Comparative proteomic analysis of sequential isolates of *Mycobacterium tuberculosis* from a patient with pulmonary tuberculosis turning from drug sensitive to multidrug resistant

Amit Singh¹, Krishnamoorthy Gopinath¹, Prashant Sharma², Deepa Bisht², Pawan Sharma³, Niti Singh⁴ & Sarman Singh¹

¹Division of Clinical Microbiology & Molecular Medicine, Department of Laboratory Medicine, All India Institute of Medical Sciences, New Delhi, ²National JALMA Institute for Leprosy & other Mycobacterial Diseases (ICMR), Agra, ³Immunology Group, International Centre for Genetic Engineering & Biotechnology, New Delhi & ⁴Department of Microbiology, National Institute of Tuberculosis & Respiratory Diseases, New Delhi, India

Received September 10, 2013

**Background & objectives:** Tuberculosis is a major health problem in India, and the emergence of multidrug resistant (MDR) and extensively drug resistant (XDR) strains of *Mycobacterium tuberculosis* (*Mtb*) has further complicated the situation. Though several studies characterizing drug sensitive and drug resistant strains are available in literature, almost all studies are done on unrelated strains. Therefore, the objective of this study was to compare the proteomic data of four sequential isolates of *Mtb* from a single patient who developed MDR-TB during the course of anti-tuberculosis therapy (ATT).

**Methods:** In this study, using two-dimensional (2D) gel electrophoresis and MALDI-TOF mass spectrometry, we compared and analyzed the cell lysate proteins of *Mtb* sequential clinical isolates from a patient undergoing anti-TB treatment. The mRNA expression levels of selected identified proteins were determined by quantitative real-time polymerase chain reaction (qRT-PCR).

**Results:** The genotypes of all four isolates remained homologous, indicating no re-infection. The initial isolate (before treatment) was sensitive to all first-line drugs, but the consecutive isolates were found to be resistant to isoniazid (INH) and rifampicin (RIF) and developed mutations in the *katG*, *inhA* and *rpoB*. The intensities of 27 protein spots were found to be consistently overexpressed in INH and RIF resistant isolates. The most prominent and overexpressed proteins found during the development of drug resistance were GarA (Rv1827), wag31 (Rv2145c), Rv1437 and Rv2970c.

**Interpretation & conclusions:** This preliminary proteomic study provides an insight about the proteins that are upregulated during drug resistance development. These upregulated proteins, identified here, could prove useful as immunodiagnostic and possibly drug resistant markers in future. However, more studies are required to confirm these findings.

**Key words** 2D gel electrophoresis - tuberculosis - MDR-TB - MALDI-TOF - proteomics
Tuberculosis (TB) is a global emergency with an estimated nine million new cases and more than 1.5 million deaths occurring annually\(^1\). The situation has worsened after AIDS epidemic and with the emergence of multidrug resistant (MDR) forms of the causative agent *Mycobacterium tuberculosis* (*Mtb*). These drug resistant strains are more infectious by virtue of their high transmissibility in the population. Therefore, identification of the reliable diagnostic, prognostic and drug resistance markers is an urgent research priority. Various *in vitro* and *in vivo* studies have identified chromosomal mutations as determinants of drug resistance\(^2,4\). For example, mutation (s) in *rpoB* allele confers rifampicin (RIF) resistance (RIF\(^\text{R}\)) in 90-95 per cent isolates\(^2\), while isoniazid (INH)-resistance (INH\(^\text{R}\)) is attributed to mutation (s) in one or more alleles viz., *katG*, *inhA*, *ahpC* and *ndh*. However, in about 20 per cent of INH\(^\text{R}\) isolates, none of these known mutations are found, suggesting the possibility of unknown mutations or mechanisms\(^2,3\). On the other hand, the genetic mutations may not necessarily correlate with phenotypic resistance; further suggesting that other factors such as drug impermeability, drug-efflux pumps, formation of survivable “persister cells” under drug pressure and several other host factors could be involved in the outcome of treatment\(^2\).

Hence, for unraveling the mechanism(s) of drug resistance, understanding the mode of action of anti-TB drugs is very crucial. Many studies have elucidated the mode of action of various anti-TB drugs using genetic analysis, mRNA expression and DNA microarray analysis\(^4,5\). Several groups have also explored the proteome of *Mtb* and provided comprehensive details about the subcellular localization and confirmed the genomic annotation\(^6,9\). In these studies two-dimensional (2D) gel electrophoresis followed by mass spectrometry (MS) identification of the differentially regulated proteins substantially helped in identifying the complex pathways and their regulatory enzymes. These studies also elucidated modes of action of various drugs and discovered new antigens that could be potential candidates for developing vaccines and diagnostics\(^6,7,9,10\). However, only a few studies are available which show differential expression of specific proteins in the drug resistant but not in drug susceptible cells\(^7,9,11\). Further, in all these studies, either the non-pathogenic mycobacteria or laboratory collections of drug sensitive and drug resistant strains of *Mtb* from different patients have been used. In the present study, protein profile of sequentially collected four clinical isolates of *Mtb* was analyzed using 2D gel electrophoresis and the differentially expressed proteins were identified by MALDI-TOF-MS analysis. All isolates were from the same patient, who developed MDR-TB during the course of chemotherapy.

### Material & Methods

The study was conducted between January 2006 and June 2010 at the TB Laboratory, Division of Clinical Microbiology and Molecular Medicine, Department of Laboratory Medicine, All India Institute of Medical Sciences (AIIMS), New Delhi, India. This study was approved by the Institutional Ethics Committee of AIIMS and written informed consent was obtained from the patient. The patient was being treated at the designated microscopy and DOTS (directly observed treatment-short course) centres of Shahpurjat, New Delhi. This patient (22 yr old male) was diagnosed as having pulmonary TB on the basis of clinical and radiological findings and sputum smear microscopy. He was prescribed with anti-TB treatment (ATT) under the DOTS programme. The thrice a week treatment regimen comprised isoniazid, rifampicin, pyrazinamide (PZA) and ethambutol (EMB) (category I treatment) in intensive phase for two months followed by four month treatment with two (isoniazid and rifampicin) drugs regimen. Pre-treatment sputum specimen was used for isolation of *Mycobacterium* sp. by BACTEC MGIT-960 (Becton Dickinson, Sparks, MD, USA), which was positive. The isolate was identified as *Mtb* by conventional phenotypic and in-house PCR method\(^12\). This culture was labelled as isolate A, and was subjected to 16sRNA gene sequencing. The patient though took full six months course of treatment but became irregular in taking drugs after initial improvement in his clinical symptoms. After three months of cessation of treatment (6+3= 9 month\(^13\), his condition again deteriorated and his sputum culture was again positive for *Mtb*. We labelled this second culture as isolate B. He was re-treated with isoniazid, rifampicin, pyrazinamide, ethambutol and streptomycin (SM) (category II regimen). Within two months his clinical condition improved but he again defaulted. After an asymptomatic period of about four months his symptoms reappeared. His sputum was again culture positive and this culture was labelled as isolate C. The patient was again prescribed with the same treatment for 12 months after counselling but he stopped treatment after six months. His condition further deteriorated and he died of multisystem failure. The fourth sample was received just before his death.
and the isolate from this sputum sample was labelled as isolate D.

All the four clinical isolates (A, B, C & D) were identified as \textit{Mtb} using standard protocols\cite{12,13}. The anti-mycobacterial drug susceptibility testing was performed on all the isolates by both BACTEC™ MGIT-960 (Becton Dickinson, Sparks, MD, USA) and proportional method using Middlebrook 7H10 (Difco, USA) agar plates containing first-line anti-TB drugs (SM 2.0 \mu g/ml, INH 0.2 \mu g/ml, RIF 1.0 \mu g/ml, EMB 6.0 \mu g/ml)\cite{13,14}. All four isolates were also genotyped by spoligotyping and identified using SITVIT-WEB database\cite{15}. The \textit{rpoB}, \textit{inhA} and \textit{katG} gene targets were sequenced using the primers as described elsewhere\cite{13}.

