

ISOLATION AND CHARACTERIZATION  
OF THE 34 KD GENE OF  
*MYCOBACTERIUM LEPRAE*.

ADDIS ABABA UNIVERSITY  
Science Library  
CIRCULATION

BY

FESEHA ABEBE

SEPTEMBER, 1993.

ISOLATION AND CHARACTERIZATION  
OF THE 34 KD GENE OF  
*MYCOBACTERIUM LEPRAE*.

A THESIS SUBMITTED TO THE  
SCHOOL OF GRADUATE STUDIES  
ADDIS ABABA UNIVERSITY,

In partial fulfillment of  
the requirements for the degree of  
Master of Science in Biology.

BY  
Feseha Abebe,  
September, 1993.

ADDIS ABABA UNIVERSITY,  
SCHOOL OF GRADUATE STUDIES.

ISOLATION AND CHARACTERIZATION  
OF THE 34 KD GENE OF  
*MYCOBACTERIUM LEPRAE*.

BY  
FESEHA ABEBE.

## TABLE OF CONTENTS

Contents.....	PAGE.
i. List of Figures .....	i.
ii. List of Tables .....	ii.
iii. List of Appendices.....	ii.
iv. Acknowledgement.....	iii.
1. Abstract.....	1
2. Introduction.....	2
2.1. The disease.....	2
2.2. Distribution and prevalence.....	3
2.3. Leprosy in Ethiopia.....	3
2.4. Transmission, Reservoirs and hosts.....	3
2.5. Clinical and immunological aspects of leprosy.....	4
2.6. The pathogen.....	6
2.6.1. Classification & morphology.....	7
2.6.2. Growth requirements and other features.....	7
2.6.3. Molecular biology of <i>M. leprae</i> .....	8
2.6.4. The aim of the project.....	13

Content.....	Page
3. Materials and methods.....	15
3.1. Recombinant DNA library and bacterial strains.....	15
3.2. Plasmid vectors used in the experiments.....	15
3.3. Probes used in the experiments.....	15
3.4. Media and antibiotics.....	16
3.5. DNA Techniques.....	16
3.6. Immunological techniques.....	20
4. Results.....	22
4.1. Selection of the 34 kd gene of <i>M. leprae</i> .....	22
4.2. Isolation and subcloning of the 34 kd gene.....	27
4.3. Expression of the recombinant 34 kd antigen.....	33
4.4. DNA sequencing of the 34 kd gene.....	35
5. Discussion.....	40
6. Conclusion.....	46
7. References.....	47
8. Abbreviations.....	58
9. Appendix.....	59

<u>List of figures</u>	Page.
Fig 1 A. Isolation of the probe DNA from gel and its purification by gene clean.	23
Fig.2. Southern Hybridization of the cosmid library with the probe DNA.	24
Fig 3. Cosmid isolation and identity confirmation.	25
Fig 4 A & B. Restriction digestion of BL 1527 with the 8 enzymes .	25
Fig 5. Autoradiography of the twin blots with discriminating, probes.	26
Fig 6. Restriction digest of Cosmid BL 1527 with the 5 enzymes for isolation of gene specific bands.	27
Fig 7. Gene clean purified DNA fragments and plasmid vectors.	28
Fig 8. Gel used for monitoring ligation efficiency.	28
Fig 9. Colony Blot Hybridization of <i>Cla</i> I recombinants with gene specific probe.	29
Fig 10. Spot Blot Hybridization of recombinant clones with gene specific probe.	30
Fig 11. Gel showing size comparison of inserts with their counterpart bands from the cosmid.	31
Fig. 12. Physical map of the four clones showing the relative size of the clones and the restriction sites of the enzymes used in the restriction analysis.	32
Fig. 13. Western blot of the four clones.	33

	page
Fig. 14. Restriction enzyme <i>Pst</i> I digests of colonies showing opposite orientation of the inserts.	34
Fig. 15. Western Blot of the selected representative sub clones with inserts in opposite orientation.	35
Fig. 16. Nucleic acid sequence of the 34 kd gene DNA of <i>M. leprae</i> showing the translation of the 34 kd ORF.3	38
Fig. 17. Alignment of the nucleotide sequences of 34 kd <i>M. leprae</i> gene and the 34 kd <i>M. tuberculosis</i> gene.	40
Fig. 18. Alignment of the deduced amino acid sequences of the 34 kd <i>M. leprae</i> and <i>M.tuberculosis</i> genes.	40

### **List of Tables.**

Table 1. Distribution of registered cases of leprosy by WHO region 1991.	3
Table 2. The immunological & clinical features of the leprosy spectrum.	6
Table 3. Fully or partially characterized <i>M. leprae</i> protein antigens.	11
Table 4. Summary of the clones selected for characterization of the 34 kd <i>M. leprae</i> gene.	31

### **List of Appendices.**

Appendix 1. $\lambda$ DNA Hind III Molecular weight marker.	59
---	----

Acknowledgement.

My heart felt thanks goes to my research advisors, Dr. Peter W.M. Hermans, for introducing me to the field of molecular biology as well as for his consistent advice and encouragement; and to Dr. Berehanu A. Gashe for his invaluable advice during my work and for his critical review of my thesis.

I am grateful to SAREC and the Armauer Hansen Research Institute (AHRI) & its sponsors, NORAD & SIDA, for funding my project. I am specially grateful to AHRI for allowing me to enjoy their superb laboratory facility. I have enjoyed the open atmosphere at AHRI and the cooperation of practically every member of the AHRI staff.

My special thanks goes to: Dr. P. W.M. Hermans for introducing to molecular biological techniques; Ato Ayenew Nurilgn for introducing me to immunological and photographic techniques; Dr. Victor Kuteyi for his indispensable technical assistance; Ato Amha-Tsion G/Hiwot and W/o Asrat Tilahun for introducing me to the use of computers.

My sincere thanks also goes to W/t Haimanot Gebre-Egziabher, Ato Demissew Beyene, Dr. Mimi Gubere-Xavier and W/t Azeb Tadesse for their constant encouragement during my research work.

My sister Workye Abebe and my friends, Eyualem Abebe, Haile-Melekot Alemu and Amelework Alemayehu, merit my deepest appreciation and thanks for giving me the incentive to engage in M.Sc. study and for their help and advice throughout the M.Sc. program.

## ABSTRACT.

An *M. leprae*::pHC79 cosmid genomic library was screened for the identification, isolation and characterization of the 34 kd antigen gene of *Mycobacterium leprae* using its homologue, the 34 kd antigen gene of *Mycobacterium tuberculosis*, as a DNA probe. Out of a total of 79 different cosmids screened, covering about 88% of the genome of *M. leprae*, one cosmid designated as BL 1527 was found to harbor the gene of interest. Restriction enzyme analysis of the cosmid BL 1527 yielded four DNA fragments - two containing the complete gene and two containing part of the gene. The four DNA segments were subcloned in plasmid vectors pUCBM21 and pBKS+ and characterized physically and immunologically. Physical characterization of the gene by nucleotide sequencing revealed that the 34 kd antigen gene has an open reading frame of 801 nucleotides encoding 266 amino acid residues bearing a molecular weight of 28,887 daltons and an isoelectric point of 4.3. Homology comparison with the 34 kd antigen gene from *M. tuberculosis* has revealed 80.6 % (633 nucleotides) identity at the DNA level and 85 % (221 amino acids) identity at the protein level. Immunological characterization by western blot analysis has shown that the gene is expressed in *E. coli* from its own promoter signals and that the recombinant protein is recognized by the monoclonal antibody F-126.2 which is specific for the mycobacterial 34 kd antigens. Crossed Immuno-electrophoresis analysis has established that the recombinant 34 kd antigen corresponds to the non secreted Antigen 84 of mycobacteria.

## 2. INTRODUCTION

### 2.1 THE DISEASE:

The word leprosy, charged with revulsion and horror in every society, describes the chronic infectious disease caused by the obligate intra cellular pathogen, *Mycobacterium leprae*. Although a number of other organs like the testis, eyes, liver etc. can be affected in advanced lepromatous cases, leprosy is essentially a disease of the skin and nerves (Pfaltzgraf & Bryceson, 1985; Job, 1971, 1983).

**2.2. Distribution and prevalence:** Although leprosy is eradicated in North America, Western and Northern Europe (except Iceland) it is still endemic at a very low level in Eastern and Southern Europe and the Cajun community in Louisiana, USA. On the contrary, leprosy remains to be a serious public health problem in most parts of Asia, Africa, Central & Latin America (Bryceson & Pfaltzgraf, 1990; Nordeen, *et al.*, 1992). At present leprosy have a world wide tropical distribution affecting in particular people of low standard of living and those suffering from poverty as well as over crowded conditions of life (Bryceson & Pfaltzgraf, 1990). The over all picture of the distribution pattern of leprosy is shown on Table 1.

The current estimate of the WHO (Nordeen, *et al.*, 1991) puts the estimated number of active leprosy cases in need of treatment at 5.5 million in contrast to the 10 - 12 million estimated in the early 1980's. This reduction in active cases of leprosy is due to the success achieved by the implementation of multi drug therapy (MDT) (Nordeen *et al.*, 1992).

WHO regions	Estimated cases (*)	Registered cases	Prevalence per 10,000	% of total
Africa	916,000(46)	482,669	9.2	12.91
Americas	391,000 (40)	301,704	4.2	8.08
S.E.Asia	3,750,000(9)	2,693,104	20.5	72.06
Europe	9,000 (14)	7,246	0.1	0.19
Eastern Mediterranean	207,000 (21)	99,913	2.6	2.67
Western Pacific	238,000 (22)	152,739	1.0	4.09
World	5,511,000 (152)	3,737,375	7.1	100.0

Table 1. Distribution of registered cases of leprosy by WHO region 1991. (Adapted from Nordeen, *et al.*, 1991 & 1992). \* Numbers in bracket = No of countries.

### 2.3. LEPROSY IN ETHIOPIA:

Leprosy is considered as one of the major health problems in Ethiopia, the higher incidence of the disease occurring in the highlands (Berhe, *et al.*, 1990). MDT was introduced in 1983 and after 5 years, prevalence has been reduced to 75 %. The number of estimated leprosy cases in Ethiopia is around 60,000 and in 1991, the registered cases were 31, 753 while the prevalence and detection rates were estimated at 6.8 and 1.1 per 10,000 persons, respectively (Daumerie, 1991).

### 2.4. Transmission, Reservoirs & Hosts:

The mode of transmission of leprosy is not established definitely. The available bits of evidences, however, implicate the upper respiratory tract as the most likely route.

Although, a number of reports (see Meyers, *et al.*, 1992) have established

that there are known extra human reservoirs of *M. leprae* in three animal species, the epidemiological evidences available are in favor of suggesting that human to human transmission is the most important, if not the sole, source of infection for man.

### **2.5. Clinical and immunological aspects of leprosy:**

Although Leprosy is principally a disease of man, the majority of the species (95 - 99 %) are refractory to the establishment of the disease in spite of their close contact to the bacilli (Newell, 1966) and clear immunological evidences that they have been actually exposed to the pathogen (Myrvang, *et al.*, 1975). Among the susceptible 1 - 5% individuals again the severity of the disease differs significantly and the incubation time required for clinical symptoms to appear may take from two months to 10 years or more. The clinical manifestations of the disease, which are the outcome of the degree of resistance mounted by the victim of infection, are depicted as a spectrum that ranges from a single, localized lesion affecting a small patch of skin (in the tuberculoid pole) to widely disseminated infiltration of many organs and occurrence of profuse skin lesions ( at the lepromatous pole) with a number of intermediate forms (the border line cases) occurring in between.

The Ridley Jopling system of classification (Ridley and Jopling, 1966) which reflects the spectral pattern of the disease, assigns leprosy to one of five classes as: LL, polar lepromatous leprosy; BL, border line lepromatous leprosy; BB, mid border line leprosy; BT, border line tuberculoid leprosy; TT, polar tuberculoid leprosy.

It has long been recognized that the position of a particular case of leprosy along the spectrum is determined by the out come of the immunological scenario that follow the individuals reaction to *M. leprae* and/or its released antigens (Ridley & Jopling, 1966). There is a clear evidence that indicate the existence of a strong inverse correlation of CMI and bacillary multiplication suggesting that the level of CMI determines the clinico-pathological manifestations of the disease (Rees & Weddle, 1968; Bloom, *et al.*, 1989).

Various immunological tests like the lepromin test, lymphocyte transformation tests, macrophage migration inhibition tests etc. have established beyond doubt that CMI and DTH are well developed in the tuberculoid pole where as they are absent in the lepromatous pole (Bjune, *et al.*, 1976; Bullock, 1978; Haregewoin, *et al.*, 1983 & 1984; Gills & Godal, 1986; Converse, *et al.*, 1988; Ottenhoff, *et al.*, 1989).

Interestingly the cellular infiltration of nerve granulomas or lesions mimic that of the skin granuloma and lesions (Turk, *et al.*, 1991). The most prominent granulomatous and lesional features of the classes in the leprosy spectrum are summarized in Table 2.

Immunological stability is one of the most crucial aspects of the leprosy spectrum as sudden shifts in the immunological balance account for many of the fatal complication of the disease (Scott, *et al.*, 1976, Klenerman, 1987) as well as spontaneous healing. Among all the classes of leprosy, BB is the most immunologically unstable state and stability increases as the disease moves towards either pole.

	<i>TTp</i>	<i>TTs</i>	<i>BT</i>	<i>BB</i>	<i>BL</i>	<i>LLs</i>	<i>LLp</i>
Epithelioid cells, mature	++	++	±	-	-	-	-
Epithelioid cells, immature	+	+	++	++	±/-	-	-
Langhans' giant cells, large	±/-	++	±/-	-	-	-	-
Macrophages	-	-	-	±	++	++	++
Lymphocytes	+++	+/-	++/±	±	++	+/-	±
Dermal nerves, maximum diameter (µm)	200	1000	400	250	200	200	80
Clear subepidermal zone	±/-	+/-	++/-	++	++	++	++
Erosion of epidermis	±/-	++/-	±/-	-	-	-	-
Bacterial index of granuloma	0	0/1	0/2½	3/4½	4/5½	5/6½	5½/6½
LIB, percentage fall in six months	...	...	100	78	23	14	5.5
Bacilli in nasal smears	-	-	-	-	±	+	++
Lepromin, Fernandez reaction	+++	++/-	++/-	+/-	-	-	-
Lepromin, Mitsuda reaction	+++	++±	++/+	-	-	-	-
LTT, percentage lymphocyte transformation	16	10	5.7	2.0	0.4	0.3	0.2
Leucocyte migration index	0.76	0.78	0.84	0.89	0.92	0.92	0.96
Antibody, anti- <i>M. leprae</i>	-/+	-/+	-/+	++	+++	+++	+++
Immunological stability	+++	+++	+	-	+	++	+++
Borderline (DH) reactions	-	+	++	+++	+	+	-
Erythema nodosum leprosum	-	-	-	-	±	++	++
Distribution of cases, approximate	4	5	31	8	13	29	10

Table 2. The clinical & immuno-histopathological features of the leprosy spectrum. (Taken from Ridley, 1988)

## 2.6. THE PATHOGEN.

*Mycobacterium leprae*, the etiologic agent of Hansen's disease or Leprosy, has proved to be the most enigmatic and paradoxical micro-organism. Enigmatic; because, in spite of intensive biochemical research, many of its properties remain obscure and when they are established they often show peculiar features that distinguish *M. leprae* from other mycobacteria. Paradoxical; because, although it was the first micro organism to be implicated as the cause of a human disease (Hansen, 1874), [discovered a decade earlier than the tubercle bacillus by Koch in 1882 (Bryceson and Pfaltzgraf, 1990)] with the successful cultivation of *Treponema pallidum* it admittedly remains to be the last of human pathogenic bacteria yet to be cultivated on artificial media (Hutchinson, 1987).

