

RESEARCH

Open Access



# MSMEG\_3955 from *Mycobacterium smegmatis* is a FMN bounded homotrimeric NAD(P)H:Flavin mononucleotide (FMN) oxidoreductase

Neha Khosla<sup>1</sup>, Seema Madhumal Thayil<sup>1</sup>, Rajinder Kaur<sup>2</sup> and Anup Kumar Kesavan<sup>1,3\*</sup>

## Abstract

**Background:** Tuberculosis (TB) remains an important public health problem since it is the major cause of elevated morbidity and mortality globally. Previous works have shown that *Mycobacterium tuberculosis* (*Mtb*); the prime causative agent of the deadly disease has dormancy survival regulator (DosR) regulon, a two-component regulatory system which controls the transcription of more than 50 genes. However, the structure and detailed functions of these DosR regulated genes are largely undetermined. Out of many DosR regulon genes, *Rv3131* gets up regulated in hypoxic conditions and was believed to encode for a nitroreductase flavoprotein. The utilization of mycobacteria-specific model systems has greatly added to our understanding of the molecular mechanisms involved in the life cycle and pathogenesis of *Mtb*.

**Results:** In this study the non-pathogenic mycobacterial model organism *Mycobacterium smegmatis* (*Msmeg*) was used to reveal the structure and function of MSMEG\_3955; which is a homologue of *Rv3131* from *Mtb*. Using chromatography and spectroscopy techniques it was revealed that cofactor flavin mononucleotide (FMN) was bound to flavoprotein MSMEG\_3955. Consistent with the homology modelling predictions, Circular Dichroism (CD) analysis indicated that the MSMEG\_3955 is composed of 39.3%  $\alpha$ -helix and 24.9%  $\beta$ -pleated sheets. In contrast to the current notions, the enzymatic assays performed in the present study revealed that MSMEG\_3955 was not capable of reducing nitro substrates but showed NADPH dependent FMN oxidoreductase activity. Also, gel permeation chromatography, dynamic light scattering and native acidic gels showed that MSMEG\_3955 exists as a homotrimer. Furthermore, the presence of NADPH dependent FMN oxidoreductase and homotrimeric existence could be an alternative function of the protein to help the bacteria survive in dormant state or may be involved in other biochemical pathways.

**Conclusion:** MSMEG\_3955 is a FMN bound flavoprotein, which exists as a trimer under in vitro conditions. There is no disulphide linkages in between the three protomers of the homotrimer MSMEG\_3955. It has a NADPH dependent FMN oxidoreductase activity.

## Introduction

*Mtb* can reside within host tissues for years or even decades without causing the disease, the mechanism known as latency or dormancy [1]. The sensitivity of dormant bacterium against antibiotics is far less than the active bacterium, which requires long course of antibiotics [2]. One third of the population has latent TB, placing them

\*Correspondence: akesav@gmail.com

<sup>1</sup> Department of Molecular Biology and Biochemistry, Guru Nanak Dev University, Amritsar, Punjab 143005, India

Full list of author information is available at the end of the article



© The Author(s) 2021. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

on risk of active TB in future [3]. Thus to fight against TB, understanding the genes involved in latency are of critical interest. *Mtb* encodes a set of 50 genes called the dormancy survival regulator (DosR) regulon that are expressed during stress, hypoxia, NO stress and nutrient starvation and helps the bacteria to survive in latency [4–6]. Genes within the DosR regulon are involved in multiple processes i.e., central metabolism, energy generation and gene regulation. However, majority of them are not functionally characterized. The gene *Rv3131* is a part of DosR regulon and it belongs to the *acr co-regulated gene (AcrG)*. *AcrG* gene is known to be essential for growth and virulence in vivo [6]. *Rv3131* gene product is also required for mycobacterial growth which has been studied using high density mutagenesis [4, 7–9]. It is considered to be a putative nitroreductase flavoprotein. Flavins are extremely versatile cofactors that undergo redox reaction by accepting either one electron or two electrons, alternating between the oxidized and reduced states. Flavoenzymes are the class of oxidizing enzymes containing an electron acceptor FMN or Flavin adenine dinucleotide (FAD). These electron carriers accommodate 90% of apo-protein through non-covalent interactions [10]. Flavin reductases are the enzymes which reduce free flavins (FMN, FAD/riboflavin) through reduced pyridine nucleotide, NADH or NADPH [11]. Dihydroflavins produced thus participate in a variety of biological prokaryotic redox reactions include hydroxylation of phenolic compounds, epoxidation of styrene, reduction and mobilization of iron from ferric complexes of siderophores etc. [12–14]. It is believed that the putative nitroreductase *Rv3131* involve in detoxification of nitrogen-containing bi-products in the host, but the structure and functional role of *Rv3131* in mycobacterial species is still unknown [3].