\textbf{Preparation of mycobacterial whole cell lysate:} All \textit{Mtb} isolates were grown without shaking in Middlebrook 7H9 medium supplemented with 0.2 per cent (v/v) glycerol, 10 per cent oleic acid, albumin-dextrose and catalase (OADC, Difco, USA) at 37°C for two weeks. Whole cell lysate was prepared according to protocol of Sharma \textit{et al}\cite{11}. Cells were washed three times with normal saline and then suspended in sonication buffer [50 mM tris-HCl containing 10 mM MgCl$_2$, 0.1\% sodium azide, 1 mM phenylmethylsulfonyl fluoride (PMSF) and 1mM ethylene glycol tetra acetic acid (EGTA); pH 7.4] at a concentration of 1g wet cell mass per 5ml, and then broken by intermittent sonication for 15 min at 4°C using sonicator (Sonic\& Materials Inc, USA). The homogenate was centrifuged at 12,000 x g for 20 min at 4°C. The pellets were discarded and supernatant was stored at -70°C until further use.

\textbf{Protein precipitation with sodium dodecyl sulphate (SDS)-trichloroacetic acid (TCA)-acetone:} The cell lysates were treated with 1 per cent SDS and then subjected to TCA-acetone precipitation procedure\cite{9}. The protein pellet was suspended in appropriate volume of two-dimensional rehydration buffer (Bio-Rad, USA), and the protein concentration was estimated using the Bradford method\cite{16}.

\textbf{Two-dimensional gel electrophoresis:} Isoelectric focusing (IEF) was done using the in-gel rehydration method (Bio-Rad, USA).

2D gels were analysed using PDQuest Advanced software (version 8.0) (Bio-Rad, USA). After acquisition, the images were analyzed using stepwise spot detection and spot matching followed by differential expression analysis. The quantity of each spot was normalized by total valid spot intensity. The expression differences for all four mycobacterial isolates were compared using the same software. Images for sensitive and resistant isolates were manually checked for artifactual spots, merged spots and missed spots, and spots with more isolate-specific variability were omitted in the downstream processing. Equal amount of protein was loaded in all gels and experiments were repeated three times with three independent biological replicates.

\textbf{In-gel digestion of protein spots with trypsin:} Protein spots of interest were excised from the coomassie brilliant blue R250 stained 2D gels using spot picker Investigator ProPic (Genomic Solutions Ltd., Huntingdondon, UK). Digestion of proteins and spotting of peptides on matrix assisted LASER desorption/ ionization-time of flight (MALDI-TOF) target plate was carried out using protein digester investigator ProPrep (Genomic Solutions, Huntingdon, UK).

For protein digestion, method of Shevchenko \textit{et al}\cite{17} was followed with slight modifications. In brief, the gel plugs were de-stained and dehydrated by washing three times (10 min) with 25 mM NH$_4$HCO$_3$-50 per cent acetonitrile (ACN) (1:1v/v) solution and treated with freshly prepared 10 mM dithiothreitol (DTT) in 50 mM NH$_4$HCO$_3$ for 45 min at 37°C. After incubation, DTT was replaced with freshly prepared 55 mM iodoacetamide for 30 min and then dehydrated with 100 per cent ACN. The dried gel pieces were incubated for 12 h at 37 \degree C with 25 mM NH$_4$HCO$_3$, containing 0.02 \mu g/\mu l of mass spectrometry grade trypsin (Promega, USA). The resulting peptides were extracted twice from the gel pieces, using peptide extraction buffer [1:1v/v mixture of 70\% ACN and 0.1\% trifluoroacetic acid (TFA)].

\textbf{Mass spectrometric analysis:} Mass spectrometry (MS) was carried out as described earlier\cite{8}. The digested samples were desalted and concentrated on C-18 ZipTips (Millipore, USA) using the manufacturer’s protocol before mass spectrometric analysis. ZipTips were eluted on MTP 384 target plate with 2 \mu l of \alpha-cyano-4-hydroxy-cinnamic acid (HCCA) (Sigma-Aldrich, USA) saturated solution dissolved in 50 per cent ACN and 0.2 per cent TFA. Mass spectra of digested proteins were acquired using Autoflex II TOF/TOF 50 (Bruker GmbH, Leipzig, Germany) in positive reflectron mode. AnchorChip target plate was placed in sample inlet of the instrument, controlled by flexControl 2.4 software (Bruker, Germany). The instrument was equipped with a 337 nm nitrogen LASER, delayed extraction
Isolation of Mtb total RNA and real-time quantitative PCR (qRT-PCR): Mtb H37Rv was grown in Middlebrook 7H9 broth containing 10 per cent OADC, and was treated with INH (0.1µg/ml), RIF (1.0µg/ml), EMB (5µg/ml) and INH (0.1µg/ml) + RIF (1.0µg/ml). Total RNA was isolated using a TRI reagent (Sigma, USA) following manufacturer’s instructions. To analyze mRNA expression, cDNA was synthesized from 1 µg of total RNA by using Superscript III (Invitrogen Life Technologies, USA) and random primers (Invitrogen, Life Technologies, USA), followed by amplification of the gene(s) by gene-specific primers, using master mix SYBR green (Applied Biological Materials Inc., Canada). The expression is represented in fold increase. 16sRNA was used as internal control for mRNA expression analysis of ahpC, Rv1827, pknA, pknB, pknG and wag31.

Each reaction was repeated thrice with three independent RNA samples in a smart cycler Cepheid machine (Cepheid, USA). RT-PCR conditions were as follows: an initial denaturation step of 10 min, followed by 40 amplification cycles of 30 sec at 95°C, 30 sec at 60°C and 30 sec at 72°C. Melting curve analysis was carried out to confirm the specificity of the amplified product. After baseline corrections and determination of threshold settings, calculation and statistical analyses were carried out using the 2^ΔΔCt Method19. The results are shown as fold increase in expression profile.

Results

Drug resistance pattern and mutations in katG, inhA and rpoB genes of the isolates: The four sequential culture isolates were identified as Mtb by conventional phenotypic and in-house PCR method. Genotyping was done by spoligotyping and the results confirmed that all the isolates belonged to the Central Asian Strain Delhi (CAS1_Delhi, ST26) genotype. The drug susceptibility test results showed that the initial isolate (isolate A) was sensitive to all the four first-line drugs (SM, INH, RIF and EMB), but the consecutive isolates (isolates B, C and D) became resistant to three drugs; INH, RIF and EMB. The minimum inhibitory concentration (MIC) of isolate B increased as compared to isolate A against INH, RIF and EMB. However, it was still sensitive to kanamycin. The isolates C and D became resistant not only to INH, RIF and EMB but also to kanamycin. The sequencing of the rpoB (RIF), katG and inhA (INH) regions revealed mutated alleles associated with resistance to the respective drugs13. Morphologically the resistant isolates were stunted, thicker and coccobacillary in shape.

Glycogen estimation: Logarithmic and stationary phase growth of Mtb sensitive (isolate A) and MDR (isolates B,C,D) isolates were collected by centrifugation (3 min at 5000 x g and at 4°C), and the pellet (15-20 mg wet weight) was re-suspended in 0.25M Na2CO3 and incubated at 95°C for 4 h. The glycogen content was estimated by following the procedure of Schulze et al.18.
Differentially expressed proteins in drug sensitive and resistant isolates: The cell lysate proteins of four *Mtb* isolates were analyzed by 2D gel electrophoresis, which showed 430 protein spots in isolates A and 495, 556 and 395 spots in isolates B, C and D, respectively (Figs 1 & 2). Quantitative analysis of 2D gel spots was carried out using PDQuest software which revealed 27 spots upregulated in MDR isolates (Table I). The spots showing more than 2-fold upregulation were further identified by MALDI-TOF/TOF MS (Table II). To rule out possibility of any artifact, proteins showing equal intensity were taken as internal control (represented as square in Fig. 1). Upregulated proteins were functionally classified according to Tuberculist web server which showed that most of the identified proteins belonged to the functional group 0, 1, 2, 3, 5, 7 and 9; corresponding to virulence, detoxification and adaptation (18.5%), lipid metabolism (11.1%), information pathway (14.8%), cell wall and cell process (3.7%), insertion sequence and phages (18.5%), and intermediary metabolism and information (29.6%); respectively (Table III). The magnified regions of upregulated proteins are shown in Fig. 3. Of the 27 upregulated proteins, eight were hypothetical protein (Rv2004c), probable glutamyl-tRNA (GLN) amidotransferase A gatA (Rv3011c), possible phosphoserine aminotransferase SerC (Rv0884c), probable lipase/esterase LipN (Rv2970c), probable phosphoglycerate kinase Pgg (Rv1437), conserved hypothetical protein with FHA domain, GarA (Rv1827), bacterioferritin (Rv1876) and conserved hypothetical protein (Rv0543) and were not found in 2D-PAGE database system accessible at http://www.mpiib-berlin.mpg.de/2D-PAGE, whereas three proteins probable iron-regulated aconitase hydratase Acn (Rv1475c), probable chaperone protein DnaK (Rv0350) and 60 kDa chaperonin 2 groEL2 (Rv0440), were found in two spots. Six proteins (gatA, serC, fbd, garA, Rv2204c and Rv0543c) in isolate B, 10 proteins (Rv685c, Rv3457c, Rv1479, Rv2970c, Rv1437, qor, and two spots each of Rv1475c and groEL2 family) in isolate C, three proteins (fadB, fabG4 and rrf) in isolate D, three proteins (Rv3075c, Rv1436 and GroES) in isolates C as well as in D were found upregulated. Only five proteins were consistently upregulated in all the three resistant isolates and these were identified as, chaperonin protein dnaK HSP70 (spots 4 and 5), hypothetical protein (Rv2004, spot 8), antigen 84 (wag31, spot 19) and bfrA (spot 24) (Fig. 1).

