Tuberculosis 108 (2018) 83–92

Contents lists available at ScienceDirect

Tuberculosis

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

Molecular Aspects

An IclR like protein from mycobacteria regulates *leuCD* operon and induces dormancy-like growth arrest in *Mycobacterium smegmatis*

Rajendra Kumar Angara[a,b,1], Suhail Yousuf[a,b,1], Shailesh Kumar Gupta[a,b], Akash Ranjan[a,*]

[a] Computational and Functional Genomics Group, Centre for DNA Fingerprinting and Diagnostics (CDFD), Hyderabad, Telangana, India
[b] Graduate Studies, Manipal University (MU), Manipal, Karnataka, India

ARTICLE INFO

Keywords:
IclR like protein
Leucine biosynthesis
Lipid droplets
Dormancy
Mycobacteria

ABSTRACT

leuCD operon encodes isopropylmalate isomerase (IPMI), an essential enzyme in leucine biosynthesis. Leucine biosynthesis is one of the essential metabolic pathways for *Mycobacterium tuberculosis* survival inside the macrophage. In this study, we identified an IclR like transcription regulator, Rv2989 involved in regulation of *leuCD* expression. Further, we have shown that the Rv2989 binding site overlaps with the promoter region of *leuCD*, indicating its direct involvement in the regulation of this operon. Ectopic expression of Rv2989 in *M. smegmatis* induced growth arrest with significantly decreased levels of *leuCD* transcript. However, supplementation with leucine could not reverse the growth arrest, suggesting the involvement of Rv2989 in the regulation of other essential pathways. Growth-arrested cells were elongated, had lost acid fastness and accumulated lipid droplets similar to a dormancy-like state. In conclusion, the Rv2989 expression has pleiotropic effects on *M. smegmatis*. It negatively regulates *leuCD* operon and induces dormancy-like growth arrest.

1. Introduction

Tuberculosis remains a major health challenge for millennia. Only 5–15% of infected individuals show active tuberculosis in their lifetime. In the rest, *Mycobacterium tuberculosis* remains dormant and poses a risk of disease development in future [1]. Understanding basic metabolism or pathways essential for survival is required to develop new drugs. One of the pathways that remains an interesting target for developing new drugs is leucine biosynthesis. *leuCD* operon encodes isopropylmalate isomerase enzyme complex (IPMI), an essential enzyme for leucine biosynthesis and survival of mycobacteria. The deletion of *leuD* leads to leucine auxotrophy and mycobacterium fails to grow *in vitro* and *in vivo* [2,3]. Since leucine is an essential amino acid for humans, IPMI involved in leucine biosynthesis of mycobacteria emerged as a prominent target for developing chemotherapy. Currently, studies are underway to crystallise the entire IPMI complex and identify drugs that have the potential to target it. However, these attempts have not been successful, because of difficulties in achieving the required purity and stability of IPMI complex [4].

The IclR family of proteins is present in both eubacteria and archaea and are known regulators of genes associated with glyoxylate shunt, multidrug resistance, aromatic compound degradation, plant cell wall degradation, quorum-sensing signals, sporulation, and amino acid metabolism [5]. LtbR in *Corynebacterium glutamicum* [6], AreB in *Streptomyces clavuligerus* [7] and NdgR in *Streptomyces coelicolor* [8] are IclR family of proteins known to be regulators of leucine biosynthesis. Although they share similar genomic organisation (i.e., divergent transcription with respect to *leuCD* operon), LtbR is involved in negative regulation, while AreB and NdgR are involved in positive regulation [6–8].

In this context, we studied the involvement of IclR like proteins in regulation of *leuCD* operon in *M. tuberculosis*. *M. tuberculosis* genome encodes three IclR like proteins *viz* Rv1719, Rv1773c and Rv2989. Among the three proteins, Rv2989 transcribes divergently to *leuCD* operon and shares similarity with LtbR, NdgR, and AreB; indicating a possible regulatory role. In a recent study by Li et al., Rv2989 is shown to be involved in isoniazid tolerance and its absence influences survival inside macrophages [9]. However, the role of Rv2989 as a transcriptional regulator of *leuCD* operon as well as its effect on mycobacterial physiology has not been explored.

In this study, we identified Rv2989 as a negative regulator of the *leuCD* operon. Overexpression of Rv2989 in *M. smegmatis* besides regulating an operon associated with leucine biosynthesis also induces growth arrest which shows dormancy-like phenotype.

