, Jacobs WR Jr, Bloom BR. 2000. Attenuation of and protection induced by a leucine auxotroph of *Mycobacterium tuberculosis*. *Infect Immun*. **68**:2888-98

Horwitz MA, Harth G, Dillon BJ, Maslesa-Galic' S. 2000. Recombinant bacillus calmette-guerin (BCG) vaccines expressing the *Mycobacterium tuberculosis* 30-kDa major secretory protein induce greater protective immunity against tuberculosis than conventional BCG vaccines in a highly susceptible animal model. *Proc Natl Acad Sci USA*. **97**:13853-8.

Horwitz MA, Harth G. 2003. A new vaccine against tuberculosis affords greater survival after challenge than the current vaccine in the guinea pig model of pulmonary tuberculosis. *Infect Immun*. **71**:1672-9.

Huang L, Nakatsukasa M, Nester E.1974. Regulation of aromatic amino acid biosynthesis in *Bacillus subtilis* 168. Purification, characterization, and subunit structure of the bifunctional enzyme 3-deoxy-D-arabinoheptulosonate 7-phosphate synthetase-chorismate mutase. *J Biol Chem*. **249**:4467-72.

Hudson GS, Wong V, Davidson BE 1984. Chorismate mutase/prephenate dehydrogenase *from Escherichia coli* K12: purification, characterization, and identification of a reactive cysteine. *Biochemistry*. **23**:6240-9.

Hur S, Bruice TC. 2002. The mechanism of catalysis of the chorismate to prephenate reaction by the *Escherichia coli* mutase enzyme. *Proc Natl Acad Sci USA*. **99**:1176-81.

Hur S, Bruice TC. 2003. The near attack conformation approach to the study of the chorismate to prephenate reaction.*Proc Natl Acad Sci USA*..**100**:12015-20.

Hur S, Bruice TC. 2003. Just a near attack conformer for catalysis (chorismate to prephenate rearrangements in water, antibody, enzymes, and their mutants). *J Am Chem Soc*.**125**:10540-2.

Ito J, Yanofsky C. 1966. The nature of the anthranilic acid synthetase complex of *Escherichia coli*. *J Biol Chem*. **241**:4112-4.

Ito J, Yanofsky C. 1969. Anthranilate synthetase, an enzyme specified by the tryptophan operon of *Escherichia coli*: Comparative studies on the complex and the subunits. *J Bacteriol*. **97**:734-42.

Ito J, Cox EC, Yanofsky C. 1969. Anthranilate synthetase, an enzyme specified by the tryptophan operon of *Escherichia coli*: purification and characterization of component I. *J Bacteriol*. **97**:725-33.

Jackson M, Phalen SW, Lagranderie M, Ensergueix D, Chavarot P, Marchal G, McMurray DN, Gicquel B, Guilhot C. 1999. Persistence and protective efficacy of a *Mycobacterium tuberculosis* auxotroph vaccine. *Infect Immun*. **67**:2867-73.

Jensen R, Nester EW. 1965. The regulatory significance of intermediary metabolites: control of aromatic acid biosynthesis by feedback inhibition in *Bacillus subtilis*. *J Mol Biol*. **12**:468-481.

Jungblut PR, Muller EC, Mattow J, Kaufmann SH. 2001. Proteomics reveals open reading frames in *Mycobacterium tuberculosis* H37Rv not predicted by genomics. *Infect Immun.* **69**:5905-7.

Kaiser A, Leistner E. 1990. Role of the entC gene in enterobactin and menaquinone biosynthesis in *Escherichia coli. Arch Biochem Biophys.* **276**:331-5

Kast P, Grisostomi C, Chen IA, Li S, Krengel U, Xue Y, Hilvert D. 2000.A strategically positioned cation is crucial for efficient catalysis by chorismate mutase. *J Biol Chem.* **275**:36832-8.

Kaushal D, Schroeder BG, Tyagi S, Yoshimatsu T, Scott C, Ko C, Carpenter L, Mehrotra J, Manabe YC, Fleischmann RD, Bishai WR. 2002. Reduced immunopathology and mortality despite tissue persistence in a *Mycobacterium tuberculosis* mutant lacking alternative sigma factor, SigH. *Proc Natl Acad Sci USA.* **99**:8330-5.

Kienhofer A, Kast P, Hilvert D. 2003. Selective stabilization of the chorismate mutase transition state by a positively charged hydrogen bond donor. *J Am Chem Soc.* **125**:3206-7.