Among the identified proteins, we were more interested in studying the possible role of Rv1827 (GarA) and wag31 in drug resistance, since these proteins have been identified as physiological substrates for protein kinases G (pknG). Our result revealed that GarA and wag31 were upregulated in the drug resistant isolates. We analyzed the mRNA expression of Rv1827 and its cognate protein kinases, pknG, pknB, pknA and wag31.

Drug induced changes in mRNA expression of protein kinases: To verify our protein expression observations, we studied the mRNA expression profile to see the effect of the four drugs on the standard strain of *Mtb* (H37Rv) which was sensitive to all anti-TB drugs. For this, the mRNA from H37Rv strain was isolated before and after exposing it to INH (0.1µg/ml), EMB (5.0 µg/ml), RIF (1.0 µg/ml) and INH+RIF (0.1 + 1.0µg/ml) for 6 h. Consistent with the proteomic data, seen in clinical isolates, Rv1827 expression was upregulated in all the tested conditions. As expected, the upregulation was 6.82 fold when the *Mtb* standard strain (H37Rv) was exposed to INH and RIF together, but other tested genes had relatively diminished expression. While combining the EMB, the expression of wag31 was higher and pknA and pknG expressions were highest (Fig. 4).

Glycogen storage: GarA, which is a glyogen regulatory protein, was found upregulated in our MDR isolates. It was found that as compared to sensitive isolates the glycogen accumulation in MDR isolates was higher. Consistent with GarA protein levels, the glycogen accumulation measured after seven days was 1.8, 2.0 and 2.1 folds higher in isolates B, C and D, respectively, as compared to sensitive isolate A. Interestingly, after 15 days the glycogen storage remained almost unchanged (Fig. 5).

Discussion

Emergence of drug resistance in *Mtb* has become a major concern for TB control programme managers and treating physicians. Though advances in genome sequencing methods have provided better opportunities to our understanding about functional genomics and proteomics of the *Mtb*, the knowledge about mechanism of drug resistance still remains limited only to the association of genetic polymorphism. Most often, the data from proteomic studies are used to understand host-pathogen interaction, virulence, drug resistance and drug tolerance. Such studies have provided a comprehensive list of *Mtb* proteins that are found differentially regulated in laboratory maintained standard H37Rv strain exposed to drug pressure.
Fig. 1. 2D gel profiles of whole cell lysate proteins of *Mtb* clinical isolates collected sequentially from a single patient. The upregulated proteins are highlighted by circles. (A) First *Mtb* isolate (before treatment) sensitive to all 4 drugs, (B) Second isolate (during treatment) acquires MDR, (C) Third isolate (after 15 months of treatment) acquires drug resistance to yet another drug kanamycin and (D) Fourth MDR isolate after 27 months. Two Proteins, A (Rv1080c) and B (Rv2140c) marked in green rectangles, were selected for observing expression variation in all four samples and showed similar level of expression in all gels.
However, in such studies protein(s) that could be modulated \emph{in vivo} during or after acquiring \emph{in vivo} drug resistance are missed out\footnote{22}. The ever-increasing evidences suggest that the expression of genes\textsuperscript{22} and/or proteins\textsuperscript{23} in clinical isolates is markedly different from the laboratory maintained H37Rv strain; suggesting that majority of these observations may not have direct impact in real life scenario.

In the present study, several proteins were identified that were upregulated in drug resistant isolates. Of the 27 upregulated proteins, five were upregulated in all sequential resistant isolates (B, C and D). Approximately half of these overexpressed proteins are reported as essential for \emph{in vitro} growth of \emph{Mtb}\textsuperscript{21-24}. Although their functional role in drug resistance is elusive, validation of these proteins as a biomarker of drug resistance will provide a scope for finding an effective candidate drug for MDR-TB. It also needs to be emphasized that identifying a protein as “upregulated” does not necessarily imply that it is a true determinant of drug resistance, because it is possible that some or all of these proteins are associated with adaptability of the \emph{Mtb} to survive longer in the host system.

Among the upregulated proteins, rpoA (Rv3457c) was found 2.9 folds upregulated in the drug resistant isolates. In a similar study compensatory mutations in \emph{rpoA} and \emph{rpoC} (Rv0668) were identified particularly in more than 30 per cent of RIF-resistant strains and the authors proposed that mutation in these alleles could also be associated with MDR\textsuperscript{24}. Though we have

\textbf{Fig. 2.} Upregulated proteins, their distributions, amount of overlap and functional classification in \emph{Mtb} clinical isolates: (A) Upregulated proteins distribution and amount of overlap in clinical drug resistant isolates. (B) Functional classification of differentially expressed proteins according to the TubercuList Server.
| Spot No. | Proteins identified                          | Open reading frame (ORF) No. | Accession No. | Mascot score | Nominal mass (Da) | Isoelectric point (pl) | No. of peptides matched | Sequence coverage (%) | Densitometric ratio of protein upregulation between sensitive vs. resistant isolates | Protein classification according to Pasteur Institute of Genomics (Tuberculist) |
|---------|---------------------------------------------|-----------------------------|---------------|--------------|------------------|-----------------------|------------------------|----------------------|----------------------------------------------------------------|---------------------------------------------------------------------|
| 1       | Aconitase hydratase*                        | Rv1475c                     | NP_215991     | 132          | 102728           | 4.95                  | 18                     | 21                   | 1:1.0 1:4.36 1:1.71 | 7                                                                 |
| 2       | Aconitase hydratase*                        | Rv1475c                     | NP_215991     | 126          | 102728           | 4.95                  | 17                     | 21                   | 1:1.22 1:4.70 1:1.58 | 7                                                                 |
| 3       | Probable fabB protein                       | Rv0860                      | NP_215375     | 171          | 76170            | 5.42                  | 24                     | 33                   | 1:1.0 1:1.0 1:2.08 | 1                                                                 |
| 4       | Chaperonin protein dnaK (HSP70)*           | Rv0350                      | NP_214864     | 99           | 66790            | 4.85                  | 11                     | 22                   | 1:6.88 1:7.88 1:4.76 | 0                                                                 |
| 5       | Chaperonin protein dnaK (HSP70)*           | Rv0350                      | NP_214864     | 165          | 66790            | 4.85                  | 22                     | 38                   | 1:156.2 1:195.6 1:66.4 | 0                                                                 |
| 6       | 60kDa Chaperonin 2 (cprn60-2, groEL2)*     | Rv0440                      | NP_214954     | 128          | 56692            | 4.85                  | 16                     | 29                   | 1:1.22 1:4.70 1:1.58 | 0                                                                 |
| 7       | 60 kDa Chaperonin 2 (groEL2)*              | Rv0440                      | NP_214954     | 231          | 56659            | 4.85                  | 19                     | 49                   | 1:1.09 1:4.24 1:1.27 | 0                                                                 |
| 8       | Hypothetical protein                        | Rv2004c                     | NP_216520     | 53           | 54959            | 5.89                  | 10                     | 29                   | 1:54.21 1:132.28 1:49.82 | 5                                                                 |
| 9       | Glutamyl tRNA (Gln) amidotransferase        | Rv3011c                     | NP_217527     | 98           | 51787            | 4.91                  | 9                      | 27                   | 1:2.26 1:1.13 1:0.93 | 2                                                                 |
| 10      | Elongation factor Tu (EF-Tu)               | Rv0685                      | NP_215199     | 248          | 43556            | 5.28                  | 24                     | 69                   | 1:1.38 1:3.07 1:0.97 | 2                                                                 |
| 11      | DNA directed RNA polymerase α-chain (rpoA) | Rv3457c                     | NP_217974     | 205          | 37740            | 4.64                  | 16                     | 43                   | 1:1.7 1:2.98 1:1.0  | 2                                                                 |
| 12      | Putative phosphoserine aminotransferase    | Rv0884c                     | NP_215399     | 90           | 40266            | 4.77                  | 7                      | 28                   | 1:2.21 1:0.70 1:1.75 | 7                                                                 |
| 13      | Probable fabG4 protein                      | Rv0242c                     | NP_214756     | 207          | 46916            | 6.04                  | 17                     | 52                   | 1:1.01 1:1.0 1:2.08 | 1                                                                 |
| 14      | Probable moxR protein                       | Rv1479                      | YP_177816     | 103          | 40738            | 5.96                  | 12                     | 37                   | 1:1.0 1:2.08 1:1.0  | 9                                                                 |
| 15      | Probable lipase                            | Rv2970c                     | NP_217486     | 82           | 43685            | 6.33                  | 6                      | 30                   | 1:1.19 1:2.28 1:1.0  | 7                                                                 |
| 16      | Hypothetical protein                        | Rv1437                      | NP_27230     | 88           | 35519.79         | 5.14                  | 10                     | 41                   | 1:1.91 1:2.82 1:1.28 | 7                                                                 |
| 17      | Hypothetical protein                        | Rv3075c                     | NP_217591     | 98           | 33194            | 4.73                  | 9                      | 28                   | 1:1.73 1:8.12 1:2.65 | 5                                                                 |
| 18      | Glyceraldehyde-3-phosphate dehydrogenase   | Rv1436                      | NP_215952     | 71           | 36105            | 5.19                  | 6                      | 17                   | 1:1.0 1:20.61 1:7.0  | 7                                                                 |
| 19      | Antigen 84 (wag31)*                        | Rv2145c                     | NP_216661     | 120          | 28260            | 4.8                   | 8                      | 43                   | 1:12.0 1:7.0 1:3.0  | 3                                                                 |

