Mycobacterial Gene Cloning and Expression, Comparative Genomics, Bioinformatics and Proteomics in Relation to the Development of New Vaccines and Diagnostic Reagents

A.S. Mustafa

Department of Microbiology, Faculty of Medicine, Kuwait University, Kuwait

Key Words

Mycobacterium tuberculosis · Antigens · Vaccine

Abstract

Recent advances in molecular and genomic techniques have facilitated research on several aspects of mycobacteriology, such as diagnosis and the identification of new vaccines and therapeutic targets for various diseases, including tuberculosis. The aim of this review was to analyze the implications of advances in molecular and genomic techniques on the development of new vaccines for tuberculosis as well as immunological reagents to diagnose the disease. Gene cloning and expression, DNA and protein sequencing, polymerase chain reaction, comparative genomics, bioinformatics, proteomics and DNA and peptide synthesis coupled with the application of cellular immunology techniques have led to the identification of several antigens of *Mycobacterium tuberculosis,* which have potential for diagnosis and vaccine applications. For example, cross-reactive mycobacterial antigens like heat shock proteins, MTB32 and MTB39, have been identified as new vaccine candidates, and antigens encoded by *M. tuberculosis*-specific genomic regions as new reagents for diagnosis.

Copyright © 2005 S. Karger AG, Basel

Introduction

Tuberculosis (TB) and leprosy are two major mycobacterial diseases that have afflicted humans throughout recorded history. However, due to multidrug therapy and active case finding efforts coordinated through the World Health Organization, impressive advances have been made towards the control of leprosy. For example, in 1970 there were 10.8 million people suffering from leprosy, but the number was reduced to less than 1 million in 1996 and further dropped to about 0.5 million by 2003, and the disease has been eliminated from at least 98 countries [1]. On the other hand, the global problem of TB, instead of improving, has deteriorated due to several factors, including the emergence and increase in the number of patients infected with multidrug-resistant strains of *Mycobacterium tuberculosis,* and TB-HIV coinfection. Currently, about 8 million people develop active disease and 2 million die of TB each year [2]. TB ranks among the top 10 causes of human mortality in the world and the World Health Organization has declared TB 'a global emergency'.

To control the worldwide problem of TB, universally efficacious anti-TB vaccines and reagents for specific diagnosis are urgently needed. With respect to anti-TB vaccines, vaccination with the attenuated *Mycobacterium bovis* Bacillus Calmette-Guérin (BCG) has been widely used for several decades, but its efficacy to protect against the most infectious form of TB (pulmonary disease) in

adults varies tremendously in different parts of the world. In particular, BCG has shown poor efficacy in the developing countries of Asia and Africa, where TB is highly endemic and vaccine-induced protection is most urgently needed [3]. In addition, the diagnostic value of the presently used skin test reagent, purified protein derivative (PPD) of *M. tuberculosis,* is low due to the presence of antigens that cross-react with environmental mycobacteria and the vaccine strains of *M. bovis* BCG. Furthermore, BCG being a live vaccine can cause disease by itself in immunocompromised subjects for whom vaccination with BCG is contraindicated, including AIDS patients, who are usually at a very high risk of developing TB [4]. Identification of *M. tuberculosis* antigens useful for the specific diagnosis of TB and the development of new TB vaccines that are protective in all parts of the world is therefore an urgent global need.

Gene Cloning and Expression to Identify Antigenic Proteins of *M. tuberculosis*

The advances in gene cloning and expression of mycobacterial genes in heterologous hosts have greatly facilitated the identification of antigenic proteins of *M. tuberculosis* recognized by antibodies and T cells. The first breakthrough in this field was reported by Bloom et al. in 1985 [5]. They used the λ gt11 phage expression system for high-level expression of mycobacterial DNA as fusion proteins (fused to the structural protein β-galactosidase) in *Escherichia coli.* The fusion protein approach assured efficient translation and transcription of DNA in *E. coli.* Another advantage of the fusion protein approach is the relative stability of the mycobacterial polypeptide. The fusion proteins are usually more resistant to the proteolytic machinery of *E. coli* than the foreign polypeptides alone [5].

The mycobacterial recombinant DNA libraries constructed by Bloom et al. [5] were first screened with antibody probes to identify recombinant phage clones expressing the antigenic proteins of mycobacteria. This resulted in the identification of DNA sequences expressing 12-, 14-, 18-, 19-, 28-, 36-, 65- and 70-kDa proteins [5]. When tested for immunological reactivity with T cell clones established from TB patients and healthy donors vaccinated with killed *Mycobacterium leprae,* the 65-, 19- and 18-kDa proteins were found reactive with human T cells [6–8]. In addition, these antigens were also recognized by polyclonal T cell lines established against whole-cell mycobacteria [9]. However, the 65-kDa antigen was found to be reactive with most T cell lines [9], and T cells responding to it were detected even 8 years after vaccination with a preparation of killed *M. leprae* [10]. These results suggested that the T cells reactive with the 65-kDa antigen were long-lasting memory T cells, and therefore vaccination with the 65-kDa antigen may be useful in providing long-term immunity. Furthermore, the 65-kDa antigen was presented to T cells promiscuously and in association with frequently expressed HLA-DR molecules [11, 12]. Thus, if used as a vaccine, this antigen will not have limitations due to HLA restriction. However, the 65-kDa antigen is not useful in the specific diagnosis of TB because it is a cross-reactive heat shock protein (hsp) shared among different mycobacterial species including the vaccine strains of BCG and environmental mycobacteria [13].