In the present study, the *Msmeg* is used as model organism to study the gene *MSMEG\_3955* which is the homologue of gene *Rv3131* from *Mtb*. The gene *Rv3131* shares 63% identity to the *MSMEG\_3955*. The characteristics of slow growing *Mtb* and fast growing *Msmeg* are quite similar, thus the use of the non-pathogenic and fast growing species as a model organism to study the virulence of *Mtb* is important [15]. The gene *MSMEG\_3955* was cloned and expressed to get its structural and functional insights.

## Experimental procedures

### Strains, reagents, plasmids, media and growth conditions

*M. smegmatis mc<sup>2</sup> 155* was grown in Middlebrook 7H9 broth (HiMedia), containing 0.5% (v/v) glycerol and 0.05% Tween 80 and appropriate amount of Oleic acid dextrose catalase (OADC) (HiMedia) for 48 h aerobically at 37°C. *E.coli* (DH5α) was used for DNA amplification

and *E.coli* (BL21) (DE3) for overexpression of recombinant proteins. Both the strains were grown in LB broth (luria-bertani) (HiMedia) for 16 h, keeping the media in shaker with constant agitation at 180 rpm at 37°C. The vector pET-28a was used for the over-expression of the recombinant protein *MSMEG\_3955*. Kanamycin 50 µg/mL was routinely used in LB media during the growth of strains having vector pET-28a. Isopropyl β-D-1-thiogalactopyranoside (IPTG) (HiMedia) was used to induce the recombinant protein *MSMEG\_3955*. Nickle nitrilotriacetate (Ni-NTA) agarose beads (Qiagen) were used for the purification of the expressed protein. Polymer chain reaction (PCR) products were gel purified through GEL/PCR DNA Fragments Extraction kit (IBI SCIENTIFIC, USA), and plasmid were isolated through High-Speed Plasmid Mini kit (IBI SCIENTIFIC, USA). Amicon Centrifugal Filter Unit (EMD Millipore) were used to concentrate and dialysed the proteins. Nitrofurantoin was purchased from MP Biomedicals. FMN, NADH and NADPH were from HiMedia. Urea, Sodium dodecyl sulfate (SDS) and Trichloroacetic acid (TCA) were obtained from HiMedia.

### Cloning and expression of recombinant protein

Genomic DNA was isolated from *Msmeg* using the hexadecyltrimethyl ammonium bromide (CTAB) method [16]. Spectrophotometrical analysis was performed to quantify the isolated genomic DNA. For the amplification of *MSMEG\_3955*, forward primer and reverse primer were designed. The primer set used; forward, 5'CGCGGATCCATGAATACGCACTTCCCGGAT3', reverse, 5'CCCAAGCTTCTACTCCGATTCCGGCTC GAA3' (underline are Bam HI and Hind III restriction sites respectively). PCR was done with Phusion High Fidelity DNA Polymerase (NEB) in the Agilent thermocycler (USA) using following conditions for amplification; hot start at 98°C for 3 min, denaturation at 98°C for 30 s, annealing at 66°C for 30 s, elongation at 72°C for 45 s (35 cycles), followed by final elongation at 72°C for 10 min. The amplified gene *MSMEG\_3955* was cloned in the expression vector pET-28a and transformed in *E. coli* (DH5α) for amplification of recombinant clones. The positive clone was confirmed with restriction double digestion. The confirmed clone was further transformed in *E.coli* (BL21) (DE3) for the over-expression of recombinant protein *MSMEG\_3955*. For the over-expression, the cells were grown in 3 mL of LB broth for overnight (16 h incubation) at 37°C and subcultured to the fresh 500 mL LB broth. The subculture allowed to grow at 37°C till the OD<sub>600</sub> reach 0.4–0.6. For the induction of recombinant protein 1 mM IPTG was added to the culture after required OD<sub>600</sub>, keeping the culture at 25°C for 24 h. All the cultures were

kept in the shaker at optimum temperatures with constant shaking at 180 rpm. The cells were harvested by centrifugation and suspended in lysis buffer containing 50 mM Tris-HCl (pH-8.0), 500 mM NaCl, 100 µg/mL Lysozyme, 1 mM PMSF and 10 mM Imidazole. One gram pellet of cells were suspended in 1 mL of lysis buffer containing 50 mM Tris-HCl (pH-8.0), 500 mM NaCl, 10 mM imidazole, 100 µg/mL lysozyme and 0.02% NaN<sub>3</sub>. The cells were lysed using SONICS vibra cell sonicator at 37% amplitude with 5 s ON and 5 s OFF for 1–2 min (for each sample) or until the sample got clear. The samples were kept in ice during the sonication to avoid denaturation of protein during extraction as the sample was getting hot when sonication proceeded. The sonicated cell lysates was centrifuged at max speed for 30 min. Supernatant containing recombinant protein MSMEG\_3955 was collected in fresh tubes.