Contd...
| Spot No. | Proteins identified                                         | Open reading frame (ORF) No. | Accession No. | Mascot score | Nominal mass (Da) | Isoelectric point (pI) | No. of peptides matched | Sequence coverage (%) | Densitometric ratio of protein upregulation between sensitive vs. resistant isolates | Protein classification according to Pasteur Institute of Genomics (TubercuList) |
|---------|-------------------------------------------------------------|-----------------------------|---------------|--------------|------------------|-----------------------|------------------------|----------------------|---------------------------------------------------------------------------------|----------------------------------------------------------------------------------|
| 20      | Probable quinone oxidoreductase (qor)                       | Rv1454c                     | NP_215970     | 80           | 34140            | 5.37                  | 7                      | 42                   | 1:1.15 1:6.42 1:1.89                                                             | 7                                                                               |
| 21      | Antigen precursor (MPT51)                                   | Rv3803c                     | YP_178017     | 64           | 31069            | 6.13                  | 5                      | 23                   | 1:2.19 1:1.12 1:1.02                                                             | 1                                                                               |
| 22      | Ribosome recycling factor (RRF)                             | Rv2882c                     | NP_217398     | 108          | 20815            | 5.71                  | 10                     | 67                   | 1:1.0 1:1.0 1:3.45                                                             | 2                                                                               |
| 23      | Hypothetical protein (GarA)                                | Rv1827                      | NP_216343     | 87           | 17240            | 4.29                  | 5                      | 45                   | 1:7.66 1:1.37 1:1.16                                                             | 5                                                                               |
| 24      | bfrA (Bacterioferritin)                                    | Rv1876                      | NP_216392     | 90           | 18443            | 4.5                   | 12                     | 82                   | 1:9.27 1:37.5 1:38.58                                                           | 7                                                                               |
| 25      | Hypothetical protein                                       | Rv2204c                     | NP_216720     | 103          | 12707            | 4.4                   | 6                      | 51                   | 1:9.11 1:1.0 1:1.04                                                           | 5                                                                               |
| 26      | Hypothetical protein                                       | Rv0543c                     | NP_215057     | 67           | 14743            | 5.2                   | 5                      | 38                   | 1:3.03 1:1.01 1:1.53                                                         | 5                                                                               |
| 27      | 10 kDa Chaperonin (cpn10, groES protein)                    | Rv3418c                     | NP_217935     | 162          | 10667            | 4.62                  | 10                     | 99                   | 1:1.0 1:10.86 1:9.28                                                        | 0                                                                               |

Cut-off limit ≥2.0 fold for overexpression of proteins. † Spot number of the protein as marked in Fig. 1. * Some proteins are having mobility difference.
| Spot No. | Peak mass (Da) | Protein identified | Rv No. | Nominal Mass (Da) | pI   | Mascot score | Sequence peptide          |
|---------|----------------|--------------------|--------|-------------------|------|--------------|---------------------------|
| 1       | 829.4455       | Aconitase hydratase | Rv1475c | 102728            | 4.95 | 39           | KSYQIYRL                  |
|         | 1132.6464      | Aconitase hydratase | Rv1475c | 102728            | 4.95 | 32           | RNNgilQVLYRNR             |
|         | 1170.5177      | Aconitase hydratase | Rv1475c | 102728            | 4.95 | 34           | RWgQgAfDDFKV              |
|         | 1299.587       | Aconitase hydratase | Rv1475c | 102728            | 4.95 | 37           | RIDTPGEADYYRNR            |
| 2       | 829.4455       | Aconitase hydratase | Rv1475c | 102728            | 4.95 | 37           | KSYQIYRL                  |
|         | 1132.6464      | Aconitase hydratase | Rv1475c | 102728            | 4.95 | 39           | RNNgilQVLYRNR             |
|         | 1170.5177      | Aconitase hydratase | Rv1475c | 102728            | 4.95 | 35           | RWgQgAfDDFKV              |
|         | 1299.587       | Aconitase hydratase | Rv1475c | 102728            | 4.95 | 35           | RIDTPGEADYYRNR            |
| 3       | 1530.73        | Probable fadB protein | Rv0860 | 76170             | 5.42 | 49           | KSGSQQPQLQMDIR           |
|         | 1836.9363      | Probable fadB protein | Rv0860 | 76170             | 5.42 | 40           | KGVDFVIEAVFENQELKH       |
|         | 2078.1801      | Probable fadB protein | Rv0860 | 76170             | 5.42 | 37           | KGQUALPEVTGLLPGGGVTRT     |
| 4       | 1568.0111      | Chaperonin protein dnaK (hsp70) | Rv0350 | 66790            | 4.85 | 50           | KLGSFELTIPAPPRG           |
|         | 1743.8923      | Chaperonin protein dnaK (hsp70) | Rv0350 | 66790            | 4.85 | 37           | RATSGDNHLGDDWDQRV         |
|         | 2109.2639      | Chaperonin protein dnaK (hsp70) | Rv0350 | 66790            | 4.85 | 40           | RNGEVLGPQAKNVNVDRT        |
| 5       | 1062.611       | Chaperone protein dnaK | Rv0350 | 66790            | 4.85 | 32           | RTTPSIJAFAARN             |
|         | 1226.711       | Chaperone protein dnaK | Rv0350 | 66790            | 4.85 | 33           | KDAGQIAGLNVI            |
|         | 1645.962       | Chaperone protein dnaK | Rv0350 | 66790            | 4.85 | 37           | RIVNEPTAAALAYGDLK         |
|         | 2613.426       | Chaperone protein dnaK | Rv0350 | 66790            | 4.85 | 55           | RSETFTADDNQPSVQVYQGERE    |
| 6       | 914.6248       | 60kDa Chaperonin 2 (cpn60-2, groEL2) | Rv0440 | 56692            | 4.85 | 55           | KGVRNVLK11               |
|         | 1223.7164      | 60kDa Chaperonin 2 (cpn60-2, groEL2) | Rv0440 | 56692            | 4.85 | 33           | KTIAYDEEARR                |
|         | 1266.754       | 60kDa Chaperonin 2 (cpn60-2, groEL2) | Rv0440 | 56692            | 4.85 | 35           | MAKTIAYDEEARR              |
|         | 1580.0454      | 60kDa Chaperonin 2 (cpn60-2, groEL2) | Rv0440 | 56692            | 4.85 | 42           | REGLRNVAAGANPLGLKR         |
| 7       | 1067.5167      | 60kDa Chaperonin 2 (cpn60-2, groEL2) | Rv0440 | 56692            | 4.85 | 34           | KTIAYDEEARR                |
|         | 1264.5909      | 60kDa Chaperonin 2 (cpn60-2, groEL2) | Rv0440 | 56692            | 4.85 | 36           | KEELEDPEK11               |
|         | 1529.7887      | 60kDa Chaperonin 2 (cpn60-2, groEL2) | Rv0440 | 56692            | 4.85 | 46           | KGVRAPITNGVIAKE           |
|         | 2075.0432      | 60kDa Chaperonin 2 (cpn60-2, groEL2) | Rv0440 | 56692            | 4.85 | 66           | KTDVVADGTTTAVLFAVQ       |
| 8       | 804.3472       | Hypothetical protein | Rv0204c | 54959            | 5.89 | 35           | RERACIRE                  |
|         | 1434.7007      | Hypothetical protein | Rv0204c | 54959            | 5.89 | 42           | RIEHVMDEFVSGRE            |
|         | 1796.9013      | Hypothetical protein | Rv0204c | 54959            | 5.89 | 45           | RIDDAAFLAMDLEFLGRK        |

