**Mycobacterium tuberculosis HspX/EsxS Fusion Protein: Gene Cloning, Protein Expression, and Purification in Escherichia coli**

Farzad Khademi#1, Arshid Yousefi-Avarvand#1, Mohammad Derakhshan1, Zahra Meshkat1, Mohsen Tafaghodi2, Kiarash Ghazvini1, Ehsan Aryan1, Mojtaba Sankian*3

Abstract

**Background:** The purpose of this study was to clone, express, and purify a novel multidomain fusion protein of *Mycobacterium tuberculosis* (*Mtb*) in a prokaryotic system.

**Methods:** An *hspX/esxS* gene construct was synthesized and ligated into a pGH plasmid, *E. coli* TOP10 cells were transformed, and the vector was purified. The vector containing the construct and pET-21b (+) plasmid were digested with the same enzymes and the construct was ligated into pET-21b (+). The accuracy of cloning was confirmed by colony PCR and sequencing. *E. coli* BL21 cells were transformed with the pET-21b (+)/hspX/esxS expression vector and protein expression was evaluated. Finally, the expressed fusion protein was purified on a Ni-IDA column and verified by SDS-PAGE and western blotting.

**Results:** The *hspX/esxS* gene construct was inserted into pET-21b (+) and recombinant protein expression was induced with IPTG in *E. coli* BL21 cells. Various concentrations of IPTG were tested to determine the optimum concentration for expression induction. The recombinant protein was expressed in insoluble inclusion bodies. Three molar guanidine HCl was used to solubilize the insoluble protein.

**Conclusions:** An HspX/EsxS *Mtb* fusion protein was expressed in *E. coli* and the recombinant protein was purified. After immunological analysis, the HspX/EsxS fusion protein might be an anti-tuberculosis vaccine candidate in future clinical trial studies.

**Keywords:** *Mycobacterium tuberculosis*, HspX/ EsxS fusion protein, Gene cloning, Protein expression, Protein purification

Introduction

Successful vaccination against intracellular organisms, such as the bacterium *Mycobacterium tuberculosis* (*Mtb*), has proven difficult (1). Now, due to the limited effectiveness of the current vaccine (*Mycobacterium bovis*, Bacillus Calmette-Guerin, or BCG vaccine) against tuberculosis (TB) infections in different regions of the world, immunodominant antigens of *Mtb*, as prophylactic protein subunit vaccines, can act as alternative vaccines or as boosters of the BCG vaccine (2, 3). Of the new TB vaccines in preclinical or clinical trials, the subunit vaccines, such as Ag85B-ESAT-6, Ag85B-TB10.4, and Mtb72F, are important (4-6). These antigens are expressed early by replicating *Mtb*; however, proteins...
expressed by non-replicating bacteria in the latent phase must also be considered. Fusions containing proteins expressed in both replicating and latent phases are referred to as multistage vaccines (7, 8). In the present study, a plasmid containing \textit{esxS}, expressed in the early replication stage, and \textit{hspX}, expressed in the latent stage, was constructed. The HspX protein (Synonym: acr, Rv2031c) (alpha-crystalline homolog), the 16-kDa heat shock protein of \textit{Mtb}, is a latent-phase protein necessary for survival within the macrophages, produced during static growth or low-oxygen conditions (anoxia) (8). In latent phase \textit{Mtb} infections, HspX promotes both cell-mediated and humoral immune responses. Therefore, HspX is an important immune system target (9). The 23 members of the Esx protein family, including EsxS, are immunodominant antigens that strongly induce cell-mediated immunity (10). In recent years, it was believed that the Esx antigens are expressed early and secreted, but several reports demonstrated that they may be expressed at different stages of TB infection (11-14). The EsxS protein (Synonym: PE28, Rv3020c) is the 10-kDa ESAT-6-like protein or PE-family-related protein with unknown function which located in the cell wall and cell processes functional category (http://tuberculist.epfl.ch). The aim of this study was to design a fusion protein composed of proteins secreted by both replicating and non-replicating \textit{Mtb} in early and late infection phases, respectively, to be used as an anti-TB vaccine.