The vaccine potential of the 65-kDa mycobacterial hsp (hsp65) was first reported by Silva and Lowrie [14], who demonstrated that immunization with hsp65 expressed in a mouse macrophage cell line J774 protects mice against challenge with *M. tuberculosis.* The adoptive transfer experiments from immunized to naïve mice further showed that antigen-specific T cells were responsible for providing protection against the growth of *M. tuberculosis* in challenged mice [15]. Vaccination with hsp65-plasmid DNA protects mice against *M. tuberculosis* challenge in a prophylactic mode [16]. In addition, vaccination with hsp65-plasmid DNA has also been shown to be effective in immunotherapy of mice preinfected with *M. tuberculosis,* by shifting Th2 response to the protective Th1 type, and providing a sterilization effect, thus preventing the disease due to reactivation [17].

Another approach to identify antigens of *M. tuberculosis* recognized by T cells was established by us in 1988, in which recombinant DNA libraries were directly screened for T cell reactivity by using mycobacterial antigen-specific T cell clones [18]. In this approach, pools of recombinant antigens were prepared from a large number of recombinant phage clones followed by testing the antigenic pools with a pool of T cell clones established against whole-cell mycobacteria [18]. These studies led to the identification and characterization of 24-kDa lipoprotein antigens of *M. tuberculosis* (Rv2945c, LppX) and *M. leprae* (ML0136, LppX), which are absent in environmental mycobacteria [19, 20]. A synthetic peptide approach identified promiscuous T cell epitopes of LppX [20] and suggested that this antigen could be useful in developing a common vaccine against tuberculosis and leprosy.

By using T cell and serological expression cloning in conjunction with peripheral blood mononuclear cells

(PBMC) from healthy donors, T cell lines derived from mice infected with *M. tuberculosis* and sera from active TB patients, a group of scientists working at Corixa Corporation, USA, have identified 8 different Th1-cell-stimulating proteins with molecular mass ranging from 8.4 to 41 kDa [21–25]. Among these proteins, vaccination with MTB32 and MTB39 as fusion proteins has shown to protect 4 different species of animals, i.e. mice, guinea pigs, rabbits and monkeys. The vaccine is safe and effective in naïve animals as well as in animals preinfected with *M. tuberculosis* [25].

Reverse Genetics to Identify Antigens of *M. tuberculosis*

The techniques of reverse genetics have helped to identify the genes of several *M. tuberculosis* antigens that are actively secreted in the culture filtrate of *M. tuberculosis.* The work to identify the secreted protein antigens of *M. tuberculosis* has received considerable attention because of their possible immunodominance and involvement in protective immunity [26–28]. The evaluation of fractionated proteins of short-term culture filtrate (ST-CF) of *M. tuberculosis* in memory immune mice identified low-molecular-weight (6–10 kDa) proteins as the most potent stimulators of the protective cytokine IFN-γ [29]. In the subsequent studies, the low-molecular-weight fraction of ST-CF was used to raise monoclonal antibodies and purify proteins for amino-terminal sequencing. These studies led to the isolation of the gene for the 6-kDa early secreted antigenic target (ESAT6), which was then expressed as a recombinant protein in *E. coli* and purified [29]. Further testing of human T cells from TB patients with recombinant ESAT6 showed that it was the major stimulator of human T cells in proliferation and IFN-γ assays [30]. In addition, the use of overlapping synthetic peptides covering the sequence of the protein showed that multiple epitopes of ESAT6 were presented to Th1 cells in association with frequently expressed HLA class II molecules [31, 32]. The observation that ESAT6 is expressed in all members of the *M. tuberculosis* complex causing human TB, but not in many other mycobacteria including vaccine strains of *M. bovis* BCG [33], indicates that this antigen is relevant to specific diagnosis of TB. This was confirmed by demonstrating that ESAT6 induced high level of IFN-γ responses in TB patients but not in BCG-vaccinated healthy donors, whereas PPD induced equivalent responses in both groups [34]. Vaccination of mice with ESAT6 has been shown to provide pro-

tection against *M. tuberculosis* challenge in the mouse model of TB [35]. In addition, vaccination with a single peptide of ESAT6 (aa 51–70) provided a level of protection that was equivalent to the protection achieved after vaccination with complete ESAT6 or *M. bovis* BCG [35]. These findings demonstrate the potential of ESAT6 and its peptide (aa 51–70) in developing subunit and peptide-based vaccines against TB.

The studies summarized above suggest that ESAT6 could be of value in both specific diagnosis as well as in the development of a new vaccine against TB. However, in order to differentiate between vaccination with ESAT6 and infection with *M. tuberculosis,* a single antigen will not suffice. Therefore, studies have been carried out to identify additional antigens of *M. tuberculosis* with diagnostic and vaccine potential.

