CHAPTER I

Cloning, expression and characterization of
*M. leprae* sHsp18 in *E. coli*

INTRODUCTION

*M. leprae* is an obligate, intracellular pathogen that has a long doubling time of approximately 11 days (Ridley, 1988). It cannot be cultured *in vitro* (Cole et al., 2001) and the infected nine-banded armadillo is the only source of appreciable amounts of live leprosy bacilli. Footpad infected normal and nude mice, also serve as sources of smaller amounts of viable bacilli (Marques et al., 2004). Under these circumstances, study of the biology of *M. leprae* and its pathogenicity had never been simple and easy.

During 1980s, researchers had adopted two major strategies for the identification and functional characterization of *M. leprae* antigens. One was to screen the *M. leprae* genomic libraries with antibodies and T-cells from patients and healthy exposed individuals (family contacts) (Young et al., 1985; Cherayil and Young, 1988; Mustafa et al., 1988; Sathish et al., 1990; Rinke de Wit et al., 1992). The other strategy was the subcellular fractionation of *M. leprae* to identify the abundant protein antigens. Genes of these major antigens were then traced back using oligonucleotide probes after partial amino acid determination. Use of the above two strategies enabled the identification of more than 40 antigens (Thole et al., 1995). The knowledge of protein antigens was further enriched by proteome studies of *M. leprae* and these studies contributed greatly towards the understanding of the *in vivo* expressed proteins in host derived *M. leprae*. In addition, it provided an insight into the minimal set of proteins essential for an exclusive intracellular survival (Marques et al., 1998; 2004).

Although the above strategies allow the identification and general characterization of many of the protein antigens, functional characterization of a gene/protein *per se* using *M. leprae* is still not possible. The non-availability of a synthetic medium for culturing, combined with an extremely slow growth rate precluded the routine genetic manipulations. Methodologies like gene knockout and complementation approaches
normally adopted for understanding the in vivo function of bacterial genes and gene products are not possible with M. leprae.

Nevertheless, a large number of M. leprae antigens have been studied and their role in immunity has been evaluated. Earlier studies were possible through cloning and expression of the gene of interest in the bacteriophage λgt11 (Young et al., 1985). However, extensive studies were carried out after the availability of E. coli as the system for cloning and overexpression.

The primary requirement for functional characterization of a protein, in vitro is its availability in a highly pure, soluble and native form. In general, natural protein sources rarely meet this requirement in terms of yield and the ease of isolation. In the current research scenario, protein production and characterization has been greatly facilitated by the development of systems for expression of proteins to high levels in homologous or heterologous host (Zhang et al., 2004). The gram-negative bacterium, E. coli is considered to be the workhorse for such recombinant protein expression. The simplicity in protein expression, fast high density cultivation, well-understood genetics, and economic feasibility makes E. coli an attractive and obvious choice for overexpression of heterologous proteins (Sorensen and Mortensen, 2005). Using this strategy, a large number of M. leprae genes have been cloned and overexpressed and a representative list is given in table 1.1.

Generally, an expression system for the production of recombinant proteins in E. coli involves a combination of a plasmid and a strain of E. coli. This is an ideal system for overproducing the protein of interest, which can further be purified by exploiting the presence of an N-terminal or a C-terminal tag encoded by the expression vector. A variety of such protein tags are now available and the most widely used tags are 6X-His tag (a contiguous stretch of 6 histidine residues) (Smith et al., 1988), MBP tag (maltose binding protein) (Bedouelle and Duplay, 1988), and GST tag (glutathione-S-transferase) (Smith and Johnson, 1988). These tags are mainly designed to be fused with the recombinant protein to enable the selective purification of the overexpressed protein, generally by affinity chromatography. The nature of tag in the protein determines the component in the affinity column used for purification. The 6X-His tagged protein is purified by metal affinity chromatography using Ni-NTA matrix to which the His-tag binds with high affinity. Similarly the MBP and GST tagged
<table>
<thead>
<tr>
<th>S. No</th>
<th>M. leprae protein</th>
<th>Swissprot ID</th>
<th>Objective of cloning M. leprae gene</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>65 kDa protein antigen, GroL2</td>
<td>P09239</td>
<td>For immunological studies</td>
<td>Nomaguchi et al., 1989; Lamb et al., 1988</td>
</tr>
<tr>
<td>2.</td>
<td>Superoxide dismutase, SOD</td>
<td>P13367</td>
<td>For characterization studies</td>
<td>Thangaraj et al., 1990</td>
</tr>
<tr>
<td>3.</td>
<td>18 kDa protein, sHsp18</td>
<td>P12809</td>
<td>For epitope studies For functional characterization</td>
<td>Hussain et al., 1992; Lini et al., 2008</td>
</tr>
<tr>
<td>4.</td>
<td>10 kDa Antigen, GroS</td>
<td>P24301</td>
<td>For characterization</td>
<td>Rivoire et al., 1994</td>
</tr>
<tr>
<td>5.</td>
<td>T5 Antigen</td>
<td>P95672</td>
<td>For molecular characterization and T-cell stimulatory study</td>
<td>Wieles et al., 1995a</td>
</tr>
<tr>
<td>6.</td>
<td>Bifunctional thioredoxin reductase/thioredoxin TrxB/A</td>
<td>P46843</td>
<td>For functional analysis</td>
<td>Wieles et al., 1995b</td>
</tr>
<tr>
<td>7.</td>
<td>Penicillin binding protein</td>
<td>Q9CBU6</td>
<td>For affinity studies with penicillin and cephalosporins</td>
<td>Basu et al., 1996</td>
</tr>
<tr>
<td>8.</td>
<td>Dihydropteroate synthase, FolP1 &amp; FolP2</td>
<td>P0C0X1</td>
<td>For studying dapsone resistance</td>
<td>Nopponpunth et al., 1999; Williams et al., 2000</td>
</tr>
<tr>
<td>9.</td>
<td>Holiday junction DNA helicase, RuvA</td>
<td>P40832</td>
<td>For studying the functional interactions with E.coli RuvB and RuvC</td>
<td>Arenas-Licea et al., 2000</td>
</tr>
<tr>
<td>10.</td>
<td>UDP-N-acetylmuramate:L-alanine ligase enzymes, MurC</td>
<td>P57994</td>
<td>For comparative study with M. tuberculosis MurC</td>
<td>Mahapatra et al., 2000</td>
</tr>
<tr>
<td>11.</td>
<td>Lipoprotein, LpK</td>
<td>O33133</td>
<td>To evaluate the role of LpK in protective immunity against M. leprae</td>
<td>Maeda et al., 2002</td>
</tr>
<tr>
<td>12.</td>
<td>Haemoglobin like oxygen carrier, GloBO (truncated)</td>
<td>Q9CC59</td>
<td>For studying its role in intracellular survival of M. leprae</td>
<td>Fabozzi et al., 2006</td>
</tr>
<tr>
<td>13.</td>
<td>DNA gyrase, GyrA DNA gyrase, GyrB</td>
<td>Q57532 Q59533</td>
<td>For developing an in vitro assay to study the activities of quinolones</td>
<td>Matrat et al., 2007</td>
</tr>
<tr>
<td>14.</td>
<td>Serine hydroxymethyl transferase, GlyA</td>
<td>Q9X794</td>
<td>For folding and stability studies</td>
<td>Sharma and Bhakuni, 2007</td>
</tr>
</tbody>
</table>