### 2.6.1. Classification & morphology.

**Classification:** Despite some atypical lineaments, *M. leprae* belongs to the family *Mycobacteriaceae* in the order Actinomycetales (Runyon, Wayne & Kubica, 1974). The mycobacteria are characterized by being non-spore forming, Gram positive bacilli often acid and alcohol fast after staining with carbol fuchsin and with an optimum growth temperature range of 28 - 45°C (Shepard, 1965).

**Morphology:** Electron and Light microscopy studies of the leprosy bacillus, both on stained and un-stained preparations, reveals it as a straight or slightly curved rod-shaped bacterium with parallel sides and rounded ends that range in size between 1-8µm length and 0.3µm diameter. The bacilli are often enhanced with irregular beaded appearance (Edwards, 1970; Draper, 1982 & 1983; Hirata, 1985a & 1985b) and divide by binary fission (Edwards, 1970; Fukunishi, 1985). In infected tissue, the bacilli are most often stacked or clumped together to form globi.

### 2.6.2. GROWTH REQUIREMENTS & OTHER GENERAL FEATURES:

Since *M. leprae* has defied all attempts to cultivate it in vitro, all our knowledge of the bacteria has to be derived from assays done on its limited growth in mouse foot pad (Shephard, 1960, 1965) from manipulations performed on bacteria purified from armadillo tissue (Kirchheimer & Storrs, 1971) and bacteria obtained from patients (Abe, 1970; Cocito & Delville, 1985; Hirata, 1985a) as well as from monitoring the onset and progress of the disease on patients.

*M. leprae* has a predilection for cooler sites in the human body reflecting its low optimum growth temperature. This fact, evidently, has led to the choice of the

mouse foot pad which maintains an average temperature of 30°C (Shepard, 1960) and the nine banded armadillo which has a core temperature of 31 - 35°C (Kirchheimer & Storrs, 1971) as experimental animals. In the former, the optimum multiplication was observed at around 30°C and was markedly reduced at 36°C (Shepard, 1965; Levy, 1976). The generation time of the bacillus is one of its features known with least certainty. In mouse foot pad assays, the doubling time was estimated to be about 11 days (Levy, 1976) in marked contrast to the 26 h average in nude mice estimated by Hastings & Morales (1982).

**2.6.3. MOLECULAR BIOLOGY OF *M. leprae*:** Prior to the early 1980's, research efforts directed at prying open the biochemical and immunological mysteries of *M. leprae* were hindered by the inability to grow the micro-organism *in vitro*. As a result, the identification of *M. leprae* antigenic determinants were confined to the use of extracts of the bacilli derived from human biopsies (Abe, *et al.*, 1970) and bacilli obtained from experimental animals like the nine banded armadillo, normal and nude mice (Closs, *et al.*, 1979; Ehrenberg & Gebre, 1987; Britton, *et al.*, 1988). Thus a limited number of antigenic determinants were identified by employing methods such as indirect fluorescent antibody test (Abe, *et al.*, 1972) immune-electrophoresis (Kronvall, *et al.*, 1975; Britton, *et al.*, 1988) SDS-PAGE (Klatser, *et al.*, 1984) and immunodiffusion (Abe, *et al.*, 1970).

However, the mere identification of antigenic determinants was not sufficient as ways and means to harness the immunological potential of the identified antigens were equally important for detailed analysis of the protective or pathologic

implication of the individual antigens. Considering the notoriously difficult task of purifying individual antigens from a complex mixture of antigens that result from cell extracts of *M. leprae* (Daniel & Janicki, 1978), the situation is even more aggravated.

Thus, the introduction of molecular and immuno-biological techniques in the early 1980's, notably the introduction of the Polymerase Chain Reaction (PCR) techniques, application of monoclonal antibody (MoAb), and recombinant DNA technologies; heralded not only better means and ways for identification of novel antigens but also afforded improved methods for their isolation and purification. The introduction of these novel techniques has also opened new vistas in the *in vitro* manipulation of the cloned antigen genes i.e in studying the outcome of the deletion, addition and modification of the genes encoding them. These advances in effect breached the barrier imposed by the failure of *in vitro* cultivation of the bacillus. The introduction of PCR techniques is proving useful in detecting *M. leprae* specifically even when it is present as scantily as a single bacillus per sample (Hartskeerl, *et al.*, 1989; Williams, *et al.*, 1990) and is being polished so as to have application in determining the viability of the organisms detected as such (Woods & Cole, 1990).

MoAb's, which are produced by immunizing laboratory mice with various forms of mycobacterial antigens and subsequently fusing them with mouse myeloma cells (de St.Groth, *et al.*, 1980; Seckle, 1985; Young, 1991), are being applied successfully in mycobacterial taxonomy, diagnosis, epidemiology and immunological detection of mycobacterial antigens, (Kolk, *et al.*, 1988).

At present, recombinant DNA technology (see Huynh, *et al.*, 1985; and Maniatis, *et al.*, 1982, Sambrook, *et al.*, 1989) which at the out set proved its versatility by allowing isolation of immunodominant antigens, is being employed to identify, characterize and produce other *M. leprae* components such as key metabolic and regulatory (eg. transcriptional and translational) enzyme targets for use in screening potential blocking agents as well as for applied purposes as in the development of immuno-diagnostic and molecular epidemiological tools (Hopwood, *et al.*, 1988; Kolk, *et al.*, 1988; Young, 1988; Colston & Lamb, 1989).

The genome of *M. leprae* which is estimated to have the size of  $2.2 \times 10^9$  daltons and G+C content of 56 % (Clark-Curtiss, *et al.*, 1985) has been manipulated by recombinant DNA techniques and at present genomic expression libraries constructed both in bacteriophages such as  $\lambda$ gt11 (Young & Davis, 1985) and cosmid libraries, constructed using Cosmid pHc79 (Clark-Curtiss, *et al.*, 1985) and using Cosmid Lorist-6 (Eiglmeier, *et al.*, 1993) do exist. These libraries are used for the purpose of identifying antigens relevant to the immunopathology of leprosy. The libraries are screened by T cell clones (Villarreal-Ramos, *et al.*, 1991) monoclonal antibodies (Young, *et al.*, 1985), patient sera (Cherayil & Young, 1988; Satish, *et al.*, 1990), healthy contact sera (Hartskeerl, *et al.*, 1990) and DNA probes (this work) as it suits each investigator. In this way many antigens have been (and are being) isolated and characterized with respect to their physical and immunological properties (Summarized in table 3).

Table 3. Fully or partially characterized *M. leprae* protein antigens<sup>(A)</sup>.

No	Subunit-KD	Other names	Structural and functional feature.	Immunological characteristics	Ref <sup>B</sup>
1.	70 KD	Dna K	HSP, role in protein folding and translocation. >50% homology with <i>E. coli</i> DnaK and Human HSP 70	Antibody response in immune mice & man, proliferative T cell responses in patients & controls. Potential target for autoimmunity	1,2
2.	65 kd	Gro EL CIE Ag82	HSP, role in protein folding and translocation, >50% homology with <i>E. coli</i> Gro EL and Human HSP 60, functions with Gro ES	Antibody response in mouse & man, proliferative & cytotoxic T cell response in patients & controls, multiple peptide epitopes mapped, autoreactive responses in rodents and humans.	1,2
3.	65 KD	2 <sup>nd</sup> Gro EL	Similar to first Gro EL above.	Similar as above.	5
4.	28 kd	Sod A	Superoxide dismutase. >50% homology with <i>E. coli</i> Sod A.	Recognized by monoclonal antibodies F116.5 & D2D.	1,2
5.	14 kd	Gro ES, MCP1.	HSP, role in protein folding and translocation, functions with Gro EL, extensive homology with <i>E. coli</i> Gro ES.	Recognized by monoclonal antibodies, induces strong proliferative T cell responses in tuberculoid leprosy and in lepromin test.	1,2
6.	18 KD	L5	Not identified.	Antibody response in mouse & man, proliferative T cell responses in patients and contacts.	1,2
7.	36 KD	proline rich ag.	Not identified.	Antibody responses in patients, proliferative & suppressive T cell response	1,2
8.	28 kd		Iron regulated protein, potential signal sequence.	Antibody response in lepromatous leprosy.	3
9.	18 KD		Member of small HSP family.	Antibody response in mouse & man, proliferative T cell responses.	1,2
10	15 KD		Expressed as free protein.	Antibody & T cell responses in leprosy patients & contacts.	4
11	12 KD			Recognized by monoclonal antibodies	1
12	10 KD	2 <sup>nd</sup> Gro ES	As Gro ES above	As Gro ES above	5

A. In addition to the ones listed above, a number of *M. leprae* antigens; the 85 complex, the 45, 35, 30/31, 30/33, and 12 KD, are under investigation with respect to their structural and immunological properties.

References: 1. In Engers, *et al.*, 1986; 2. In Young, *et al.*, 1992. 3. Cherayil & Young, 1988.

4. Sela, *et al.*, 1991, 5. Rinke de Wit, *et al.*, 1992. B. Reference 1 & 2 are compilations of papers.

The overall objective of such antigen searches is to look for potent antigens that (at least in combination) might result in producing reliable immuno-diagnostic reagents and candidate vaccine components which, in the long run, may be used to develop a vaccine for the disease.

By far the most important land mark in the short history of the molecular biology of *M. leprae* is the Genome Project initiated and being pursued by the group of Stewart Cole in Paris (Honore, *et al.*, 1993).

Initial results of the genome project (Eiglmeier, *et al.*, 1993) have generated substantial volume of information about the *M. leprae* genome with respect to, precise localization of genes and partial sequences, the nature of the regulatory apparatus in transcription and translation, G + C content, codon usage, etc.. These findings would undoubtedly have applications in improving existing expression vectors or in the development of new ones for efficient expression of *M. leprae* genes or gene clusters comprising entire metabolic pathways.

It is expected that by the end of 1993, 30 % of the genome will be sequenced (S.Cole, personal communication). The completion of the project would usher the era of surrogate molecular biology of *M. leprae*. That is, scientists interested in the study of the pathogen will focus on the manipulation of genes already sequenced by the project and be limited to doing applied research on the consequences of their deletion, addition, modification, transfer to and from between *M. leprae* and other bacteria including mycobacteria.

#### 2.6.4. The aim of the project.

Earlier in 1990 the *M. tuberculosis* 34 kd antigen encoding gene was identified by screening a  $\lambda$ gt 11 genomic expression library with monoclonal antibody F-126.2. Later on the presence of a homologous gene in all other mycobacteria was established by performing Southern hybridization on genomic DNA of several mycobacteria including *M. leprae* (Hermans, manuscript in preparation).

Western blot analysis of *M. tuberculosis* and other mycobacterial lysates using tuberculosis patient sera have indicated that there is strong antibody reaction to protein antigens located around the 34 kd area indicating that this group of antigens are among those that are recognized by the immune system of human patients (A.H.J. Kolk, personal communications). Further more the 34 kd antigens of the mycobacteria were shown to correspond to the non secreted antigen 84 by crossed immuno-electrophoresis (CIE)(M. Harboe, personal communication).

As *M. leprae* is of medical importance, the isolation of its 34 kd gene which was also shown to be immunogenic by its reaction with patient sera, was found to be relevant to the study of the antigenic repertoire of the bacillus. The decision to carry out the isolation and characterization of the *M. leprae* 34 kd gene culminated in the realization of this work.

Thus, the aim of this thesis work was to screen a genomic library of *Mycobacterium leprae* constructed in the vector pHc79 using the *Mycobacterium*

*tuberculosis* 34 kilo dalton antigen encoding gene, as a DNA probe in order to be able:

1. to identify, isolate, and clone the 34 kd antigen encoding gene of *Mycobacterium leprae* into plasmid vectors for characterization,
2. to characterize the gene by DNA sequencing,
3. to determine the homology of the *M. leprae* 34 kd gene at the level of the nucleotide and amino acid sequences with the 34 kd gene of *M. tuberculosis*.
4. to investigate the expression of the recombinant 34 kd *M. leprae* protein in *Escherichia coli* and,
5. to test the reactivity of the recombinant protein encoded by the gene with the monoclonal antibody F-126.2 which recognize the 34 kd antigen of mycobacteria,

### 3. MATERIALS AND METHODS.

#### 3. 1. Recombinant DNA library and Bacterial Strains.

**Recombinant DNA library:** The recombinant DNA library used in this thesis project was derived from the original Cosmid library constructed by Clark-Curtiss, *et al.*, (1985). It was made available to our use by courtesy of Dr. J.E.R. Thole (Leiden University hospital, Department of Immunohematology & Blood Bank, Leiden, The Netherlands) as Southern Blots on two Nylon membranes. The two blots carry *Pst* I digests of 79 different cosmids that were chosen on the basis of their difference in restriction patterns.

**Bacterial Strains.** The bacterial strains used are:

- i. *E. coli* XL1-Blue [F'::Tn10 *proA*<sup>+</sup>*B*<sup>+</sup> *lacI*<sup>n</sup>  $\Delta$ (*lacZ*)M15/*recA1 endA1 gyrA96 (Nal*<sup>r</sup>) *thi hsdR17 (r*<sub>k</sub><sup>-</sup>*m*<sub>k</sub><sup>+</sup>) *supE44 relA1 lac*]
- ii. *E. coli* DH5 $\alpha$ F' [F'/*endA1 hsdR17 (r*<sub>k</sub><sup>-</sup>*m*<sub>k</sub><sup>+</sup>)*supE44 thi-1 recA gyrA(Nal*<sup>r</sup>) *relA1  $\Delta$ (lacZYA-argF)U169 (phi80 $\Delta$  (lacZ) M15)*]

#### 3.2. Plasmid vectors used for subcloning:

- i. pUC BM21: *Col E1 Origin, Lac Z, Amp*<sup>r</sup> (Boehringer Mannheim, GMBH. 1992 Catalogue).
- ii. pBluescript KS+ (pBKS+): *Col E1 Origin, Fl Origin, Lac Z, Amp*<sup>r</sup> (Stratagene, Paris. 1992 Catalogue).