The need to identify additional antigens present in the low-molecular-weight fractions of ST-CF is further supported by our recent work aimed to evaluate the immune response of TB patients treated with a multidrug regimen for 6 months [36]. The results showed that, in treated patients, antigen-induced proliferation and IFN-γ responses of PBMC in response to the low-molecular-weight fractions F1 and F2 were severalfold higher than the responses obtained after stimulation of PBMC with ESAT6 [36]. Thus, this suggests that the low-molecular-weight fractions of ST-CF contain additional immunologically important proteins of *M. tuberculosis.*

The experiments carried out to identify the proteins in the low-molecular-weight fraction of ST-CF by raising monoclonal antibodies and screening the λ gt11 recombinant DNA library of *M. tuberculosis* with these antibodies led to the identification of 3 other immunogenic proteins termed as CFP10, TB7.3 and TB10.4 [33]. The sequence analyses demonstrated that TB10.4 and CFP10 belong to the ESAT6 family of low-mass proteins, whereas TB7.3 is a low-molecular-mass protein outside this family. TB10.4 and CFP10 have almost identical size (10 kDa) and genes coding for both have approximately 40% identity to *esat6* [33]. The proteins CFP10, TB7.3 and TB10.4 were expressed as recombinant proteins in *E. coli,* and the immunological reactivity of the recombinant proteins was assessed with PBMC from TB patients and *M. bovis* BCG-vaccinated and nonvaccinated healthy donors. In TB patients, the recombinant TB10.4 and CFP10 were strongly recognized and induced IFN-γ release at the same level (CFP10) as or at an even higher level (TB10.4) than ESAT6 [33], whereas TB10.4 showed strong recognition in *M. bovis* BCG-vaccinated healthy subjects as well and TB7.3 induced low level of responses

Table 1. Description of *M. tuberculosis*-specific ORFs predicted in RD1

ORF	Rv gene designation	Encoded protein	Protein length, number of aa
ORF2	*Rv3871*	unknown	591
ORF3	*Rv3872*	PE35	98
ORF4	unknown	unknown	239
ORF5	*Rv3873*	PPE68	371
ORF6	*Rv3874*	CFP10	100
ORF7	*Rv3875*	ESAT6	95
ORF8	unknown	unknown	139
ORF9	*Rv3876*	unknown	666
ORF10	*Rv3877*	unknown	552
ORF11	*Rv3878*	unknown	280
ORF14	unknown	unknown	262
ORF15	unknown	unknown	95

in all the subjects [33]. Thus, CFP10 was found to distinguish TB patients from *M. bovis* BCG-vaccinated donors and could therefore be, together with ESAT6, an interesting candidate for the specific diagnosis of active TB. A comparative evaluation of ESAT6 and CFP10 responses in an enzyme-linked immunospot IFN-γ assay with tuberculin skin testing has revealed that the enzyme-linked immunospot IFN-γ assay with ESAT6 and CFP10 is more accurate than tuberculin skin testing for identification of individuals who have latent TB [37]. Therefore, in addition to the diagnosis of active TB, the use of ESAT6 and CFP10 could improve TB control by more precise targeting of preventive treatment in people latently infected with *M. tuberculosis.*

The studies conducted to determine the distribution of *tb7.3, tb10.4* and *cfp10* genes in *M. tuberculosis, M. bovis* BCG and environmental mycobacteria have shown that all of these genes are present in *M. tuberculosis* but, like *esat6, cfp10* is absent in BCG and most environmental mycobacteria [33]. In addition to specific diagnosis, these antigens could also be useful in the development of new vaccines against TB due to lack of interference from most environmental mycobacteria. Further studies have shown that ESAT6 and CFP10 are expressed from a single operon and genes encoding these low-molecular-weight proteins are transcribed together in *M. tuberculosis* [38]. The 9.5-kb genomic region of *M. tuberculosis* encoding ESAT6 and CFP10 has been termed as RD1 [38]. The analysis of RD1 for open reading frames (ORFs) has shown that, in addition to ESAT6 (ORF7, Rv3875) and CFP10 (ORF6, Rv3874), this genomic region can potentially encode 12 other *M. tuberculosis*-specific proteins from an equal number of ORFs (table 1) [39]. As described below, further studies were conducted to identify other major antigens encoded by RD1 ORFs using recombinant DNA and overlapping synthetic peptide approaches.

Identification of Major *M. tuberculosis* Antigens Encoded by RD1 Using Recombinant DNA Approach

To immunologically characterize the putative proteins encoded by RD1 ORFs, we first attempted to clone and express the putative proteins encoded by RD1 as fusion proteins in *E. coli* using appropriate plasmids as cloning and expression vectors [39, 40]. Subsequently, the recombinant proteins were purified and isolated free of the fusion partner (glutathione *S*-transferase of *Schistosoma japonicum*) by affinity purification on glutathione Sepharose and/or Ni-NTA-agarose affinity matrix and cleavage of the purified fusion proteins by treatment with thrombin protease [39, 40]. However, these procedures could result in the expression and purification of only 6 of the targeted 14 proteins, i.e. ORF3 (PE35, Rv3872), ORF5 (PPE68, Rv3873), ORF6 (CFP10, Rv3874), ORF7 (ESAT6, Rv3875), ORF11 (Rv3878) and ORF14 [39, 40]. The problems with other ORFs included low level of expression, degradation of the mycobacterial proteins and the presence of contaminating *E. coli* components in purified preparations. When tested for antibody reactivity using sera from different groups of donors, the purified ORF14 protein was frequently recognized by antibodies present in the sera of TB patients, but not healthy *M. bovis* BCG-vaccinated donors [39, 41], whereas other proteins showed either low frequency of recognition or were recognized by sera from TB patients as well as healthy donors [39–41]. These results suggest that ORF14 protein could be useful in the serodiagnosis of TB. However, when tested with PBMC from TB patients and healthy donors for Th1 cell responses, ORF14 protein showed weak reactivity in antigen-induced proliferation and IFN-γ assays [42], thus suggesting that the ORF14 protein is not important as a new vaccine candidate against TB.