Table 1.1: Representative list of *M. leprae* genes cloned and overexpressed in *E. coli*. 
recombinant proteins can be purified using maltose and glutathione column, respectively.

In our lab, this technology was employed for cloning, overexpression and purification of one of the major T-cell antigen of *M. leprae*, sHsp18 (Mustafa *et al.*, 1986; Dockrell *et al.*, 1989; Doherty *et al.*, 1993). In this part of the study, overexpression and localization of sHsp18 in the heterologous host, *E. coli* are described. During the course of these experiments, sHsp18 was found to exist in multiple forms, which is also discussed in detail.

**RESULTS**

**Cloning of shsp18 gene of *M. leprae***

*shsp18* gene (ML1795) was cloned in an *E. coli* overexpression vector as described by Shabaana (Ph.D Thesis, 2002) and the strategy used for cloning this 447 bp gene (figure 1.1A) is shown in figure 1.1B. In brief, *M. leprae* was isolated from the biopsy of a leprosy patient. Using the mRNA from the bacilli, cDNA was prepared and was used as the template for amplifying the *shsp18* gene. Exploiting the A-tail in the PCR amplicon, this gene was cloned into pGEM-T plasmid, a T-tailed vector to obtain the recombinant plasmid, pSA6. Subsequently, the *shsp18* gene was released from pSA6 using *Sph*I and *Pst*I restriction enzymes and subcloned in *Sph*I and *Pst*I restricted pQE31 vector. The recombinant plasmid, pH18 was used for transforming *E. coli*, M15/pREP4 (Villarejo and Zabin, 1974) and ampicillin resistant transformants were selected. The pREP4 plasmid encoded lac repressor in the M15 strain prevents the expression of *shsp18* gene under non-inducible conditions. The recombinant *E. coli*, M15/pREP4/pH18 which overexpresses sHsp18 will hence forth be referred as M15/pH18. The *shsp18* gene was cloned such that it is expressed as a fusion protein with a 6X His tag in the N-terminus. Presence of this tag enables the protein to be purified by affinity chromatography.

**Overexpression of sHsp18 at different growth temperatures**

The level of expression of sHsp18 was determined at four time points after IPTG induction of *E. coli* cultures grown at 30 °C and 37 °C. The level of sHsp18 at these
Figure 1.1A: Sequence of *M. leprae* shsp18. The 148 residue protein sequence is also shown here. Arrows indicate the position of the primers used for amplifying the shsp18 gene.
Biopsy tissue from a leprosy patient

Isolate *M. leprae*

Isolate RNA

Reverse Transcriptase reaction

cDNA preparation

PCR amplification using Taq polymerase

**shsp18**

Ligation

pGEM-T vector

SphI & PstI double digestion

pSA6

pQE31 vector *(SphI & PstI digest)*

Transform to *E. coli*, M15/pREP4

Fig 1.1B: Schematic representation of the strategy used for cloning *shsp18* in an *E. coli* overexpression vector.
conditions are shown in figure 1.2A. Comparison of the lanes in this figure clearly indicates that the level of sHsp18 increases progressively with the increase in time of induction. This trend was observed irrespective of the growth temperature. Image analysis of the gel shown in figure 1.2A was performed using ImageQuant™ software as described in materials and methods. The scan trace (figure 1.2B) obtained from this analysis shows the sHsp18 levels in each lane of the gel which further confirms the above conclusion. From these results, it is apparent that the level of expression of sHsp18 in cells grown at 37 °C is higher than the levels achieved at 30 °C.

Quantification of sHsp18 expression levels

To quantify the levels of sHsp18 observed in the above experiment, the gel image was analyzed using the ImageQuant™ software as described in materials and methods. Results of this analysis are summarized in table 1.2 and figure 1.3A. In consensus with the observation made from figure 1.2, the quantification results shown in table 1.2 also indicate that the sHsp18 expression level increases with the increase in induction time. And, higher levels of sHsp18 were achieved when the cells were grown at 37 °C. At 37 °C, sHsp18 contributed to as much as 29% of the total E. coli protein while at 30 °C it was only 21%. As a control and for comparison, an E. coli protein of approximately 35 kDa (indicated with an * in figure 1.2A) was chosen arbitrarily and its concentration at different times of IPTG induction was determined as shown in figure 1.3B. It is apparent that the level of this E. coli protein was not altered with IPTG induction. Thus, this experiment clearly explains the dependence of sHsp18 expression on the growth temperature as well as the time of induction. Also, it gives a clear idea about the parameters that can be modified to obtain the desired level of sHsp18. Therefore, this experiment served as the basis for the ensuing experiments.