Materials and Methods

Design of the \textit{HspX/EsxS} plasmid

The nucleotide sequences of the \textit{hspX} and \textit{esxS} genes of 435 and 294 base pairs (bp), respectively, were obtained from the NCBI website. \textit{HindIII} and \textit{XhoI} restriction sites were added at the 5’ and 3´ ends, respectively, using Gene Runner software version 5. After software translation of the \textit{hspX} and \textit{esxS} DNA sequences, the two proteins were fused using a 2x GGGGS flexible linker. Finally, using online software of the Codon Adaptation Tool (JCAT), the codon adaptation was used to obtain maximum expression in the prokaryotic host. The desired fragment was outsourced for synthesis in the pGH cloning vector (Generay Biotech, China).

Bioinformatics analysis of fusion protein

The tertiary structure of the HspX/EsxS fusion protein was predicted by uploading several models to the I-TASSER server. Uploaded models included HspX and EsxS of 144 and 97 amino acids, respectively, connected via 15-amino-acid linkers. Five models were designed and selected the best model (Fig. 1). In each model HspX and EsxS are shown in green and blue, respectively. The fifteen amino acid linker is shown in yellow.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure1.png}
\caption{A) Schematic view of the \textit{hspX/esxS} construct. \textit{HindIII} and \textit{XhoI} restriction sites were added at the 5’ and 3´ ends, respectively, the \textit{hspX} and \textit{esxS} genes are 435 and 294 bps in length, respectively, and the 15-amino acid flexible linker was 2x GGGGS. B) HspX/EsxS fusion protein profile. Amino acid composition; length, molecular weight, and tertiary structure are shown. The fusion protein tertiary structure was predicted by I-TASSER; the blue band (LYS1-THR145) is HspX, the green band (GLY159-GLU260) is EsxS, and the yellow band (THR145-GLY159) is the linker.}
\end{figure}
**Cloning and Expression of *Mycobacterium tuberculosis***

**Subcloning of hspX/esxS**
Lyophilized pGH (2907 bp), containing the hspX/esxS gene construct (780 bp) was diluted in 40 µL of UltraPure™ DNase and RNase-free distilled water (Invitrogen, USA) and stored at -20 °C. *E. coli* Top10 cells were grown in 5 mL of Luria-Bertani (LB) broth medium (HIMEDIA, India) until mid-log phase (an optical density (O.D.) of ~0.4-0.6 at 600 nm). Then, MgCl2-CaCl2 (0.1M) and CaCl2 (0.1M) were respectively used to make the cells competent as described previously (15). For transformation, the pGH/hspX/esxS vector was added to the competent *E. coli* cells and cultured on LB solid medium supplemented with 100 µg/mL of ampicillin overnight at 37 °C. To identify colonies containing pGH/hspX/esxS, colony PCR was performed using the T7 universal primer. After the colony confirmation, pGH/hspX/esxS and pET-21b (+) (5442 bp) vectors were extracted using a Prime Prep Plasmid DNA Isolation Kit (GeNet Bio, Korea) and digested with HindIII and XhoI (Thermo Fisher Scientific, USA) at 37 °C for 4 h. The hspX/esxS gene construct was electrophoresed on a 1% agarose gel, purified using a Prime Prep Gel Purification Kit (GeNet Bio, Korea) and ligated into the multiple cloning site (MCS) of the T7lac promoter expression plasmid pET-21b(+) with T4 DNA ligase (5 U/µL) (Thermo Fisher Scientific, USA). To confirm the cloning of hspX/esxS into pET-21b (+), the recombinant plasmid was extracted and sequenced (Macrogen, Korea). Finally, *E. coli* BL21 cells were transformed with pET-21b (+)/hspX/esxS for expression of the hspX/esxS gene construct.

**Expression of hspX/esxS**
*E. coli* BL21 cells, containing pET-21b (+)/hspX/esxS, were cultured in 5 mL of LB broth supplemented with 100 µg/mL of ampicillin for 3-4 hr. When growth reached an O.D. of ~0.4-0.6 at 600 nm, 0.1, 0.5, 1 and 2 mM isopropyl-β-D thiogalactopyranoside (IPTG) (Sigma, USA) were added to induce expression of the HspX/EsxS fusion protein. The *E. coli* BL21 cells were incubated overnight at 37 °C. The culture medium containing the bacteria was centrifuged at 2700 x g for 10 min and the cell pellet was suspended in 1.5 ml of lysis buffer (50 mM K2HPO4, 50 mM H3KPO4, 150 mM NaCl, 10% glycerol, 1% Triton X100, pH 7.8). The cells were sonicated for 5 min on ice followed by centrifugation at 2700 x g for 10 min. Protein expression in the supernatant and pellet were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE).