Overlapping Synthetic Peptides to Identify Antigens Encoded by RD1 Genes

To overcome the problems associated with the recombinant DNA approach in expression and purification of RD1-encoded proteins and to identify additional anti-

gens relevant for diagnosis and vaccine development against TB, we have used pools of overlapping synthetic peptides covering the sequence of each protein and tested these pools for T cell reactivity with PBMC from TB patients and BCG-vaccinated healthy donors in antigen-induced proliferation and IFN-γ assays [43]. The results showed that the RD1 pool was as good in stimulating Th1 cells from both donor groups as complex antigens of *M. tuberculosis,* i.e. whole cells and culture filtrate antigens and whole-cell *M. bovis* BCG [44]. The studies further showed that ORF5 (PPE68, Rv3873), ORF6 (CFP10, Rv3874) and ORF7 (ESAT6, Rv3875) are the major antigens recognized by Th1 cells from TB patients, whereas only ORF5 was the major antigen recognized by Th1 cells from *M. bovis* BCG-vaccinated healthy donors [44]. Similar experiments in the cattle model of TB have shown that PPE68, CFP10 and ESAT6 are also the major stimulators of Th1 cell reactivity in *M.-bovis*-infected but not in noninfected or BCG-vaccinated animals [45].

Comparative Genomics and Bioinformatics to Identify *M. tuberculosis*-Specific Genomic Regions

The availability of the complete genome sequence of *M. tuberculosis* in 1998 [46] has opened up new frontiers of research in TB, including the identification of antigens useful as new vaccine candidates and diagnostic reagents. The comparative genomics and bioinformatics have shown that in addition to RD1, 10 other genomic regions present in *M. tuberculosis* H37Rv (RD4–RD13) are deleted in all strains of BCG [47–49]. These regions cover 108 ORFs of *M. tuberculosis* H37Rv and are deleted from all 13 BCG substrains currently used as anti-TB vaccines in different parts of the world. By using an overlapping synthetic peptide approach, we are in the process of screening these regions to identify additional *M. tuberculosis*-specific antigens useful for diagnosis and vaccine applications [50].

In silico comparisons between *M. tuberculosis* and *M. leprae* genome sequences have revealed that among the 1,439 genes common to both mycobacteria, 219 genes code for proteins that show no similarity with proteins from other organisms [51]. Macroarray and bioinformatic analyses revealed that among the 219 'core' genes conserved among *M. tuberculosis* and *M. leprae,* only few were not present in other mycobacteria, i.e. *Mycobacterium avium, Mycobacterium marinum* and *Mycobacterium smegmatis.* Interestingly, some of these missing genes code for proteins of the ESAT6 family [51].

Diagnostic and Vaccine Potentials of RD1 and Other RD Regions

A number of recent studies demonstrate the diagnostic and vaccine potential of proteins encoded by RD1 and other RD regions. The study of Liu et al. [52] using synthetic peptides covering the sequence of 4 gene products encoded by RD2 (Rv1989c) and RD1 (Rv3873, Rv3878, and Rv3879c) showed that 45 and 53% of the TB patients responded to Rv3879c and Rv3873, respectively, identifying these proteins as major *M. tuberculosis* T cell antigens in humans, while 35 and 25% of the patients responded to Rv3878 and Rv1989c, respectively. Of the 38 BCG-vaccinated donors, 2.6% responded to peptides from Rv3878 and Rv3879c, 7.9% responded to Rv3873, and none responded to Rv1989c [52]. Exclusion of cross-reactive peptides encoded in conserved motifs of Rv3873 increased its specificity to 97.4%. The high specificity of Rv3879c peptides and nonconserved Rv3873 sequences, together with their moderate sensitivity in TB patients, identifies these peptides as candidates for inclusion in new T-cell-based tests for *M. tuberculosis* infection [52].

Using pools of synthetic peptides, Demangel et al. [53] have examined the potential immunogenicity of 4 RD1 products, i.e. Rv3872, Rv3873, Rv3878, and Rv3879c. Rv3873 was the only one to elicit IFN-γ-producing cells in C57BL/6 mice infected with *M. tuberculosis.* Anti-Rv3873 T cells were predominantly raised against an epitope mapped in the N-terminal end of the protein. Their results further show that Rv3873 is an immunogenic product of the RD1 region, which does not interfere with the secretion and immunogenicity of CFP10 and ESAT6. When tested as a His-tagged recombinant protein, Rv3873 stimulated high levels of IFN-γ secretion in peripheral blood mononuclear cells isolated from TB patients, as well as from healthy tuberculin PPD-positive subjects and *M. bovis* BCG-vaccinated donors [54]. Epitope mapping performed with overlapping peptides revealed a broad pattern of T cell recognition comprising both TB-specific epitopes and epitopes also recognized by BCG-vaccinated donors [54]. The immunological evaluation of Rv3878 by Agger et al. [55] showed a prominent T cell response to this antigen in TB patients, but very limited responses were seen in BCG-vaccinated donors.