#### Purification of recombinant protein

For the purification of recombinant protein, Ni<sup>2+</sup>-NTA agarose Quick-Star (Qiagen) was used. One milliliter of Ni-NTA slurry (0.5 mL bed volume) was added in the column and washed thrice with lysis buffer. After washing, 4 mL of clear cell lysate was added to the slurry in the column and mixed gently by shaking at 200 rpm at 4 °C for 1 h to get the bound protein. The flow-through was collected by removing bottom cap of the column. The bound protein was washed thrice with 5 bed volume (2.5 mL) wash buffer containing 50 mM Tris-HCl (pH-8.0), 300 mM NaCl, 40 mM imidazole and 0.02% NaN<sub>3</sub>. The washed fractions were collected for 10% SDS-PAGE analysis. The purified protein MSMEG\_3955 was eluted 5 times with 0.5 mL of elution buffer containing 50 mM Tris-HCl (pH-8.0), 300 mM NaCl, 300 mM imidazole and 0.02% NaN<sub>3</sub>. The eluted protein was dialysed and concentrated using Amicon Centrifugal Filter Unit (EMD Millipore). Quantification of protein was done by Bradford method using bovine serum albumin (BSA) as standard [17].

#### Estimation of reduced and non-reduced protein

SDS-PAGE was carried out to analyse the size of reduced and non-reduced protein MSMEG\_3955. To determine the reduced form, the protein was treated with reducing buffer containing 12% SDS, 30% glycerol, 0.05% coomassie blue dye, 150 mM Tris-HCl (pH-7.0) and 6% β-mercaptoethanol and run on 12% acidic SDS-PAGE. To determine the non-reduced form of recombinant protein, it was treated with non-reducing buffer containing 12% SDS, 30% glycerol, 0.05% coomassie blue dye, 150 mM Tris-HCl (pH-7.0) and run on 12% acidic SDS-PAGE.

#### Estimation of oligomeric state of protein

##### NATIVE-PAGE analysis

The oligomeric state of purified native protein was estimated by NATIVE-PAGE analysis. In this, 10 µg of the protein sample was loaded directly to the 12% NATIVE-PAGE, without any treatment. The loading buffer contains a coloured dye bromophenol blue and a density agent glycerol only.

##### Size exclusion chromatography

Oligomeric state of recombinant protein MSMEG\_3955 was confirmed by gel permeation chromatography (GPC) using NGC Chromatography system on Enrich SEC 70 10/300 column (Bio-Rad) pre-equilibrated with wash buffer containing 50 mM NaH<sub>2</sub>PO<sub>4</sub>, 150 mM NaCl, 0.02% NaN<sub>3</sub> (pH-8.0). The retention volume of MSMEG\_3955 was compared with the standards BSA (132 kDa) and Lysozyme (14 kDa). The flow rate was maintained at 0.5 mL/min.

##### Dynamic light scattering (DLS)

Oligomeric state and purity of protein MSMEG\_3955 was determined by peak analysis and overall polar dispersity using DLS. Purified 20 µM protein sample was analysed in disposable sizing cuvette using a path length of 1 cm at 20 °C in Malvern Zetasizer instrument.

#### Homology modelling

The homology model of MSMEG\_3955 was generated to explore its structural similarity with the conformation of MSMEG\_3955 from CD. MSMEG\_3955 shows 35% identity to protein MSMEG\_5246, therefore it was used as a template to generate the 3 dimensional structure of MSMEG\_3955. The FASTA format of amino acid sequence of protein was retrieved from Mycobrowser (<http://mycobrowser.epfl.ch>) for structure prediction. BLAST2 was used to obtain homologous entries of protein MSMEG\_3955 from Protein Data Bank in NCBI for template search. Top hit template MSMEG\_5246 PDB ID: 2ymv was used for homology modelling using SWISS-MODEL [18] a fully automated protein structure homology-modelling server (<https://swissmodel.expasy.org/interactive>). Generated homology model of protein MSMEG\_3955 was monomer as template protein MSMEG\_5246 used was monomer [19] PyMOL program was used to visualize the generated protein models [20]. The quality of generated protein was validated with quality factors such as RAMACHANDRAN PLOT as PROCHECK [21] (<http://www.ebi.ac.uk/thornton-srv/software/PROCHECK/>) to evaluate backbone and side chains of protein, Verify3D [22] ([http://services.mbi.ucla.edu/Verify\\_3D/](http://services.mbi.ucla.edu/Verify_3D/)) was performed to measure the compatibility

of the protein's structure with its respective sequence and ERRAT [23] (<http://services.mbi.ucla.edu/ERRAT/>) was performed to find out overall reliability of generated 3D structure of protein.

#### Prediction of secondary structure

The secondary structure of protein MSMEG\_3955 was predicted by CD spectroscopy. The spectra provides an index of structure. The scanning measurements were made from 200 nm to 260 nm wavelength, with JASCO J-1500 Circular Dichroism spectrometer. MSMEG\_3955 protein concentration used was 8  $\mu$ M in 10 mM HEPES using a 0.1 cm path length quartz cuvette at 25 °C. The solvent spectrum was subtracted from the sample spectrum.