Contd...
| Spot No. | Peak mass (Da) | Protein identified | Rv No. | Nominal Mass (Da) | pI | Mascot score | Sequence peptide |
|---------|----------------|--------------------|--------|-------------------|----|--------------|------------------|
| 9       | 1094.6241      | Glutamyl tRNA (Gln | Rv3011c | 51787             | 4.91 | 43            | RSPYDATLTARL     |
|         |                | amidotransferase   |        |                   |     |              |                  |
|         |                | subunit A (gatA)   |        |                   |     |              |                  |
| 9       | 1881.9284      | gatA               |        | 51787             | 4.91 | 35            | RYGLVACASSLDQGP  |
|         |                |                    |        |                   |     |              | CART             |
| 9       | 2017.1693      | gatA               |        | 51787             | 4.91 | 36            | RQPAALTATVGKTPY  |
|         |                |                    |        |                   |     |              | GTVSRY           |
| 10      | 1413.8456      | Ef-tu              | Rv0685 | 43556             | 5.28 | 32            | RQVGVPYILVALNK  |
|         |                |                    |        |                   |     |              |                  |
| 10      | 1693.8359      | Ef-tu              |        | 43556             | 5.28 | 40            | RHYAHVDAPGHADYIK|
|         |                |                    |        |                   |     |              | KN                |
| 10      | 2091.0286      | Ef-tu              |        | 43556             | 5.28 | 42            | KADAVDDEELLEL  |
|         |                |                    |        |                   |     |              | VEMRE             |
| 11      | 1085.5899      | DNA directed RNA   | Rv3457c | 37740             | 4.64 | 35            | KLEVELVERG       |
|         |                | polymerase α-chain |        |                   |     |              |                  |
| 11      | 1485.8157      | DNA directed RNA   |        | 37740             | 4.64 | 27            | RTLSSIPGAAVTSIRI|
|         |                | polymerase α-chain |        |                   |     |              |                  |
| 11      | 1611.8113      | DNA directed RNA   |        | 37740             | 4.64 | 25            | RIDGVLHEFTTVPGV | K                 |
|         |                | polymerase α-chain |        |                   |     |              | KE                  |
| 12      | 1352.6949      | Putative phosphoserine | 40266 | 4.77              | 34  |               | RSLHLTYGEFSAKF  |
|         |                | aminotransferase   |        |                   |     |              |                  |
| 12      | 1900.0249      | Putative phosphoserine | 40266 | 4.77              | 33  |               | MADQLTPHLEIPTA   |
|         |                | aminotransferase   |        |                   |     |              | IKPRD             |
| 12      | 1928.0345      | Putative phosphoserine | 40266 | 4.77              | 38  |               | RWVPDFLSLPIA      |
|         |                | aminotransferase   |        |                   |     |              | VENSLKN           |
| 13      | 1237.69        | Probable fabG4 protein | Rv0242c | 46916             | 6.04 | 44            | RQLGVPQPETLRR    |
| 13      | 1393.7932      | Probable fabG4 protein |        | 46916             | 6.04 | 35            | RQLGVPQPETLRRY   |
| 13      | 1565.862       | Probable fabG4 protein |        | 46916             | 6.04 | 39            | RAGEEPSTGLSIGAG  |
| 13      | 1425.9539      | Probable moxR      | Rv1479 | 40738             | 5.96 | 41            | KRIIVGQDQLVERM   |
| 13      | 1786.142       | Probable moxR      |        | 40738             | 5.96 | 32            | RQIQTDPVLPTD     |
| 13      | 1983.2737      | Probable moxR      |        | 40738             | 5.96 | 33            | RDYVIPQDVIEVPI   |
| 13      | 1550.7799      | Probable lipase protein | Rv2970c | 34146             | 4.83 | 34            | RVVDLADGPGPPIG  |
| 14      | 1699.7705      | Probable lipase protein |        | 34146             | 4.83 | 37            | RQHAVGADAIIVSV   |
| 14      | 1740.8621      | Probable lipase protein |        | 34146             | 4.83 | 22            | RIAVAGDSAGGTIA   |
| 14      | 1315.7195      | Phosphoglycerate kinase | Rv1437 | 42600             | 4.83 | 43            | RGLLETYHDVLR    |
| 15      | 1420.7409      | Phosphoglycerate kinase |        | 42600             | 4.83 | 40            | KGAFSVVGGDSAA   |
| 15      | 1683.9461      | Phosphoglycerate kinase |        | 42600             | 4.83 | 38            | RAEGLTGGDILLLEN | IRF               |
| 16      | 1016.503       | Hypothetical protein  | Rv3075c | 33194             | 4.73 | 25            | KEFFAEFARD       |
| 16      | 1322.619       | Hypothetical protein  | Rv3075c | 33194             | 4.73 | 60            | RWFGDGNADWVRI    |
| 16      | 1583.78        | Hypothetical protein  | Rv3075c | 33194             | 4.73 | 59            | RDTGFGEKPATLAYARS|
| Spot No. | Peak mass (Da) | Protein identified | Rv No. | Nominal Mass (Da) | pI | Mascot score | Sequence peptide |
|---------|---------------|--------------------|--------|------------------|----|--------------|------------------|
| 18      | 1085.7325     | Hypothetical protein Rv3075c |        | 33194            | 4.73 | 43           | KRLPNVPIVALVETARG |
| 1134.7153 | G-3-P dehydrogenase |            |        | 36105            | 5.19 | 37           | KVLDDFEFIGVK |
| 1384.8723 | G-3-P dehydrogenase |            |        | 36105            | 5.19 | 34           | RAAALNIVPTSTGAAKA |
| 19      | 1088.5844     | ag84/wag31        | Rv2145c | 28260            | 4.8  | 22           | RLIEENSDLRQ |
| 1171.6518 | ag84/wag31 |            |        | 28260            | 4.8  | 31           | RANAEOILGEARH |
| 1413.7358 | ag84/wag31 |            |        | 28260            | 4.8  | 36           | KHEISMGTINQRA |
| 20      | 1011.5087     | qor               | Rv1454c | 34140            | 5.37 | 43           | RTGEEFSWRA |
| 1615.8965 | qor            |            |        | 34140            | 5.37 | 40           | KAEIAGVFNDITYFRS |
| 2195.0742 | qor            |            |        | 34140            | 5.37 | 34           | KDAGADVVLDPADQFARGV |
| 21      | 1037.4687     | Antigen precursor (MPT51) Rv3803c |        | 31069            | 6.13 | 34           | RMFYNYQRS |
| 2044.93 | Antigen precursor (MPT51) |            |        | 31069            | 6.13 | 42           | KWHDQPWHASLLAQNTRV |
| 2132.9473 | Antigen precursor (MPT51) |            |        | 31069            | 6.13 | 45           | KQWDTFLSAELPDWLAANRG |
| 22      | 1655.8488     | Ribosome recycling factor (RRF) Rv2882c |        | 20815            | 5.71 | 29           | RNSDLGVNPNDGALIRV |
| 1674.88 | Ribosome recycling factor (RRF) |            |        | 20815            | 5.71 | 37           | KTTHQYVTQIDELVKH |
| 2146.0955 | Ribosome recycling factor (RRF) |            |        | 20815            | 5.71 | 38           | KDLDKTHQYVTQIDELVKH |
| 23      | 1291.7384     | Hypothetical protein (GarA) Rv1827 |        | 17240            | 4.29 | 21           | RFLLDQAITSAGRH |
| 1715.8809 | Hypothetical protein (GarA) |            |        | 17240            | 4.29 | 43           | RHPDSIFLDDVTVSRR |
| 1840.9915 | Hypothetical protein (GarA) |            |        | 17240            | 4.29 | 28           | REPVDSAVALANGDEVQIGF |
| 24      | 1046.5234     | Bfr (Bacterioferritin) Rv1876 |        | 18443            | 4.5  | 40           | MQGDPDVLRL |
| 1414.7907 | Bfr (Bacterioferritin) |            |        | 18443            | 4.5  | 37           | RILLLDGLPNYQRI |
| 1935.8135 | Bfr (Bacterioferritin) |            |        | 18443            | 4.5  | 29           | RAESFDEMRAEETDRI |
| 25      | 1089.5305     | Hypothetical protein Rv2204c |        | 12707            | 4.4  | 22           | RYNLFFDDRT |
| 1281.7399 | Hypothetical protein Rv2204c |            |        | 12707            | 4.4  | 28           | KTHGVILTEAAAAKA |
| 1198.665 | Hypothetical protein Rv2204c |            |        | 12707            | 4.4  | 35           | RIAVQPGGCAGLRY |
| 26      | 965.4423      | Hypothetical protein Rv0543c |        | 14743            | 5.2  | 26           | RDDAPYWAKY |