Based on the above quantification, parameters for the subsequent set of experiments were decided as discussed below

- **Purification of sHsp18**

  From the comparison of sHsp18 levels in lane 6 and 11 of figure 1.2 combined with the quantification data in table 1.2, it is clear that the highest level of sHsp18 can be achieved at the 4th hour of IPTG induction when E. coli is grown at 37 °C. Hence, this growth and induction conditions were employed
Figure 1.2: Expression levels of sHsp18 at 30 °C and at 37 °C after IPTG induction. A. Independent cultures of M15/pH18 cells were grown to $A_{600}=0.6$ at 30 °C or at 37 °C and induced with IPTG. At each time point, one OD equivalent of cells were pelleted, resuspended in 200 µl of SDS-lysis buffer and lysed by heating at 100 °C for 3 min. Equal volumes (20 µl) of the total cell lysate from all the samples were resolved on a 12% SDS-polyacrylamide gel and visualized after colloidal coomassie staining. Lanes correspond to the total cell lysate of *E. coli* grown at 30 °C (2-6) or at 37 °C (7-11). Lanes indicate 1, protein molecular weight marker; 2 & 7, 0 h of IPTG addition; 3 & 8, 1 h after IPTG addition; 4 & 9, 2 h after IPTG addition; 5 & 10, 3 h after IPTG addition; 6 & 11, 4 h after IPTG addition. B. Scan trace depicting the sHsp18 level in each lane of the gel shown in A as obtained from ImageQuant™ analysis.
Percentage of sHsp18 present in the *E. coli* cell with respect to the total protein

<table>
<thead>
<tr>
<th>Growth temperature</th>
<th>Duration of IPTG induction in hours</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>30 °C</td>
<td>11.76%</td>
</tr>
<tr>
<td>37 °C</td>
<td>18.49%</td>
</tr>
</tbody>
</table>

Table 1.2: Percentage of sHsp18 with respect to the total *E. coli* protein is given for cultures grown with IPTG induction for different time points at 30 °C and 37 °C. Quantification of sHsp18 level was done using the ImageQuant™ software as described in materials and methods. The software gives the intensity of a protein as band volume. Percentage of sHsp18 was calculated as (volume of sHsp18/total volume of all the proteins present in a lane) x 100.

**Figure 1.3:** Level of overexpressed sHsp18 in *E. coli*. The gel image of the total cell lysate of M15/pH18 at different times after IPTG addition was taken for analysis using ImageQuant™ as described in materials and methods. Known concentration of the molecular weight marker proteins (97 kDa, 435.5 ng; 66 kDa, 539.5 ng; 30 kDa, 539.5 ng) were used for calibration and concentration of the proteins in the gel were determined. Comparison of the protein concentration at different times of IPTG induction at 30 °C and 37 °C is shown for sHsp18 (A) and for an arbitrarily chosen *E. coli* protein (B).
to grow *E. coli* for sHsp18 purification under denaturing conditions. This resulted in the higher yield of pure sHsp18. Such high concentrations of protein were essential for some purposes such as raising polyclonal antibodies in rabbit.

- **Thermotolerance studies**

For evaluating the protective role of sHsp18 during thermal stress, sHsp18 expression was induced for 2 hours based on the method described by Muchowski and Clark (1998). From the ImageQuant™ data shown in table 1.2, it is apparent that in *E. coli* grown at 37 °C, the level of sHsp18 at the 2nd hour of IPTG induction is 75% of that obtained at the 4th hour. Thus, it is clear that under the conditions used for thermotolerance assay, a high level of sHsp18 can be achieved in lesser induction time and hence was used in subsequent heat shock experiments.

- **Purification of sHsp18 in native form**

For the *in vitro* characterization studies, it was essential to purify sHsp18 in a functionally active and native state. Growth temperature greatly influences the soluble state of the recombinant protein on overexpression and it has been shown that lower temperature favors greater solubility (Schein and Noteborn, 1988; Schein, 1989). The extent of sHsp18 that can be recovered in the soluble fraction of *E. coli* grown at 30 °C is shown in figure 1.4. Comparison of lanes 2 and 3 of figure 1.4 shows that ~ 50% of the overexpressed sHsp18 was present in the soluble fraction. In addition, from table 1.2 it is clear that 75% of the maximum sHsp18 level (at 37 °C) can be obtained at the 4th hour of IPTG induction when *E. coli* was grown at 30 °C. Therefore, for purifying sHsp18 in the native form, *E. coli* was grown at 30 °C and IPTG induction was given for 4 hours.

It is clear from figure 1.4 that a significant portion of sHsp18 is present in the 12K supernatant on overexpression. Nevertheless, a major portion still remains in the insoluble pellet fraction, as seen in lane 3 of figure 1.4. Three probabilities exist which can explain the appearance of sHsp18 in this pellet fraction. One possibility is that sHsp18 moves into the high speed pellet because it forms inclusion bodies. An equally plausible alternative explanation is that the overexpressed protein is
Figure 1.4: Analysis of the soluble form of sHsp18. M15/pH18 cells were grown to $A_{600}=0.6$ at 30 °C and induced with IPTG. Four hours after IPTG induction, the cells were pelleted, resuspended and lysed by sonication. After centrifugation at 12000 rpm (12K), the supernatant and pellet was recovered and 10 µg of the protein was separated on a 12% SDS-polyacrylamide gel. Lanes 1, 12K supernatant of uninduced cells; 2, 12K supernatant of sHsp18 induced cells; 3, 12K pellet of sHsp18 induced cells. Dotted arrow indicates the position of sHsp18.
associated with the membranes by virtue of which it is pelleted. Third possibility is the formation of loose aggregates that can be easily solubilized. The above mentioned possibilities were tested one by one to determine the nature of sHsp18 in the pellet fraction.

**Analysis of inclusion body formation**

A very well established and the most commonly encountered problem during overexpression of recombinant proteins in *E. coli* is the accumulation of the overexpressed proteins as insoluble aggregates known as “inclusion bodies” (Thomas and Baneyx, 1996; Sorensen and Mortensen, 2005). Inclusion bodies are refractile aggregates with porous structures that have high density. They are known to be in the non-native form and hence, biologically inactive (Betiku, 2006). To determine the state of sHsp18 in the 12K pellet, it was essential to determine whether sHsp18 forms inclusion bodies on overexpression. For this analysis, a construct from our lab which overexpresses the N-terminal truncated form of a “mammalian cell entry” protein (Mce*) was used as the control. This is also a *M. leprae* protein and its expression in *E. coli* has already been shown to be segregating into inclusion bodies (Sathosh Ph.D., Thesis, 2005).