**Purification of HspX/EsxS fusion protein**
The recombinant protein, with an affinity tag of six consecutive histidine residues, was purified on a Ni-IDA resin chromatography column (5 mL) (Parstous Biotechnology, Iran). Initially, to eliminate preservative materials and equilibrate the resin, the column was washed with 150 mM NaCl and then the start buffer (50 mM K2HPO4, 50 mM H3KPO4, 150 mM NaCl, 0.5% Triton X100, and 1.5 M guanidine HCl, pH 7.8). The protein solution was passed through the column with a flow rate of 0.25-0.5 mL/min. In the next step, to remove any non-bound and non-specific proteins, the column was washed with 100 mL of 50 mM K2HPO4, 50 mM H3KPO4, 150 mM NaCl, 0.1% Triton X114, 1.5 M guanidine HCl, and 500 mM imidazole pH 7.8. In the last step, the fusion protein was eluted with 15 mL of elution buffer (50 mM K2HPO4, 50 mM H3KPO4, 150 mM NaCl, 1.5 M guanidine HCl, and 500 mM imidazole pH 7.8). Fractions of interest were pooled and dialyzed using 500 mL of 1 M phosphate-buffered saline (PBS) followed by 500 mL of 400 mM PBS for 10 h each. The protein purity was evaluated by SDS-PAGE and confirmed by western blot. Briefly, first HspX/EsxS fusion protein was transferred from the gel onto polyvinylidene difluoride (PVDF) membrane (BioRad, USA) using transfer buffer, 25 mM Tris base, 190 mM glycine and 20% methanol pH 8.3, and under electrical current. The time and voltage of transfer were 0.35 A for 40 min at 25 °C. The membrane was washed 4 times with washing buffer, PBS, 0.05% Tween 20, 10 min each. Next the membrane was blocked with 2% (w/v) bovine serum albumin (BSA) in PBS for 16-18 hours at 4 °C and then incubated for 1 hour with monoclonal anti-polyHistidine (mouse IgG2a), peroxidase conjugate antibody (Sigma-Aldrich, USA), at a 1:5000 dilution. The membrane was washed 4 times, 10 min each, and incubated with substrate for 2 minutes. Finally, read of protein band was performed using Axygen Gel Documentation System (Coming Life Sciences, USA). The protein concentration was determined with a BCA Protein Quantification Kit (Parstous Biotechnology, Iran).
Results

Colony-PCR of E. coli BL21 cells transformed with pGH/hspX/esxS amplified a 1035 bp fragment using the T7 universal primer (Fig. 2). Sequencing results were consistent with the target and the size of inserted fragment was confirmed. HspX/EsxS fusion protein expression was induced with IPTG. IPTG concentrations between 0.01 and 5.0 mM were utilized to optimize protein expression. An IPTG concentration of 1 mM resulted in optimum expression; and concentration 2 mM did not increase protein expression (data not shown). The solubility of the fusion protein was calculated using recombinant protein solubility prediction online software (http://biotech.ou.edu). The solubility of our fusion protein was 0.0% when overexpressed in E. coli. So, the expressed protein was insoluble and precipitated in inclusion bodies (Fig. 3A). Three and four M guanidine HCl were used to refold the expressed protein (Fig. 3B). After guanidine HCl treatment, the protein was purified on a Ni-IDA resin chromatography column. The purified protein was then analyzed by SDS-PAGE and western blotting (Figs 4A and 4B).

Fig. 2. Agarose gel electrophoresis of the colony PCR product of hspX/esxS. Lane 1: negative control; lane M: 1kb DNA ladder; lane 2: hspX/esxS amplicon.

Fig. 3. SDS-PAGE of HspX/EsxS fusion protein before and after renaturation with 3 and 6 M guanidine HCl. Gels were stained with Coomassie Brilliant Blue. A) HspX/EsxS fusion protein expression in E. coli strain BL21. Lane M: protein size marker; lanes 1 to 4: supernatant of BL21 lysate in different concentration of IPTG (0.1, 0.5, 1 and 2 mM); lanes 5 to 8: sediment of BL21 lysate in different concentration of IPTG (0.1, 0.5, 1 and 2 mM); lane 9: E. coli strain BL21 without HspX/EsxS fusion protein. B) Refolded fusion protein using 3 and 6 M guanidine HCl. Lane M: protein size marker; lanes 1 and 2: supernatant of BL21 lysate treated with 6 M guanidine HCl; lanes 3 and 4: supernatant of BL21 lysate treated with 3 M guanidine HCl; lanes 5 and 6: pellet of BL21 lysate treated with 6 M guanidine HCl; lanes 7 and 8: pellet of BL21 lysate treated with 3 M guanidine HCl.
Cloning and Expression of Mycobacterium tuberculosis