A study conducted in the cattle model of TB used synthetic peptides covering 13 ORFs from the RD1, RD2, and RD14 regions of the *M. tuberculosis* genome [56]. Pools of overlapping peptides spanning these ORFs were tested in *M. bovis*-infected, BCG-vaccinated and unvaccinated cattle. All the peptide pools were recognized in

infected cattle, with responder frequencies varying between 16 and 86%. In particular, 8 highly immunogenic antigens were identified whose potentials as diagnostic reagents or as subunit vaccines warrant further study (Rv1983, Rv1986, Rv3872, Rv3873, Rv3878, Rv3879c, Rv1979c, and Rv1769) [56].

For functional characterization of genes within the RD1 segment, *M. tuberculosis* H37Rv deleted for RD1 (H37Rv:δRD1) was assessed for virulence after infections of the human macrophage-like cell line THP-1, human peripheral blood monocyte-derived macrophages, and C57BL/6 mice. In each of these systems, the H37Rv:δRD1 strain was strikingly less virulent than wild-type *M. tuberculosis* H37Rv and was very similar to BCG controls [57]. These results led to the conclusion that genes within or controlled by RD1 are essential for *M. tuberculosis* virulence and that loss of RD1 was important in BCG attenuation [57]. In addition to *M. bovis* BCG, loss of RD1 has also been suggested to contribute to the attenuation of the other live TB vaccine, i.e. *Mycobacterium microti* [58]. To increase the immunogenicity and protective capacity of these vaccines, both were complemented with different constructs containing the genes encoding ESAT6 and CFP10, as well as a variable number of flanking genes. Only reintroduction of the complete RD1 locus led to full secretion of the antigens and resulted in specific ESAT6-dependent immune responses; this suggests that the flanking genes encode a secretory apparatus [59]. Mice and guinea pigs vaccinated with the recombinant strain BCG: RD1-2F9 were better protected against challenge with *M. tuberculosis,* showing less severe pathology and reduced dissemination of the pathogen, as compared with control animals immunized with BCG alone [59].

Proteomics to Identify *M. tuberculosis* Antigens

The availability of the complete genome sequence of *M. tuberculosis* has revolutionized many areas of tuberculosis research and facilitated functional genomics studies, including the identification of antigens useful in diagnosis and vaccine applications. Proteomics has been widely applied to develop two-dimensional polyacrylamide gel electrophoresis maps and databases, evaluate gene expression profiles under different environmental conditions, assess global changes associated with specific mutations, and define drug targets of bacterial pathogens. Genomics has revealed the sequence of about 4,000 genes in *M. tuberculosis* [46]. Proteomics complements genomics in showing which genes are really expressed and could be useful as drug targets and new vaccine candidates [60]. Moreover, proteomics may identify the expression of genes not predicted by genomics [61].

Classical proteomics combines two-dimensional gel electrophoresis (2-DE) for the separation and quantification of proteins in a complex mixture with mass spectrometric identification of selected proteins. Using these techniques, Jungblut et al. [61] demonstrated the expression of 6 genes of *M. tuberculosis* not predicted by genomics. When coupled to immunological assays, proteomics may also be used to identify B cell and T cell antigens within complex protein mixtures [62]. In a study conducted by Covert et al. [63], large quantities of subcellular protein fractions of *M. tuberculosis* were resolved by two-dimensional liquid-phase electrophoresis, resulting in 355 and 299 fractions of culture filtrate and cytosolic proteins, respectively. Analysis of these fractions against splenocytes of C57Bl/6 mice infected with *M. tuberculosis* resulted in the identification of 37 fractions that stimulated a dominant T cell response, as measured by the production of IFN-γ. Molecular characterization of the proteins contained in each fraction by liquid chromatography-mass spectrometry and liquid chromatography-tandem mass spectrometry resulted in the identification of 30 individual proteins. Although many of these represented previously defined antigens, 17 of these proteins were novel T cell antigens [63]. These data demonstrate that proteomics offers a rapid and facile approach for elucidation of immunodominant T cell antigens of *M. tuberculosis.*

A comprehensive proteomics analysis of culture supernatant (CSN) proteins of *M. tuberculosis* H37Rv was accomplished by combination of 2-DE, mass spectrometry, and N-terminal sequencing by Edman degradation [64]. Analytical 2-DE gels resolved approximately 1,250 protein spots from CSN of *M. tuberculosis* H37Rv, 381 of which were identified by mass spectrometry and/or Edman degradation. The study revealed 137 different proteins, 42 of which had previously been described as secreted [64]. Comparative proteome analysis of CSN from virulent *M. tuberculosis* H37Rv and attenuated *M. bovis* BCG (Copenhagen) identified 39 *M. tuberculosis*-specific spots containing 27 different proteins. These included 5 proteins encoded by ORFs absent from *M. bovis* BCG, e.g. ESAT6, as well as 22 novel differential proteins, such as acetyl-CoA C-acetyltransferase (Rv0243) and 2 putative Esat6-like proteins (Rv1198, Rv1793) [64]. These proteins could provide additional antigens useful in the specific diagnosis and/or development of new vaccines against TB.