#### 3.3. Probes used in the experiments:

- i. A 2 kb *Eco* RI DNA fragment of *M. tuberculosis* containing the 34 kd gene rescued from recombinant plasmid pPH 5220. (Courtsey of Dr P.W.M. Hermans).

- ii. Plasmid pBKS+ (Stratagene, Paris).
- iii. Monoclonal Antibody F-126.2 (Courtesy of Dr. A.H.J. Kolk, Royal Tropical Institute, Amsterdam).
- iv. Goat  $\alpha$ -mouse IgG- Alkaline phosphatase conjugate (Sigma Chem. Co)

### 3.4. Media and antibiotics.

The medium used for propagating bacteria was Luria-Bertany (LB) broth [1% w/v Tryptone (Difco) 1% w/v NaCl (Merck) 0.5% w/v Yeast extract (Sigma Chem. Co.) , and LB agar (LB broth , 1.5% Agar (Gibco, Scotland). Antibiotics used were ampicillin (Sigma, Chem. Co) 100  $\mu$ g/ml on plates and 200  $\mu$ g/ml for liquid cultures. Tetracycline (Sigma, Chem. Co) was used at a concentration of 25  $\mu$ g/ml and 50  $\mu$ g/ml for solid and liquid media respectively.

### 3.5. DNA TECHNIQUES.

**Plasmid isolation (Birnboim):** Plasmid isolations were carried out by the Birnboim method following the modified protocol of Sambrook, *et al.*, (1989). Medium scale plasmid isolation from cultures of 30 - 150 ml and maxi plasmid isolation from cultures of 250 - 500 ml cultures were done by the Qiagen column elution using the protocol outlined in the Qiagenologist manual (Qiagen USA) and by CsCl density gradient centrifugation procedures (Sambrook, *et al.*, 1989) respectively.

**Restriction enzyme digestion:** Restriction enzyme digestions were carried out as follows: 1  $\mu$ l (= 1  $\mu$ g) of plasmid or cosmid DNA, 1  $\mu$ l restriction enzyme (2 - 4 units), 2  $\mu$ l of 10X restriction enzyme buffer and 16  $\mu$ l of dd-H<sub>2</sub>O to make a final reaction volume of 20  $\mu$ l. The reactions were carried out at 37°C for 1 hr.

**Ligation of DNA fragments:** Ligation reactions were performed as follows: pre-determined amounts of the two DNA's of interest treated appropriately, 1  $\mu$ l (2.5 units) of T4 DNA ligase (Amersham, Int.), 1  $\mu$ l 10X ligation buffer, dd-H<sub>2</sub>O to make final reaction volume 10  $\mu$ l. The reactions were carried out for 2 - 16 hrs at 16 °C in a Thermostatic Circulator (LKB, Bromma). Aliquots of 1  $\mu$ l were taken before and after ligation to monitor the efficiency of ligation by running them on 1 % agarose gel.

**Transformation of competent *E. coli* cells:** Appropriate strains of *E. coli* cells were made competent before hand by treatment with ice cold 50 mM CaCl<sub>2</sub> and stored at -70°C (Sambrook, *et al.*, 1989). Transformation proceeded as follows: competent cells were thawed on ice, 100  $\mu$ l of the thawed cells was added to plasmid DNA and kept on ice for 1 hr, the cell plasmid mixture was heat shocked for 2 min at 42°C, 500  $\mu$ l of liquid LB medium, pre-equilibrated at 37°C, was added to each tube and incubated at 37°C for 30 min. The transformants were then plated on LB agar supplemented with 100  $\mu$ g/ml ampicillin and containing IPTG (100  $\mu$ m) plus X-Gal (24 ng/ $\mu$ l) and incubated overnight at 37°C for blue/white screening.

**Agarose gel electrophoresis:** DNA of interest, plasmid, cosmid or restriction enzyme digested DNA, were loaded on a 1 % agarose gel mixed with DNA sample buffer and electrophoresed in 1X TAE buffer (2 M Tris-HCl, 5.72 % Glacial Acetic Acid, 50 mM EDTA) and electrophoresed at the desired voltage for varying lengths of time as necessary. The gels were photographed by a Polaroid Land Camera (Brand Manufacturers) .

**Extraction of DNA from agarose gel slices (Gene Clean):** DNA of interest were extracted from gel slices after agarose gel electrophoresis according to the protocol supplied with the Gene Clean Kit (Bio 101, La Jolla, CA.)

**Radioactive labelling:** Probes were labelled by random priming using Amersham Mega Prime DNA labelling Kit (Amersham Int.) following the instruction manual supplied with the kit.

**Southern transfer of DNA:** Gels that were intended for blotting were treated as follows, 5 min UV irradiation, 10 min in 0.25 M HCl, 2 X 20 min in 0.4 M NaOH. Blotting was carried out by HSI-TE80 vacuum blotter (Hoffer Scientific Inst. San Francisco, CA) on a Hybond N+ membrane (Amersham Int.) with 10X SSPE buffer. After blotting the gels are restained in ethidium bromide (EtBr) to see how efficient the blotting was.

**Dot blotting of plasmid DNA:** Plasmids isolated by mini-preps were denatured by boiling for 5 min and immediately cooling in ice water. Then the denatured plasmids were spotted onto Hybond N+ nylon membrane and air dried.

**Colony blotting:** Colony blotting was done following the Colony/Plaque Screen instruction manual supplied by manufacturers (Dupont, USA).

**Southern hybridization:** Southern blots were pre-hybridized at 65°C for 30 min. in 25 ml 5X Denhardt's solution [hybridization buffer(40 mM Tris-Cl pH 7.5, 1 M NaCl, 1 % SDS), 0.1 % BSA, 0.1 % Poly Vinyl Pyrollidone (P.V.P.) and 0.1 % Ficoll-400] in a hybridization incubator (Robbins Scientific). Then labelled DNA probes and Herring Sperm DNA (Serva Biochemica) were denatured by boiling for 5 min. and immediately cooling on ice water. The denatured probe DNA and 100  $\mu$ l of 10 mg/ml Herring Sperm DNA (blocking agent) were added into the pre hybridization mix and hybridization was carried out over night at 65°C. The following morning, the blots were rinsed once with 5X SSPE (0.05 M Na<sub>2</sub>HPO<sub>4</sub>.2H<sub>2</sub>O, 0.9 M 5 mM EDTA, 1 % SDS) and washed in the same buffer for 30 min. at 65°C. After the washing was over, the blots were taken out of the hybridization tubes, placed DNA side up on intensifying screens and wrapped in saran wrap (Dow Chem. co.). The residual radioactivity was measured by a Mini Monitor (Mini Instruments Ltd.) and based on the strength of the signals, the blots were exposed for varying lengths of time on Kodak X-OMAT-AR X ray films (Sigma Chem. co.). The films were processed in Kodak GBX developer and fixer (Sigma Chem. Co).

**Restriction pattern analysis:** DNA chosen for restriction analysis were subjected to restriction enzyme digestion by using the selected enzymes and digested by single or double enzymes in concert. The digests were then loaded on gel and the size of each fragment estimated by comparison with the standard DNA molecular weight marker.

**DNA sequencing:** The sequencing procedure employed was the Sanger deoxy chain termination method (Sanger, 1977) using  $\alpha$ -<sup>35</sup>S-dATP (Amersham, Int.) and a DNA sequencing kit (Promega Biotech., Madison, WI). The products of the sequencing reaction were separated on a 6 % polyacrylamide - 7 % Urea , 1X TBE (89 mM Tris base, 89 mM Boric Acid, 2.5 mM EDTA) gels poured with 0.4 mm flat spacers on 60 cm long glass plates and running them using an BRL vertical electrophoresis apparatus(BRL, USA). After electrophoresis the gels were fixed in 10 % methanol, 10 % glacial acetic acid solution for 45 min rinsed in water and dried for 90 min at 65°C in an oven (Gallenkamp Medium incubator Size C). Finally the dried gels were exposed overnight on either X-OMAT-AR (Sigma Chem. co.) or Hyperfilm  $\beta$ -Max (Amersham Int.) X-ray films in the dark at room temperature for 12 - 16 hrs.. Once the open reading frame was identified, the largest clone i.e. pPF-1000 was used to conduct the sequencing to completion. Oligonucleotides were developed for progressive sequencing on both strands and the sequence results were compiled to give the 34 kd *M. leprae* antigen encoding gene sequence (Fig.16). The nucleotide sequence for both strands were determined separately and the results were analyzed by the PC/Gene Nucleic acid and protein sequence analysis soft ware (IntelliGenetics Inc. Switzerland) and the Beckman Microgenie Sequence analysis program (SciSoft Inc. USA).

### 3.6. IMMUNOLOGICAL TECHNIQUES.

**Induction of clones for expression and cell lysate preparation:** Cell lysates for Western blotting procedure were prepared as follows: 5 ml cultures of each

clone were initiated in the morning and induced by IPTG, final concentration of 24 ng/ml, when they reach the log phase (late afternoon) and allowed to grow overnight. The clones induced as such were concentrated by centrifugation, sonicated in 1 % SDS in the presence of 20 mM Phenyl Methyl Sulfonyl Fluoride (PMSF), in ice bath using a sonifier (Branson Sonic Power Co.) for six periods of 30 seconds at full power, centrifuging after each burst of sonication.

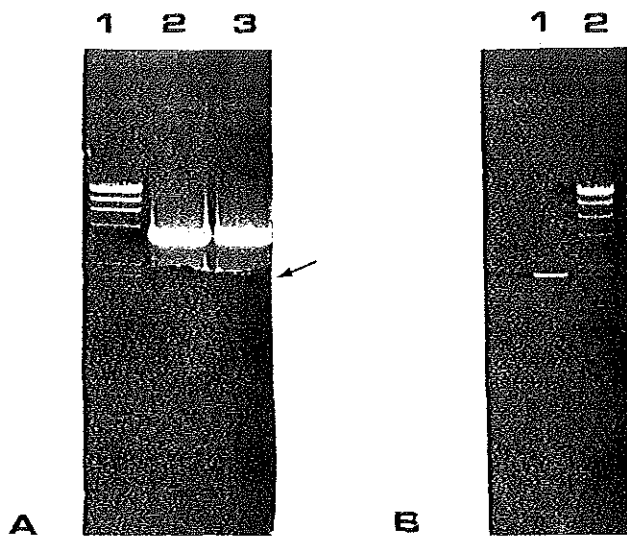
**SDS-PAGE and Western Blotting:** Samples to be blotted were run on a 12 % acrylamide gel (Laemmli, 1970) and electrophoresed at 200 V for 4 hrs using a Protean-II XI Cell vertical electrophoresis apparatus (Biorad, Richmond, CA.) in 1X SDS-PAGE running buffer (2.5 mM Tris base, 1.92 M Glycine, 0.1 % SDS). The gels were blotted on NC paper in a Trans-Blot™ Cell (Biorad) for 2.5 hrs at 36 V in Western blotting buffer (30 mM Tris-Glycine, pH 8.3). The resulting blot was blocked by blocking buffer (40 mM  $\text{KH}_2\text{PO}_4$ , 160 mM  $\text{Na}_2\text{HPO}_4$ , 1.15 M NaCl; pH 7.2, 1 % BSA, 0.5 % Gelatin) for 30 min rinsed and washed by PBS-Tween 3 x 10 min, and incubated with primary antibody, Monoclonal antibody F-126.2 (A.H.J. Kolk), for 1 hr. The blot was rinsed and washed as before and incubated with secondary antibody,  $\alpha$ -mouse IgG Alkaline phosphatase conjugate (Sigma, Chem. Co), for 1 hr. Finally the blot was rinsed and washed for a third time and developed with substrates. The substrates were BCIP (0.5 mg/ml) and NBT (1 mg/ml) . After a few minutes the reaction was stopped with 0.5 M EDTA.

## 4. RESULTS.

### 4.1. SELECTION OF THE 34 kd GENE OF *M. leprae*:

The search for the identification of the 34 kd antigen coding gene of *M. leprae* was carried out by screening a genomic library of *M. leprae* with the homologous *M. tuberculosis* 34 kd antigen gene. The recombinant DNA library screened was derived from the original Cosmid library constructed by Clark-Curtiss, *et al.*, (1985). It was made available for our use as Southern Blots on two Nylon membranes. The two blots carry *Pst* I digests of 79 different cosmids chosen on the basis of their dissimilarity in restriction pattern. Since each of the cosmids contain an insert of 35 - 50 kb *M. leprae* genomic DNA the two blots are estimated to represent more than 88% of the genome.

The DNA probe used in the screening procedure was the previously characterized *M. tuberculosis* 34 kd antigen gene which was made available as recombinant plasmid pPH 5220 - a pNGS20+ (N.Stocker) recombinant containing the gene and its flanking region. The 2 kb *Eco* RI insert DNA was then isolated first by digesting pPH 5220 with the restriction enzyme *Eco* RI and cutting the fragment out of gel and purifying it by Gene clean system (Fig 1. A & B). About 2 micro gram of probe DNA was obtained in this way.



**Fig 1.** Isolation of the probe DNA from gel and its purification by gene clean. A) Lanes; 1,  $\lambda$  DNA *Hind* III marker; 2 & 3, pPH 5220 digested with *Eco* RI, arrow show the 2 kb fragments cut position, B. Lanes; 1, the Gene clean purified probe DNA; 2,  $\lambda$  DNA *Hind* III marker.

Screening of the *M. leprae* genomic library by probing with the *Mycobacterium tuberculosis* 34 kd antigen gene resulted in the detection of the *Mycobacterium leprae* equivalent in cosmid designate BL 1527 (lane 27 of the blot) (Fig. 2).

Once the cosmid of interest was identified, adequate amount of DNA was isolated from it by small scale plasmid isolation procedure. This was later on subjected to restriction enzyme digestion by *Pst* I, the enzyme used originally to construct the library, to confirm, by comparison with the original polaroid picture of the Southern blots (not shown), the identity of cosmid BL 1527 (Fig. 3).