King RD, Karwath A, Clare A, Dehaspe L. 2000. Accurate prediction of protein functional class from sequence in the *Mycobacterium tuberculosis* and *Escherichia coli* genomes using data mining. *Yeast.* **17**:283-93.

Kishore GM, Shah DM. 1988. Amino acid biosynthesis inhibitors as herbicides. *Annu Rev Biochem.* **57**:627-63.

Knochel T, Ivens A, Hester G, Gonzalez A, Bauerle R, Wilmanns M, Kirschner K, Jansonius JN. 1999. The crystal structure of anthranilate synthase from *Sulfolobus solfataricus*: functional implications. *Proc Natl Acad Sci USA.* **96**:9479-84.

Krappmann S. *et. al.* 1999. The aroC gene of *Aspergillus nidulans* codes for a monofunctional, allosterically regulated chorismate mutase. *J. Biol. Chem.* **274**: 22275-22282.

Lambert KN. *et. al.* 1999. Cloning and characterization of an esophageal-gland-specific chorismate mutase from the phytoparasitic nematode *Meloidogyne javanica. Mol Plant Microbe Interact.* **12**:328-3

Lane SJ. *et. al.* 1998. Isolation and charecterization of carboxymycobactins as the second extracellular siderophores in *Mycobacterium smegmatis. Biometals.* **11**: 13-20.

Lee, A.Y., Karplus, P.A., Ganem, B. & Clardy, J. 1995 Atomic structure of the buried catalytic pocket of *Escherichia coli* chorismate mutase. *J. Am. Chem. Soc.* **117**:3627-3628

Levitskaya J, Coram M, Levitsky V, Imreh S, Steigerwald-Mullen PM, Klein G, Kurilla MG, Masucci MG. 1995. Inhibition of antigen processing by the internal repeat region of the Epstein-Barr virus nuclear antigen. *Nature.* **375**:685-8.

Lewis KN, Liao R, Guinn KM, Hickey MJ, Smith S, Behr MA, Sherman DR. 2003. Deletion of RD1 from *Mycobacterium tuberculosis* mimics bacille Calmette-Guerin attenuation. *J Infect Dis.* **187**:117-23.

Liljeqvist JA, Trybala E, Svennerholm B, Jeansson S, Sjogren-Jansson E, Bergstrom T. 1998. Localization of type-specific epitopes of herpes simplex virus type 2 glycoprotein G recognized by human and mouse antibodies. *J Gen Virol.* **79**:1215-24.

Litwin CM, Calderwood SB.1993. Role of iron in regulation of virulence genes. *Clin Microbiol Rev.* **6**:137-49. Review

Liu J, Quinn N, Berchtold GA, Walsh CT. 1990. Overexpression, purification, and characterization of isochorismate synthase (EntC), the first enzyme involved in the biosynthesis of enterobactin from chorismate. *Biochemistry.* **29**:1417-25.

Llewellyn DJ, Daday A, Smith GD.1980. Evidence for an artificially evolved bifunctional 3-deoxy-D-arabinoheptulosonate-7-phosphate synthase-chorismate mutase in *Bacillus subtilis*. *J Biol Chem.* **255**:2077-84

Lowrie DB, Tascon RE, Bonato VLD, Lima Valeria MF , Faccioli LH, Stavropoulos E, Colston MJ, Hewinson RG, Moelling K, Silva CL. 1999. Therapy of tuberculosis in mice by DNA vaccination. *Nature.* **400**: 269 - 271.

MacBeath G, Kast P, & Hilvert D. 1998. A small, thermostable, and monofunctional chorismate mutase from the archeon *Methanococcus jannaschii. Biochemistry.* **37**: 10062-10073.

Madan Babu M. 2003. Did the loss of sigma factors initiate pseudogene accumulation in *M. leprae*? *Trends Microbiol.* **11**:59-61.

Maddock JR, Shapiro L.1993. Polar location of the chemoreceptor complex in the *Escherichia coli* cell. *Science.* **259**:1717-23.

Majeed AA, Ahmed N, Rao KR, Ghousunnissa S, Kauser F, Bose B, Nagarajaram HA, Katoch VM, Cousins DV, Sechi LA, Gilman RH, Hasnain SE. 2004. AmpliBASE MT: a *Mycobacterium tuberculosis* diversity knowledgebase.*Bioinformatics.* **20**:989-92.