Conclusions

The advances in molecular and genomic techniques have led to the identification of several antigens of *M. tuberculosis* including hsp, low-molecular-weight proteins actively secreted in the culture filtrate of *M. tuberculosis* and proteins encoded by regions deleted in BCG and other mycobacteria. However, the antigens belonging to the ESAT6 family of low-molecular-weight proteins appear to be more useful for vaccine and diagnostic applications because of their immunodominance for recognition by protective Th1 cells and specificity for *M. tuberculosis.*

References

1 Elimination of leprosy as a public health problem. http://www.who.int/lep/.
2 Frieden TR, Sterling TR, Munsiff SS, Watt CJ, Dye C: Tuberculosis. Lancet 2003;362:887–899.
3 Fine PE: Variation in protection by BCG: Implications of and for heterologous immunity. Lancet 1995;346:1339–1345.
4 Schlossberg D: Tuberculosis and nontuberculous mycobacterial infections, ed 4. Philadelphia, WB Saunders Company, 1999, p 3.
5 Bloom BR, Mehra V, Young RA: Genes for the protein antigens of the tuberculosis and leprosy bacilli. Biosci Rep 1985;5:839–845.
6 Mustafa AS, Gill HK, Nerland A, Britton WJ, Mehra V, Bloom BR, Young RA, Godal T: Human T-cell clones recognize a major *M. leprae* protein antigen expressed in *E. coli.* Nature 1986;319:63–66.
7 Mustafa AS, Oftung F, Gill HK, Natvig I: Characteristics of human T-cell clones from BCG and killed *M. leprae* vaccinated subjects and tuberculosis patients. Recognition of recombinant mycobacterial antigens. Lepr Rev 1986;57(suppl 2):123–130.
8 Oftung F, Mustafa AS, Husson R, Young RA, Godal T: Human T cell clones recognize two abundant *Mycobacterium tuberculosis* protein antigens expressed in *Escherichia coli.* J Immunol 1987;138:927–931.
9 Mustafa AS: Identification of T-cell-activating recombinant antigens shared among three candidate antileprosy vaccines, killed *M. leprae, M. bovis* BCG, and mycobacterium w. Int J Lepr Other Mycobact Dis 1988;56:265–273.
10 Mustafa AS, Oftung F: Long-lasting T-cell reactivity to *Mycobacterium leprae* antigens in human volunteers vaccinated with killed *M. leprae.* Vaccine 1993;11:1108–1112.
11 Mustafa AS, Lundin KE, Oftung F: Human T cells recognize mycobacterial heat shock proteins in the context of multiple HLA-DR molecules: Studies with healthy subjects vaccinated with *Mycobacterium bovis* BCG and *Mycobacterium leprae.* Infect Immun 1993;61:5294–5301.

12 Mustafa AS, Lundin KE, Meloen RH, Shinnick TM, Oftung F: Identification of promiscuous epitopes from the mycobacterial 65-kilodalton heat shock protein recognized by human CD4(+) T cells of the *Mycobacterium leprae* memory repertoire. Infect Immun 1999;67:5683–5689.
13 Mustafa AS: HLA-restricted immune response to mycobacterial antigens: Relevance to vaccine design. Hum Immunol 2000;61:166–171.
14 Silva CL, Lowrie DB: A single mycobacterial protein (hsp 65) expressed by a transgenic antigen-presenting cell vaccinates mice against tuberculosis. Immunology 1994;82:244–248.
15 Silva CL, Silva MF, Pietro RC, Lowrie DB: Protection against tuberculosis by passive transfer with T-cell clones recognizing mycobacterial heat-shock protein 65. Immunology 1994;83:341–346.
16 Tascon RE, Colston MJ, Ragno S, Stavropoulos E, Gregory D, Lowrie DB: Vaccination against tuberculosis by DNA injection. Nat Med 1996;2:888–892.
17 Lowrie DB, Tascon RE, Bonato VL, Lima VM, Faccioli LH, Stavropoulos E, Colston MJ, Hewinson RG, Moelling K, Silva CL: Therapy of tuberculosis in mice by DNA vaccination. Nature 1999;400:269–271.
18 Mustafa AS, Oftung F, Deggerdal A, Gill HK, Young RA, Godal T: Gene isolation with human T lymphocyte probes. Isolation of a gene that expresses an epitope recognized by T cells specific for *Mycobacterium bovis* BCG and pathogenic mycobacteria. J Immunol 1988;141:2729–2733.
19 Oftung F, Wiker HG, Deggerdal A, Mustafa AS: A novel mycobacterial antigen relevant to cellular immunity belongs to a family of secreted lipoproteins. Scand J Immunol 1997;46:445–451.
20 Al-Attiyah R, Mustafa AS: Computer-assisted prediction of HLA-DR binding and experimental analysis for human promiscuous Th1-cell peptides in the 24 kDa secreted lipoprotein (LppX) of *Mycobacterium tuberculosis.* Scand J Immunol 2004;59:16–24.