* Corresponding author.
E-mail addresses: rajendra@cdfd.org.in (R.K. Angara), suhailyousufs@gmail.com (S. Yousuf), shaileshgupta@cdfd.org.in (S.K. Gupta), akash@cdfd.org.in (A. Ranjan).
[1] Authors contributed equally.

https://doi.org/10.1016/j.tube.2017.10.009
Received 27 April 2017; Received in revised form 14 October 2017; Accepted 22 October 2017

2. Materials and methods

2.1. Bacterial strains, plasmids and culture media

The Bacterial strains and plasmids used in this study are listed in Table S1 and S2. *M. smegmatis* mc^{2}155 and its modified strains were grown in Middlebrook 7H9 broth (Difco) containing 10% (v/v) oleic acid-albumin-dextrose-catalase (OADC) supplement (Himedia), 0.2% (v/v) glycerol, and 0.05% (v/v) Tween80 or on 7H10 agar medium with the same supplements without Tween80. *Mycobacterium bovis* BCG Pasteur 1173P2 strain was grown in complete Middlebrook 7H9 broth for RNA isolations. Unless otherwise mentioned, *E. coli* was grown in Luria-Bertani (LB) medium at 37 °C on shaking incubator. Kanamycin (50 μg ml^{-1}), Ampicillin (100 μg ml^{-1}) and Hygromycin (50 μg ml^{-1}) were added to the media as and when required. For induction studies, 0.2% (w/v) acetamide was included in the culture for specified time. Leucine (50 μg ml^{-1}) and casamino acids (0.2% (w/v)) was included in the culture in the required studies.

2.2. Plasmid construction

Standard molecular biology techniques were followed for DNA manipulations as described earlier [10]. ORF and intergenic regions of *leuCD-Rv2989* loci are 100% conserved (at the nucleotide level) between *M. tuberculosis* and *M. bovis* BCG Pasteur 1173P2 strain. Hence, we used *M. bovis* BCG Pasteur 1173P2 strain genomic DNA to amplify Rv2989 ORF, *Rv2989* upstream and *leuC* upstream regions. For protein purification, the homologous ORF of Rv2989 was amplified using 2989FP and 2989RP primers (Table S3) from *M. bovis* BCG Pasteur 1173P2 strain genomic DNA and cloned into pET21b vector between *Nde*I and *Xho*I sites to obtain the pETRv2989 plasmid. For studying promoter activity, 250bps of upstream and 50bp downstream of *leuC* and *Rv2989* start codon were amplified using UPleuCFP, UPleuCRP and UP2989FP, UP2989RP primers respectively, cloned between *Xba*I and *Hin*dIII sites of pEJ414 plasmid (Kind gift by Dr. Roger Buxton) to generate pEJleuCwt and pEJ2989wt. Site-directed mutagenesis was carried out to generate pEJleuCmt and pEJ2989 mt using overlapping extension PCR method. Primers used for the SDM are listed in Table S3.

For constitutive overexpression, pVVRv2989 was generated by cloning Rv2989 ORF between *Nde*I and *Hin*dIII sites of pVV16 vector (Table S3). pVVSM2386 was generated by cloning SMEG2386 ORF (amplified from *M. smegmatis* mc^{2}155 genomic DNA) between *Nde*I and *Hin*dIII sites of pVV16 vector (Table S3). For inducible expression, pJVRv2989 was generated by modifying pJV53 plasmid (kind gift by Dr. Graham Hatfull) [11] where the Rv2989ORF has been cloned between *Nde*I and *Nhe*I sites replacing Che9c genes 60–61. While amplifying Rv2989ORF with pjv2989FP and pjv2989flagRP primers (Table S3), Flag sequence was included in the reverse primer to incorporate Flag tag at the C-terminal end of expressed Rv2989. The pJVRv2989 plasmid contains acetamidase promoter, using which the expression of the protein can be regulated either by including or excluding acetamide in the media. pJVSM2386 plasmid was generated by cloning SMEG2386 ORF from *M. smegmatis* mc^{2}155 genomic DNA in the similar fashion as described above using pjvsmeg2386FP and pjvsmeg2386RP primers (Table S3). pJV used as vector control was generated by removing Che9c genes 60–61 from pJV53 vector using *Nde*I and *Nhe*I restriction enzymes. The overhangs of digested vector are filled by End-filling using DNA Polymerase I Large (Klenow) fragment (NEB -M0210) and self-ligated by blunt end ligation. All constructs were verified by DNA sequencing. Details of primers (Eurofins Genomics and Xcelris Labs Ltd.) used to generate different clones are mentioned in Table S3.