Manabe YC, Saviola BJ, Sun L, Murphy JR, Bishai WR.1999. Oct Attenuation of virulence in *Mycobacterium tuberculosis* expressing a constitutively active iron repressor. *Proc Natl Acad Sci USA.* **96**:12844-8.

Manganelli R, Dubnau E, Tyagi S, Kramer FR, Smith I.1999. Differential expression of 10 sigma factor genes in *Mycobacterium tuberculosis. Mol Microbiol.* **31**:715-24.

Manganelli R, Voskuil MI, Schoolnik GK, Smith I. 2001. The *Mycobacterium tuberculosis* ECF sigma factor sigmaE: role in global gene expression and survival in macrophages. *Mol Microbiol.* **41**:423-37.

Manganelli R, Voskuil MI, Schoolnik GK, Dubnau E, Gomez M, Smith I. 2002. Role of the extracytoplasmic-function sigma factor sigma(H) in *Mycobacterium tuberculosis* global gene expression. *Mol Microbiol.* **45**: 365-74.

Marshall BJ, Ratledge C. 1972. Salicylic acid biosynthesis and its control in *Mycobacterium smegmatis. Biochim Biophys Acta.*. **264**: 106-16.

Marshall BJ, Ratledge C. 1971. Conversion of chorismic acid and isochorismic acid to salicylic acid by cell-free extracts of *Mycobacterium smegmatis. Biochim Biophys Acta.* **230**: 643-5.

McCue LA, McDonough KA, Lawrence CE. 2000. Functional classification of cNMP-binding proteins and nucleotide cyclases with implications for novel regulatory pathways in *Mycobacterium tuberculosis. Genome Res.* **10**: 204-19.

McKinney JD, Honer zu Bentrup K, Munoz-Elias EJ, Miczak A, Chen B, Chan WT, Swenson D, Sacchettini JC, Jacobs WR Jr, Russell DG. 2000. Persistence of *Mycobacterium tuberculosis* in macrophages and mice requires the glyoxylate shunt enzyme isocitrate lyase. *Nature.* **406**: 735-8.

McLendon MM, Shinnick TM. I-TRAP: 2003. A method to identify transcriptional regulator activated promoters. *BMC Infect Dis.* **3**: 15.

Miguez J, Laferte J, Tejero Y, Gonzalez G, Otero AJ, Rivero J, Duarte C. 1998. Evaluation of the serologic response against two consensus V3 loop peptides from human immunodeficiency virus-1 in Cuban patients. *Int J Infect Dis.* **2**: 221-5.

Monod J, Wyman J, Changeux Jp. 1965. On The Nature Of Allosteric Transitions: A Plausible Model. *J Mol Biol.* **12**: 88-118

Morollo AA, Eck MJ. 2001. Structure of the cooperative allosteric anthranilate synthase from *Salmonella typhimurium*. *Nat Struct Biol.* **8**: 243-7.

Muller A, MacCallum RM, Sternberg MJ. 1999. Benchmarking PSI-BLAST in genome annotation. *J Mol Biol.* **293**: 1257-71.

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

Neilands JB. Siderophores: 1995. Structure and function of microbial iron transport compounds. *J Biol Chem.* **270**: 26723-6.

Ochsner UA, Vasil ML. 1996. Gene repression by the ferric uptake regulator in *Pseudomonas aeruginosa*: cycle selection of iron-regulated genes. *Proc Natl Acad Sci USA.* **93**: 4409-14

Parish T, Gordhan BG, McAdam RA, Duncan K, Mizrahi V, Stoker NG. 1999. Production of mutants in amino acid biosynthesis genes of *Mycobacterium tuberculosis* by homologous recombination. *Microbiology.* **145** : 3497-503.

Parish T. 2003. Starvation survival response of *Mycobacterium tuberculosis*. *J Bacteriol.* **185**:6702-6.

Parish T, Stoker NG. 2002. The common aromatic amino acid biosynthesis pathway is essential in *Mycobacterium tuberculosis*. *Microbiology.* **148**: 3069-77.

Park HD, Guinn KM, Harrell MI, Liao R, Voskuil MI, Tompa M, Schoolnik GK, Sherman DR. 2003. Rv3133c/dosR is a transcription factor that mediates the hypoxic response of *Mycobacterium tuberculosis*. *Mol Microbiol.* **48**: 833-43.

