Contents lists available at ScienceDirect

Asian Pacific Journal of Tropical Medicine

journal homepage:www.elsevier.com/locate/apjtm

Document heading doi:

Cloning, expression, identification and bioinformatics analysis of Rv3265c gene from *Mycobacterium tuberculosis* in *Escherichia coli*

Qiang Wu[1,2,3], Peng Zhou[1,2*], Shiyun Qian[3], Xi Qin[4], Zhigang Fan[3], Qiongyao Fu[3], Zhinong Zhan[3], Hua Pei[3]

[1]*Agriculture School of Hainan University, Haikou 570228, P. R. China*
[2]*Institute of Tropical Bioscience and Biotechnology/ Analysis & Testing Center 454, Chinese Academy of Tropical Agricultural Sciences, Haikou 571101, P. R. China*
[3]*The School of Troptical and Laboratory Medicine, Hainan Medical University, Haikou 571101, P. R. China*
[4]*Affiliated Hospital of Hainan Medical University, Haikou 571101, P. R. China*

ARTICLE INFO

Article history:
Received 1 December 2010
Received in revised form 14 February 2011
Accepted 15 March 2011
Available online 20 April 2011

Keywords:
Mycobacterium tuberculosis
L-rhamnosyltransferase
Expression
Identification
Bioinformatics

ABSTRACT

Objective: To clone and express Rv3265c gene of *Mycobacterium tuberculosis* in *Escherichia coli (E. coli)* under optimistic conditions, obtain and identify protein expressed, analyze the structure and characteristics of the protein using bioinformatics methods for future applications. **Methods:** Rv3265c gene from *Mycobacterium tuberculosis* H37Rv was amplified by polymerase chain reaction, and was cloned into the pET-30a vector after purification and recovery. The recombinant plasmid was sequenced and expressed in *E. coli* BL21(DE3), and then purified and identified by western blotting. The essential physical-chemical properties of the protein were predicated by bioinformatics tools, including subcellular location, secondary structure, domains, antigenic epitopes, *etc*. Tertiary structure of the protein based on homology modeling was established, while multi-sequence homological alignment and phylogenetic analysis were proformed. **Results:** The recombinant protein was obtained in soluble fraction from expression system in *E.coli* BL21(DE3) carrying pET30- Rv3265c plasmid, and Rv3265c gene was expressed correctly. Bioinformatics analysis showed the protein contained no signal peptide and transmembrane helices, located outside of membrane. Secondary structure analysis revealed it contained α-helix, extended strand and random coil, 46.8%, 14.6%, 38.6%, respectively. Furthermore, it possessed six potential antigenic epitopes, one glycosyl transferase domain. A simple three-dimensional model of this protein was constructed by Swiss-model sever. Both sequences and structures were conservative and especial either in gene or in protein. **Conclusions:** Rv3265c gene might be a desirable molecular target for anti-tuberculosis drug and vaccine. The purified protein from expression will be utilized to study the kinetics of L-rhamnosyltransferase and to develope an enzyme assay for screening vaccine or drug.

1. Introduction

Mycobacterium tuberculosis is one of the most threatening of human infectious diseases, and 2 million people die from tuberculosis every year[1]. No new drug has been licensed for tuberculosis for half a century and no new vaccine has been licensed for more than three-quarters of a century. The cell wall of *Mycobacterium tuberculosis* composed of peptidoglycan, arabinogalactan and mycolic acids complex is essential for proliferation and growth[2]. Some of the first-line tuberculosis drugs target cell wall synthesis in *Mycobacterium tuberculosis*, but their specific targets and mechanisms of inhibition are not well defined. The rhamnose-GlcNAc disaccharide linker connects the galactan of arabinogalactan to the peptidoglycan. The donor dTDP-L-rhamnose is synthesized by four enzymes (RmlA, RmlB, RmlC, and RmlD) beginning with glucose-1-phosphate (Glu-1-p) and dTTP[3,4]. The function of L-rhamnosyltransferase encoded by Rv3265c gene is transfering dTDP-L-rhamnose residue to GlcNAc to form