2.3. Expression and purification of his-tagged Rv2989 protein

For purification of recombinant (His)6-Rv2989, *E. coli* BL21 (DE3) cells were transformed with pETRv2989. The cells were grown in LB medium at 37 °C to an A_{600} - 0.6, and expression was induced with 1 mM isopropyl 1-thio-β-D-galactopyranoside (IPTG). After the addition of IPTG, the cells were incubated with shaking at 18 °C overnight. The cells were harvested by centrifugation (4000x*g*, 15min, and 4 °C). Recombinant protein was purified using Ni-NTA agarose (Qiagen) following standard protocols. Cell lysis buffer contains 500 mM NaCl, 5% (v/v) glycerol, 50 mM HEPES (pH 7.5), 10 mM imidazole. Whereas, wash buffer and elution buffer contain similar composition with 40 mM imidazole and 300 mM imidazole respectively. The purity of the protein was analyzed by SDS-polyacrylamide gel electrophoresis.

2.4. Electrophoretic mobility shift assays (EMSA)

EMSA studies were carried out to study the interaction of the purified (His)6-Rv2989 recombinant protein and the intergenic region of *leuC* and *Rv2989*. The 239bp region covering the first 127bp of *Rv2989*, first 40bp of *leuC* and 72bp intergenic region was generated by PCR using UPleuCFP and UPleuCRP primers (Table S3). The reaction mixture [30 μl contains 0.05 pmol of labeled DNA, EMSA buffer (25 mM HEPES pH-7.9, 0.1 mM EDTA, 10 mM $MgCl_2$, 20 mM KCl, 5% (v/v) glycerol), 500 ng ml^{-1} poly (dI-dC) as non-specific competitor DNA and purified Rv2989 recombinant protein (0–800 pmol)] was incubated for 30 min at room temperature before resolving on a 4% non-denaturing polyacrylamide gel buffered with 0.5X TBE. To verify the specificity of binding, cold competition assay using 50 fold excess unlabeled DNA (specific competitor) was carried out. We also tested its binding against a nonspecific target – *sigA*. 200bp upstream region of *sigA* was amplified with sigAupFPand sigAupRP (Table S3).

To identify an exact binding site in the intergenic region, synthesized overlapping oligonucleotides were used. The sequence of oligonucleotides used is listed in Table S4 and Fig. 1E. Binding assays were performed as described above.

2.5. Primer extension

The primer extension as described in Cold Spring Harbor protocols was used to identify the (+1) transcription start site (TSS) [12]. ORF and intergenic regions of *leuCD-Rv2989* loci is 100% conserved (at the nucleotide level) between *M. tuberculosis* and *M. bovis* BCG. As the regulatory elements associated with gene expression are usually present in intergenic region and are conserved, we believe the *M. bovis* BCG shares similar regulation events with *M. tuberculosis*. Hence RNA isolated from *M. bovis* BCG was used for primer extension. Briefly, total RNA was isolated from *M. bovis* BCG log phase cultures using PureLink RNA Mini kit (Ambion) as per the manufacturer's instructions. Primers were 5′- end labeled with γ^{-32P} ATP using T4 Polynucleotide kinase (Thermo scientific). Labeled primers (10 pmol) were annealed to 10 μg of total RNA at 55 °C for 90 min. The primer extension was carried out using SuperScript III RT (Invitrogen) at 37 °C for 60 min. The primer extension products were ethanol precipitated, suspended in formamide dye and denatured at 90 °C for 5 min before being separated on a 6% polyacrylamide sequencing gel containing 8 M urea. Primer extension without BCG-RNA was used as negative control (reaction with no template). The sizes of the primer extension products were determined by comparison with an unrelated DNA sequencing ladder generated by sequencing double strand pUC19 bacteriophage DNA using standard sequencing primer according to the manufacturer's guidelines (USB sequenase quick-denature plasmid sequencing kit, Product No - 70140). Using the length of primer extension product, the TSS was identified from the first base of the reverse primer used in the assay. Positions of reverse primers used for primer extension are underlined in Fig. 2A and B and listed in Table S5.

2.6. Growth curve

M. smegmatis pJVRv2989 and *M. smegmatis pJV* were grown in