Pavelka MS Jr, Chen B, Kelley CL, Collins FM, Jacobs Jr WR Jr. 2003. Vaccine efficacy of a lysine auxotroph of *Mycobacterium tuberculosis*. *Infect Immun.* **71**: 4190-2.

Pittard AJ. 1987. Biosynthesis of the aromatic amino acids. In *Escherichia coli* and *Salmonella typhimurium*: Cellular and Molecular Biology, pp. 368-394. Edited by JL Ingraham, KB Low, B Magasanik, M Schaechter & HE Umbarger. Washington, DC: American Society for Microbiology.

Pohl E, Holmes RK, Hol WG. 1999. Crystal structure of the iron-dependent regulator (IdeR) from *Mycobacterium tuberculosis* shows both metal binding sites fully occupied. *J Mol Biol.* **285**: 1145-56.

Pohnert, G., Zhang, S., Husain, A., Wilson, D. B. and Ganem, B. Regulation of phenylalanine biosynthesis. 1999. Studies on the mechanism of phenylalanine binding and feedback inhibition in the *Escherichia coli* P-protein. *Biochemistry.* **38**: 12212-12217

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

Pressler U, Staudenmaier H, Zimmermann L, Braun V. 1988. Genetics of the iron dicitrate transport system of *Escherichia coli*. *J Bacteriol.* **170**: 2716-24.

Primm TP, Andersen SJ, Mizrahi V, Avarbock D, Rubin H, Barry CE 3rd. 2000. The stringent response of *Mycobacterium tuberculosis* is required for long-term survival. *J Bacteriol.* **182**: 4889-98.

Pym AS, Domenech P, Honore N, Song J, Deretic V, Cole ST. 2001. Regulation of catalase-peroxidase (KatG) expression, isoniazid sensitivity and virulence by furA of *Mycobacterium tuberculosis*. *Mol Microbiol.* **40**: 879-89.

Pym AS, Brodin P, Brosch R, Huerre M, Cole ST. 2002. Loss of RD1 contributed to the attenuation of the live tuberculosis vaccines *Mycobacterium bovis* BCG and *Mycobacterium microti*. *Mol Microbiol.* **46**: 709-17.

Pym AS, Brodin P, Majlessi L, Brosch R, Demangel C, Williams A, Griffiths KE, Marchal G, Leclerc C, Cole ST. 2003. Recombinant BCG exporting ESAT-6 confers enhanced protection against tuberculosis. *Nat Med.* **9**: 533-9.

Quadri LE, Sello J, Keating TA, Weinreb PH, Walsh CT. 1998. Identification of a *Mycobacterium tuberculosis* gene cluster encoding the biosynthetic enzymes for assembly of the virulence-conferring siderophore mycobactin. *Chem Biol.* **5**: 631-45.

Ramakrishnan L, Federspiel NA, Falkow S. 2000. Granuloma-specific expression of Mycobacterium virulence proteins from the glycine-rich PE-PGRS family. *Science.* **288**: 1436-9.

Raman LA, Siddiqi N, Shamim M, Deb M, Mehta G, Hasnain SE. 2000. Molecular characterization of *Mycobacterium abscessus* strains isolated from a hospital outbreak. *Emerg Infect Dis.* **6**: 561-2.

Raman S, Song T, Puyang X, Bardarov S, Jacobs WR Jr, Husson RN. 2001.The alternative sigma factor SigH regulates major components of oxidative and heat stress responses in *Mycobacterium tuberculosis. J Bacteriol.* **183**: 6119-25.

Ratledge C. 2004. Iron, mycobacteria and tuberculosis.*Tuberculosis (Edinb).* **84**: 110-30.

Raymond KN, Dertz EA, Kim SS. 2003. Enterobactin: an archetype for microbial iron transport. *Proc Natl Acad Sci USA.* **100**: 3584-8.

Riepl RG, Glover GI. 1979. Regulation and state of aggregation of *Bacillus subtilis* prephenate dehydratase in the presence of allosteric effectors. *J Biol Chem.* **254**: 10321-8.

Rodriguez GM, Smith I. 2003. Mechanisms of iron regulation in mycobacteria: role in physiology and virulence. *Mol Microbiol.* **47**: 1485-94. Review

Rodriguez GM, Gold B, Gomez M, Dussurget O, Smith I. 1999. Identification and characterization of two divergently transcribed iron regulated genes in *Mycobacterium tuberculosis. Tuber Lung Dis.* **79**: 287-98.