*Corresponding author: Peng Zhou, PhD, Agriculture School of Hainan University, Haikou 571101, China.
Tel: +86-0898-66890687
Fax: +86-0898-66988559
E-mail: zhp6301@126.com

Foundation project: Supported by Research Program in higher educational institutes of The Education Department of Hainan Province(No. Hj2010-21), Research Program of The Health Department of Hainan Province(No. 2007-44), Natural Science Fund of Hainan Province(No. 2008-30837), Cultivation Program of Hainan Medical University (HY2010-006).

disaccharide linker, which is essential for *Mycobacterium tuberculosis* viability[5,6]. The pathway of biosynthetic L-rhamnose and rhamnose-GlcNAc which is highly conserved in Gram-positive and Gram-negative bacteria does not present in humans, so enzymes of the pathways are very attractive targets for antimycobacterial drugs and possible vaccines[7]. In order to study the function of this enzyme, we expressed Rv3265c gene in *Escherichia coli (E. coli)* strain and predicted the characteristics and structure of L-rhamnosyltransferase using bioinformatics methods.

2. Materials and methods

2.1. Bacterial strains, plasmids, cultivation and enzymes

Genomic DNA from *Mycobacterium tuberculosis* H37Rv was kindly provided by Lab Director ZY Zhu of Affiliated Xinhua Hospital of Hainan Medical University. *E. coli* DH5 α (TaKaRa Biotechnology) was used for cloning, transformation and propagation. *E.coli* BL21 (DE3) (Merck) and plasmid pET30-a (Merck) were used for protein expression. All *E. coli* strains were grown in LB broth or on LB plate with 1.5% (wt/vol) agar 37 ℃ or 30 ℃, respectively. When needed, kanamycin was used at final concentration of 50 μg/mL for *E. coli*. Restriction enzymes (*Bam*HI, *Sac*I), PrimeSTAR™ HS DNA Polymerase, T4 ligase and Taq DNA Polymerase were purchased from TaKaRa Biotechnology.

2.2. Amplification of Rv3265c gene

The nucleotide sequence of Rv3265c gene was obtained from *Mycobacterium tuberculosis* H37Rv genome (http://genolist.pasteur.fr/TubercuList/). Genomic DNA from *Mycobacterium tuberculosis* H37Rv was used as template in the PCR. A pair of primers, TB Rv3265c-F (5′-CGCGGATCCATGGTAGCGGTGACCTACTCGC-3′) and TB Rv3265c-R (5′-TCCGAGCTCTCAGTGCCGCCCTTCTACCAGCT-3′) contained *Bam*HI and *Sac*I endonuclease restriction sites (underlined) were used for amplification. PCR was performed by using PrimeSTAR™ HS DNA polymerase as follow conditions: 94 ℃ for 5min, 30 cycles (94 ℃ 30 s; 65 ℃ 40 s; 72 ℃ 1 min), and then 72 ℃ for 10 min.

2.2. Construction of expression plasmid

The PCR product was purified by using an agarose gel DNA fragment recovery kit (Tiangen, Beijing, China), according to the manufacturer protocol. And then, purified PCR products digested with *Bam*HI and *Sac*I was cloned into the pET30-a vector resulting in plasmid pET30-Rv3265c, and chemically transformed into the *E. coli* DH5 α cells using standard procedures. Recombinant plasmid pET30-Rv3265c was identified by PCR and restriction endonuclease, and then sequenced to confirm that no base changes had been introduced during amplification of the DNA.