Contd...
not screened our resistant isolate for these mutations, this study supports our proteomic approach and identification of an upregulated rpoA protein in drug resistant isolates. Dussurget et al\textsuperscript{25} reported that in \textit{M. smegmatis}, IdeR negatively controls iron-uptake and expression of BfrA and BfrB. In our study, BfrA was upregulated in drug resistant isolates, suggesting a role of this protein in inducing resistance to INH. However, such conclusions could not be validated by deletion mutant of \textit{bfrA} and \textit{bfrB} in \textit{Mtb}\textsuperscript{22}.

We found that the identified proteins spots (Rv1475c, dnaK, groEL2, groES and wag31) had different electrophoretic mobilities in resistant and sensitive isolates. Similar observations have been reported by Mattow et al\textsuperscript{21} when the protein profile of intra-phagosomal \textit{Mtb} H37Rv was analyzed. It has also been suggested that this shift is determined by the protein modifications rendered during sample preparation and growth conditions adopted by various laboratories\textsuperscript{8}. However, this may be the unlikely factor in our study, as we strictly followed the same protocol throughout the study. Further, our analysis was stringent and we considered a particular protein to be “upregulated” only if the observation was consistent in three independent experiments. Even then, we were cautious not to conclude whether the changes in protein mobility necessarily reflect protein modification. Further studies are warranted to establish the role of post-translation modifications in these proteins during drug resistance.

Antigen 84 (Wag31 / Rv2145c) has been demonstrated to be involved in the regulation of cell morphology and \textit{pknG} mediates the survival of \textit{Mtb} inside the host macrophages, and genome-wide transcriptional analysis reveals that the \textit{pknG} is upregulated during the exposure of INH\textsuperscript{26}. Further, \textit{pknG} is associated with the intrinsic resistance of \textit{Mtb} to various anti-TB drugs\textsuperscript{26}. Wag31 protein has been found to be overexpressed and involved in regulation of cell morphology. It also plays important role in survival of mycobacteria under oxidative stress\textsuperscript{27} and provides optimal substrate for \textit{pknA} and \textit{pknB}. In addition, mRNA expression of wag31 was increased by 15.7 folds during the INH and RIF exposure. Earlier also, wag31 has been reported to be over expressed in the MDR isolates\textsuperscript{8}, supporting our hypothesis that wag31 plays an important role in drug resistance.

Of the five hypothetical proteins (Rv2004c, Rv1437, Rv3075c, Rv2204c and Rv1827) overexpressed in MDR isolates, three (Rv2004c, Rv2204c and Rv1437) could not be assigned to any function in survival or pathogenesis of the bacteria, though Rv3075c has been reported to be overexpressed in streptomycin resistant isolates\textsuperscript{11}. In our study, Rv1827 (GarA) was found 7.6 folds upregulated in drug resistant isolates as compared to that in the susceptible isolate A. It has been identified as an optimal substrate for PknB and PknG. The protein is also reported to act as a phosphorylation-dependent molecular switch in mycobacterial signalling process mediated by protein kinases\textsuperscript{28-30}. Further, GarA has been found to be predominantly expressed under the exponential growth phase and has been suggested as a regulatory model for glycogen degradation and glutamate metabolism\textsuperscript{28-30}. To infer whether the protein overexpression of Rv1827 in drug resistant isolates facilitated increased glycogen accumulation, we quantified the glycogen content of drug resistant and sensitive isolates. The findings were consistent with this hypothesis and the glycogen content was relatively

| Spot No.\textsuperscript{†} | Peak mass (Da) | Protein identified | Rv No. | Nominal Mass (Da) | pI | Mascot score | Sequence peptide |
|----------------------------|---------------|--------------------|--------|------------------|----|--------------|------------------|
| 1218.6172                  |               | Hypothetical protein Rv0543c | 14743  | 14743            | 5.2 | 25           | MSVELTQEVSARL     |
| 1595.7916                  |               | Hypothetical protein Rv0543c | 14743  | 14743            | 5.2 | 19           | RLTSDLYGWLTTVARS  |
| 27                        | 1034.4976     | 10kDa Chaperonin groES | Rv3418c | 10798            | 4.62| 40           | RWDEDGEKRI         |
| 1523.8874                  |               | 10kDa Chaperonin groES | Rv3418c | 10798            | 4.62| 21           | KEKPOEGTVVAVGPRW   |
| 1776.0298                  |               | 10kDa Chaperonin groES | Rv3418c | 10798            | 4.62| 29           | KRIPLDVAEGDTVIYSKY |

\textsuperscript{†}Spot number of the proteins marked as in Fig. 1.
| S. No. | Spot No. | Proteins identified | Open reading frame (ORF) No. | Accession No. | Mascot score | Nominal mass (Da) | pI | No. of peptides matched | Sequence coverage (%) | Densitometric ratio of protein expression between sensitive vs. resistant isolates | Protein classification according to Pasteur Institute of Genomics (TubercuList) |
|-------|---------|---------------------|-----------------------------|---------------|--------------|-----------------|----|------------------------|----------------------------|-----------------------------------------------------------------|---------------------------------------------------------------|
|       |         | Proteins overexpressed in MDR isolate (Isolate B) |                             |               |              |                 |     |                        |                           | B: 1:2.26  C: 1:1.13  D: 1:0.93 |                                                              |
| 1     | 9       | Glutamyl tRNA (Gln amidotransferase) | Rv3011c | NP_217527 | 98          | 51787          | 4.91 | 9                      | 27                         | 2                                |
| 2     | 12      | Putative phosphoserine aminotransferase | Rv0884c | NP_215399 | 90          | 40266          | 4.77 | 7                      | 28                         | 7                                |
| 3     | 21      | Antigen precursor (MPT51) | Rv3803c | YP_178017 | 64          | 31069          | 6.13 | 5                      | 23                         | 1                                |
| 4     | 23      | Hypothetical protein (GarA) | Rv1827 | NP_216343 | 87          | 17240          | 4.29 | 5                      | 45                         | 5                                |
| 5     | 25      | Hypothetical protein | Rv2204c | NP_216720 | 103         | 12707          | 4.4  | 6                      | 51                         | 5                                |
| 6     | 26      | Hypothetical protein | Rv0543c | NP_215057 | 67          | 14743          | 5.2  | 5                      | 38                         | 5                                |
|       |         | Proteins overexpressed in MDR isolate (Isolate C) |                             |               |              |                 |     |                        |                           | B: 1:1.0  C: 1:4.36  D: 1:1.71 |                                                              |
| 1     | 1       | Aconitase hydratase | Rv1475c | NP_215991 | 132         | 102728         | 4.95 | 18                     | 21                         | 7                                |
| 2     | 2       | Aconitase hydratase | Rv1475c | NP_215991 | 126         | 102728         | 4.95 | 17                     | 21                         | 7                                |
| 3     | 6       | 60kDa Chaperonin 2 (spn60-2, groEL2) | Rv0440 | NP_214954 | 128         | 56692          | 4.85 | 16                     | 29                         | 0                                |
| 4     | 7       | 60 kDa Chaperonin 2 (groEL2) | Rv0440 | NP_214954 | 231         | 56659          | 4.85 | 19                     | 49                         | 0                                |
| 5     | 10      | Elongation factor Tu (EF-Tu) | Rv0685 | NP_215199 | 248         | 43556          | 5.28 | 24                     | 69                         | 2                                |
| 6     | 11      | DNA directed RNA polymerase α-chain (ip0A) | Rv3457c | NP_217974 | 205         | 37740          | 4.64 | 16                     | 43                         | 2                                |