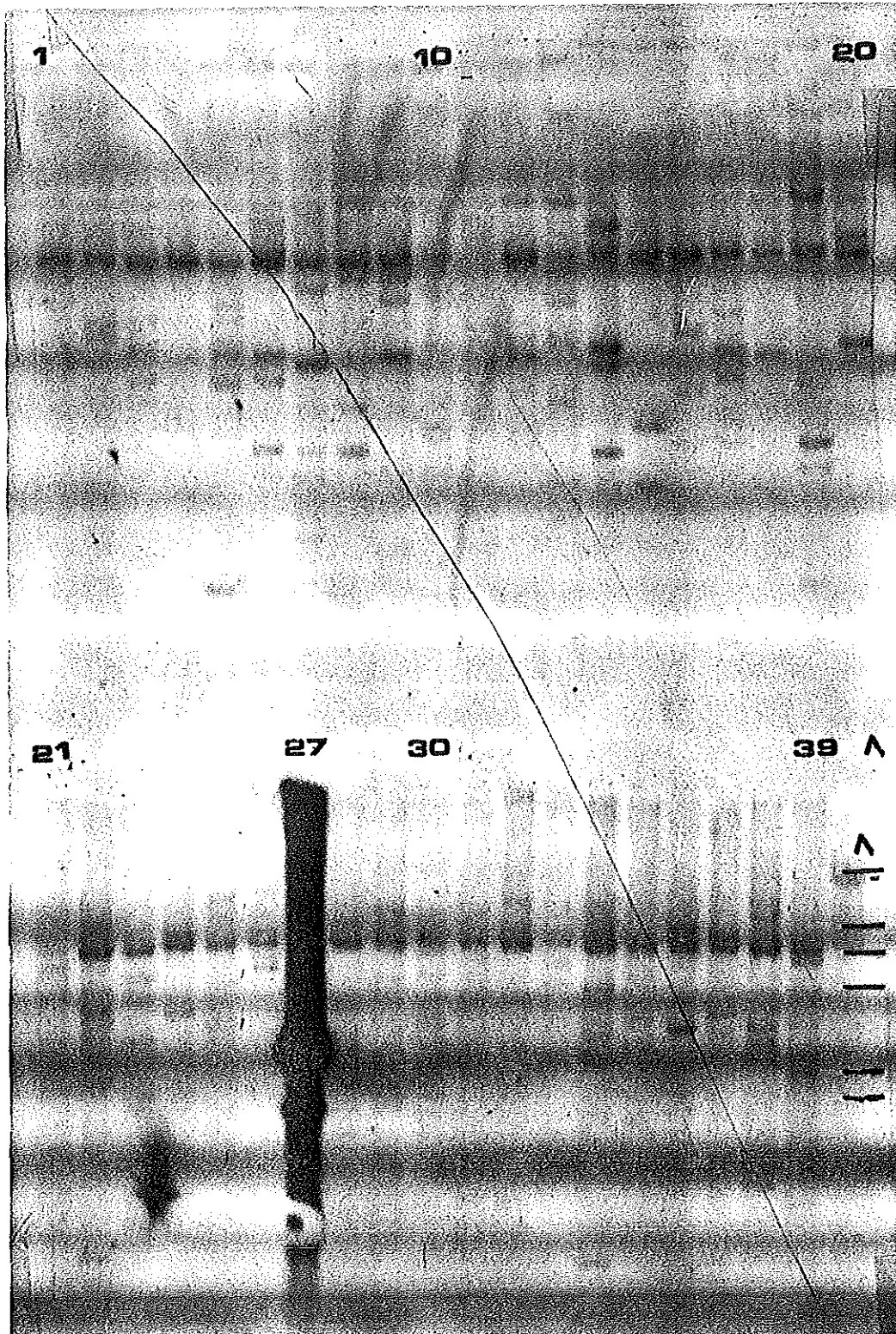


Fig.2. Southern Hybridization of the cosmid library with the probe DNA. Lanes; 1 -39, cosmids BL 1501 - 39, in order of loading.(The other blot not shown).



**Fig 3.** Cosmid isolation and identity confirmation. Lanes: 1.  $\lambda$  DNA *Hind* III Marker 2 & 3 *Pst* I digest of cosmid BL 1527 DNA.

To isolate the gene of interest, the cosmid BL 1527 was digested by 8 restriction enzymes selected on the basis that they cut only once on the vector DNA - with the idea of limiting the number of vector bands to only two so that the rest of the bands that appear will be of insert origin (Fig.4).



**Fig 4 A & B.** Restriction digestion of BL 1527 with the 8 enzymes .**A.** Lanes; 1,  $\lambda$  DNA *Hind* III marker; 2, *Eco* RI; 3, *Cla* I; 4, *Bam* HI; 5, *Sal* I; 6, *Pst* I; 7, *Hind* III; 8, *Ava* I; 9, *Pvu* II. **B.** Duplicate of A.

As the restriction enzyme digests were loaded in duplicate (Fig.4) they were probed with two discriminating probes, the 34 kd *M. tuberculosis* DNA (the Gene Specific Probe) and pBKS+ DNA (the Vector Specific Probe ) after labelling by random priming. One of the blots, designated Blot A was hybridized with the gene specific probe while the second one, designated Blot B, was hybridized with the vector specific probe as described earlier (Fig.5).

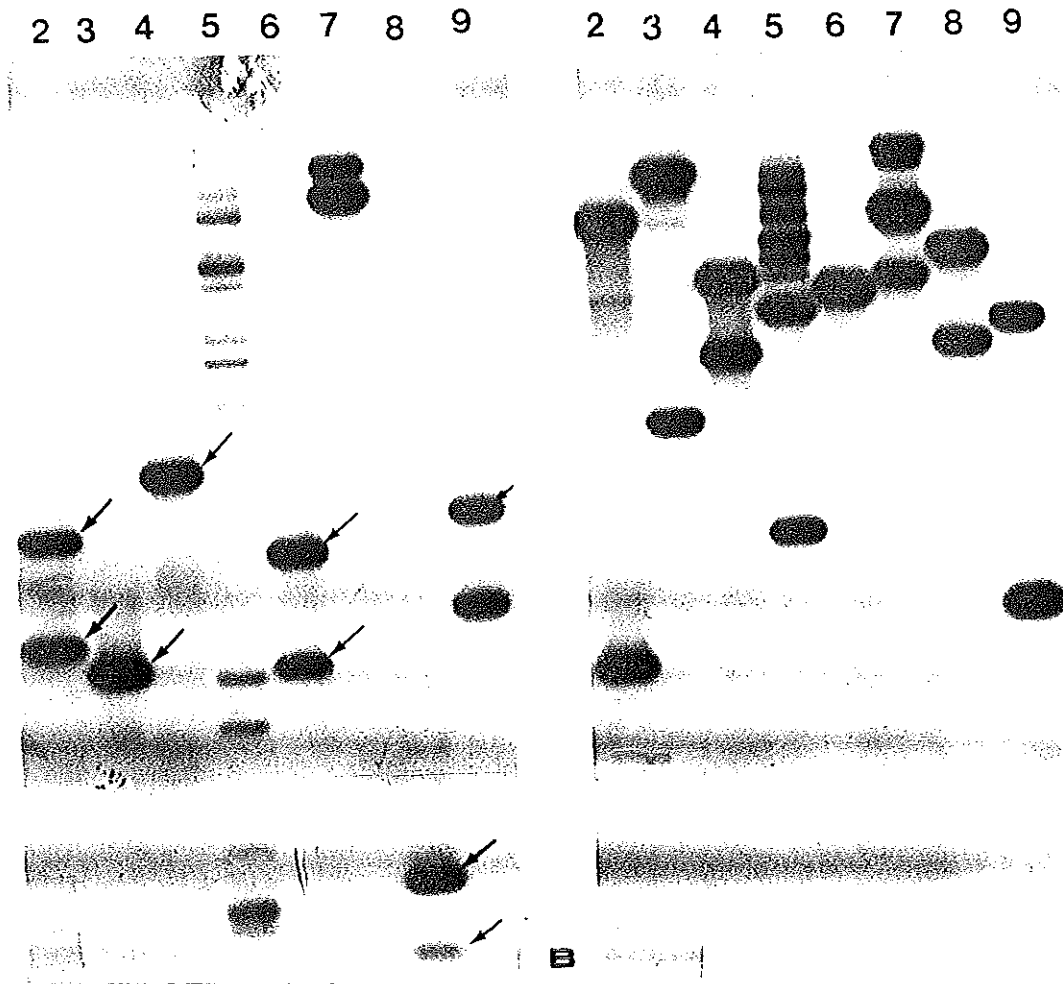
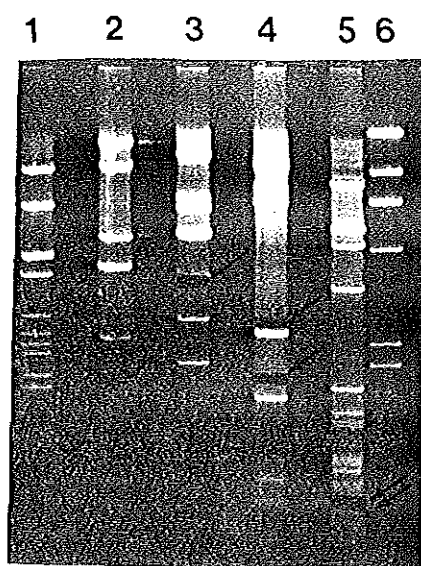


Fig 5. Autoradiograms of the twin blots with discriminating, probes. A. Blot A, hybridized with the gene specific probe, and B). Blot B, hybridized with vector specific probe. (Order of loading as in Fig. 4).

Careful comparison of the twin blots by superimposition on each other allowed the identification of gene specific bands in five lanes: that of *Eco* RI (a 3.8 & 2.1 kb fragments), *Bam* HI (a 3.5 kb fragment), *Cla* I (a 1.9 kb fragment), *Pst* I (a 2.6 & a 2.0 kb fragments) and *Ava* I (a 0.8 & 0.5 kb fragments). In the rest, (*Hind* III, *Pvu* II & *Sal* I) digests, it was difficult to obtain gene specific bands.

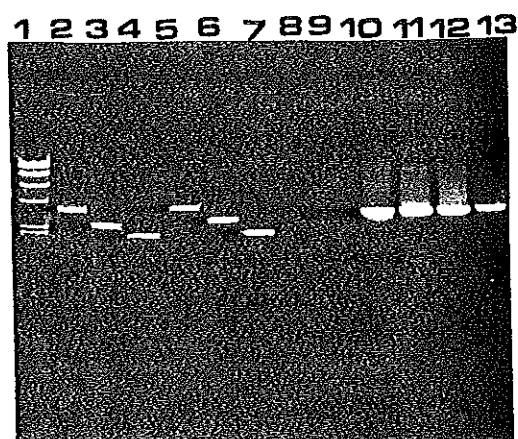
#### 4.2. Isolation and Subcloning of the 34 kd *M. leprae* gene.

Following identification of the gene specific bands, a second restriction enzyme digestion was performed, this time using the restriction enzymes that gave gene specific bands only. Following the digestion of the cosmid and separation of fragments by electrophoresis, the gene specific bands marked earlier from the result of the autoradiograms were cut from the gel and purified by Gene clean (Fig. 6).



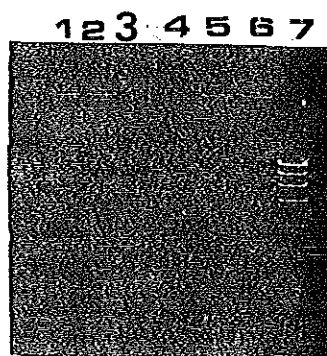
**Fig 6.** Restriction digest of Cosmid BL 1527 with the 5 enzymes for isolation of gene specific bands. Lanes; 1, *Eco* RI; 2, *Cla* I; 3, *Bam* HI; 4, *Pst* I & 5, *Ava* I. 6.  $\lambda$  DNA *Hind* III marker. (Arrows indicate position of band cut out for subcloning).

The purified DNA from each band were loaded on gel together with pUCBM21 and pBKS+ plasmids cut with the restriction enzymes used to generate the DNAs of interest to determine the amount that will be needed for a proper ligation (Fig. 7).



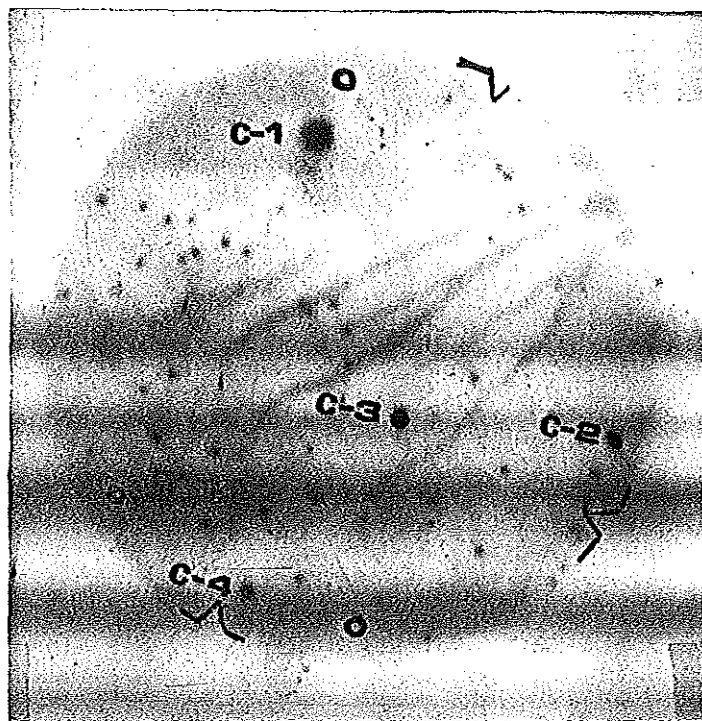
**Fig 7.** Gene clean purified DNA fragments and plasmid vectors. Lanes; 1,  $\lambda$  DNA *Hind* III marker; 2, *Eco* RI 3.8 kb; 3, *Eco* RI, 2.1 kb; 4, *Cla* I 1.9 kb; 5, *Bam* HI, 3.5 kb; 6, *Pst* I, 2.8 kb; 7, *Pst* I, 2.0 kb; 8 & 9 blank; 10, pUCBM21-*Eco* RI; 11, pBKS+-*Cla* I; 12 pUCBM21-*Bam* HI; 13, pUCBM21-*Pst* I.

The bands generated by the enzymes *Eco* RI, *Bam* HI, and *Pst* I were sub-cloned by sticky end ligation process in pUCBM21 where as the one generated by *Cla* I was sub-cloned in pBKS+ in the same way and the efficiency of ligation was monitored by loading 1  $\mu$ l of samples after ligation on gel (Fig. 8).



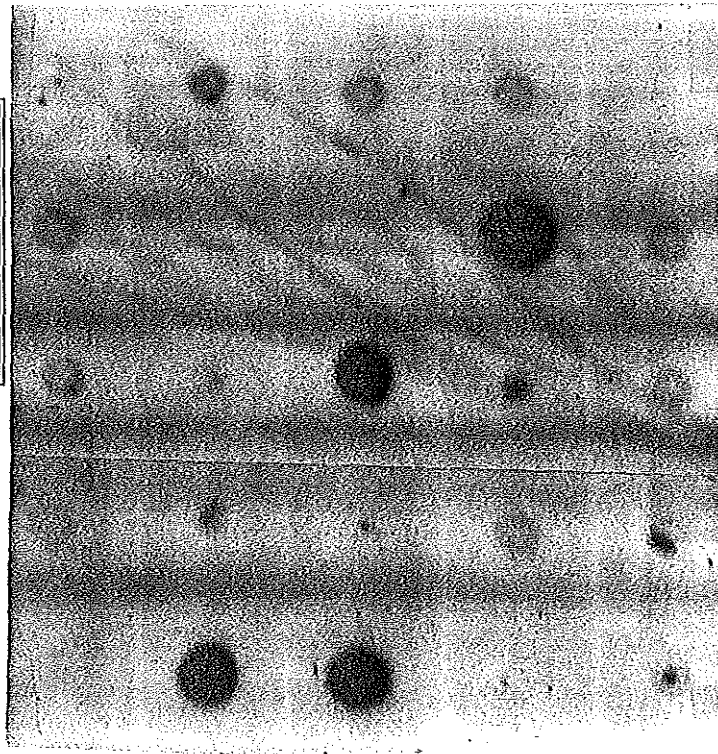
**Fig 8.** Gel used for monitoring ligation efficiency. Lanes; 1, *Eco* RI 3.8 kb; 2, *Eco* RI, 2.1 kb; 3, *Cla* I 1.9 kb; 4, *Bam* HI, 3.5 kb; 5, *Pst* I, 2.8 kb; 6, *Pst* I, 2.0 kb; 7,  $\lambda$  DNA *Hind* III marker.