2.3. Expression of Rv3265c gene in E. coli

The recombinant plasmid pET30-Rv3265c was transformed into competent *E. coli* BL21 (DE3) cells and screened on LB agar plates containing kanamycin. Individual colonies of the transformants were grown in LB broth at 37 ℃. The overnight grown cultures were subcultured in fresh medium at 30 ℃ to the density of OD_{600}≈0.6. Cells were induced with IPTG (TaKaRa Biotechnology) at final concentration of 0.1 mM for various induction periods of 1, 2 and 3 h at 30 ℃. After induction, cells were harvested by centrifugation, and then whole-cell proteins of the induced and control cells were analyzed on 12% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). To visualize proteins, gels were stained 30 min with 0.05% Coomassie blue, and then destained with 10% acetic acid and 25% isopropanol. In order to examine the extent of aggregation of the produced, induced cells were disrupted by sonication for 5 min on ice and soluble as well as insoluble fractions were analyzed using SDS-PAGE as described previously.

2.4. Protein purification and Western blotting analysis

The soluble fractions were applied to column volume HIS-Select HF affinity gel (Sigma), and then the column was washed with wash buffer. The protein was eluted with 15 mL elute buffer with 1 μM leupeptin, 1 μM pepstatin A, and 1mM PMSF as the manufacturer's protocol. The purified protein was run on 12% SDS-PAGE and transferred to nitrocellulose membrane. After blocking with 5% nonfat dry milk, the membrane was probed with 1:3000 dilution of mouse anti-polyhistidine monoclonal HIS-1 antibody (Sigma) followed by 1:5000 dilution of horseradish peroxidase(HRP) -conjugated goat anti-mouse IgG antibody (Sigma).The membrane was checked by utilizing the 3,3′-diaminobenzidine liquid substrate system.

2.5. Bioinformatics analysis of Rv3265c gene and protein

All of nucleotide and amino acid sequences in this study were from the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov). The multiple protein sequence homological alignment was carried out using the ClustalW program (http://www.ebi.ac.uk/clustalw). Signal peptides, cleavage sites and subcellular location were predicted by the SignalP program, TargetP program and TMHMM program, respectively (http://www.cbs.dtu.dk/). Conserved domains and motif were predicted by Motif scan program (http://myhits.isb-sib.ch/cgi-bin/motif_scan), InterProScan program (http://www.ebi.ac.uk/Tools/InterProScan/) and Pfam program (http://pfam.sanger.ac.uk/) together. Antigenic epitopes were predicted by Predicting Antigenic Peptides program (http://bio.dfci.harvard.edu/Tools/antigenic.html) and Abie Pro 3.0 program (http://www.changbioscience.com/abie/abie.html). Secondary structure and topology was predicted by PredictProtein program (http://www. Predictprotein.org/), ANTHEPROT 2000 software (http://antheprot-pbil.ibcp.fr/), COILS program (http://www.ch.embnet.org/software/COILS_form.html). Tertiary structure was established using SWISS-MODEL (http://swissmodel.expasy.org/) and RasTop 2.2 software (http://www.geneinfinity.org/rastop/) based on homology modeling. Basic physical and chemical properties analysis and phylogenetic analysis was analyzed using Vector NTI Advance 10 software (Invitrogen, USA) and Geneious Pro 3.8.5 software (Biomatters, USA).

3. Results

3.1. Molecular cloning of Rv3265c gene

Rv3265c gene was amplified by PCR (Figure 1A). DMSO has been added to PCR reaction buffer to enhance denaturation of GC rich template DNA and to aid polymerase extension through the secondary structures in our testing. The resulting fragment contained about 915 bp and had an open reading frame encoding 301 amino acids. The Rv3265c gene was then cloned into pET30-a. The recombinant plasmid pET30-Rv3265c was identified by PCR and restriction endonuclease (Figure 1B). The expression plasmid pET30-a was chosen because it provides the N-terminal 6×His Tag and S. Tag sequence to the expressed proteins, which should facilitate protein purification by affinity purification methods. Tag-encoded protein sequences can be removed from the resulting fusion proteins by cleavage with enterokinase. The sequencing results demonstrated that Rv3265c gene been cloned in this study was identical to the published sequence and there was no any change of base pair (GenBank Accession No.BX842582.1).