21 Coler RN, Skeiky YA, Vedvick T, Bement T, Ovendale P, Campos-Neto A, Alderson MR, Reed SG: Molecular cloning and immunologic reactivity of a novel low molecular mass antigen of *Mycobacterium tuberculosis.* J Immunol 1998;16:2356–2364.
22 Dillon DC, Alderson MR, Day CH, Lewinsohn DM, Coler R, Bement T, Campos-Neto A, Skeiky YA, Orme IM, Roberts A, Steen S, Dalemans W, Badaro R, Reed SG: Molecular characterization and human T-cell responses to a member of a novel *Mycobacterium tuberculosis* mtb39 gene family. Infect Immun 1999;67:2941–2950.
23 Skeiky YA, Lodes MJ, Guderian JA, Mohamath R, Bement T, Alderson MR, Reed SG: Cloning, expression, and immunological evaluation of two putative secreted serine protease antigens of *Mycobacterium tuberculosis.* Infect Immun 1999;67:3998–4007.
24 Alderson MR, Bement T, Day CH, Zhu L, Molesh D, Skeiky YA, Coler R, Lewinsohn DM, Reed SG, Dillon DC: Expression cloning of an immunodominant family of *Mycobacterium tuberculosis* antigens using human CD4(+) T cells. J Exp Med 2000;191:551–560.
25 Reed S: TB vaccine development: Preclinical activities. Abstract session 5. Novel vaccines/vaccine strategies for TB. First Int Conf on TB Vaccines for the World, Montreal, September 2003.
26 Andersen P: Host responses and antigens involved in protective immunity to *Mycobacterium tuberculosis.* Scand J Immunol 1997;45:115–131.
27 Mustafa AS, Shaban FA, Abal AT, Al-Attiyah R, Wiker HG, Lundin KE, Oftung F, Huygen K: Identification and HLA restriction of naturally derived Th1-cell epitopes from the secreted *Mycobacterium tuberculosis* antigen 85B recognized by antigen-specific human CD4(+) T-cell lines. Infect Immun 2000;68:3933–3940.

28 Al-Attiyah R, Shaban FA, Wiker HG, Oftung F, Mustafa AS: Synthetic peptides identify promiscuous human Th1 cell epitopes of the secreted mycobacterial antigen MPB70. Infect Immun 2003;71:1953–960.

29 Andersen P, Andersen AB, Sorensen AL, Nagai S: Recall of long-lived immunity to *Mycobacterium tuberculosis* infection in mice. J Immunol 1995;154:3359–3372.

30 Mustafa AS, Amoudy HA, Wiker HG, Abal AT, Ravn P, Oftung F, Andersen P: Comparison of antigen-specific T-cell responses of tuberculosis patients using complex or single antigens of *Mycobacterium tuberculosis.* Scand J Immunol 1998;48:535–543.

31 Mustafa AS, Oftung F, Amoudy HA, Madi NM, Abal AT, Shaban F, Rosen Krands I, Andersen P: Multiple epitopes from the *Mycobacterium tuberculosis* ESAT-6 antigen are recognized by antigen-specific human T cell lines. Clin Infect Dis 2000;30(suppl 3):S201–S205.

32 Mustafa AS, Shaban FA, Al-Attiyah R, Abal AT, El-Shamy AM, Andersen P, Oftung F: Human Th1 cell lines recognize the *Mycobacterium tuberculosis* ESAT-6 antigen and its peptides in association with frequently expressed HLA class II molecules. Scand J Immunol 2003;57:125–134.

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

34 Ravn P, Demissie A, Eguale T, Wondwosson H, Lein D, Amoudy HA, Mustafa AS, Jensen AK, Holm A, Rosenkrands I, Oftung F, Olobo J, von Reyn F, Andersen P: Human T cell responses to the ESAT-6 antigen from *Mycobacterium tuberculosis.* J Infect Dis 1999;179: 637–645.

35 Olsen AW, Hansen PR, Holm A, Andersen P: Efficient protection against *Mycobacterium tuberculosis* by vaccination with a single subdominant epitope from the ESAT-6 antigen. Eur J Immunol 2000;30:1724–1732.

36 Al-Attiyah R, Mustafa AS, Abal AT, Madi NM, Andersen P: Restoration of mycobacterial antigen-induced proliferation and interferon-gamma responses in peripheral blood mononuclear cells of tuberculosis patients upon effective chemotherapy. FEMS Immunol Med Microbiol 2003;38:249–256.

37 Ewer K, Deeks J, Alvarez L, Bryant G, Waller S, Andersen P, Monk P, Lalvani A: Comparison of T-cell-based assay with tuberculin skin test for diagnosis of *Mycobacterium tuberculosis* infection in a school tuberculosis outbreak. Lancet 2003;361:1168–1173.

38 Berthet FX, Rasmussen PB, Rosenkrands I, Andersen P, Gicquel B: A *Mycobacterium tuberculosis* operon encoding ESAT-6 and a novel low-molecular-mass culture filtrate protein (CFP-10). Microbiology 1998; 144: 3195–3203.

39 Ahmad S, Amoudy HA, Thole JE, Young DB, Mustafa AS: Identification of a novel protein antigen encoded by a *Mycobacterium tuberculosis*-specific RD1 region gene. Scand J Immunol 1999;49:515–522.