#### Determination of cofactor

##### Thin layer chromatography (TLC)

TLC was performed to analyse the cofactor bound to the recombinant protein. The purified protein was denatured at 70 °C for 20 min. The samples were centrifuged at maximum speed for 20 min. The supernatant was spotted on thin layer of silica gel along with 5  $\mu$ M of FMN, Riboflavin and FAD as control. The mobile phase used was 1-butanol-acetic acid-water (4:1:1, by volume) and the developed plate was visualized under UV light [24, 25].

##### UV-visible spectroscopy

The spectrophotometric confirmation of co-factor bound to the protein MSMEG\_3955 was analysed using SDS treatment of holoprotein (Protein + cofactor) [26, 27]. UV-Visible spectra were recorded between 400 to 500 nm. The absorbance spectrum of the 100  $\mu$ M protein in 10 mM Tris-HCl buffer, pH-8.0 was recorded. Twenty microliter fresh 10% SDS solution was added to the cuvette containing protein. After mixing, spectra were recorded every 5 min until no further changes were observed.

#### Site directed mutagenesis (SDM)

Mutant was generated using PCR-based SDM described by Hoa et.al [28]. There are three cysteine residues present at three different sites 63, 70, 263 in the amino acid sequence of MSMEG\_3955. Plasmid pET-28a containing required gene *MSMEG\_3955* was used as a template for PCR. The amino acid cysteine (TGC) at position 63 was replaced with alanine (GCC). Phusion High Fidelity DNA Polymerase (NEB) was used in the Agilent thermocycler (USA). Two step PCR was performed using set of primer forward, 5'GATCTGCTCCTCAGCGCCGGCGCGGCGCTGCAC3' and reverse, 5'GTGCAGCGCCGCGCCGCCGCTGAGGAGCAGATC3'. The following conditions were followed for amplification; hot start at 98 °C

for 1 min, denaturation at 98 °C for 15 s, annealing at 72 °C for 3 min, elongation at 72 °C for 10 min (30 cycles). The amplified cDNA was treated with DpnI to digest methylated parental DNA and the newly synthesised cDNA was transformed in *E. coli* (DH5 $\alpha$ ) for amplification of the product of mutagenesis. The isolated positive clones were confirmed with sequencing and then transformed in *E. coli* (BL21) (DE3) for its overexpression.

#### FMN reductase assay

##### UV-visible spectroscopy

Flavin reductases activities were carried out as described by Rakhi et al. with modification [29]. In the modified protocol, 2 mL of reaction volume with increased concentration of electron donor in Tris-HCl buffer was used. Two reaction mixtures were setup; for both the reactions, NADPH was used as an electron donor. For each reaction, the oxidation of electron donor was observed by decrease in the absorbance at 340 nm at 25 °C under aerobic conditions. In one experiment, FMN (50  $\mu$ M), NADPH (200  $\mu$ M) with 1.6  $\mu$ M of enzyme/protein MSMEG\_3955 in 50 mM Tris-HCl pH-7.6 in 2 mL reaction volume was used. NADPH was added just before placing the tube in spectrophotometer (Systronics double beam spectrophotometer 2202). In another experiment, NADPH (100  $\mu$ M) with 1.6  $\mu$ M of enzyme/protein MSMEG\_3955 in 50 mM Tris-HCl pH-7.6 in 1 mL reaction volume was used keeping the reaction at 25 °C for 48 h. A control experiment was performed without enzyme. Both the experiments were performed in triplicates.

##### <sup>1</sup>H NMR

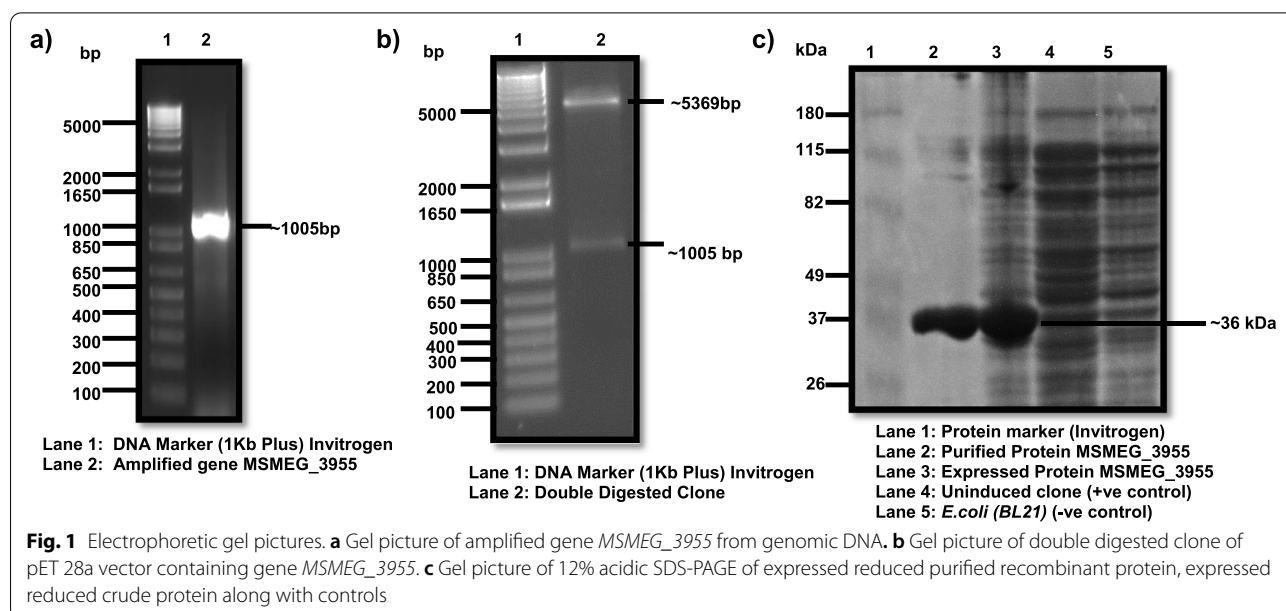
The NADPH dependent FMN reductase activity of protein MSMEG\_3955 was confirmed with <sup>1</sup>H NMR spectroscopy. The 2 mL volume of reaction mixture was set that contain FMN (500  $\mu$ M), NADPH (15 mM) with 30  $\mu$ M of enzyme/protein MSMEG\_3955 as catalyst in 50 mM Tris-HCl pH-7.6. The <sup>1</sup>H NMR spectrum of FMN after 2 h incubation at 25 °C was recorded and compared with its oxidized form.