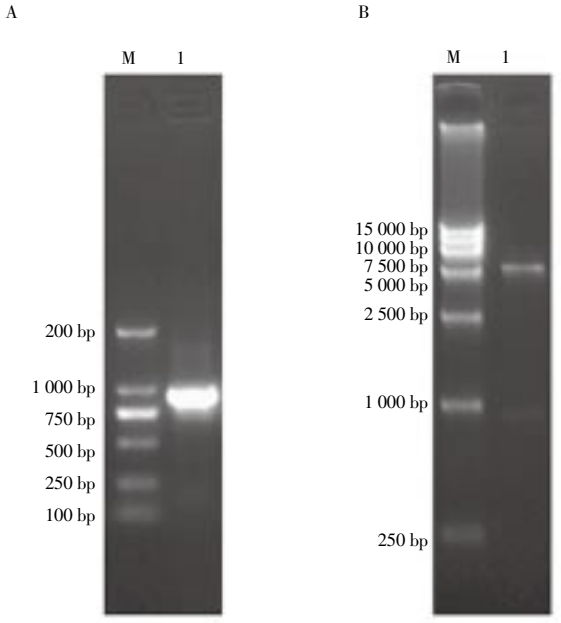

Figure 1. Amplification of Rv3265c gene and restriction endonuclease analysis of recombinant plasmid pET30-Rv3265c.
(A) Rv3265c gene was amplified from *Mycobacterium tuberculosis* H37Rv genomic DNA by PCR. Lane M, DNA ladder; lane 1, PCR product. (B) Recombinant plasmid pET30-Rv3265c was identified by restriction endonuclease. Lane M, DNA ladder; Lane 1, pET30-Rv3265c was digested by *Bam*H Ⅰ and *Sac* Ⅰ.

3.2. Expression and identification of L-rhamnosyltransferase protein

L-rhamnosyltransferase was produced in *E. coli* BL21 (DE3) cells carrying pET30-Rv3265c plasmid. Our attempts to induce the protein at a lower incubation temperature (30 ℃), lower concentration of IPTG (0.1 mM) and short induction time (1h) in order to avoid inclusion bodies formation.

L-rhamnosyltransferase fusing protein was observed on Coomassie blue R250 stained gel as expected 38 kDa band (Figure 2A). Analysis of the total and soluble fractions of the expressed L-rhamnosyltransferase fusing protein by SDS-PAGE demonstrated that the protein was localized premodinantly in the soluble fraction (Figure 2B). The L-rhamnosyltransferase was purified from the supernatant by passing through HIS-Select HF affinity gel. SDS-PAGE analysis of the purified sample showed single band slightly about 38 kDa (Figure 2C). Western blotting analysis of the same samples confirmed that the observed protein band reacted with the specific antibodies (Figure 2D).