40 Ahmad S, Ali MM, Mustafa AS: Construction of a modified vector for efficient purification of recombinant *Mycobacterium tuberculosis* proteins expressed in *Escherichia coli.* Protein Expr Purif 2003;29:167–175.

41 Mustafa AS, Al-Attiyah R: Tuberculosis: Looking beyond BCG vaccines. J Postgrad Med 2003;49:134–140.

42 Mustafa AS: Biotechnology in the development of new vaccines and diagnostic reagents against tuberculosis. Curr Pharm Biotechnol 2001;2:157–173.

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

44 Mustafa AS: Progress towards the development of new anti-tuberculosis vaccines; in: Progress in Tuberculosis Research. Kuwait. New York, Nova Science Publishers, Inc., Editor: LC Smithe, 2005, p 47–76.

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

46 Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, et al: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. Nature 1998; 393:537–544.

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

48 Gordon SV, Brosch R, Billault A, Garnier T, Eiglmeier K, Cole ST: Identification of variable regions in the genomes of tubercle bacilli using bacterial artificial chromosome arrays. Mol Microbiol 1999;32:643–655.

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

50 Mustafa AS, Al-Attiyah R: *Mycobacterium tuberculosis* antigens and peptides as new vaccine candidates and immunodiagnostic reagents against tuberculosis. Kuwait Med J 2004;36:171–176.

51 Marmiesse M, Brodin P, Buchrieser C, Gutierrez C, Simoes N, Vincent V, Glaser P, Cole ST, Brosch R: Macro-array and bioinformatic analyses reveal mycobacterial 'core' genes, variation in the ESAT-6 gene family and new phylogenetic markers for the *Mycobacterium tuberculosis* complex. Microbiology 2004;150: 483–496.

52 Liu XQ, Dosanjh D, Varia H, Ewer K, Cockle P, Pasvol G, Lalvani A: Evaluation of T-cell responses to novel RD1- and RD2-encoded *Mycobacterium tuberculosis* gene products for specific detection of human tuberculosis infection. Infect Immun 2004;72:2574–2581.

53 Demangel C, Brodin P, Cockle PJ, Brosch R, Majlessi L, Leclerc C, Cole ST: Cell envelope protein PPE68 contributes to *Mycobacterium tuberculosis* RD1 immunogenicity independently of a 10-kilodalton culture filtrate protein and ESAT-6. Infect Immun 2004;72:2170–2176.

54 Okkels LM, Brock I, Follmann F, Agger EM, Arend SM, Ottenhoff TH, Oftung F, Rosenkrands I, Andersen P: PPE protein (Rv3873) from DNA segment RD1 of *Mycobacterium tuberculosis:* Strong recognition of both specific T-cell epitopes and epitopes conserved within the PPE family. Infect Immun 2003;71: 6116–6123.

55 Agger EM, Brock I, Okkels LM, Arend SM, Aagaard CS, Weldingh KN, Andersen P: Human T-cell responses to the RD1-encoded protein TB27.4 (Rv3878) from *Mycobacterium tuberculosis.* Immunology 2003;110:507–512.

56 Cockle PJ, Gordon SV, Lalvani A, Buddle BM, Hewinson RG, Vordermeier HM: Identification of novel *Mycobacterium tuberculosis* antigens with potential as diagnostic reagents or subunit vaccine candidates by comparative genomics. Infect Immun 2002;70:6996–7003.

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

58 Pym AS, Brodin P, Brosch R, Huerre M, Cole ST: Loss of RD1 contributed to the attenuation of the live tuberculosis vaccines *Mycobacterium bovis* BCG and *Mycobacterium microti.* Mol Microbiol 2002;46:709–717.

59 Pym AS, Brodin P, Majlessi L, Brosch R, Demangel C, Williams A, Griffiths KE, Marchal G, Leclerc C, Cole ST: Recombinant BCG exporting ESAT-6 confers enhanced protection against tuberculosis. Nat Med 2003;9: 533–539.

60 Betts JC: Transcriptomics and proteomics: Tools for the identification of novel drug targets and vaccine candidates for tuberculosis. IUBMB Life 2002;53:239–242.

61 Jungblut PR, Muller EC, Mattow J, Kaufmann SH: Proteomics reveals open reading frames in *Mycobacterium tuberculosis* H37Rv not predicted by genomics. Infect Immun 2001;69: 5905–5907.

62 Dobos KM, Spencer JS, Orme IM, Belisle JT: Proteomic approaches to antigen discovery. Methods Mol Med 2004;94:3–17.

63 Covert BA, Spencer JS, Orme IM, Belisle JT: The application of proteomics in defining the T cell antigens of *Mycobacterium tuberculosis.* Proteomics 2001;1:574–586.

64 Mattow J, Schaible UE, Schmidt F, Hagens K, Siejak F, Brestrich G, Haeselbarth G, Muller EC, Jungblut PR, Kaufmann SH: Comparative proteome analysis of culture supernatant proteins from virulent *Mycobacterium tuberculosis* H37Rv and attenuated *M. bovis* BCG Copenhagen. Electrophoresis 2003; 24: 3405–3420.