The recombinant plasmids were then transformed into competent *E. coli* XL1-Blue cells and subjected for blue/white screening. The result showed that in the transformations approximately 95 % were blue and the rest 5 % were white. After transformation and blue/white screening, DNA from five white (recombinant) colonies from each of the pUCBM21 transformations were spot blotted on Hybond membrane where as the pBKS+ transformed colonies were blotted on colony/plaque blot membranes. The two classes of blots were then hybridized with the original gene specific probe, in order to select those transformants that contain gene insert (Fig. 9 & Fig. 10).



**Fig 9.** Colony Blot Hybridization of *Cla* I recombinants with gene specific probe. Numbers indicate the selected positive clones.

1	2	3	4	5
6	7	8	9	10
11	12	13	14	15
16	17	18	19	20
21	22	23	24	25



**Fig 10.** Dot Blot Hybridization of recombinant clones with gene specific probe. Spots: 1-5 *Eco* RI 2.8 kb fragment, 6-10 *Eco* RI 2.1 kb fragment; 11-15 *Bam* HI 3.5 kb fragment; 16-20 *Pst* I, 2.6 kb fragment; 21 - 25 *Pst* I 2.0 kb fragment transformants. (Box on left show order of numbering).

The positives thus selected were finally tested for size conformity with the original bands on the cosmid digested with the respective enzyme used to generate them. This was done by digesting plasmid DNA from all the selected clones with the corresponding restriction enzyme used for their isolation and running them side by side with a digest of cosmid BL 1527 by a similar enzyme (Fig. 11). Four subclones that satisfied the last condition were eventually verified as true positives, given proper designations and their characterization pursued (Table 4).

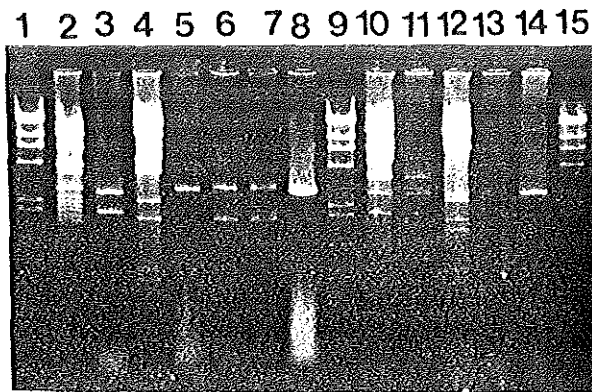


Fig 11. Gel showing size comparison of inserts with their counterpart bands from the cosmid. Lanes 1, 9 & 15  $\lambda$  DNA *Hind* III markers; 2 & 3, BL 1527 & pPF 3000 *Eco* RI digest, 4, BL 1527 & 5,6,7,8 clones pPF 2000, *Cla* I digests; 10, BL 1527 & 11, pPF1000 clone digested with *Bam* HI; 12, BL 1527 13 & 14 pPF 4000 in duplicate digested with *Pst* I. (Numbers on side show size of MW markers).

No	Enzyme used & size of DNA (A)	Plasmid vector	<i>E. coli</i> strain	Name
1.	<i>Bam</i> HI, 3.5 kb (C)	pUCBM21	XL1-Blue	pPF 1000
2.	<i>Cla</i> I, 1.9 kb (C)	pBKS+	XL1-Blue	pPF 2000
3.	<i>Eco</i> RI, 2.1 kb (P)	pUCBM21	XL1-Blue	pPF 3000
4.	<i>Pst</i> I, 2.0 kb (P)	pUCBM21	XL1-Blue	pPF 4000

Table 4. Summary of the clones selected for characterization of the 34 kd *M. leprae* antigen gene. A. (P) = Partial, (C) = Complete gene.

Restriction Pattern analysis was conducted on the four clones selected for characterization, which were derived from the *Bam* HI, *Cla* I, *Eco* RI and the *Pst* I fragments and designated pPF-1000, pPF-2000, pPF-3000, pPF-4000 respectively. They were subjected to restriction enzyme digestion by using the enzymes *Eco* RI and *Pst* I singly or in concert. The digests performed (not shown) were loaded on gel and analyzed carefully. The over all results gave the composite map shown in Fig. 12.

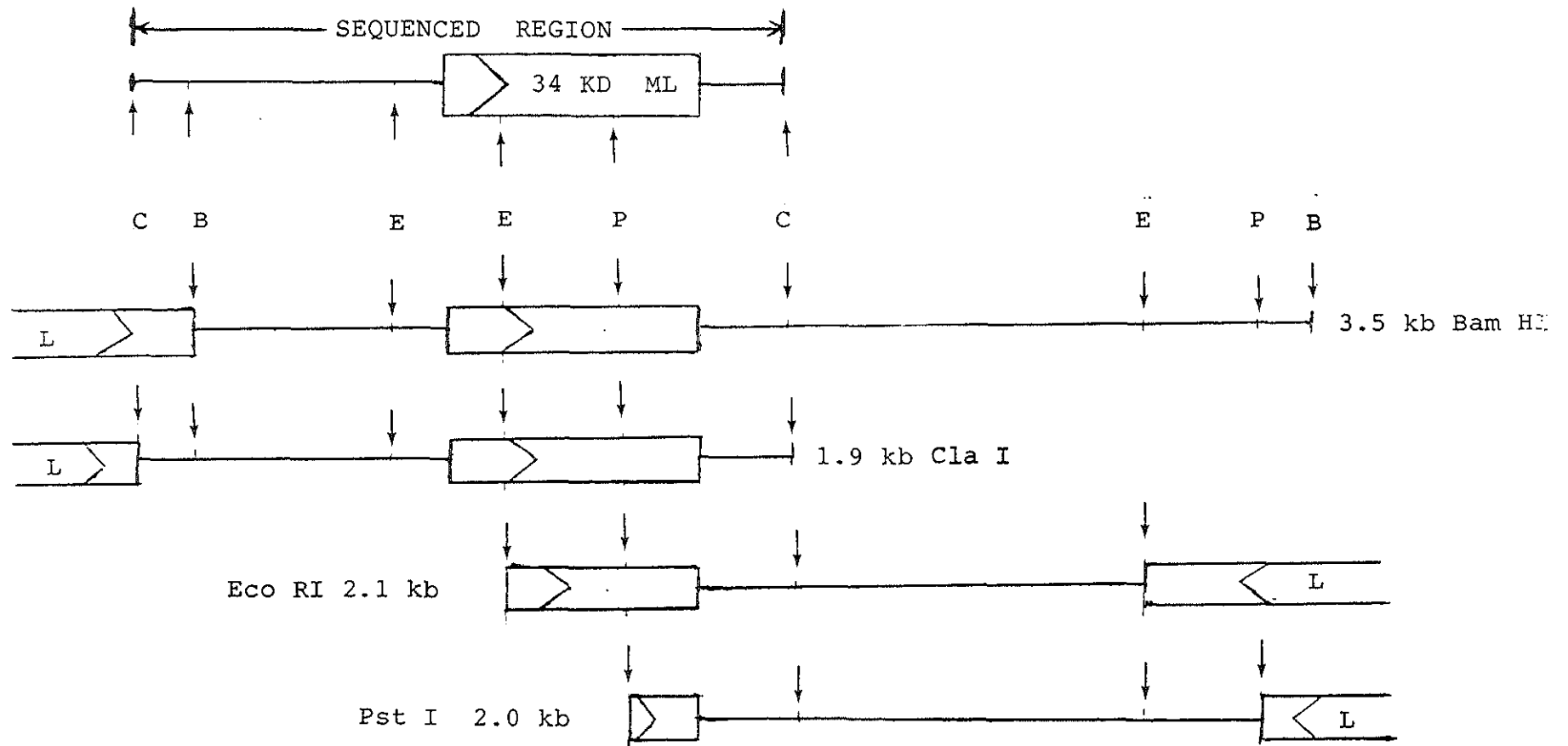
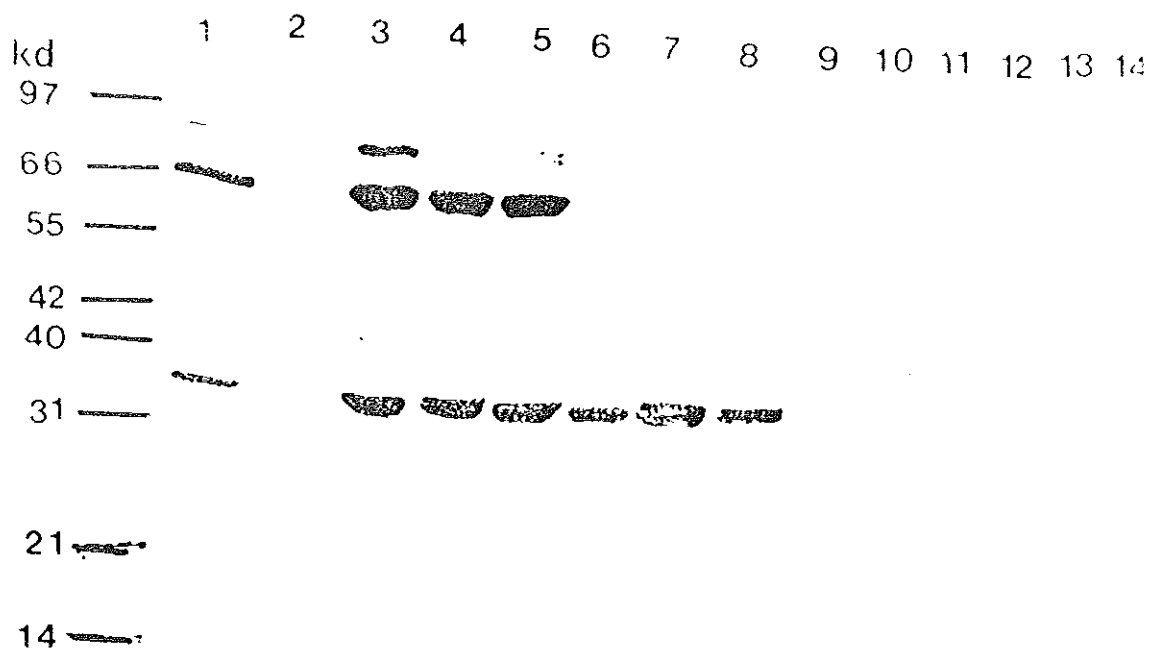


Fig. 12. Physical map of the four clones showing the relative size of the clones and the restriction sites of the enzymes used in the restriction analysis.

### 4.3. Expression of the 34 kd recombinant antigen.

Following the restriction analysis, all the four clones and the cosmid BL 1527 were tested for expression of the target protein by Western blotting procedure. The result showed that the cosmid BL 1527 and two of the clones, pPF 1000 and pPF 2000, expressed the 34 kd antigen in *E. coli*.



**Fig. 13.** Western blot of the four clones. Lanes: 1, Cosmid BL 1527; 2, *E. coli* XL1-Blue lysate (Negative control); 3, 4 & 5, pPF 1000; 6, 7 & 8, pPF 2000; 9, 10, & 11, pPF 3000; 12, 13, & 14, pPF 4000; (Numbers on the side are MW markers).

Since two of the four clones tested, pPF-1000 and pPF-2000, which contain the complete gene were shown to give expression, we were interested to know whether expression was under the strict regulation of the *E. coli*  $\beta$ -Galactosidase gene, down stream of which both were cloned. To this end the two clones were further subcloned with the objective of obtaining the same subclones in opposite orientation. The result of restriction enzyme digestion on 8 randomly selected colonies from both subclonings showed that in both cases, the expected opposite orientations were achieved which allowed selection of representative clones in both orientations (Fig. 14).



**Fig.14.** Restriction enzyme *Pst* I digests of colonies showing opposite orientation of the inserts. A) Lanes; 1,  $\lambda$  *Hind* III Marker; 2, 3 & 4, pPF 1000, 5, 6 & 7, pPF 1020; 8, 9 & 10, pPF 2000; 11, 12 & 13, pPF 2020; 14,  $\lambda$  *Hind* III marker.

The selected clones (both orientations) were tested by Western Blotting. The result showed that there was expression in both orientations in each of the two groups (Fig. 15).

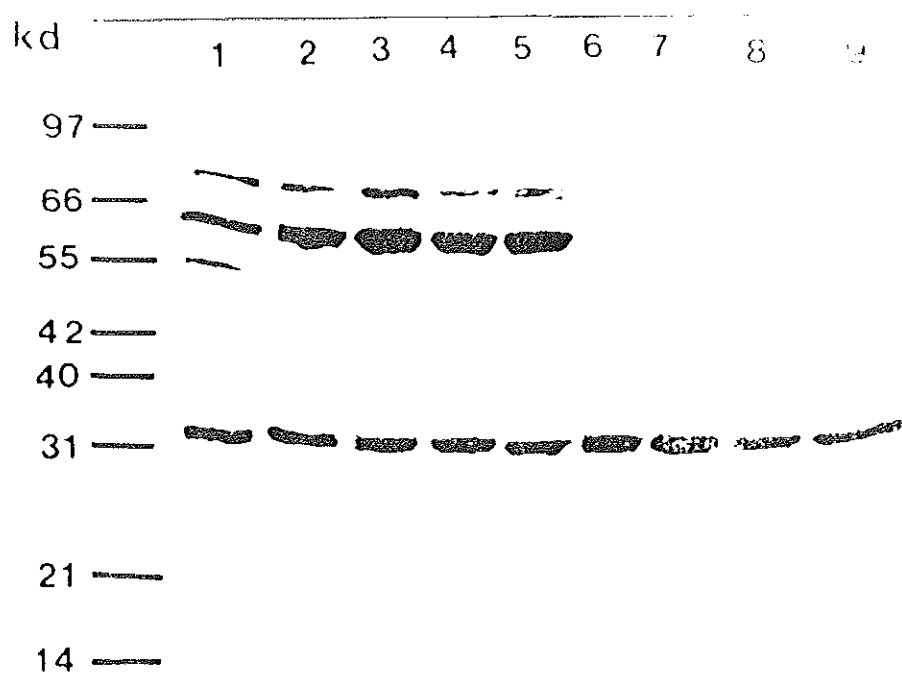


Fig. 15. Western Blot of the selected representative sub clones with inserts in opposite orientation. Lanes:1, BL 1527; 2 & 3, pPF 1000; 4 & 5, pPF 1020; 6 & 7, pPF 2000; 8 & 9, pPF 2020. (Numbers on the side are sizes of MW standards).