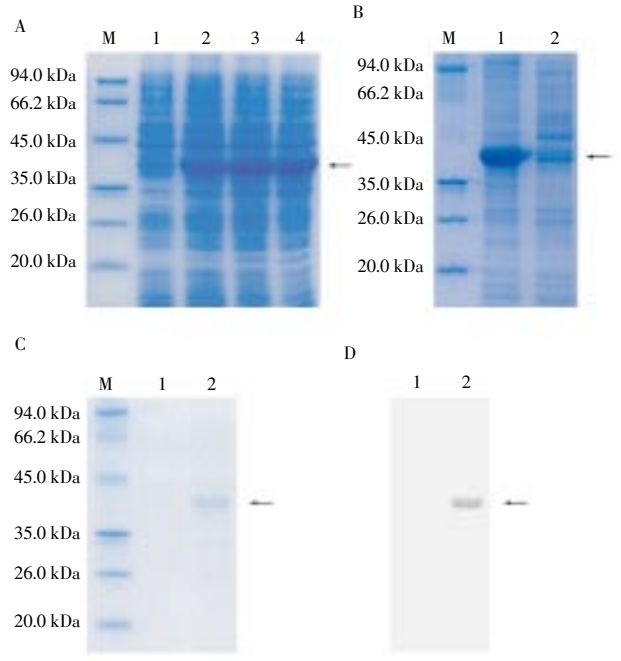

Figure 2. SDS-PAGE analysis and Western blotting identification of L-rhamnosyltransferase fusing protein expressed in *E. coli* BL21 (DE3) cells.
(A) SDS-PAGE analysis of total proteins expressed. Lane M, protein molecular mass standards; lane 1, cell lysate of bacteria transformed with pET30-Rv3265c without induction; lane 2, cell lysate of bacteria transformed with pET30-Rv3265c under 0.1 mM IPTG induction for 1 h; lane 3, cell lysate of bacteria transformed with pET30-Rv3265c under 0.1 mM IPTG induction for 2 h; lane 4, cell lysate of bacteria transformed with pET30-Rv3265c under 0.1 mM IPTG induction for 3 h. (B) SDS-PAGE analysis of soluble and insoluble fractions proteins expressed. Lane M, protein molecular mass standards; lane 1, induced pET30-Rv3265c in *E. coli* BL21 (DE3) by 0.1 mM IPTG for 1 h, supernatant; lane 2, induced pET30-Rv3265c in *E. coli* BL21 (DE3) by 0.1 mM IPTG for 1 h, pellet. (C) SDS-PAGE analysis of purified L-rhamnosyltransferase fusing protein expressed. Lane M, protein molecular mass standards; lane 1, negative control; lane 2, the protein purified by affinity chromatography from supernatant induced pET30-Rv3265c in *E. coli* BL21 (DE3) by 0.1 mM IPTG for 1 h. (D) Western blotting analysis of L-rhamnosyltransferase fusing protein. lane 1, negative control; lane 2, Western blotting analysis of L-rhamnosyltransferase fusing protein. The arrow indicates the location of the recombinant fusion protein.

3.3. Characteristics and structures of L-rhamnosyltransferase protein from bioinformatics analysis

L-rhamnosyltransferase from *Mycobacterium tuberculosi* composes 301 amino acids with an estimated molecular mass of 33.3kDa and a pI value of 9.98. This protein has an N-terminal glycosyl transferase domain, no signal peptide and transmembrane helices, locates outside of membrane (Figure 3). The amino acids sequence of L-rhamnosyltransferase from *Mycobacterium tuberculosis*

H37Rv had the highest genetic relationship and showed 100% homology with *Mycobacterium tuberculosis* H37Ra, *Mycobacterium tuberculosis* F11 and *Mycobacterium tuberculosis* CDC1551, but lower homology with other Mycobacterium (<80%) and lowest homology with other bacteria except *Mycobacterium* (<50%) (Figure 4). Secondary structure analysis revealed that the protein contained α -helix (46.8%), extended strand (14.6%) and random coil (38.6%) (Figure 5). It possesses six potential antigenic epitopes, VTYSPGPHLERFLASLSLATERPVSVLLAD (4-33), PQAAVQRYPNVRLLPT (41-56), PVGWLSGSCLLVRR (173-186), GWLSVYVPSAEVLHHK (216-237), PASHLAAHHKSTYIFLAD (239-256), LRGSLALRSHLMVRS (270-284), respectively(Figure 6).