Rodriguez GM, Voskuil MI, Gold B, Schoolnik GK, Smith I. 2002. *ide*R, An essential gene in *Mycobacterium tuberculosis*: role of IdeR in iron-dependent gene expression, iron metabolism, and oxidative stress response. *Infect Immun.* **70**: 3371-81

Rosenkrands I, Slayden RA, Crawford J, Aagaard C, Barry CE 3rd, Andersen P. 2002. Hypoxic response of *Mycobacterium tuberculosis* studied by metabolic labeling and proteome analysis of cellular and extracellular proteins. *J Bacteriol.* **184**: 3485-91.

Sambandamurthy VK, Wang X, Chen B, Russell RG, Derrick S, Collins FM, Morris SL, Jacobs WR Jr. 2002. A pantothenate auxotroph of *Mycobacterium tuberculosis* is highly attenuated and protects mice against tuberculosis. *Nat Med.* **8**: 1171- 4.

Sampson SL, Lukey P, Warren RM, van Helden PD, Richardson M, Everett MJ. 2001. Expression, characterization and subcellular localization of the *Mycobacterium tuberculosis* PPE gene Rv1917c. *Tuberculosis (Edinb).* **81**: 305-17.

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

Sbai H, Mehta A, DeGroot AS. 2001. Use of T cell epitopes for vaccine development. *Curr Drug Targets Infect Disord.* **1**: 303-13.

Schnappauf G. *et. al.* 1997. A glutamate residue in the catalytic center of the yeast chorismate mutase restricts enzyme activity to acidic conditions. *Proc Natl Acad Sci USA.* **94**: 8491-6.

Schmidheini T, Mosch HU, Evans JN, Braus G. 1990. Yeast allosteric chorismate mutase is locked in the activated state by a single amino acid substitution. *Biochemistry.* **29**: 3660-8.

Schmitt MP, Holmes RK. 1991. Iron-dependent regulation of diphtheria toxin and siderophore expression by the cloned *Corynebacterium diphtheriae* repressor gene *dtx*R in *C. diphtheriae* C7 strains. *Infect Immun.* **59**: 1899-904

Schmitt MP, Predich M, Doukhan L, Smith I, Holmes RK. . 1995. Characterization of an iron-dependent regulatory protein (IdeR) of *Mycobacterium tuberculosis* as a functional homolog of the diphtheria toxin repressor (DtxR) from *Corynebacterium diphtheriae. Infect Immun.* **63**: 4284-9.

Schuller DJ, Grant GA, Banaszak LJ. 1995. The allosteric ligand site in the V-max-type cooperative enzyme phosphoglycerate dehydrogenase. *Nat Struct Biol.* **2**: 69-76.

Serino L, Reimmann C, Baur H, Beyeler M, Visca P, Haas D. 1995. Structural genes for salicylate biosynthesis from chorismate in *Pseudomonas aeruginosa. Mol Gen Genet.* **249**: 217-28.

Serino L, Reimmann C, Visca P, Beyeler M, Chiesa VD, Haas D. 1997. Biosynthesis of pyochelin and dihydroaeruginoic acid requires the iron-regulated pchDCBA operon in *Pseudomonas aeruginosa. J Bacteriol.* **179**: 248-57.

Sharma V, Sharma S, Hoener zu Bentrup K, McKinney JD, Russell DG, Jacobs WR Jr, Sacchettini JC. 2000. Structure of isocitrate lyase, a persistence factor of *Mycobacterium tuberculosis. Nat Struct Biol.* **7**: 663-8.

Sherman DR, Voskuil M, Schnappinger D, Liao R, Harrell MI, Schoolnik GK. 2001. Regulation of the *Mycobacterium tuberculosis* hypoxic response gene encoding alpha -crystallin. *Proc Natl Acad Sci USA.* **98**: 7534-9.

Siddiqi N, Shamim M, Jain NK, Rattan A, Amin A, Katoch VM, Sharma SK, Hasnain SE. 1998. Molecular genetic analysis of multi-drug resistance in Indian isolates of *Mycobacterium tuberculosis. Mem Inst Oswaldo Cruz.* **93**: 589-94.