#### 4.4. DNA Sequencing of the 34 kd gene of *M.leprae*.

In order to characterize the 34 kd gene of *M. leprae*, sequence analysis was performed. The sequencing results were then compared with the homologous 34 kd gene of *M. tuberculosis* whose sequence has been worked out by Dr. P.W.M. Hermans (Hermans, manuscript in preparation). Totally the sequence of 1884 nucleotides were read. Computer aided analysis revealed that the open reading frame of the 34 kd gene was 801 nucleotides long encoding a protein of 266 amino acids with a calculated MW of 28,887 daltons and bearing an isoelectric point of 4.3. The nucleotide sequence and the predicted amino acid sequences are shown below (Fig. 16).

10            20            30            40            50            60  
 |            |            |            |            |            |  
 ATCGATTTGGTGTGCGATGGACAATGCTGACGCTAAACGCCTGGTTCGATTTGCCGCGGGC  
 70            80            90            100           110           120  
 |            |            |            |            |            |  
 TTGGCCTTTGCGTTGCGTGGATCCTTCGACAAGGTCGCAACCAAGGTGTTCCCTGCTGTGC  
 130           140           150           160           170           180  
 |            |            |            |            |            |  
 CCAGCCGACGTCGACGTATCCCCTGAGGAGCGTCGCCGGATCGCCGAAACCGGCTTTTAC  
 190           200           210           220           230           240  
 |            |            |            |            |            |  
 GCCTACCAATAGATCGGTTTATAATCAGATCCACCGCAACATATGGGTCCGTCTCCGACC  
 250           260           270           280           290           300  
 |            |            |            |            |            |  
 TCCGAGTCGGTATTCTGGCGGTAACCGTACGTGTAGGCGGTGGTTGGTATCTTCATTGGT  
 310           320           330           340           350           360  
 |            |            |            |            |            |  
 AAGGTCGGGCTTTAGTTTGGCGTTGTTTTATCAGATCCTTGGATTGGCGCTGTTGCTCTT  
 370           380           390           400           410           420  
 |            |            |            |            |            |  
 TTGGCTGCTGCTGATCGCCCGGGTTGTCGTCGAGTTCATACGTTTCATTTAGCCGTGACTG  
 430           440           450           460           470           480  
 |            |            |            |            |            |  
 GCGCCCAACGGTGTCACTGTTGTGATCTTGGAGACTATCATGTCGATCACCGATCCGCC  
 490           500           510           520           530           540  
 |            |            |            |            |            |  
 GGTGAAGCTGTTGCGGTTGATTCCGCAACTCACCATTGGCGCGGTCCGCTTCGACTTGTC  
 550           560           570           580           590           600  
 |            |            |            |            |            |  
 CATCATGGTGCTGTTGCTGGTTGCGTTCATAGGCATGCAATTGGCTTTAAGCGCTGCGGC  
 610           620           630           640           650           660  
 |            |            |            |            |            |  
 TTAAGCCATGCCCCGACGCCCGGGAGGAGGCGACGGTTCGGCCACGTTTGGCGACGGT  
 670           680           690           700           710           720  
 |            |            |            |            |            |  
 TTGCTCCGTGTGATTTCAAATCTAAGGGTTGGGATTTAATTGGCCTTACAGTTTGAATTC  
 730           740           750           760           770           780  
 |            |            |            |            |            |  
 TTTACTTAATTTATTAGATAACGGCAACAGCCGGTCTGGTGTGACAGGATGGACGGTAG  
 790           800           810           820           830           840  
 |            |            |            |            |            |  
 TTGCGCGATGGCTACCACGCCGTTCTACACTTTCCAAATAATTTGACAAGGGACTTGAGG

Tata box

S.d.

850	860	870	880	890	900
GGACGAACA	ATGCCGCTTACACCAGCCGACGTCCACAACGTGGCGTTCAGTAAGCCGCC				
	<u>MET</u> ProLeuThrProAlaAspValHisAsnValAlaPheSerLysProPro				
910	920	930	940	950	960
ATCGGCAAGCGTGGTTATAACGAAGATGAGGTCGACGCTTTCCTCGACCTGGTTGAGAAC					
IleGlyLysArgGlyTyrAsnGluAspGluValAspAlaPheLeuAspLeuValGluAsn					
970	980	990	1000	1010	1020
GAGCTGACTCAGCTCATTGAAGAGAATTCCGACCTACGCCAGCGGATCGAGGAGCTGGAT					
GluLeuThrGlnLeuIleGluGluAsnSerAspLeuArgGlnArgIleGluGluLeuAsp					
1030	1040	1050	1060	1070	1080
CATGAACTAGCTGCGGGTGGCGGTACTGGTGCTGGCCCCGTTATCGCCGTGCAGCCCACT					
HisGluLeuAlaAlaGlyGlyGlyThrGlyAlaGlyProValIleAlaValGlnProThr					
1090	1100	1110	1120	1130	1140
CAGGCGCTCTCCACTTTCGAACCCGAAGTGTCTGCCAAGCAGGCGCCAGTGGCAGCT					
GlnAlaLeuSerThrPheGluProGluLeuValSerAlaLysGlnAlaProValAlaAla					
1150	1160	1170	1180	1190	1200
GTTGCCGAGACGGCCGAAGAATTGGCGATGAAGGCTACTAGGGTGTGAGCCTGGCGCAG					
ValAlaGluThrAlaGluGluLeuAlaMETLysAlaThrArgValLeuSerLeuAlaGln					
1210	1220	1230	1240	1250	1260
GATACCGCTGACCAGCTCACCAGTACGGCTAAGGTTCGAATCGGACAAGATGTTGGCCGAT					
AspThrAlaAspGlnLeuThrSerThrAlaLysValGluSerAspLysMETLeuAlaAsp					
1270	1280	1290	1300	1310	1320
GCGCGTGTCAATGCCGACCAAATTCTCGGCGAGGCTCGCCTCACGGCGGAGGCTACCGTC					
AlaArgValAsnAlaAspGlnIleLeuGlyGluAlaArgLeuThrAlaGluAlaThrVal					
1330	1340	1350	1360	1370	1380
GCCGAAGCCCAGCAACGCGCCGACGCGATGTTGGCCGACGCCAAACTCGCTCTGAGGTC					
AlaGluAlaGlnGlnArgAlaAspAlaMETLeuAlaAspAlaGlnThrArgSerGluVal					
1390	1400	1410	1420	1430	1440
CAGTCACGTCAAGCCCAGGAGAAGGCCGACGCCCTGCAGGCCGAAGCCGAGCGTAAGCAT					
GlnSerArgGlnAlaGlnGluLysAlaAspAlaLeuGlnAlaGluAlaGluArgLysHis					
1450	1460	1470	1480	1490	1500
TCTGAGATTATGGGAGCCATCAGCCAGCAGCGCACGGTGTGGAAGGCCGACTCGAGCAG					
SerGluIleMETGlyAlaIleSerGlnGlnArgThrValLeuGluGlyArgLeuGluGln					

```

      1510      1520      1530      1540      1550      1560
      |         |         |         |         |         |
CTACGTACCTTCGAACGCGAGTATCGGACCCGGCTCAAGACCTATCTGGAATCTCAGCTT
LeuArgThrPheGluArgGluTyrArgThrArgLeuLysThrTyrLeuGluSerGlnLeu

      1570      1580      1590      1600      1610      1620
      |         |         |         |         |         |
GAGGAACTCGGTCAGCGTGGGTCGGCTGCACCGGTTGATTCCAATGCGGATGCTGGTGGGA
GluGluLeuGlyGlnArgGlySerAlaAlaProValAspSerAsnAlaAspAlaGlyGly

      1630      1640      1650      1660      1670      1680
      |         |         |         |         |         |
TTCGACCAGTTCAATCGGGTAATAACTAGCTGAGAAGTGTGCCGGACTCCACGGTTCGAC
PheAspGlnPheAsnArgGlyAsnAsn---

      1690      1700      1710      1720      1730      1740
      |         |         |         |         |         |
ACACTAGGTGGCAATTCTAACTTGTGGGCTGGAGGATCACCAATGCTGATCATTGCACTG

      1750      1760      1770      1780      1790      1800
      |         |         |         |         |         |
GTGTTGGCCCTCATTGGTCTCGTTGTTCTGGTGTGTTGCGGTTGCCAACCAGAATTTGCTA

      1810      1820      1830      1840      1850      1860
      |         |         |         |         |         |
ATGGCCTGGGTATGTATTGGTGCAAGTGCCTGGGCGTGCTGCTGTTAATCGTTGATGCA

      1870      1880
      |         |
GTCCGGGAACACCAGTGCATCGAT

```

**Fig. 16.** Nucleic acid sequence of the 34 kd gene DNA of *M. leprae* showing the translation of the 34 kd Open Reading Frame.

S.D. = The deduced Shine-Delgarno sequence (Ribosome binding site).

Comparison between *M. leprae* & *M. tuberculosis* 34 kd genes have depicted that the former is longer by 18 base pairs (6 amino acids). Homology comparison shown below in Fig.17 reveals 80.8% (633 nucleotides) identity.

```

34KDML- ATGCCGCTTACACCAGCCGACGTCCACAACGTGGCGTTTCAGTAAGCCGCC -50
          ::::::::::::::: ::::::::::::::: :::::::::::::::
34KDMT- ATGCCGCTTACACCTGCCGACGTCCACAATGTGGCGTTTCAGTAAGCCGCC -50

34KDML- CATCGGCAAGCGTGGTTATAACGAAGATGAGGTCGACGCTTTCCTCGACC -100
          ::::::: ::::: : ::::::::::::::: :::::::::::::::
34KDMT- TATCGGCAAACGTGGGTACAACGAAGATGAGGTCGACGCTTTCCTCGACC -100

```

```

34KDML- TGGTTGAGAACGAGCTGACTCAGCTCATTGAAGAGAATTCCGACCTACGC -150
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
34KDMT- TGGTGGAAAACGAGCTGACCCGCCTGATCGAAGAGAACTCCGATCTGCGT -150

34KDML- CAGCGGATCGAGGAGCTGGATCATGAACTAGCTGCGGGTGGCGGTACTGG -200
      : : : : : : : : : : : : : : : : : : : : : : : : : :
34KDMT- CAGAGGATCAACGAGCTGGATCAAGAGCTCGCCGCGGG-GGCGGTGCCGG -199

34KDML- TGCTGGCCCCGTTATCGCCGTGCAGCCCACTCAGGCGCTCTCCACTTTCG -250
      : : : : : : : : : : : : : : : : : : : : : : : : :
34KDMT- -----CGTTA-CGCC--GCAGGCCACCCAGGCAATCCCGGCGTATG -237

34KDML- AACCCGAAGTAGTGTCTGCCAAGCAGGCGCCAGTGGCAGCTGTTGCCGAG -300
      : : : : : : : : : : : : : : : : : : : : : : : : :
34KDMT- AGCCCGAAC-----CGGGCAAGCCGGCGCCGGCGGGTCTCGGCGGG -281

34KDML- -ACGGCCGAAGAATTGGCGATGAAGGCTACTAGGGTGTGAGCCTGGCGC -349
      : : : : : : : : : : : : : : : : : : : : : : : : :
34KDMT- GATGAACGAGGAACAGGCCCTGAAGGCGGCGAGTGTGAGTCTGGCCC -331

34KDML- AGGATACCGCTGACCAGCTCACCAGTACGGCTAAGGTCAATCGGACAAG -399
      : : : : : : : : : : : : : : : : : : : : : : : : :
34KDMT- AAGACACCGCCGACCGGCTTACAAACACCGCCAAAGCCGAGTCCGACAAG -381

34KDML- ATGTTGGCCGATGCGCGTGTCAATGCCGACCAAATCTCGGCGAGGCTCG -449
      : : : : : : : : : : : : : : : : : : : : : : : : :
34KDMT- ATGCTGGCCGATGCCCGTGCCAATGCGGAGCAGATCCTCGGTGAGGCCG -431

34KDML- CCTCACGGCGGAGGCTACCGTCGCCGAAGCCAGCAACGCGCCGACGCGA -499
      : : : : : : : : : : : : : : : : : : : : : : : : :
34KDMT- ACACACCGCCGACGCCACGGTCGCCGAGGCCCGCCAGCGTGCCGATGCCA -481

34KDML- TGTTGGCCGACGCCAAACTCGCTCTGAGGTCCAGTCACGTCAAGCCCAG -549
      : : : : : : : : : : : : : : : : : : : : : : : : :
34KDMT- TGCTGGCCGATGCCCAATCCCGATCCGAGGCCAGTTGCGCCAGGCGCAG -531

34KDML- GAGAAGGCCGACGCCCTGCAGGCCGAAGCCGAGCGTAAGCATTCTGAGAT -599
      : : : : : : : : : : : : : : : : : : : : : : : : :
34KDMT- GAGAAGGCCGATGCCTTACAGGCCGATGCGGAACGCAAGCACTCCGAGAT -581

34KDML- TATGGGAGCCATCAGCCAGCAGCGCACGGTGTGGAAGGCCGACTCGAGC -649
      : : : : : : : : : : : : : : : : : : : : : : : : :
34KDMT- CATGGGAACCATCAACCAGCAGCGCGGGTGCTTGAAGGCCGCCTCGAGC -631

34KDML- AGCTACGTACCTTCGAACGCGAGTATCGGACCCGGCTCAAGACCTATCTG -699
      : : : : : : : : : : : : : : : : : : : : : : : : :
34KDMT- AGCTGCGTACCTTCGAACGTGAGTACCGCACCAGGCTCAAGACCTACCTG -681

34KDML- GAATCTCAGCTGAGGAACTCGGTGAGCGTGGGTGGCTGCACCGGTTGA -749
      : : : : : : : : : : : : : : : : : : : : : : : : :
34KDMT- GAATCGCAGCTGGAGGAACTCGGCCAGCGTGGATCGGCGGCGCCGGTCA -731

34KDML- TTCCAATGCGGATGCTGGTGGATTGACCCAGTTCAATCGGGGTAATAACT -799
      : : : : : : : : : : : : : : : : : : : : : : : : :
34KDMT- TTCCAATGCGGATGCCGGTGGCTTCGATCAATCAACCGGGGCAAAAAC -781

34KDML- AG -801
      : :
34KDMT- AG -783

```

Fig. 17. Alignment of the nucleotide sequences of 34 kd *M. leprae* gene and the 34 kd *M. tuberculosis* gene. The character ':' show that two aligned residues are identical.