## Results

### Cloning, expression of protein

The gene *MSMEG\_3955* shares 63% identity to the gene *Rv3131* in *Mtb*. (Both the sequences were aligned using BLASTP as shown in Fig. S1). The gene specific primers for PCR resulted ~1005 bp amplicon (Fig. 1a). The gel purified PCR product was cloned into the expression vector pET28a and confirmed by restriction double digestion (Fig. 1b) and DNA sequencing. The recombinant reduced protein was identified by SDS-PAGE with an expected protein size of ~36 kDa (Fig. 1c). The





non-reduced protein gave the similar size of band in 12% acidic SDS-PAGE (Fig. 2b).

#### Protein MSMEG\_3955 is homotrimeric

The oligomeric state of protein is homotrimeric, since on NATIVE- PAGE the band of 10 µg protein corresponding to ~ 115 kDa was seen (Fig. 3a). This was also confirmed by performing size exclusion chromatography of the given protein. In order to compare NATIVE size of protein obtained from acidic NATIVE-PAGE electrophoresis, size exclusion chromatography was performed (Fig. 3c). Consideration and comparison of the size exclusion elution profile of MSMEG\_3955 and the results of SDS-PAGE and NATIVE-PAGE suggests that the protein is an oligomer and it exists as homotrimer in the solution. Dynamic light scattering measurements exhibits sharp monodisperse peak of 20 µM protein centred around 8.7 diameter (nm) with molecular weight calculated at  $105.6 \pm 12.2$  kDa [30–32], which also suggests the trimeric existence of MSMEG\_3955 (Fig. 3b).

#### Site directed mutagenesis

One of the cysteine residue was mutated at position 63 with alanine to see the change in oligomeric structure. Sanger sequencing of amplified mutated clone confirmed replacement of cysteine to alanine in the gene of *MSMEG\_3955* at amino acid position 63. Nucleotide BLAST was performed between mutated gene sequence and wild type gene sequence of *MSMEG\_3955* to confirm the mutation. The mutated protein was overexpressed and purified by Ni-NTA method. To determine