Siddiqi N, Shamim M, Amin A, Chauhan DS, Das R, Srivastava K, Singh D, Sharma VD, Katoch VM, Sharma SK, Hanief M, Hasnain SE. 2001. Typing of drug resistant isolates of *Mycobacterium tuberculosis* from India using the IS6110 element reveals substantive polymorphism. *Infect Genet Evol.* **1**:109-16.

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. 2002. Molecular characterization of multidrug-resistant isolates of *Mycobacterium tuberculosis* from patients in North India. *Antimicrob Agents Chemother.* **46**: 443-50.

Siddiqi N, Das R, Pathak N, Banerjee S, Ahmed N, Katoch VM and Hasnain SE. 2004..*Mycobacterium tuberculosis* isolate with a distinct genomic identity overexpresses a TAP like efflux pump. *Infection.* **32**: 109-11.

Singh KK, Zhang X, Patibandla AS, Chien P Jr, Laal S. 2001. Antigens of *Mycobacterium tuberculosis* expressed during preclinical tuberculosis: serological immunodominance of proteins with repetitive amino acid sequences. *Infect Immun.* **69**: 4185-91.

Smith DA, Parish T, Stoker NG, Bancroft GJ. 2001. Characterization of auxotrophic mutants of *Mycobacterium tuberculosis* and their potential as vaccine candidates. *Infect Immun.* **69**: 1142-50.

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

Sogo, S. G., Widlanski, T. S., Hoare, J. H., Grimshaw, C. E., Berchtold, G. A., & Knowles, J. R. 1984. *J. Am. Chem. Soc.* **106**: 2701-2703.

Stover CK, Warrener P, VanDevanter DR, Sherman DR, Arain TM, Langhorne MH, Anderson SW, Towell JA, Yuan Y, McMurray DN, Kreiswirth BN, Barry CE, Baker WR. 2000. A small-molecule nitroimidazopyran drug candidate for the treatment of tuberculosis. *Nature.* **405**: 962-6.

Strater N, Hakansson K, Schnappauf G, Braus G, Lipscomb WN. 1996. Crystal structure of the T state of allosteric yeast chorismate mutase and comparison with the R state. *Proc Natl Acad Sci USA.* **93**: 3330-4.

Strater N, Schnappauf G, Braus G, Lipscomb WN. 1997. Mechanisms of catalysis and allosteric regulation of yeast chorismate mutase from crystal structures. *Structure.* **5**:1437-52.

Su WJ. 2002. Recent advances in the molecular diagnosis of tuberculosis. *J Microbiol Immunol Infect.* **35**: 209-14.

Tao X, Boyd J, Murphy JR. 1992. Specific binding of the diphtheria tox regulatory element DtxR to the tox operator requires divalent heavy metal ions and a 9-base-pair interrupted palindromic sequence. *Proc Natl Acad Sci USA.* **89**: 5897-901.

Tatusov,R.L., Natale,D.A., Garkavtsev,I.V., Tatusova,T.A., Shankavaram,U.T., Rao,B.S., Kiryutin,B., Galperin,M.Y., Fedorova,N.D. and Koonin,E.V. 2001. The COG database: new developments in phylogenetic classification of proteins from complete genomes. *Nucleic Acids Res.* **29**: 22–28.

Taylor JL, Turner OC, Basaraba RJ, Belisle JT, Huygen K, Orme IM. 2003. Pulmonary necrosis resulting from DNA vaccination against tuberculosis. *Infect Immun.* **71**: 2192-8.

Tazi L, Kreiswirth B, Carriere C, Tibayrenc M. 2002. Molecular epidemiology of *Mycobacterium tuberculosis* and its relevance to the surveillance and control of TB: an e-debate. *Infect Genet Evol.* **2**:153-8.

Tekaia F, Gordon SV, Garnier T, Brosch R, Barrell BG, Cole ST. 1999. Analysis of the proteome of *Mycobacterium tuberculosis* in silico. *Tuber Lung Dis.* **79**: 329-42.

Unniraman S, Prakash R, Nagaraja V. 2001. Alternate paradigm for intrinsic transcription termination in eubacteria. *J Biol Chem.* **276**: 41850-5.

Van de Loosdrecht AA, Beelen RH, Ossenkoppele GJ, Broekhoven MG, Langenhuijsen MM. 1994. A tetrazolium-based colorimetric MTT assay to quantitate human monocyte mediated cytotoxicity against leukemic cells from cell lines and patients with acute myeloid leukemia. *J Immunol Methods.* **174**: 311-20.