Reischl (2003) showed that proteins which form inclusion bodies can be solubilized only at 6-8 M urea while other aggregates that do not form inclusion bodies can be easily solubilized at lower concentrations of urea. Thus, this method was followed to determine if sHsp18 forms true inclusion bodies or just aggregates which can be easily solubilized.

The experiment to analyze the formation of inclusion bodies was carried out as described in materials and methods. For both the proteins, sHsp18 and Mce*, the high speed centrifugation pellet fraction (12K pellet) was obtained from the sonicated cell lysate and subjected to sequential washes with increasing molar concentration of urea (2 M, 4 M, 6 M and 8 M). Protein solubilized in each step was then analyzed by SDS-PAGE and the results are shown in figure 1.5A and B.

In the case of Mce*, the overexpressed protein was present exclusively in the 12K pellet fraction (lane 2 of figure 1.5 B). When the proteins in this fraction was
Figure 1.5: Analysis of 12K pellet of recombinant cells. *E. coli*, M15/pREP4 overexpressing sHsp18 (A) or Mce* (B) was cultured as discussed in materials and methods. Cells were pelleted after four hours of IPTG induction, resuspended in 0.1 volume of sonication buffer and lysed by sonication. The soluble (lane 1) and insoluble proteins (lane 2) of the total cell lysate were separated by centrifugation at 12000 rpm. The 12K pellet fraction was resuspended in sonication buffer and the proteins were solubilized by sequential washes with 2 M urea (lane 3), 4 M urea (lane 4), 6 M urea (lane 5) and 8 M urea (lane 6). The dotted and solid arrow indicates the position of sHsp18 and Mce*, respectively. C. Comparison of the recovery of sHsp18 and Mce* from the 12K pellet during urea washes. Quantification was done using *ImageQuant™* analysis as described in materials and methods.
given sequential urea washes, it is apparent that only a small fraction of Mce* was solubilized at lower urea concentration. Even at 8M urea wash, Mce* still continues to be solubilized as seen in lane 6 of figure 1.5B. This is the typical solubilization behavior of a protein in inclusion bodies as described by Reischl (2003). Unlike Mce*, sHsp18 in the 12K pellet was solubilized efficiently at 2 M and 4 M urea (lane 3 and 4 of figure 1.5A). Only negligible levels of sHsp18 was seen in 8M urea wash indicating that the protein has been completely solubilized in the earlier urea washes itself. Comparison of the recovery of these two proteins with urea washes is shown in figure 1.5C which further reinforces the observations made from the gel. Thus, from the above results, it can be inferred that sHsp18 does not form inclusion bodies on overexpression in *E. coli*.

**sHsp18 and membrane association**

Since sHsp18 does not form inclusion bodies, the next step was to determine whether the sHsp18 in the 12K pellet is bound to membranes or present as loose aggregates. To assess this, the insoluble pellet fraction was washed with a solubilization buffer containing EGTA, a chelating agent and Tween-20, a detergent. Many membranes contain calcium activated neutral proteases which can be inhibited by the chelation of free calcium. Hence, addition of EGTA has been employed to scavenge both heavy metals and free calcium in an effective fashion (Hjelmeland, 1990). Non-ionic detergents have often been applied to membrane solubilization and proteins can be selectively solubilized with Tween-20 (Johansson *et al.*, 1975). Hence, if sHsp18 had moved into the 12K pellet by virtue of its association with membranes, it can be solubilized with a detergent.

Proteins solubilized from the 12K pellet using EGTA and Tween-20 containing solubilization buffer is shown in lane 3 of figure 1.6. It is apparent that sHsp18 can be solubilized from the pellet fraction with Tween-20. The ability to be solubilized by a detergent indicates that sHsp18 is indeed associated with membranes and it is by virtue of this association that a proportion of the overexpressed protein gets pelleted.

In a parallel experiment carried out in our lab, majority of the overexpressed sHsp18 was shown to be localized to the outer membrane and to a lesser extent to the
Figure 1.6: EGTA and Tween-20 solubilizes sHsp18 in the 12K pellet. M15/pH18 was grown and induced with IPTG for four hours as described earlier. The 12K pellet was prepared as described in materials and methods. This pellet was resuspended in the solubilization buffer containing 0.1 mM EGTA and 0.25% Tween-20. After vortexing, the proteins were incubated at room temperature for 10 min after which the solubilized proteins were recovered by centrifugation. Fifteen micrograms of protein were separated on a 12% SDS-polyacrylamide gel and visualized after colloidal coomassie staining. Lanes 1, Molecular weight marker; 2, 12K pellet fraction; 3, EGTA and Tween-20 solubilized proteins; 4, proteins retained in the 12K pellet after solubilization. Dotted arrow indicates the position of sHsp18.
perioplasm of *E. coli* (Lini et al., 2008). Hence, the above observation on the membrane association of sHsp18 is indicative of its membrane localization.

**Multiple forms of sHsp18**

When the soluble and insoluble fractions of *E. coli* overexpressing sHsp18 were analyzed by SDS-PAGE, two bands corresponding to molecular mass of 22 kDa and 18 kDa were observed consistently. These two proteins are shown in figure 1.7. Based on the earlier experiments (figure 1.2), it is known that the 22 kDa band corresponds to the His-tagged sHsp18. However, the identity of the lower 18 kDa protein was not known.

The lower 18 kDa protein was observed irrespective of the duration of IPTG induction, both in the soluble and insoluble fractions. Hence, this protein along with the 22 kDa protein was excised out for identification by MALDI-TOF mass spectrometry. As expected, the 22 kDa band was identified as *M. leprae* sHsp18. Surprisingly, the 18 kDa protein was also identified as *M. leprae* sHsp18. These two forms of sHsp18 appear to differ in molecular mass by approximately 4 kDa in a SDS-polyacrylamide gel. Therefore, to determine the actual molecular mass difference between these two forms, purified sHsp18 was analyzed by MALDI-TOF mass spectrometry operated in the linear mode. The average molecular mass of the two forms of sHsp18 determined from this analysis is shown in figure 1.8.