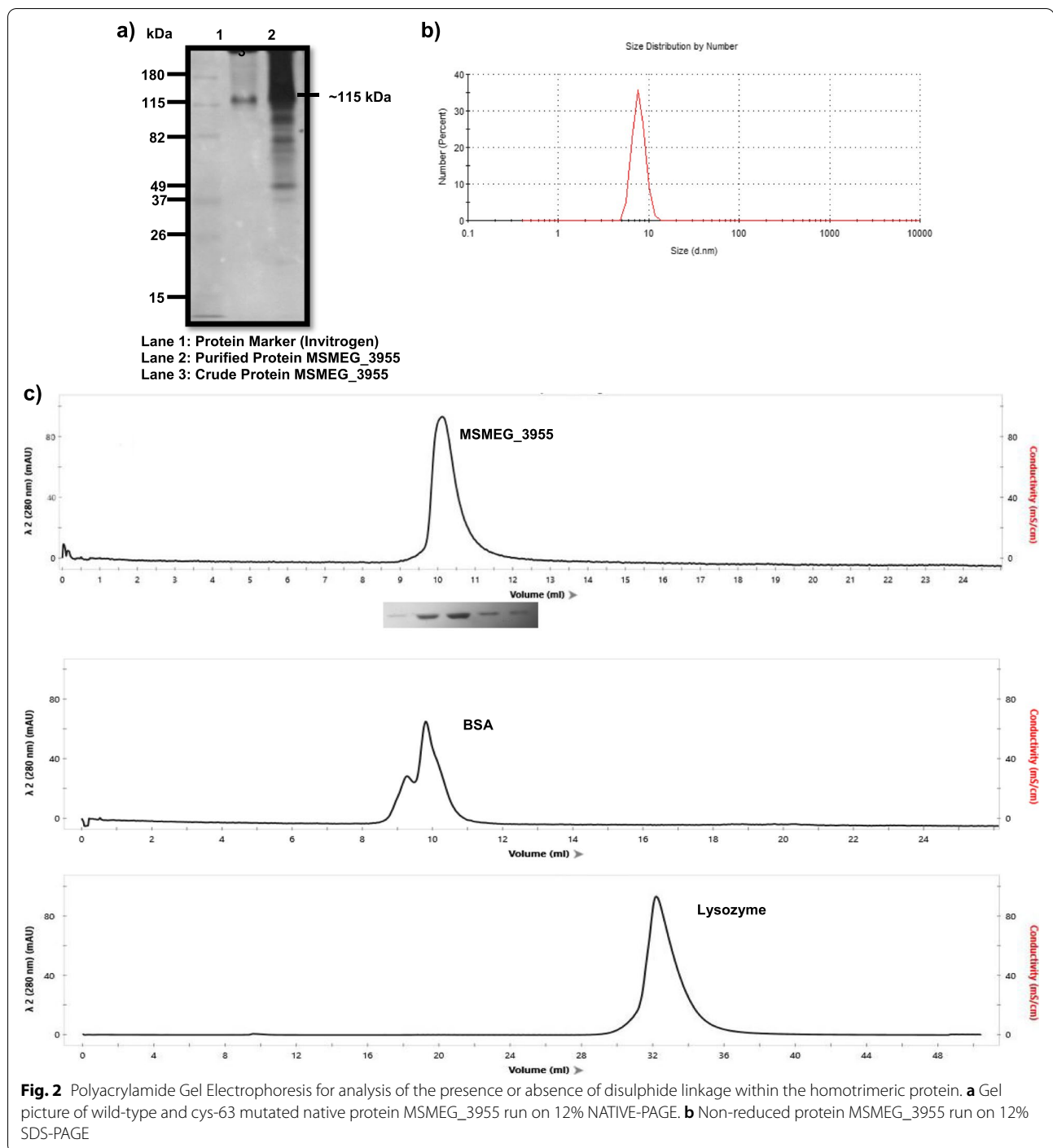
any oligomeric alterations in the mutated native protein, 20 µg of mutated protein and 20 µg of wild type protein were run on NATIVE-PAGE. On SDS-PAGE (data not shown) and NATIVE-PAGE, the wild-type protein and the mutant protein had identical sizes (Fig. 2a).

#### Homology modelling and CD analysis

Three dimensional structure of monomeric MSMEG\_3955 was obtained from homology modelling. The generated model MSMEG\_3955 composed of 40% α-helix and 22% β-pleated sheets analysed using STRIDE programme [33]. The quality of monomeric protein was validated with PROCHECK, Verify 3D and ERRAT to assess their acceptability and were found suitable for structural analysis as seen in Fig. S4(a), S4(b), S4(c) resp. CD spectroscopy was performed to validate experimentally the secondary structural conformation of protein MSMEG\_3955 predicted by homology modelling. Analysis of CD spectra suggested that MSMEG\_3955 is composed of 39.3% α-helix, 24.9% β-pleated sheets, 19.1% turn and remaining part assumed to be randomly coiled (Fig. 4a). The CD spectra was analysed as described by Yang's et.al [34]. These results are in close agreement with MSMEG\_3955 homology model prediction for α-helix and β-pleated sheets contents (Fig. 4b).