**Discussion**

It was demonstrated that supplementing BCG with *Mtb* secreted protein subunit vaccines enhanced protective immune responses against TB in animal models (4, 16, and 17). Several studies investigated the efficacy of fusion proteins as subunit vaccines. They suggested that fusion proteins of Ag85B-ESAT-6 (4, 17, and 18), Mtb10.4-HspX (7), and Ag85B-TB10.4 (5) are excellent candidates for clinical investigations and enhance BCG-primed immunity. Because ESAT-6 has been examined in numerous applications, including interferon gamma release assays, and because low expression of Ag85B has been reported, we chose to study EsxS, which has been shown to stimulate the immune system more than ESAT-6 or TB10.4 (Rv0288, esxH) (14, 19, and 20). EsxH belongs to the ESAT6 family and has been proposed as a vaccine candidate. EsxH is similar to EsxS in that it is located in the cell wall, its function is unknown, and it is the same length and molecular mass as EsxS (http://tuberculist.epfl.ch). It is expected that EsxS, in combination with HspX as a multistage fusion protein, has high potential to induce an antimycobacterial immune response in the host. Here, early and latent phase proteins of *Mtb* were expressed as a fusion protein. An (AGGGGS)ₙ flexible linker was used to join the fusion protein domains. These flexible linkers contain the small polar residues serine and glycine, which contribute to flexibility, stability, and folding of the fusion protein. To further improve protein solubility the polar amino acids lysine and glutamine can also be used in the linker. On the basis of protein solubility prediction online software, the solubilities of HspX and EsxS proteins alone were 100%, but they became insoluble when they were connected by a linker (21). Because it is easily grown, expresses recombinant proteins at high levels, and is tolerant of foreign proteins, *E. coli* strain BL21 was used as the host expression system. However, a problem with protein expression in these cells is the formation of inclusion bodies and instability of the recombinant protein (22). As expected, our results were similar to those of other studies in which the expressed protein accumulated in inclusion bodies and was insoluble (4, 23, and 24).

Numerous methods are used to obtain soluble and active protein; these include addition of 3 to 6 M guanidine HCl, 6 to 8 M urea, detergents, including SDS, and alkaline pH (22). We used 3 M and 6 M guanidine HCl in the lysis buffer to solubilize the fusion protein. Guanidine HCl is a chaotropic agent that converts aggregated and denatured proteins into their native conformations (22). After dialysis of the solubilized fusion protein against PBS containing 1 M and then 400 mM NaCl, pH 7.8, respectively, aggregation of the protein in the dialysis tubing was observed. In this step, to refold the protein into an active form, it was dialyzed in PBS containing 3 M guanidine HCl, and 25% glycerol, pH 7.8 for 48 h. Glycerol is an important protein refolding mediator that inhibits protein aggregation. Slow removal of guanidine HCl and glycerol from the aggregated fusion protein by dialysis leads to protein refolding (22). 0.1% Triton X114 was used in the wash buffer to remove bacterial endotoxin from the fusion protein. Endotoxin can indirectly affect host cells and organs via induction of immune responses and confuse the evaluation of the protein’s immunogenic effect. (25). Zhang *et al.* showed that Triton X114 reliably eliminates 98 to 99% of...
bacterial endotoxin from fusion protein solutions (26).

In the present study, an Mtb HspX/EsxS fusion protein was expressed in E. coli and purified. Detection of multistage immunodominant antigens of Mtb such as HspX/EsxS is important for the design of subunit vaccines in order to upgrade the current TB vaccine arsenal. We believe the multistage fusion protein is a good vaccine candidate. Further studies to evaluate the immunogenic potential of the HspX/EsxS fusion protein are needed.

Acknowledgement

The current study was the result of Ph.D. theses of Farzad Khademi and Arshid Yousefi-Avarvand and they are equally first authors. This study has been financially supported by Grant No. 940964-5 from Mashhad University of Medical Sciences.

References

14. Villarreal DO, Walters J, Laddy DJ, Yan J, Weiner DB. Multivalent TB vaccines targeting the esx gene family generate potent and broad cell-
mediated immune responses superior to BCG. Hum Vaccin Immunother 2014; 10(8):2188-98.