The average molecular mass of the two forms of sHsp18 was determined to be 19330.95 Da and 16716.71 Da. The 19330.95 Da matches with the calculated mass of the recombinant sHsp18 (with His-tag) based on the amino acid composition and correspond to the protein which migrates in the SDS-PAGE at approximately 22 kDa. The lower form of sHsp18 which migrates as the 18 kDa band has a molecular mass of 16716.71 Da. From figure 1.8, it can be observed that the height of the 16.7 kDa form of sHsp18 could be matched with the peak height of the 19.3 kDa form only after a 50-fold magnification (indicated in the figure as X50). This indicates that under the conditions in which sHsp18 was purified for this analysis, the 19.3 kDa sHsp18 protein is the predominant form. The results from this analysis suggest that the lower form of sHsp18 (18 kDa form) could be the truncated form of the recombinant His-tag sHsp18.
Figure 1.7: Resolution of recombinant sHsp18. M15/pH18 was grown and sHsp18 was induced as described in materials and methods. Cells were lysed by sonication and the total cell lysate was centrifuged at 12000 rpm to recover the soluble and insoluble protein fractions. Ten microgram of proteins was resolved on a 12% SDS-polyacrylamide gel. Lanes 1, molecular weight marker; 2, 12K supernatant; 3, 12K pellet. Dotted arrow indicates the position of the 22 kDa and 18 kDa proteins.
Figure I.8: Analysis of the molecular mass of the two forms of sHsp18 by MALDI-TOF mass spectrometry. Hundred picomoles of purified sHsp18 was mixed with sinnapinic acid matrix (10 mg/ml) and spotted. The protein was analyzed by MALDI-TOF mass spectrometer operated in the linear mode as described in materials and methods. X50 indicates that the peak height of the 18 kDa form of sHsp18 was magnified 50 times (X50) to match the peak height of the 22 kDa form.
Sample preparation method and its effect on sHsp18 isoforms

Formation of the lower form of sHsp18 could be due to the activation of proteolytic activity, perhaps an autolytic activity. It was speculated that during sonication the recombinant sHsp18 is probably subjected to the action of a protease that cleaves this protein. To check this possibility, the cells after overexpression were lysed by one of the following methods:

1. boiling using SDS lysis buffer
2. sonication in the absence of protease inhibitor cocktail
3. sonication in the presence of protease inhibitor cocktail
4. sonication in the presence of orthovanadate, a phosphatase inhibitor.

The total cell lysate (after lysis by method 1) or the 12K supernatant and 12K pellet (after lysis by methods 2, 3 or 4) was analyzed by SDS-PAGE and the protein profile obtained is shown in figure 1.9. Comparing lane 2 with the rest of the lanes in figure 1.9, it is evident that the lower form of sHsp18 appear only when the cells were lysed by sonication. In lane 2, which corresponds to the total cell lysate obtained by boiling the cells, only the high molecular mass form of sHsp18 could be seen. This confirms that the lower form of sHsp18 does not result from premature termination of protein translation and comes into existence due to protein processing during cell lysis. Moreover, it is clear from figure 1.9 that the formation of the lower form of sHsp18 remained unaffected in the presence of protease inhibitor cocktail (lanes 5 & 6) as well as in the presence of orthovanadate (lanes 7 & 8). As the protease inhibitor cocktail includes inhibitors for the asp, cys and ser proteases, role of these three types of proteases in the truncation of sHsp18 can be ruled out. Moreover orthovanadate, a known inhibitor of phosphatase also did not prevent the formation of the lower form of sHsp18 indicating that phosphatase too does not have a role in sHsp18 truncation.

In this study, for purification of sHsp18 under non-denaturing conditions, the cells were usually lysed by sonication and the soluble fraction was taken. Since the cells were lysed by sonication, the two forms of sHsp18 were usually seen in the 100-150 mM imidazole eluates (lane 1 of figure 1.10A, B and C). But, sHsp18 in the later eluates (200-500 mM imidazole) was seen as a single band which corresponds to the His-tag sHsp18 (19.3 kDa upper form). However, if the sonication time was increased during cell lysis, the lower form of sHsp18 could be seen in the later
Figure 1.9: Protein profile of *E. coli* total cell lysate prepared using different methods. M15/pH18 cells overexpressing sHsp18 were lysed by different methods as described earlier and 10 µg of protein from all the samples were resolved on a 12% SDS-polyacrylamide gel. Lanes indicate molecular weight marker (lane 1); total cell lysate prepared by boiling with SDS lysis buffer (lane 2); 12K supernatant and 12K pellet from cells lysed by sonication in the absence of protease inhibitor cocktail (lane 3 & 4); in the presence of protease inhibitor cocktail (lanes 5 & 6); in the presence of orthovanadate (lanes 7 & 8).
Figure 1.10: Sonication results in more than two forms of sHsp18. M15/pH18 cells were grown and sHsp18 was induced as described earlier. Cells were lysed by sonication and the number of cycles (a pulse of 5 sec ON; 5 sec OFF for 1 min) given was varied – 5 cycles (A), 10 cycles (B) and 15 cycles (C). Soluble proteins obtained from the total cell lysate were taken for sHsp18 purification under non-denaturing conditions as discussed in materials and methods. Lanes 1, 150 mM imidazole eluate; 2, 500 mM imidazole eluate. Solid and dotted arrow indicates the position of 19.3 kDa and 16.7 kDa forms of sHsp18, respectively. Dashed arrow indicates the position of an extra band seen during sHsp18 purification. Concentration of imidazole used to elute sHsp18 is indicated in the figure above the lane number.
eluates, as well. From figure 1.10B and C, it is apparent that increase in sonication
time not only increases the level of the lower 18 kDa form, but also gives rise to yet
another form, indicated in the figure with a dashed arrow. Hence, from the above
result, it is evident that sonication indeed results in the lower form(s) of sHsp18.
However, the exact cause for the formation of this lower form is not known at this
point.

**Multiple isoforms of sHsp18**

In tune with the above experiments, the purified sHsp18 was analyzed by two
dimensional gel electrophoresis, also. sHsp18 is usually observed as a single band if
purified under denaturing conditions whereas, it exists as two (or three) bands when
purified under native conditions. Hence, sHsp18 purified under both denaturing and
native conditions were resolved by 2DE as described in materials and methods.
Resolution of the purified sHsp18 in a 2D gel is shown in figure 1.11.