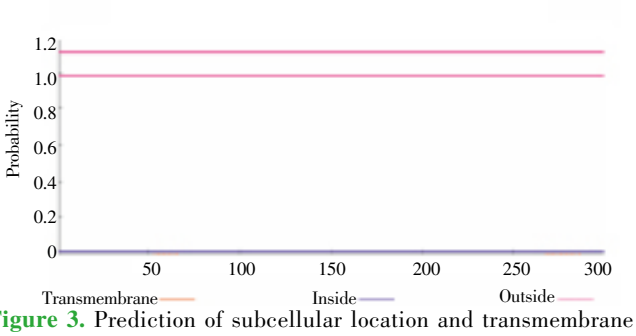

Figure 3. Prediction of subcellular location and transmembrane helices of L-rhamnosyltransferase.

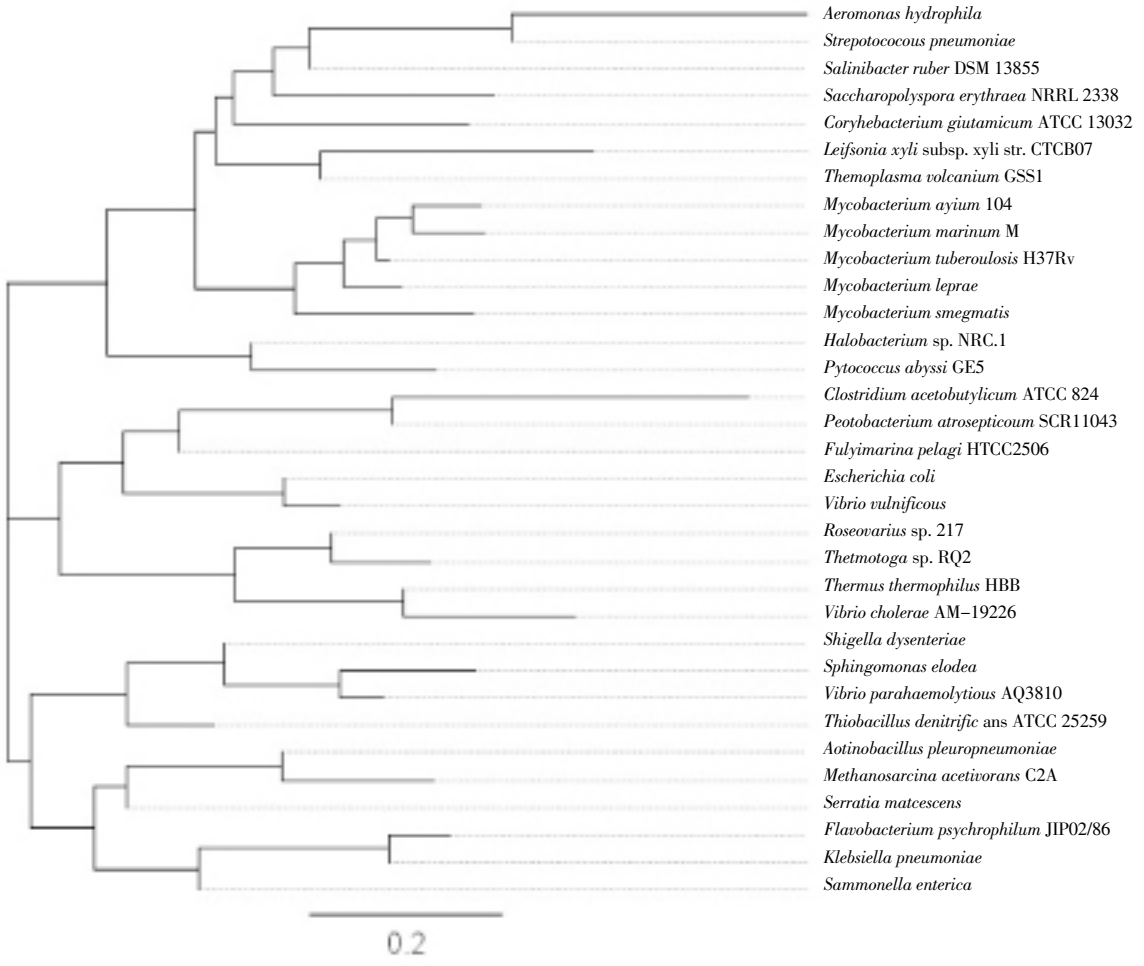

Figure 4. Evolution tree according to L-rhamnosyltransferase from other species.