Contd...
| S. No. | Spot No. | Proteins identified                  | Open reading frame (ORF) No. | Accession No. | Mascot score | Nominal mass (Da) | pI | No. of peptides matched | Sequence coverage (%) | Densitometric ratio of protein expression between sensitive vs. resistant isolates | Protein classification according to Pasteur Institute of Genomics (TubercuList) |
|--------|----------|------------------------------------|-----------------------------|--------------|--------------|------------------|----|------------------------|----------------------|----------------------------------------------------------------|------------------------------------------------------------------------|
| 7      | 14       | Probable moxR protein              | YP_177816                   | 103          | 40738        | 5.96             | 12 | 37                     | 1:1.0    | 1:2.08                   | 1:1.0                        | 9                                                                   |
| 8      | 15       | Probable lipase                    | NP_217486                   | 82           | 43685        | 6.33             | 6  | 30                     | 1:1.19   | 1:2.28                   | 1:1.0                        | 7                                                                   |
| 9      | 16       | Hypothetical protein               | NP_217230                   | 88           | 35519.79     | 5.14             | 10 | 41                     | 1:1.91   | 1:2.82                   | 1:1.28                       | 7                                                                   |
| 10     | 20       | Probable quinone oxidoreductase (qor) | NP_215970                  | 80           | 34140        | 5.37             | 7  | 42                     | 1:1.15   | 1:6.42                   | 1:1.89                       | 7                                                                   |
|        |          |                                    |                             |              |              |                  |     |                        |                      |                                                                     |                                                                      |
|        |          | Proteins overexpressed in MDR isolate (Isolate D) |                             |              |              |                  |     |                        |                      |                                                                     |                                                                      |
| 1      | 3        | Probable fadB protein              | NP_215375                   | 171          | 76170        | 5.42             | 24 | 33                     | 1:1.0    | 1:1.0                    | 1:2.08                       | 1                                                                   |
| 2      | 13       | Probable fabG4 protein             | NP_214756                   | 207          | 46916        | 6.04             | 17 | 52                     | 1:1.01   | 1:1.0                    | 1:2.08                       | 1                                                                   |
| 3      | 22       | Ribosome recycling factor (RRF)    | NP_217398                   | 108          | 20815        | 5.71             | 10 | 67                     | 1:1.0    | 1:1.0                    | 1:3.45                       | 2                                                                   |
|        |          |                                    |                             |              |              |                  |     |                        |                      |                                                                     |                                                                      |
|        |          | Proteins overexpressed only in MDR isolates (Isolates C and D) |                             |              |              |                  |     |                        |                      |                                                                     |                                                                      |
| 1      | 17       | Hypothetical protein               | NP_217591                   | 98           | 33194        | 4.73             | 9  | 28                     | 1:1.73   | 1:8.12                   | 1:2.65                       | 5                                                                   |
| 2      | 18       | Glyceraldehyde-3-phosphate         | NP_215952                   | 71           | 36105        | 5.19             | 6  | 17                     | 1:1.0    | 1:20.61                  | 1:7.0                        | 7                                                                   |
| 3      | 27       | 10 kDa chaperonin (cpn 10, groES)  | NP_217935                   | 162          | 10667        | 4.62             | 10 | 99                     | 1:1.0    | 1:10.86                  | 1:9.28                       | 0                                                                   |
|        |          |                                    |                             |              |              |                  |     |                        |                      |                                                                     |                                                                      |
|        |          | Proteins overexpressed only in all MDR isolates (Isolates B, C and D) |                             |              |              |                  |     |                        |                      |                                                                     |                                                                      |
| 1      | 4        | Chaperonin protein dnaK (HSP70)    | NP_214864                   | 99           | 66790        | 4.85             | 11 | 22                     | 1:6.88   | 1:7.88                   | 1:4.76                       | 0                                                                   |
| 2      | 5        | Chaperonin protein dnaK (HSP70)    | NP_214864                   | 165          | 66790        | 4.85             | 22 | 38                     | 1:156.2  | 1:195.6                  | 1:66.4                       | 0                                                                   |
| 3      | 8        | Hypothetical protein               | NP_216520                   | 53           | 54959        | 5.89             | 10 | 29                     | 1:54.21  | 1:132.28                 | 1:49.82                      | 5                                                                   |
| 4      | 19       | Antigen 84 (wag31)                 | NP_216661                   | 120          | 28260        | 4.8              | 8  | 43                     | 1:12.0   | 1:7.0                    | 1:3.0                        | 3                                                                   |
| 5      | 24       | bfrA (Bacterioferritin)            | NP_216392                   | 90           | 18443        | 4.5              | 12 | 82                     | 1:9.27   | 1:37.5                   | 1:38.58                      | 7                                                                   |
more in resistant isolates (B, C and D) than the sensitive isolate A. However, the difference in accumulation was significant up to 7th day, but not after 15th day of growth, suggesting that the expression of Rv1827 might be an important marker of dormancy. While the precise role of glycogen storage in mycobacteria is not known, glycogen stores may serve as a reservoir of carbon and energy that can be mobilized by mycobacteria for survival during periods of carbon starvation. Such observations have been made in other bacteria such as Vibrio cholerae during transition stage between host and aquatic environments.

In our study, the identified proteins cannot be categorically classified as drug resistance-specific alterations, as this can also happen due to host immune response or stresses encountered by the individual strain in vivo. Most of these proteins are essential for the survival of mycobacteria in phagosome or as virulence factor (unpublished observation), but their role in drug resistance cannot be ruled out completely.

The 2D gel electrophoresis followed by MS-based proteomic analysis on sequential isolates showed approximately 500 proteins per gel. This resolution is
Fig. 4. Detection of mRNA expression in drug treated *Mtb* isolates. Isolates treated with isoniazid (INH, 0.1µg/ml), rifampicin (RIF, 1.0µg/ml), ethambutol (EMB, 5µg/ml) and INH (0.1µg/ml) +RIF (1.0µg/ml). The expression is represented in fold increase. The 16sRNA was used as internal control for mRNA expression analysis. mRNA expression of (A) *ahpC*, (B) Rv1827, (C) *pknA*, (D) *pknB*, (E) *pknG* and (F) *wag31*. Calculation and statistical analysis carried out by using the 2^−ΔΔC_T method and results presented in fold increase.

much less than expected, as there are 4000 predicted genes in *Mtb*. The poor sensitivity observed in the present study could be attributed to various reasons such as the extraction protocol; low resolution power of the coomassie brilliant blue stain used or due to the IEF-strips (pH 4-7). These strips resolve only the proteins having isoelectric point in the range of 4-7.

In conclusion, our study highlights the intricacies associated with sequential clinical isolates of *Mtb*, a rare opportunity for any laboratory, which is a natural phenomenon and cannot be generated artificially in the laboratory. The sequential isolation of four isolates from the same patient during the treatment period showed a phenomenon where a sensitive isolate turned
to a multidrug resistant isolate. It is possible that some of the upregulated proteins identified from MDR clinical isolates of *Mtb* in the present study may prove as potential biomarkers of drug resistance in future.