From figure 1.11A, it can be observed that sHsp18 purified under denaturing
conditions that appears as a single band in the 1D gel resolves to four protein spots
(D1-D4) in a 2D gel. Similarly, the three bands of the native purified sHsp18 resolve
to as many as 9 spots (N1-N9) in a 2D gel (figure 1.11B). Thus, it is evident that
sHsp18 exists as multiple isoforms that differ not only in their molecular mass but
also in pI. The different isoforms of sHsp18 shown in figure 1.11B, labeled as N1 to
N9, were excised out and processed for identification by MALDI-TOF mass
spectrometry. The results of the database search are tabulated in table 1.3. Six out of
the 9 proteins were identified to be sHsp18 of *M. leprae* further confirming the
existence of sHsp18 in multiple isoforms.

**DISCUSSION**

sHsp18 of *M. leprae* is a major T-cell antigen (Dockrell *et al.*, 1989) and this protein
belongs to the family of small heat shock proteins. When all other mycobacterial
species have multiple sHsps (table 1.4), *M. leprae* alone has only one. *shsp18* is the
only small heat shock protein gene retained by *M. leprae* during the course of massive
gene decay (Cole *et al.*, 2001). This gene has been shown to be activated specifically
during the intracellular growth and hence, suggested to be involved in the survival of
*M. leprae* in macrophages (Dellagostin *et al.*, 1995). These data indicate that sHsp18
Figure 1.11: Resolution of purified sHsp18 by 1DE and 2DE. sHsp18 purified by two different methods were separated by one dimensional and two dimensional electrophoresis and stained with colloidal coomassie blue. A. sHsp18 purified by denaturing method. B. sHsp18 purified by non-denaturing method. The spot IDs are prefixed with D and N indicating the denaturing and non-denaturing method by which sHsp18 was purified.
<table>
<thead>
<tr>
<th>S. No</th>
<th>Spot ID</th>
<th>Experimental pI</th>
<th>Database searched</th>
<th>MASCOT score</th>
<th>Variable modifications</th>
<th>Sequence coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>N1</td>
<td>4.54</td>
<td>NCBI nr 20100924</td>
<td>83</td>
<td>Gln to pyro-Glu (N-term Q)</td>
<td>58%</td>
</tr>
<tr>
<td>2.</td>
<td>N2</td>
<td>4.70</td>
<td>NCBI nr 20100924</td>
<td>86</td>
<td>Gln to pyro-Glu (N-term Q)</td>
<td>58%</td>
</tr>
<tr>
<td>3.</td>
<td>N3*</td>
<td>4.96</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4.</td>
<td>N4</td>
<td>5.20</td>
<td>Swissprot 2010_09</td>
<td>57</td>
<td>Gln to pyro-Glu (N-term Q)</td>
<td>39%</td>
</tr>
<tr>
<td>5.</td>
<td>N5*</td>
<td>6.36</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>6.</td>
<td>N6*</td>
<td>4.89</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>7.</td>
<td>N7</td>
<td>5.11</td>
<td>NCBI nr 20100924</td>
<td>89</td>
<td>None</td>
<td>76%</td>
</tr>
<tr>
<td>8.</td>
<td>N8</td>
<td>6.46</td>
<td>NCBI nr 20100924</td>
<td>104</td>
<td>Gln to pyro-Glu (N-term Q) Glu to pyro-Glu (N-term E)</td>
<td>77%</td>
</tr>
<tr>
<td>9.</td>
<td>N9</td>
<td>6.65</td>
<td>NCBI nr 20100924</td>
<td>89</td>
<td>Gln to pyro-Glu (N-term Q)</td>
<td>70%</td>
</tr>
</tbody>
</table>

*unidentified

Table 1.3: Details of the MALDI-TOF identification of sHsp18 isoforms.
is essential for the survival and/or virulence of this pathogen. Although a great deal of information is available on the immunological properties, the functional aspects of sHsp18 are yet to be deciphered. Hence, the objective of this study was to characterize this protein and as a first step the shsp18 gene was cloned and overexpressed in the E. coli expression system.

Expression of the shsp18 gene in the recombinant plasmid is under the control of a repressor protein (encoded by the pREP4 plasmid) in the M15 strain. Hence, there would be no leaky expression and the production of sHsp18 can be induced only after the addition of IPTG. In an earlier study from our lab, it was shown that the optimal concentration of IPTG needed for a high level expression of sHsp18 was 0.4 mM (Shabaana Ph.D., Thesis, 2002). However, as with the overexpression of any recombinant protein, the level of sHsp18 on IPTG induction was determined in the first place. There are two factors which play a pivotal role in determining the levels of the overexpressed protein. One is the growth temperature and the other is the duration of IPTG induction. E. coli was grown at 30 °C and 37 °C to determine the sHsp18 levels at four time points after IPTG addition. From the results it was evident that higher levels of sHsp18 were achieved in cells grown at 37 °C than at 30 °C. Further, highest level of the protein was obtained by the fourth hour of induction, irrespective of the growth temperature. After 4 hours of induction at 37 °C, sHsp18 contributes to as much as 29% of the total protein in E. coli.