Van Pinxteren LA, Cassidy JP, Smedegaard BH, Agger EM, Andersen P. 2000. Control of latent *Mycobacterium tuberculosis* infection is dependent on CD8 T cells. *Eur J Immunol.* **30**: 3689-98.

Vega-Lopez F, Brooks LA, Dockrell HM, De Smet KA, Thompson JK, Hussain R, Stoker NG. 1993. Sequence and immunological characterization of a serine-rich antigen from *Mycobacterium leprae. Infect Immun.* **61**: 2145-53.

Vinella, D., Gagny, B., JoseleauPetit, D., Dari, R. & Cashel, M. 1996. Mecillinam resistance in *Escherichia coli* is conferred by loss of a second activity of the AroK protein. *J Bacteriol.* **178**: 3818-3828.

Walker-Simmons M, Atkinson DE. 1977. Functional capacities and the adenylate energy charge in *Escherichia coli* under conditions of nutritional stress. *J Bacteriol.***130**: 676-83.

Washio T, Sasayama J, Tomita M. 1998. Analysis of complete genomes suggests that many prokaryotes do not rely on hairpin formation in transcription termination. *Nucleic Acids Res.* **26**: 5456-63.

Wayne,L.G. 1994. Dormancy of *Mycobacterium tuberculosis* and latency of disease. *Eur J Clin Microbiol Infect Dis.* **43**: 908-914.

Weiss, U. & Edwards, J. M. 1980. in The Biosynthesis of Aromatic Amino Compounds. John Wiley & Sons, Inc., New York.

Wiest O , Houk KN. 1994. On the Transition State of the Chorismate-Prephenate Rearrangement. *J. Org. Chem.* **59**: 7582-7584.

Wildermuth MC, Dewdney J, Wu G, Ausubel FM. 2001. Isochorismate synthase is required to synthesize salicylic acid for plant defence. *Nature.* **414**: 562-5. Erratum in: Nature. 2002;417:571.

Wilkinson RJ, Vordermeier HM, Wilkinson KA, Sjolund A, Moreno C, Pasvol G, Ivanyi J. Peptide-specific T cell response to *Mycobacterium tuberculosis*: clinical spectrum, compartmentalization, and effect of chemotherapy. 1998. *J Infect Dis.* **178**:760-8.

Wilson M, DeRisi J, Kristensen HH, Imboden P, Rane S, Brown PO, Schoolnik GK. 1999. Exploring drug-induced alterations in gene expression in *Mycobacterium tuberculosis* by microarray hybridization. *Proc Natl Acad Sci USA.* **96**:12833-8.

Wong DK, Gobin J, Horwitz MA, Gibson BW. 1996. Characterization of exochelins of *Mycobacterium avium*: evidence for saturated and unsaturated and for acid and ester forms. *J Bacteriol.* **178**: 6394-8.

Wu QL, Kong D, Lam K, Husson RN. 1997. A mycobacterial extracytoplasmic function sigma factor involved in survival following stress. *J Bacteriol.* **179**: 2922-9.

Wyborn NR, Stapleton MR, Norte VA, Roberts RE, Grafton J, Green J. 2004. Regulation of *Escherichia coli* hemolysin E expression by H-NS and Salmonella SlyA. *J Bacteriol.* **186**: 1620-8.

Xia T, Song J, Zhao G, Aldrich H, Jensen RA. 1993. The aroQ-encoded monofunctional chorismate mutase (CM-F) protein is a periplasmic enzyme in *Erwinia herbicola. J Bacteriol.* **175**: 4729-37.

Xie G, Keyhani NO, Bonner CA, Jensen RA. 2003. Ancient origin of the tryptophan operon and the dynamics of evolutionary change. *Microbiol Mol Biol Rev.* **67**: 303-42. Review

Xue Y, Lipscomb WN, Graf R, Schnappauf G, Braus G. 1994. The crystal structure of allosteric chorismate mutase at 2.2-A resolution. *Proc Natl Acad Sci USA.* **91**: 10814-8.

Xue Y, Lipscomb WN. 1995. Location of the active site of allosteric chorismate mutase from *Saccharomyces cerevisiae*, and comments on the catalytic and regulatory mechanisms. *Proc Natl Acad Sci USA.***92**:10595-8.