Figure 5. Secondary structure prediction of L-rhamnosyltransferase. H: α -helix, E: extended strand, C: random coil.

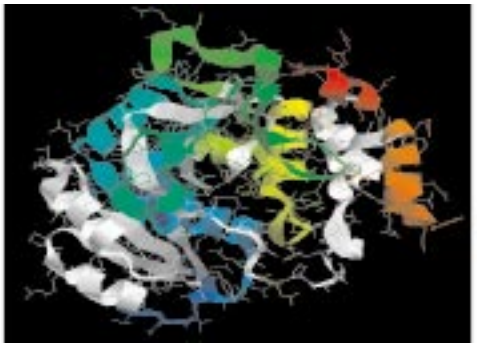

Figure 6. The 3D structure of L-rhamnosyltransferase and the spacial localization of possible epitopes (white segments).

4. Discussion

Although tuberculosis remains a serious global public health threat due to the serious problem of drug resistance, particularly multi-drug resistant-tuberculosis, and tuberculosis associated with HIV infection, technologies for diagnosis, treatment, and prevention for tuberculosis are old and inadequate[8,9]. There is concurrent agreement that new anti-tuberculosis drugs are needed to shorten or simplify treatment. Despite ten compounds have being evaluated in clinical trials in the past decade, the global drugs pipeline for tuberculosis is still insufficient to address the unmet needs of treatment[10]. *Mycobacterium bovis* Bacillus Calmette-Guérin (BCG) vaccine is the only currently available licensed vaccine against tuberculosis now[11]. Although effective against development of tuberculosis in some countries, its effectiveness has been questioned in several studies, most notably in India, where very limited or no protection has been reported[12]. Therefore, the development of new therapeutic vaccine as well as new prophylactic vaccine against tuberculosis is urgently needed.

Looking for novel candidates or targets is first step of development of drugs and vaccines against tuberculosis. Better understanding of *Mycobacterium tuberculosis* cell wall biosynthesis is required in order to elucidate the targets of existing drugs and vaccines to find new ones[13]. Essential genes of *Mycobacterium tuberculosis* growth *in vitro* and survival *in vivo* are main targets for anti-tuberculosis drugs development now[14]. The possibility of the protein encoded by Rv3265c gene of *Mycobacterium tuberculosis* is a putative rhamnosyltransferase which is needed for growth of mycobacteria[15-17].

The completion of *Mycobacterium tuberculosis* genome sequences has greatly facilitated identification of a large numbers of targets and antigens with potential in anti-tuberculosis drugs and vaccines[18,19]. Bioinformatics analysis is a powerful tool for protein identification, study of its localization, modification, function and possible interaction. Both sequences and structures of L-rhamnosyltransferase are conservative and especial. It located outside of membrane and posses potential antigenic epitopes. Accordingly, L-rhamnosyltransferase is a potential and suitable molecular target to develop new anti-tuberculosis drugs and vaccines.

Conflict of interest statement

We declare that we have no conflict of interest.

References

[1] Sutherland JS, Young JM, Peterson KL, Sanneh B, Whittle HC, Rowland-Jones SL, et al. Polyfunctional $CD4^+$ and $CD8^+$ T cell responses to tuberculosis antigens in HIV-1-infected patients before and after anti-retroviral treatment. *J Immunol* 2010; **184**(11): 6537-6544.

[2] Wolfe LM, Mahaffey SB, Kruh NA, Dobos KM. Proteomic definition of the cell wall of *Mycobacterium tuberculosis*. *J Proteome Res* 2010; **9**(11): 5816-5826.