Many of the recombinant proteins overexpressed in E. coli tends to form inclusion bodies. In our lab, five proteins have been cloned and overexpressed using the same E. coli overexpression system. Of these five proteins, two are from M. leprae (sHsp18 and Mce*) and the other three are from Streptomyces peucetius (ChiS, ChiR and ChiC). Both, M. leprae and S. peucetius, are GC rich organisms with a high codon bias. Genes encoding these proteins have been cloned utilizing the pQE vector system and induced with IPTG. Interestingly, except sHsp18, all the other four proteins on overexpression formed inclusion bodies. The expression levels of all the four proteins were comparable to that of sHsp18. However, only low levels of these proteins were detected in the soluble fraction and the major fraction of the protein moved into inclusion bodies (Santhosh Ph.D., Thesis, 2005; Singh and Dharmalingam, unpublished data). On the contrary, ~50% of sHsp18 protein was
<table>
<thead>
<tr>
<th>Organism</th>
<th>Genome Size (Mb)</th>
<th>Total Proteins</th>
<th>No of sHsps</th>
<th>Protein Identifier</th>
<th>Locus ID</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>M. leprae</em> TN</td>
<td>3.3</td>
<td>1605</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>M. abscessus</em> ATCC 19977</td>
<td>5.12</td>
<td>4920</td>
<td>2</td>
<td>Hsp20</td>
<td>MAB_4402</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Hsp</td>
<td>MAB_3467c</td>
</tr>
<tr>
<td><em>M. bovis</em> BCG str. Pasteur 1173P2</td>
<td>4.4</td>
<td>3949</td>
<td>2</td>
<td>Hsp</td>
<td>BCG_0289c</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>HspX</td>
<td>BCG_2050c</td>
</tr>
<tr>
<td><em>M. gilvum</em> PYR-GCK</td>
<td>5.96</td>
<td>5241</td>
<td>2</td>
<td>Hsp20</td>
<td>MMAR_0515</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Hsp20</td>
<td>MMAR_3484</td>
</tr>
<tr>
<td><em>M. marinum</em> M</td>
<td>6.62</td>
<td>5423</td>
<td>2</td>
<td>Hsp</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>HspX_1</td>
<td></td>
</tr>
<tr>
<td><em>M. tuberculosis</em>, H37Rv</td>
<td>4.4</td>
<td>3988</td>
<td>2</td>
<td>Hsp</td>
<td>Rv0251c</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>HspX</td>
<td>Rv2031c</td>
</tr>
<tr>
<td><em>M. ulcerans</em> Agy99</td>
<td>5.8</td>
<td>4160</td>
<td>2</td>
<td>Hsp</td>
<td>MUL_1178</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>sHsp</td>
<td>MUL_2232</td>
</tr>
<tr>
<td><em>M. smegmatis</em> str. mc²155</td>
<td>7.0</td>
<td>6716</td>
<td>3</td>
<td>Hsp20</td>
<td>MSMEG_0424</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Spore Protein</td>
<td>MSMEG_5611</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>14 kDa Ag</td>
<td>MSMEG_3932</td>
</tr>
<tr>
<td><em>Mycobacterium</em> sp. JLS</td>
<td>6.0</td>
<td>5739</td>
<td>3</td>
<td>Hsp20</td>
<td>Mjls_0247</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Hsp20</td>
<td>Mjls_1109</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Hsp20</td>
<td>Mjls_4761</td>
</tr>
<tr>
<td><em>M. avium</em> subsp. Paratuberculosis K-10</td>
<td>4.8</td>
<td>4350</td>
<td>4</td>
<td>Hsp18_1</td>
<td>MAP1696c</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Hsp18_2</td>
<td>MAP1698c</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Hsp18_3</td>
<td>MAP3268</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Hsp</td>
<td>MAP3701c</td>
</tr>
<tr>
<td><em>M. vanbaalenii</em> PYR-1</td>
<td>6.5</td>
<td>5979</td>
<td>4</td>
<td>Hsp20</td>
<td>Mvan_0296</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Hsp20</td>
<td>Mvan_1374</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Hsp20</td>
<td>Mvan_3941</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Hsp20</td>
<td>Mvan_4943</td>
</tr>
<tr>
<td><em>M. avium</em> 104</td>
<td>5.5</td>
<td>5120</td>
<td>6</td>
<td>Hsp20</td>
<td>MAV_271</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Hsp20</td>
<td>MAV_4106</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Hsp20</td>
<td>MAV_4906</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>18kDa Ag 2</td>
<td>MAV_2721</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>18kDa Ag 2</td>
<td>MAV_1990</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>18kDa Ag 2</td>
<td>MAV_1988</td>
</tr>
</tbody>
</table>

Table 1.4: Number of sHsp genes present in different mycobacteria. This table includes all the mycobacteria whose genome was completely sequenced as on August, 2010 and listed in [http://www.ncbi.nlm.nih.gov/genomes/lproks.cgi](http://www.ncbi.nlm.nih.gov/genomes/lproks.cgi).
found in the soluble fraction and analysis of sHsp18 in the insoluble pellet clearly indicated that the protein was not in inclusion bodies.

Thus, a greater fraction of overexpressed sHsp18 remained in the soluble form. Increased solubility of this protein could be either because this protein has a natural tendency to go into solution or because it is a chaperone. sHsp18 has been demonstrated to have an inherent chaperone activity (Lini et al., 2008). And, function of molecular chaperones are to direct other polypeptides to fold properly but they themselves are not components of the final functional structure (Thomas and Baneyx, 1996; Bhutani and Udgaonkar, 2009). In a highly crowded milieu of the cell cytoplasm, proper folding of the overproduced proteins is a seemingly difficult operation and molecular chaperones are known to interact with folding intermediates to facilitate their proper folding. (Thomas and Baneyx, 1996) Moreover, it has been suggested that the co-expression of molecular chaperones prevent or decrease protein aggregation. Thomas and Baneyx (1996) have shown that the co-expression of DnaK, DnaJ, GroEL and GroES enhances the enzymatic recovery of overexpressed β-galactosidase. Many such evidences are available for the direct correlation between the levels of molecular chaperones and the increased solubility of the overexpressed protein in E. coli (Betiku, 2006; Chen et al., 2002). Based on this, it is very tempting to speculate that the chaperone nature of sHsp18 might be responsible for keeping more of its own protein species in the soluble form.

A fraction of the overexpressed sHsp18 was found to associate with membranes, indicating its translocation from the cytosol into membrane compartments. Subcellular localization of sHsp18 has been reported earlier in M. leprae. Marques et al. (1998) have initially shown the presence of sHsp18 in more than one subcellular fraction in the armadillo-derived M. leprae. They have shown clearly that sHsp18 localizes to both the membrane and cell wall fraction. Later, this protein was identified in the membrane and cytosolic protein fraction of M. leprae by proteomics approach (Marques et al., 2004). In addition, studies from our lab have also shown the localization of the overexpressed sHsp18 to the outer membrane and periplasm of E. coli (Lini et al., 2008).