Homology comparison between the putative amino acid sequences of the *M. leprae* and the *M. tuberculosis* 34 kd antigens show that there is 85 % (221 amino acids) identity and 13 % (34 amino acids) conserved substitutions (Fig. 18).

```

      10      20      30      40      50      60
34k-ML MPLTPADVHNVAFSKPPIGKRGYNEDEVDAFLDLVENELTQLIEENSDLRQRIEELDHEL
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
34k-MT MPLTPADVHNVAFSKPPIGKRGYNEDEVDAFLDLVENELTRLIEENSDLRQRINELDQEL
      10      20      30      40      50      60

      70      80      90      100     110     120
34K-ML AAGGGTGAGPVIQPTQALSTFEPELVSAKQAPVAAVAETAELAMKATRVLSLAQDTA
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
34K-MT AAGGGGAGVTP----QATQAI PAYEPE--PGKPAPAAVSAGMNEEQALKAARVLSLAQDTA
      70      80      90      100     110

      130     140     150     160     170     180
34K-ML DQLTSTAKVESDKMLADARVNADQILGEARLTAEATVAEAQQRADAMLADAQTRSEVQSR
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
34K-MT DRLTNTAKAESDKMLADARANAQILGEARHTADATVAEARQRADAMLADAQSRSEAQLR
      120     130     140     150     160     170

      190     200     210     220     230     240
34K-ML QAQEKADALQAEARKHSEIMGAISSQRTVLEGRLEQLRTFEREYRTRLKTYLESQLEEL
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
34K-MT QAQEKADALQADAERKHSEIMGTIQQRAVLEGRLEQLRTFEREYRTRLKTYLESQLEEL
      180     190     200     210     220     230

      250     260
34K-ML GQRGSAAPVDSNADAGGFDQFNRGNN
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
34K-MT GQRGSAAPVDSNADAGGFDQFNRGKN
      240     250     260

```

**Fig. 18.** Alignment of the deduced amino acid sequences of the 34 kd *Mycobacterium leprae* and *Mycobacterium tuberculosis* genes.

**N.B.** ':' = Identical residues; '.' = conserved substitutions.

## 5. DISCUSSION.

The search for the 34 kd antigen encoding gene of *Mycobacterium leprae* was carried out by screening an *M. leprae* cosmid library using the earlier isolated 34 kd antigen encoding gene of *Mycobacterium tuberculosis* as a DNA probe.

Screening of the cosmid library showed that, out of a total of 79 cosmids screened, only one cosmid, designated as BL 1527, carry the gene. Since the library is estimated to represent more than 88% of the *M. leprae* genome, this was interpreted as an indication that the gene exists as a single copy in the genome, which was in agreement with the result of hybridization on total genomic DNA mentioned previously.

Following the detection of the gene on cosmid BL 1527, differential hybridization of the cosmid vector and the gene of interest (using the 34 kd antigen gene of *M. tuberculosis* against the gene and plasmid pBKS+ against the vector) allowed the identification and cloning of four different DNA fragments.

Two of the four fragments, a 3.5 kb *Bam* HI fragment and a 1.9 kb *Cla* I fragment, were shown to contain the complete gene as inferred from the single hybridizing bands they gave. The other two, a 2.1 kb *Eco* RI fragment and 2.0 kb *Pst* I fragment of DNA, were shown to contain only part of the gene as they were derived from one band of two positively hybridizing signals in each case (Fig. 6).

The four clones and the cosmid BL 1527 were tested for the expression of the recombinant protein by western blotting using the monoclonal antibody F-126.2 as a probe. The result (Fig. 13) depicted expression of the protein in the cosmid BL 1527 and two of the clones that contain the complete gene, pPF-1000 and pPF-

2000, where as the other two clones that contain partial fragment of the gene, pPF-3000 & pPF-4000, did not show any expression. In the cosmid BL 1527 and the two clones that gave expression, the molecular weight of the expressed protein was of the same size and it was about 34 kd.

In addition to the band around the 34 kd area, reactive bands were seen around the 60 - 70 kd area in cosmid BL 1527 and in pPF 1000 which were absent in pPF 2000. As BL 1527 and pPF 2000 contain more *M. leprae* DNA than pPF 2000, the higher bands observed are believed to be results of dimer or oligomer formations mediated by down stream genes. The fact that the observed higher reacting bands were present only in BL 1527, pPF 1000 and pPF 1020 which contain substantial downstream DNA, as opposed to pPF 2000, pPF 2020 and pPH 5273 (the *M. tuberculosis* 34 kd gene) which do not contain down stream regions, indicates that the phenomenon is apparently mediated by products of these down stream DNA.

When the two clones showed expression, we were interested to see whether expression was under the control of the *E. coli*  $\beta$ -galactosidase gene promotor, down stream of which they were cloned. To address this question, the two clones were re-subcloned in the same vectors and clones with inserts in opposite orientation were obtained (Fig. 14). Western blot analysis of the clones in both orientation (Fig. 15) revealed that the two clones were expressible in both orientations showing that their expression is not under strict control from the  $\beta$ -galactosidase gene promotor but from their own transcriptional and translational signals. Furthermore, the fact that the protein products were of the same size when

expressed from both orientations was an indication that the protein in question was expressed free and not in fusion with  $\beta$ -galactosidase.

These observations may look suspicious in view of the existing consensus on the recognition of mycobacterial regulatory sequences by *E. coli*. Early experiments in cloning mycobacterial genes into *E. coli* were suggestive of poor or no recognition of the mycobacterial regulatory elements especially the transcriptional sequences (Clark-curtiss, *et al.*, 1990; Lamb and Colston, 1988). In addition it was believed that only mycobacterial heat shock promoters, which belong to the genes encoding the stress proteins which are evolutionarily highly conserved through out the living world, are recognized by *E. coli* (Young, *et al.*, 1988, 1990; Thangaraj, *et al.*, 1990). However, recent reports on the expression of mycobacterial proteins in *E. coli* are indicating that the above statements are not universal and a number of mycobacterial proteins such as the 38 kd gene of *M. leprae* (Andersen, *et al.*, 1988); the biotin carrier proteins of several species of mycobacteria (Collins, *et al.*, 1987) ; the 15 kd *M. leprae* gene contained in a cosmid and evidently far from *E. coli* promoters (Sela, *et al.*, 1991), have been shown to be expressed as free proteins apparently from their own promoter signals.

In fact analysis of published mycobacterial DNA sequences have revealed the presence of typical consensus prokaryotic promoter sites in other genes in addition to those above (Dale & Patki, 1990). Given that the mycobacterial initiation elements are recognized by *E. coli*, the expression of such recombinant clones would be independent of their orientation and reading frame (Young, *et al.*, 1987). Hence there is ample illustration in the literature to assume that the

transcriptional and translational sequences of the 34 kd gene of *M. leprae* could be among those recognized by *E. coli*, thus accounting for the observed expression of the gene in both orientations. As the two clones which show expression are long enough to contain intact *M. leprae* transcription and translation regulating elements, the above consideration seems likely.

Crossed Immuno-electrophoresis (CIE) analysis performed on two of the clones (pPF 1000 & pPF 2000) by our collaborator (Dr. M. Harboe, personal communication) has shown that the *M. leprae* and *M. tuberculosis* 34 kd recombinant antigens correspond to antigen 84 of mycobacteria.

DNA analysis of the 34 kd gene of *M. leprae* was carried out both by Southern hybridization as well as sequencing procedures. The DNA sequencing strategy employed was to start sequencing all four clones using the pUC forward and reverse universal primers and search for clues on the open reading frame. Once the open reading frame was located, with the help of sequence homology comparison with that of the previously known sequence of the *M. tuberculosis* equivalent, the largest clone, pPF 1000, was used to conduct the sequencing of the 801 base pairs long open reading frame of the 34 kd *M. leprae* gene plus 849 base pairs upstream of the gene and 234 base pairs downstream of the gene, giving the nucleotide sequence of a total of 1884 base pairs.

Computer aided analysis of the sequenced DNA of *M. leprae* 34 kd antigen encoding gene revealed the presence of three open reading frames of which only the second ORF, which stretches for 801 base pairs from an ATG start codon located at the 850<sup>th</sup> base down to a TAG stop codon identified at the position of the

1648<sup>th</sup> base pair, corresponds to the 34 kd gene. This open reading frame gives a protein with a deduced amino acid sequence of 266 amino acid residues that is estimated to bear a molecular weight of 28,887 Daltons and an isoelectric point of 4.3 (Fig. 16).

Comparison of this DNA sequence data with that of the *M. tuberculosis* 34 kd sequence shows a very high degree of homology between the two genes. The DNA sequence analysis comparison have revealed 80.8 % homology at their open reading frames which differ in length by 18 base pairs, the *M. leprae* 34 kd gene being the longer one. At the protein level 85 % homology was found between the two genes harboring 221 amino acid identity out of a total of 266 residues of the *M. leprae* 34 kd gene and 260 residues of *M. tuberculosis* homologue compared. From the remaining 39 residues, 34 were found to be conserved substitutions, leaving only five residues unmatched.

## 6. CONCLUSION.

A brief scanning glance at the literature of the 1980s and the 1990s will illustrate the speed with which the molecular genetics of *M. leprae* is being pried open with the application of modern immunological and molecular biological tools. By the day, the enigma that surrounded *M. leprae* in the past is being eroded little by little and will continue to be so.

The obscurities that surround the immunology of leprosy, however, can only be brought to light by a comprehensive and exhaustive knowledge of the antigenic repertoire of the bacillus. This fact provides the impetus for conducting basic research in molecular biology for antigen identification and characterization.

The work done in this project involved the identification of an immunologically recognized 34 kd antigen encoding gene of the leprosy bacillus. The research, which was carried out with the aim of establishing the physical properties of the gene as a prelude to detailed immunological study, has culminated in establishing some of the physical and immunological properties of the 34 kd gene of *M. leprae* by: 1). working out the partial restriction map of the gene, 2). determining the nucleotide sequence of the gene, 3). showing that the recombinant protein encoded by the gene is expressed in recombinant *Escherichia coli*, 4). establishing the fact that the expression of the protein is signalled from its own regulatory sequences and 5). confirming that the recombinant 34 kd antigen of *Mycobacterium leprae* corresponds to the non secreted Antigen 84 of mycobacteria.

**7. REFERENCES.**

- Abe, M. (1970). Studies on the antigenic specificity of *Mycobacterium leprae*. I. Demonstration of soluble antigens in leprosy nodules by immunodiffusion. **Int. J. Lep.** 38: 113 - 125.
- Abe, M., Minagawa, F. & Yoshino, Y. (1972). Indirect fluorescent antibody test for the detection of *M. leprae* and corresponding antibodies. Abstract. **Int. J. Lep.**, 40: 54.
- Andersen, A.B., Worsaae, A. & Chaparas, S.D. (1988). Isolation and characterization of recombinant  $\lambda$ gt11 bacteriophages expressing eight different mycobacterial antigens of potential immunological relevance. **Infect. Immun.**, 56: 1344-1351.
- Berhe, D., T-Haimanot, R., Tedla, T. & Tadesse, T. (1990). Epidemiological pattern of leprosy in Ethiopia: a review of the control programs. **Lep. Rev.** 61: 258 - 266.
- Bjune, G., Barnetson, R. StC., Ridley, D.S. & Kronvall, G. (1976). Lymphocyte transformation test in leprosy: correlation of the response with inflammation of the lesions. **Clin. Exp. Immunol.** 25: 85 - 94.
- Bloom, B.R., Salgame, P., Mehra, V., *et al.*, (1989). Vaccine development . On relating immunology to the third world: some studies on leprosy. **Immunology.** 2: 87 - 90.
- Britton, W.J., Helliquest, L., Gorsia, R.J., & Basten, A. (1988). Antigens of *Mycobacterium leprae* identified by immunoprecipitation with sera from leprosy and tuberculosis patients. **Clin. Exp. Immunol.**, 71: 394 - 398.

- Bryceson, A. & Pfaltzgraff, R.E. (1990). Leprosy. 3<sup>rd</sup> Edition. Churchill Livingstone Inc. pp 240.
- Bullock, W.E. (1978). Leprosy: A model of immunological perturbation in chronic infection. **J. Infect. Dis.**, **137**: 341 - 354.
- Cherayil, B.j. & Young, R.A. (1988). A 28-KD protein from *Mycobacterium leprae* is target of the human antibody response in lepromatous leprosy. **J. Immunol.** **141**: 4370 - 4375.
- Clark-Curtiss, J.E. (1990). Genome structure of mycobacteria. In Molecular Biology of the Mycobacteria. (McFadden, J. Ed.). Surrey University Press. pp 77 - 96.
- Clark-Curtiss, J.E., Jacobs, W.R., Docherty, M.A, Ritche, L.R. & Curtiss III, R. (1985). Molecular analysis of DNA and construction of cosmid libraries of *Mycobacterium leprae*. **J. Bacteriol.** **161**: 1093 - 1102.
- Closs, O., Mshana, R.N. & Harboe, M. (1979). Antigenic analysis of *Mycobacterium leprae*. **Scand. J. Immunol.** **9**: 297 - 302.
- Cocito, C. & Delville, J. (1985). Biological, chemical, and staining properties of bacteria isolated from tissues of leprosy patients. **Eur. J. Epidemiol.** **1**: 202 - 231.
- Collins, M.E., Moss. M.T., Wall, S. & Dale, J.W. (1987). FEMS Microbiol. Lett. **43**: 53-56. (Cited by Dale & Patky, 1990).
- Colston, M. J. & Lamb, F. I. (1989). Molecular biology of the mycobacteria. **Lep. Rev.**, **60**: 89 - 93.

- Converse, P.J., Ottenhoff, T.H.M., Gebre, N., Ehrenberg, J.P. & Kiessling, R. (1988). Cellular, humoral and gamma interferon response to *Mycobacterium leprae* and BCG antigens in healthy individuals exposed to leprosy. **Scand. J. Immunol.**, **27**: 515 - 525.
- Dale, J.W. & Patki, A. (1990). Mycobacterial gene expression and regulation. In *Molecular Biology of the Mycobacteria*, (McFadden, J. Ed.). Surrey University Press. pp173-198.
- Daniel, T.M. & Janicki, B.W. (1978). Mycobacterial antigens: a review of their isolation, chemistry and immunological properties. **Microbiol. Rev.**, **42**: 84 - 113.
- Daumerie, D. (1991). Leprosy in the WHO African region. **World Health. Statistics Quarterly.** **44**: 16 -22.
- de St.Groth, S.F. & Scheidegger, M. (1980). Production of monoclonal antibodies: strategies and tactics. **J. Immunol. Methods.** **35**: 1 - 21.
- Draper, P. (1982). The anatomy of mycobacteria. In *the Biology of the Mycobacteria*. Vol I. (Ratledge, C. & Stanford, J.L. Eds) Academic Press. London. pp 9 - 52.
- Draper, P. (1983).The bacteriology of *Mycobacterium leprae* **Lep. Rev.**, **51**: 563 - 575.
- Edwards, R.P. (1970). Electron microscope observations of division in *M. leprae* **J. Med. Microbiol.** **3**: 493 - 499.