[3] Li W, Xin Y, McNeil MR, Ma Y. *rml*B and *rml*C genes are essential for growth of mycobacteria. *Biochem Biophys Res Commun* 2006; **342**(1):170-178.

[4] Ma Y, Stern RJ, Scherman MS, Vissa VD, Yan W, Jones VC, et al. Drug targeting *Mycobacterium tuberculosis* cell wall synthesis: genetics of dTDP-rhamnose synthetic enzymes and development of a microtiter plate-based screen for inhibitors of conversion of dTDP-glucose to dTDP-rhamnose. *Antimicrob Agents Chemother* 2001; **45**(5): 1407-1416.

[5] Mills JA, Motichka K, Jucker M, Wu HP, Uhlik BC, Stern RJ, et al. Inactivation of the mycobacterial rhamnosyltransferase, which is needed for the formation of the arabinogalactan-peptidoglycan linker, leads to irreversible loss of viability. *J Biol Chem* 2004; **279**(42): 43540-43546.

[6] Grzegorzewicz AE, Ma Y, Jones V, Crick D, Liav A, McNeil MR. Development of a microtitre plate-based assay for lipid-linked glycosyltransferase products using the mycobacterial cell wall rhamnosyltransferase WbbL. *Microbiology* 2008; **154**(Pt 12): 3724-3730.

[7] Qu H, Xin Y, Dong X, Ma Y. An rmlA gene encoding d-glucose-1-phosphate thymidylyltransferase is essential for mycobacterial growth. *FEMS Microbiol Lett* 2007; **275**(2): 237-243.

[8] Lönnroth K, Castro KG, Chakaya JM, Chauhan LS, Floyd K, Glaziou P, et al. Tuberculosis control and elimination 2010-50: cure, care, and social development. *Lancet* 2010; **375**(9728): 1814-1829.

[9] Nuermberger EL, Spigelman MK, Yew WW. Current development and future prospects in chemotherapy of tuberculosis. *Respirology* 2010; **15**(5): 764-778.

[10] Ma Z, Lienhardt C, McIlleron H, Nunn AJ, Wang X. Global tuberculosis drug development pipeline: the need and the reality. *Lancet* 2010; **375**(9731): 2100-2109.

[11] Russell DG, Barry CE 3rd, Flynn JL. Tuberculosis: what we don't know can, and does, hurt us. *Science* 2010; **328**(5980): 852-856.

[12] Verma I, Grover A. Antituberculous vaccine development: a perspective for the endemic world. *Expert Rev Vaccines* 2009; **8**(11): 1547-1553.

[13] Crick DC, Brennan PJ, McNeil MR. The cell wall of *Mycobacterium tuberculosis*. In: *Tuberculosis*. 2nd edition. 2004; 115-134.

[14] Sassetti CM, Boyd DH, Rubin EJ. Genes required for mycobacterial growth defined by high density mutagenesis. *Mol Microbiol* 2003; **48**(1): 77-84.

[15] Ma Y, Pan F and McNeil M. The formation of dTDP-rhamnose is essential for growth of mycobacteria. *J Bacteriol* 2002; **184**(12): 3392-3395.

[16] Zhang JH, Shang ZG, Zhang XH, Zhang YT. Modeling and analysis of Schistosoma Argonaute protein molecular spatial conformation. *Asian Pac J Trop Biomed* 2011; **1**(4): 275-278.

[17] Garberi J, Labrador J, Garberi F, Garberi JE, Peneipil J, Garberi M, et al. Diagnosis of Mycobacterium tuberculosis using molecular biology technology. *Asian Pac J Trop Biomed* 2011; **1**(2): 89-93.

[18] Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, et al. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* 1998; **393**(6685): 537-544.

[19] Camus JC, Pryor J, Medigue C, Cole ST. Re-annotation of the genome sequence of *Mycobacterium tuberculosis* H37Rv. *Microbiology* 2002; **148**(10): 2967-2973.