Yellaboina S, Ranjan S, Chakhaiyar P, Hasnain SE, and Ranjan A. 2004. Prediction of DtxR regulon: Identification of binding sites and operons controlled by Diphtheria toxin repressor in *Corynebacterium diphtheriae. BMC Microbiology.* **00**:000-000 [In press]

Yellaboina S., Seshadri J., Senthil Kumar M and Ranjan A. 2004. Predictregulon: A webserver for the prediction of the regulatory protein binding sites and operons in prokaryote genomes. *Nucleic Acids Research.* **32**: W318-320.

Young D, Kent L, Rees A, Lamb J, Ivanyi J. 1986. Immunological activity of a 38-kilodalton protein purified from *Mycobacterium tuberculosis. Infect Immun.* **54**: 177-83.

Young DB, Garbe TR. 1991. Heat shock proteins and antigens of *Mycobacterium tuberculo sis. Infect Immun.* **59**: 3086-93.

Young IG, Batterham TJ, Gibson F 1969. The isolation and properties of isochorismic acid, an intermediate in the biosynthesis of 2,3-dihydroxybenzoic acid. *Biochim Biophys Acta.* **177**: 389-400.

Zahrt TC, Song J, Siple J, Deretic V. 2001. Mycobacterial FurA is a negative regulator of catalase-peroxidase gene katG. *Mol Microbiol.* **39**: 1174-85.

Zhang S, Pohnert G, Kongsaeree P, Wilson DB, Clardy J, Ganem B. 1998. Chorismate mutase-prephenate dehydratase from *Escherichia coli.* Study of catalytic and regulatory domains using genetically engineered proteins. *J Biol Chem.* **273**: 6248-53.

Zhang S, Wilson DB, Ganem B. 2003. An engineered chorismate mutase with allosteric regulation. *Bioorg Med Chem.* **11**: 3109-14.

Zimmermann L, Hantke K, Braun V. 1984. Exogenous induction of the iron dicitrate transport system of *Escherichia coli* K-12. *J Bacteriol.* **159**: 271-7.

VITAE

PRACHEE

Laboratory of Molecular and Cellular Biology,
Centre for DNA Fingerprinting and Diagnostics,
HYDERABAD, INDIA

EDUCATIONAL QUALIFICATIONS

Ph.D (Thesis submitted to University of Hyderabad), 2004
Centre for DNA Fingerprinting and Diagnostics, Hyderabad, INDIA
Title of thesis: Studies on biomedically important *Mycobacterium tuberculosis* genes involved in nutrient assimilation and immune response

Master of Science (M.Sc), 1999
Department of Botany,
University of Delhi, Delhi, INDIA
Specialization: Genetics and Plant Breeding
First Division with 75% Marks

Bachelor of Science (B.Sc), 1997
University of Delhi, Delhi, INDIA
Subjects; Botany (Hons), Zoology and Chemistry
First Division with 77% Marks

PUBLICATIONS

Chakhaiyar Prachee, Nagalakshmi Y, Aruna B, Murthy KJ, Katoch VM, Hasnain SE. 2004. Regions of high antigenicity within the hypothetical PPE Major Polymorphic Tandem Repeat Open-Reading Frame, Rv2608, show a differential humoral response and a low T cell response in various categories of patients with tuberculosis. ***Journal of Infectious Diseases***. **190**:1237-44.

Chakhaiyar Prachee and Hasnain SE. 2004. Defining the mandate of tuberculosis research in a post genomic era. ***Medical Principles and Practice.*** **13:**177-184.

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

Yellaboina S., Ranjan S., **Chakhaiyar Prachee**, Hasnain SE, and Ranjan A. 2004. Prediction of DtxR regulon: Identification of binding sites and operons controlled by Diphtheria toxin repressor in *Corynebacterium diphtheriae*. ***BMC Microbiology.*** 00:000-000 [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. 2002. Molecular characterization of multidrug-resistant isolates of *Mycobacterium tuberculosis* from patients in North India. ***Antimicrobial Agents And Chemotherapy***. **46**:443-50.

Ghosh S, Rasheedi S, Rahim SS, Banerjee S, Choudhary RK, **Chakhaiyar Prachee**, Ehtesham NZ, Mukhopadhyay S and Hasnain SE. 2004. A novel method for enhancing solubility of the expressed recombinant proteins in *E. coli*. ***BioTechniques.*** **37**:000-000 [In press].