**Acknowledgment**

This work was supported by grant from Department of Biotechnology, Government of India to the last author (SS) (BT/PR7880/ MED/ 14/1161/2006). The first author (AS) received financial assistance in the form of senior research fellowship from the Indian Council of Medical Research, Government of India. The authors thank Shri Sandeep Kadian for his help in spoligotyping data analysis and Dr V.M. Katoh, Director-General of Indian Council of Medical Research, New Delhi, for allowing us to use his laboratory facilities for MALDI-TOF/TOF MS at JALMA, Agra.

**Patent Applied:** Sarman Singh, Gopinath K., Amit Singh and Niti Singh. Novel protein markers of drug resistance in *Mycobacterium tuberculosis* (1752/DEL/2008).

**References**

1. World Health Organization. *Global Tuberculosis Report* 2014. Available from: http://apps.who.int/iris/bitstream/10665/137094/1/9789241564809_eng.pdf?ua=1, accessed on November 3, 2014.

2. Rattan A, Kalia A, Ahmad N. Multidrug-resistant *Mycobacterium tuberculosis*: molecular perspectives. Emerg Infect Dis 1998; 4: 195-209.

3. Ramaswamy SV, Reich R, Dou SJ, Jaspere L, Pan X, Wanger A, *et al*. Single nucleotide polymorphisms in genes associated with isoniazid resistance in *Mycobacterium tuberculosis*. Antimicrob Agents Chemother 2003; 47: 1241-50.

4. Gryadunov D, Mikhailovich V, Lapa S, Roudinskii N, Donnikov M, Panikov S, *et al*. Evaluation of hybridization on oligonucleotide microarrays for analysis of drug-resistant *Mycobacterium tuberculosis*. Clin Microbiol Infect 2005; 11: 531-9.

5. Louw GE, Warren RM, Gey van Pittius NC, McEvoy CR, Van Helden PD, Victor TC. A balancing act: efflux/influx in mycobacterial drug resistance. Antimicrob Agents Chemother 2009; 53: 3181-9.

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

7. Wang R, Marcotte EM. The proteomic response of *Mycobacterium smegmatis* to anti-tuberculosis drugs suggests targeted pathways. J Proteome Res 2008; 7: 855-65.

8. Jiang X, Zhang W, Gao F, Huang Y, Lu C, Wang H. Comparison of the proteome of isoniazid-resistant and -susceptible strains of *Mycobacterium tuberculosis*. Microb Drug Resist 2006; 12: 231-8.

9. Sharma P, Kumar B, Singhal N, Katoch VM, Chauhan DS, *et al*. Streptomycin induced protein expression analysis in *Mycobacterium tuberculosis* by two-dimensional gel electrophoresis & mass spectrometry. Indian J Med Res 2010; 132: 400-8.

10. Jungblut PR, Schaible UE, Mollenkopf HJ, Zimny-Arndt U, Raupach B, Mattow J, *et al*. Comparative proteome analysis of *Mycobacterium tuberculosis* and *Mycobacterium bovis* BCG strains: towards functional genomics of microbial pathogens. Mol Microbiol 1999; 33: 1103-17.

11. Sharma P, Kumar B, Gupta Y, Singhal N, Katoch VM, Venkatesan K, *et al*. Proteomic analysis of streptomycin resistant and sensitive clinical isolates of *Mycobacterium tuberculosis*. Proteome Sci 2010; 8: 59.

12. Gopinath K, Singh S. Multiplex PCR assay for simultaneous detection and differentiation of *Mycobacterium tuberculosis*, *Mycobacterium avium* complexes and other Mycobacterial species directly from clinical specimens. J Appl Microbiol 2009; 107: 425-35.

13. Singh A, Gopinath K, Singh N, Singh S. Deciphering the sequential events during in vivo acquisition of drug resistance in *Mycobacterium tuberculosis*. Int J Mycobacteriol 2014; 3: 36-40.

14. Patricia TK, Cubica GP. *Public health mycobacteriology: a guide for the level III laboratory*. Atlanta, Ga: U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control; 1995.

15. Demay C, Liens B, Burguière T, Hill V, Couvin D, Millet J, *et al*. SITVITWEB - a publicly available international multilocus database for studying *Mycobacterium tuberculosis* genetic diversity and molecular epidemiology. Infect Genet Evol 2012; 12: 755-66.
16. Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principles of protein-dye binding. *Anal Biochem* 1976; 72: 248-54.

17. Shevchenko A, Tomas H, Havlis J, Olsen JV, Mann M. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat Protoc* 2006; 1: 2856-60.

18. Schulze U, Larsen ME, Villadsen J. Determination of intracellular trehalose and glycogen in *Saccharomyces cerevisiae*. *Anal Biochem* 1995; 228: 143-9.

19. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2−ΔΔC method. *Methods* 2001; 25: 402-8.

20. Reddy PV, Puri RV, Khera A, Tyagi AK. Iron storage proteins are essential for the survival and pathogenesis of *Mycobacterium tuberculosis* in THP-1 macrophages and the guinea pig model of infection. *J Bacteriol* 2012; 194: 567-75.

21. Mattow J, Siejak F, Hagens K, Becher D, Albrecht D, Krah A, et al. Proteins unique to intraphagosomally grown *Mycobacterium tuberculosis*. *Proteomics* 2006; 6: 2485-94.

22. Homolka S, Niemann S, Russell DG, Rohde KH. Functional genetic diversity among *Mycobacterium tuberculosis* complex clinical isolates: delineation of conserved core and lineage-specific transcriptomes during intracellular survival. *PLoS Pathog* 2010; 6: e1000988.

23. Pheiffer C, Betts JC, Flynn HR, Lukey PT, van Helden P. Protein expression by a Beijing strain differs from that of another clinical isolate and *Mycobacterium tuberculosis* H37Rv. *Microbiology* 2005; 151: 1139-50.

24. Comas I, Borrell S, Roetzer A, Rose G, Malla B, Kato-Maeda M, et al. Whole-genome sequencing of rifampicin-resistant *Mycobacterium tuberculosis* strains identifies compensatory mutations in RNA polymerase genes. *Nat Genet* 2011; 44: 106-10.

25. Dussurget O, Rodriguez M, Smith I. Protective role of the *Mycobacterium smegmatis* IdeR against reactive oxygen species and isoniazid toxicity. *Tuber Lung Dis* 1998; 79: 99-106.

26. Wolff KA, Nguyen HT, Cartabuke RH, Singh A, Ogwang S, Nguyen L, et al. Protein kinase G is required for intrinsic antibiotic resistance in mycobacteria. *Antimicrob Agents Chemother* 2009; 53: 3515-9.

27. Mukherjee P, Sureka K, Datta P, Hossain T, Barik S, Das KP, et al. Novel role of Wag31 in protection of mycobacteria. *Antimicrob Agents Chemother* 2009; 53: 3515-9.

28. O'Hare HM, Durán R, Cerveñansky C, Bellinzoni M, Wehenkel AM, Pritsch O, et al. Regulation of glutamate metabolism by protein kinases in mycobacteria. *Mol Microbiol* 2008; 70: 1408-23.

29. Kang CM, Abbott DW, Park ST, Dascher CC, Cantley LC, Hussen RN. The *Mycobacterium tuberculosis* serine/threonine kinases PknA and PknB: substrate identification and regulation of cell shape. *Genes Dev* 2005; 19: 1692-704.

30. Englund P, Wehenkel A, Martins S, Hoos S, André-Leroux G, Villarino A, et al. The FHA-containing protein GarA acts as a phosphorylation-dependent molecular switch in mycobacterial signaling. *FEBS Lett* 2009; 583: 301-7.

31. Bourassa L, Camilli A. Glycogen contributes to the environmental persistence and transmission of *Vibrio cholerae*. *Mol Microbiol* 2009; 72: 124-38.

*Reprint requests:* Dr Sarman Singh, Division of Clinical Microbiology & Molecular Medicine, Department of Laboratory Medicine, All India Institute of Medical Sciences, Ansari Nagar, New Delhi 110 029, India

e-mail: sarman_singh@yahoo.com, sarman.singh@gmail.com