- Ehrenberg, J.P. & Gebre, N. (1987). Analysis of the antigenic profile of *Mycobacterium leprae* : Cross-reactivity and unique specificities of human and rabbit antibodies. **Scand. J. Immun.**,**26**: 673 - 681.
- Eiglmeier, K., Honore, N., Woods, S.A., Caudron, B. & Cole, S.T. (1993). Use of an ordered cosmid library to deduce the genomic organization of *Mycobacterium leprae*. **Mol. Microbiol.** **7**: 197 - 206.
- Engers, H., Abe, M., Bloom, B.R. *et al.*, (1985). Results of a World Health Organization sponsored workshop on monoclonal antibodies to *Mycobacterium leprae*. **Infect. Immun.**, **48**: 603 - 605.
- Fukunishi, Y. (1985). Electron microscopic findings of *M. leprae* by freeze etching methods. **Int. J. Lep.** **53**: 247 - 250.
- Gills, H.K. & Godal, T. (1986). Deficiency of cell mediated immunity in leprosy. **Prog. Allergy.**, **37**: 377 - 390.
- Hansen, G.A. (1874). *Spedalskhedens arsager* (Causes of leprosy). Reprinted in **Int. J. Lep.** **1955**; **23**: 307 - 309.
- Harboe, M. (1985). The immunology of *M. leprae*. In *Leprosy* (Hastings, R.C. Ed.). Churchill Livingstone Inc. pp 53 - 87.
- Haregewoin, A., Godal, T., Mustafa, A.S., Belehu, A. & Yemaneberhan, T. (1983). T-cell conditioned media reverse unresponsiveness in lepromatous leprosy. **Nature**, **303**: 342 - 344.
- Haregewoin, A., Mustafa, A.S., Helle, I., Waters, M.F.R., Leiker, D.L. & Godal, T. (1984). Reversal by interleukin-2 of the T cell unresponsiveness of lepromatous leprosy to *Mycobacterium leprae*. **Immunol. Rev.**, **80**: 77 - 86.

- Hartskreel, R.A., de Wit, M.Y.L., & Klatser, P.R. (1989). Polymerase chain reaction for the detection of *Mycobacterium leprae*. **J. Gen. Microbiol.**, **135**: 2357-64.
- Hartskreel, R.A., van Rens, R.M., Stable, L.E.F., de Wit, M.Y.L. & Klatser, P.R. (1990). Selection and characterization of recombinant clones that produce *Mycobacterium leprae* antigens recognized by antibodies in sera from house hold contacts of leprosy patients. **Infect. Immun.** **58**: 2821-2827.
- Hastings, R.C. & Morales, M.J. 1982. Observations, calculations and speculations on the growth & death of *M.leprae in vivo*. **Int. J. Lep.** **50**: 579 - 592.
- Hirata, T. (1985a). Electron microscopic observations of acid fast bacilli in nasal mucosal biopsies of lepromatous leprosy patients. **Int. J. Lep.** **53**: 577 - 581.
- Hirata, T. (1985b). Electron microscopic observations of cell wall and cytoplasmic membrane in murine and human leprosy bacilli. **Int. J. Lep.** **53**: 433 - 440.
- Hohn, B. & Collins, J. (1980). A small cosmid for cloning large DNA fragments. **Gene**, **11**: 291 - 298.
- Honore, N., Bergh, S., Chanteau, s., Cole, S.T., *et al.*, (1993). Nucleotide sequence of the first cosmid from the *Mycobacterium leprae* genome project: structure and function of the Rif-Str regions. **Mol. Microbiol.**, **7**: 207-214.
- Hopwood, D.A., Kieser, T., Colston, M.J. & Lamb, F.I. (1988). Molecular biology of mycobacteria. **Brit. Med. Bull.**, **44**: 528- 546.

- Hutchinson, J. (1987). The *in vitro* cultivation of *Mycobacterium leprae*. In Essays in leprosy. (Ryan, T & McDougall, A.C., Eds) Alden press, Oxford. pp 30 - 51.
- Huynh, T.V., Young, R.A. & Davis, R.W. (1985). Constructing and screening cDNA libraries in  $\lambda$ gt10 &  $\lambda$ gt 11. In DNA cloning: A practical approach. (Glover, D.M., Ed.) Vol I. IRL press, Washington DC. pp.49 - 78.
- Job, C.K. (1971). Pathology of peripheral nerve lesions in lepromatous leprosy - a light and electron microscope study. **Int. J. Lep.** **39**: 251 - 268.
- Job, C.K. (1983). Pathogenesis of nerve lesions in leprosy. In immunological aspects of leprosy, tuberculosis and leishmaniasis. (Humber, D.P. Ed) International congress series 574. Excerpta Medica. pp 53 - 57.
- Kircheimer, W.F. & Storrs, E.E. (1971). Attempts to establish the armadillo as a model for the study of leprosy: report of lepromatous leprosy in an experimentally infected armadillo. **Int. J. Lep.** **39**: 693 - 702.
- Klatser, P.R., van Rens, M.M., and Egglete, T.A. (1984). Immunochemical characterization of *Mycobacterium leprae* antigens by the SDS-Polyacrylamide gel electrophoresis immunoperoxidase technique using patient sera. **Clin. Exp. Immunol.**, **56**: 537-544.
- Klenerman, P. (1987). Etiological factors in delayed type hypersensitivity reactions in leprosy. In Essays on leprosy.(Ryan, T.J. & McDougall, A.C., Eds.). Alden press, Oxford. pp 52 - 63.

- Kolk, A.H.J. van Leeuwen, J., Gilis, H., *et al.*, (1988). Use of monoclonal antibodies in the identification of mycobacterial antigens. Health Cooperation Papers - 7. (Nunzi, E., ED.) Associazione Italiana "Amici di R. Follereau" O.C.S.I. Bologna. pp101 - 106.
- Kronvall, G., Bjune, G. Stanford, J., Menzel, S. & Samuel, D. (1975). Mycobacterial antigens in antibody responses of leprosy patients. **Int. J. Lep.**, **43**: 299-306.
- Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. **Nature.**, **226**: 680 - 685.
- Lamb, F.I. & Colston, M.J. (1986). Cloning of *Mycobacterium leprae* genes in streptomycetes. *Lep Rev.*, 57 (Suppl. 2): 225 - 229.
- Levy, L. (1976). Studies of the mouse foot-pad technique for cultivation of *M. leprae*. 3. Doubling time during logarithmic multiplication. **Lep. Rev.** **47**: 103 - 106.
- Maniatis, T., Fritsch, E.F. & Sambrook, J. (1982). Molecular Cloning: A laboratory manual. Cold Spring Harbor Press p545.
- Meyers, W.M., Gromus, B.J. & Walsh, G.P. (1992). Non human sources of leprosy. **Int. J. Lep.** **60**: 477 - 480.
- Myrvang, B., Negassi, K., Lofgren, M. & Godal, T. (1975). Immune responsiveness to *M. leprae* of healthy humans. **Acta. Pathol. Microbiol. Scand.(C)**. **83**: 43 - 51.
- Newell, K.W. (1966). An epidemiologists view of leprosy. **Bull. W.H.O.**, **34**: 827 - 857.

- Nordeen, S.K., Lopez-Bravo, L. & Daumerie, D. (1991). Global review of multi drug therapy (MDT) in leprosy. **World Health Statistics Quarterly**, **44**: 2 - 14.
- Nordeen, S.K., Lopez-Bravo, L. & Sundaresan, T.K. (1992). Estimated cases of leprosy in the world. **Indian. J. Lep.** **64**: 521-528.
- Ottenhoff, T.H.M., Converse, P.J., Gebre, N., Wondimu, A., Ehrenberg, J.P. & Kiessling, R. (1989). T-cell response to fractionated *Mycobacterium leprae* antigens in leprosy. The lepromatous non-responder defect can be overcome *in vitro* by stimulation with fractionated *M. leprae* components. **Eur.J. Immunol.** **19**: 707 - 713.
- Pfaltzgraf, R.E. & Bryceson, A. (1985). Clinical leprosy. In *Leprosy* (Hastings, R.C.Ed.) Churchill Livingstone Inc. pp134 - 176.
- Rees, R.J.W. & Weddille, A.G.M. (1968). Experimental model for studying leprosy. **Ann. NY. Acad. Sci. USA.**, **154**: 214 - 236.
- Ridley, D.S. (1988). *Pathogenesis of leprosy and related diseases*. Butterworth & Co. (Publishers) Ltd. pp 250.
- Ridley, D.S. & Jopling, W.H. (1966). Classification of leprosy: a five group system. **Int. J. Lep.** **34**: 255 - 273.
- Rinke de Wit, T.F., Bekelie, S., Osland, A., *et al.*, (1992). Mycobacteria contain two Gro EL genes: The *Mycobacterium leprae* Gro EL gene is arranged in an operon with Gro ES. **Mol. Microbiol.**, **6**: 1995 - 2005.

- Runyon, E.H., Wayne, L.G. & Kubica, G.P. (1974). Actinomycetes and related organisms. Family II. Mycobacteriaceae Chester (1897). In Bergey's manual of determinative bacteriology. (Buchanan, R.E. & Gibbons, N.E., Eds.), 8<sup>th</sup> Edition. Baltimore, Williams & Wilkins Co. pp 681.
- Sambrook, J., Fritsch, E.F. & Maniatis, T. (1989). Molecular Cloning: A laboratory manual. 2<sup>nd</sup> Edition, Volumes I, II & III. Cold Spring Harbor Press.
- Sanger, F., Nicklen, S. & Coulson, A.R. (1977). DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA.*, **74**: 5463 - 5467.
- Satish, M., Esser, R.E., Thole, J.E.R., & Clark-Curtiss, J.E. (1990). Identification and characterization of antigenic determinants of *Mycobacterium leprae* that react with antibodies in sera of leprosy patients. *Infect. Immun.*, **58**: 1327 - 1336.
- Scott, G.C., Russel, D.A., Boughton, C.R., & Vincin, D.R. (1976). Untreated leprosy: probability for shifts in Ridley-Jopling classification. Development of 'flares' or disappearance of clinically apparent disease. *Int J. Lep.* **44**: 110 - 122.
- Seckl, M. (1985). Monoclonal antibodies and recombinant DNA technology: present and future uses in leprosy and tuberculosis. *Int. J. Lep.* **55**: 618 - 640.
- Sela, S., Thole, J.E.R., Ottenhoff, T.H.M., & Clark-Curtiss, J.E. (1991). Identification of *Mycobacterium leprae* antigens from a cosmid library: Characterization of a 15-kilodalton antigen that is recognized by both the

- humoral and cellular immune systems in leprosy patients. **Infect. Immun.** **59**: 4117-4124.
- Shepard, C.C. (1960). The experimental disease that follows injection of human leprosy bacilli into foot-pads of mice. **J. Exp. Med.** **112**: 445 - 452.
- Shepard, C.C. (1965). Temperature optimum of *Mycobacterium leprae* in mice. **J. Bacteriol.** **90**: 1271 - 1275.
- Thangaraj, H.S. Lamb, F.I., Davis, E.O. et, al., (1990). Identification, sequencing, and expression of *Mycobacterium leprae* superoxide dismutase, a major antigen. **Infect. Immun.**, **58**: 1937 - 1942.
- Turk, J.L. Curtis, J. & de Blaquiére, G. (1991). Immunopathogenesis of nerve involvement in leprosy. **Ind. J. Lep.** **63**: 483 - 491.
- Villarreal-Ramos, B., Sanchez-Gracia, J. Stocker, N. *et al.*, (1991). Screening gene expression libraries for epitopes recognized in *Mycobacterium leprae* by mouse T-cells. **Eur. J. Immunol.**, **21**: 2621-2624.
- Williams, D.L., Gillis, T P. & Protaels, F. (1990). Geographically distinct isolates of *Mycobacterium leprae* exhibit no genotypic diversity by restriction fragment-length polymorphism analysis. **Mol. Microbiol.**, **4**: 1653 - 1659.
- Woods, S.A. & Cole, S.T. (1989). A rapid method for the detection of potentially viable *Mycobacterium leprae* in human biopsies: a novel application of PCR. **FEMS Microbiol. Lett.**, **65**: 305 - 310.
- Young, D.B. (1988). Structure of mycobacterial antigens. **Brit. Med. Bull.**, **44**: 562 - 583.
- Young, D.B. (1991). What next in basic research? **Int. J. Lep.** **59**: 95 - 101.

- Young, D.B., Garbe, T., Lathriga, R. & Abou-Zied, C. (1990). Protein antigens: Structure, function & regulation. In *Molecular Biology of Mycobacteria* (McFadden, J., Ed.). Surrey University Press. pp 1 - 35.
- Young, D.B., Kaufmann, S.H.E., Hermans, P.W.M. & Thole, J.E.R. (1992). Mycobacterial protein antigens.: a compilation. **Mol. Microbiol.**, **6**: 133 - 145.
- Young, D.B., Kent, L. & Young, R.A. (1987). Screening of a recombinant mycobacterial DNA library with polyclonal antiserum and molecular weight analysis of expressed antigens. **Infect. Immun.** **55**: 1421 - 1425.
- Young, D.B., Lathriga, R., Hendrix, R., Sweetser, D. & Young, R.A. (1988). Stress proteins are immune targets in leprosy and tuberculosis. **Proc. Natl. Acad. Sci. USA.**, **85**: 4267 - 4270.
- Young, R.A. & Davis, R.W. (1985). Efficient isolation of genes by using antibody probes . **Proc. Natl. Acad. Sci.** **80**: 1194-8.
- Young, R.A., Mehra, V., Sweetser, D., *et al.*, (1985). Genes for the major antigens of the leprosy parasite *Mycobacterium leprae*. **Nature**, **316**: 450 - 452.

**Abbreviations.**

Ag.	Antigen
BCIP.	5-Bromo-4-Chloro-3-Indolyl Phosphate
CMI	Cell Mediated Immunity
DTH	Delayed Type Hypersensitivity
EDTA	Ethylene Diamine Tetra Acetic Acid
HSP	Heat Shock Protein
IPTG	Isopropyl Thio- $\beta$ -D-Galactopyranoside
kb	Kilo base
kd	Kilo dalton
NBT	Nitro Blue Tetrazolium
NC	Nitro Cellulose
PAGE	Polyacrylamide Agarose Gel Electrophoresis
PBS	Phosphate Buffered Saline
PMSF	Phenyl Methyl Sulfonyl Fluoride
SDS	Sodium Dodecyl Sulphate
Tris	Hydroxymethyl-Aminomethane
X-Gal	5-Bromo-4-Chloro-3-Indolyl- $\beta$ -D-Galactoside

Appendices.

Appendix 1.

The DNA molecular weight standard used throughout the text is the  $\lambda$  DNA *Hind* III marker. The molecular sizes of the fragments in base pairs in descending order

is :

1. 23,130

2. 9,416

3. 6,557

4. 4,361

5. 2,322

6. 2,027

7. 564.