#### Cofactor FMN is bound to MSMEG\_3955

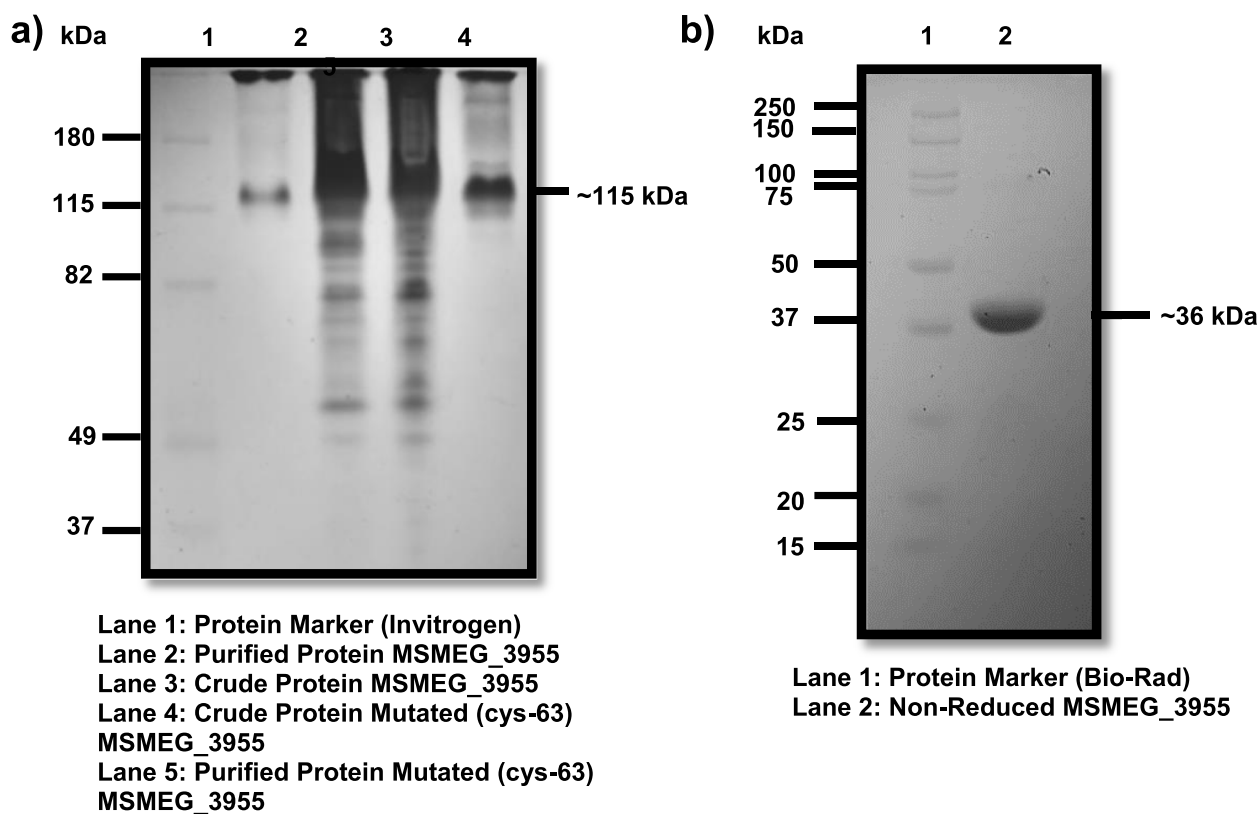
The absorption spectra recorded using UV-Visible spectroscopy gave the peak at 446 nm after 10 min of SDS treatment to the protein MSMEG\_3955. Since the FMN absorbs at 447 nm, UV-visible spectroscopy identified



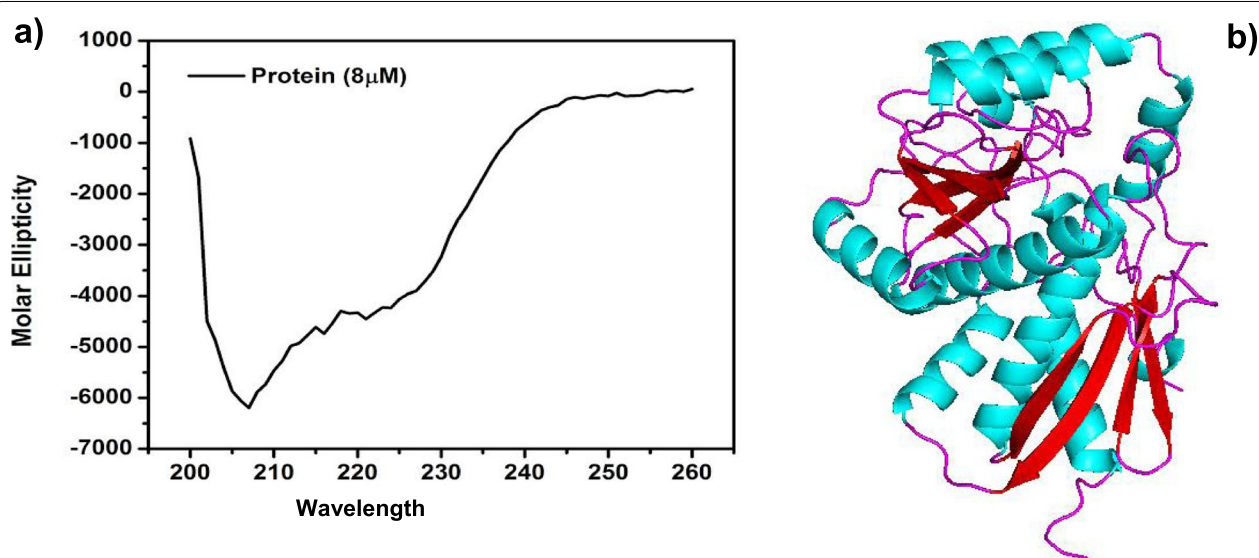
the cofactor bound to the protein MSMEG\_3955 is FMN (Fig. 5a). The cofactor extracted by thermal denaturation of protein MSMEG\_3955 co-migrated with FMN (control) on the silica gel, when kept in the mobile phase confirmed the cofactor FMN is bound to the protein MSMEG\_3955. The chromatogram was seen under UV torch (Fig. 5b).