The sHsp18 sequence was taken for prediction of localization/signal sequence using the servers PSORT, SignalP and SecretomeP. This analysis rules out the presence of
any characterized signal peptide in the sHsp18 sequence. Thus, in the absence of a
signal sequence and transmembrane domains, the exact mechanism of translocation of
sHsp18 to membrane compartments is not known. Many of the bacterial pathogens
are known to possess specialized export systems, which transport/export proteins
lacking the classical signal sequence. It has been shown in *M. tuberculosis* that
SecA2 is one such specialized export pathways for the export of a specific subset of
proteins such as SodA and KatG, both of which lack the classical signal sequence.
This work also reports a decreased secretion of HspX, an α-crystallin type of sHsp in
ΔSecA2 mutant (Braunstein *et al.*, 2003). Another protein, glutamine synthetase
which lacks a leader peptide, was found to be secreted abundantly in pathogenic
mycobacteria, *M. tuberculosis* and *M. bovis*. And, the recombinant enzyme was also
found to be secreted in *M. smegmatis* (Harth and Horwitz, 1997). Thus, these studies
support the utilization of secretion systems by bacterial pathogens to export virulence
factors in a non-classical pathway. The ability of *M. lepraes* Hsp18 to be localized to
the outer membrane of *E. coli* provides a clue to its ability to utilize such specialized
export systems and/or the existence of such systems in *M. lepraes*.

The association of sHsp18 to *E. coli* membranes was demonstrated in this study.
Zhang *et al.*, (2005) have shown that Hsp16.3 of *M. tuberculosis* interacts with the
MTB plasma membrane, reversibly. Similar membrane associations have been
demonstrated for α-crystallin and *Synechocystis* HSP17, and this association has been
shown to maintain the membrane integrity under thermal stress (Tsvetkova *et al*.,
2002). HSPB2, expressed in the heart and skeletal muscles, have also been shown to
associate with the outer membrane components of mitochondria and has a role in
stress response (Nakagawa *et al*., 2001). For a pathogen, the plasma membrane
would be the first cellular structure to be exposed to harmful conditions when it
encounters a stress and the sHsps are probably involved in counteracting such stress
(Zhang *et al*., 2005). Therefore, this kind of membrane association, particularly for a
protein from an intracellular pathogen such as *M. lepraes*, might have relevance to the
pathogens virulence mechanisms. And, the appearance of antibodies against sHsp18
in humans also implies that this might be a surface exposed antigen, accessible to the
human immune system.

Purified sHsp18 was observed as two bands in SDS-polyacrylamide gel and both the
forms were confirmed to be *M. lepraes* sHsp18 by mass spectrometry. The average
molecular mass of the two forms as determined from mass spectrometry analysis clearly indicates that the lower form is indeed a truncated version of the recombinant sHsp18. Since the shsp18 gene is of prokaryotic origin, splicing can be ruled out as the reason for the generation of the lower form. Therefore, the size difference might arise as a result of either N-terminal or C-terminal truncation.

In this study, sHsp18 was expressed with a His tag at the N-terminus of the protein. An N-terminal truncation leading to the loss of ~2614 Da would obviously mean the loss of this N-terminal His tag. However, experiments performed with the anti-His antibodies indicated that this N-terminal His tag region is indeed intact (Lini Ph.D., Thesis, 2008). Therefore, it is speculated that the lower form of sHsp18 results from the C-terminal truncation of the 19.3 kDa form. And, this truncation occurs only when the cells are lysed by sonication. It was also observed that increasing the sonication time during cell lysis not only increases the level of the lower form but also generates an additional form. Since the use of protease and phosphatase inhibitors did not prevent the formation of the lower form, it is evident that neither protease nor phosphatase has any role in sHsp18 truncation. Although, truncation is the probable reason for the lower forms of sHsp18, the exact cause and point of cleavage is not known at this point.

Interestingly, when purified sHsp18 was resolved in a 2D gel, multiple isoforms differing in pI was observed indicating the presence of post-translational modifications, as well. Earlier, Marques et al., (2004) reported the presence of multiple spots for sHsp18 when the different subcellular fractions of armadillo-derived M. leprae were resolved on a 2D gel. Seven spots in the membrane fraction and four spots in the soluble/cytosolic fraction were identified as sHsp18 by this group. Small heat shock proteins are known to undergo various post-translational modifications (PTMs). Phosphorylation of specific serine residues is a feature which has been shown for shHsps such as α-crystallins (Kantorow and Piatigorsky, 1994), and human hsp27 (Landry et al., 1992) and mouse hsp25 (Gaestel et al., 1991). Also, α-crystallins have been shown to be modified in vivo with O-linked N-acetyl-glucosamine (Roquemore et al., 1992) and post-translationally by transglutaminase (Groenen et al., 1992). Although, phosphorylation is a possible modification for a protein expressed in a prokaryotic system, the probability of the other modifications needs to be analyzed carefully.
At this point, it is apparent that sHsp18 protein is subjected to post-translational modification(s) which probably results in multiple isoforms that differ in pI. Post-translational modification (PTM) is a common biological mechanism for regulating protein functions and hence, has a critical role in modulating cellular functions. Examples of the varied biological effects of some of the protein modifications include phosphorylation for signal transduction, ubiquitination for proteolysis, attachment of fatty acids for membrane anchoring and association, glycosylation for protein half-life, targeting, cell:cell and cell:matrix interactions (Lee et al., 2009). Considering the functional versatility provided by the PTMs, identification of the nature of modifications in sHsp18 becomes important particularly from the mycobacterial perspective. This analysis will further throw light on the different functional capabilities of this important protein antigen which might be essential for the *M. leprae* life cycle. At this point, phosphorylation is definitely one of the PTM involved in sHsp18 and the possibilities of other PTMs need to be explored further. Phosphorylation of sHsp18 and its significance with respect to host-pathogen interaction will be discussed in more detail in the last part of this study.

**CONCLUSION**

In this part, the overexpression characteristics of sHsp18 in *E. coli* were analyzed. Studies indicated that highest level of sHsp18 could be achieved by four hours of IPTG induction at 37 °C. And, even at a high level of expression, the protein did not form inclusion bodies. The overexpressed sHsp18 was demonstrated to be associated with membranes despite the absence of signal sequence. Overexpressed sHsp18 was found to exist in 9 forms, which differed in both molecular mass and pI. While C-terminal truncation is speculated for the difference in molecular mass of sHsp18 forms, post-translational modifications, particularly phosphorylation is one modification responsible for isoforms differing in pI. The membrane association of sHsp18 combined with its existence as multiple isoforms might have significance with respect to its function in *M. leprae*. 