#### MSMEG\_3955 is a FMN reductase

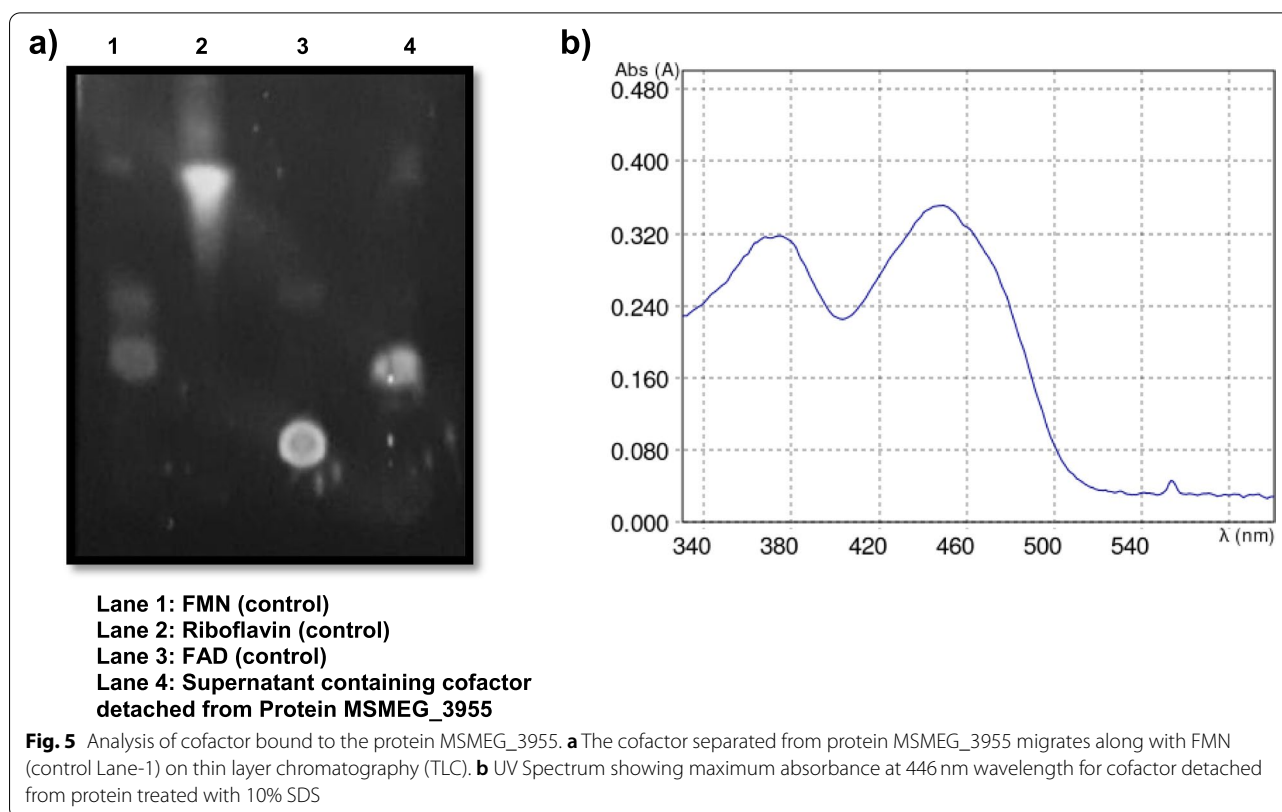
The MSMEG\_3955 protein/enzyme showed oxidative NADPH dependent FMN reductase activity as oxidation of NADPH was seen. The decrease in NADPH absorbance seen in first experiment showed the enzyme MSMEG\_3955 is reducing FMN provided. Readings were taken after addition of NADPH every 2 min interval for



**Fig. 3** Oligomeric state analysis of protein MSMEG\_3955. **a** Gel picture of native protein run on 10% NATIVE-PAGE. **b** Percentage particle size distribution graph of protein showing diameter of molecule ~7.8 nm, analysed by DLS. **c** Gel permeation chromatography showing elution profile of protein (~115 kDa) with BSA (132 kDa) and Lysozyme (14 kDa) as standards



**Fig. 4** Secondary structure analysis of MSMEG\_3955. **a** Graphical representation of Circular Dichroism spectrum of MSMEG\_3955 protein (8 μM) in 10 mM HEPES. **b** Modelled structure of MSMEG\_3955, Helices are colored blue, sheets are red and coils are purple



10 min Fig. 6a). The absorbance of NADPH was zero which shows the whole FMN bound to the protein was reduced. Fig. S5 is a plot of absorbance vs time between 0 and 600 s for a protein in the first experiment. By Beer's law, the absorbance of the solution is directly proportional to the concentration of the protein in the solution, so observing the absorbance as a function of time is basically the same as observing the concentration versus time.

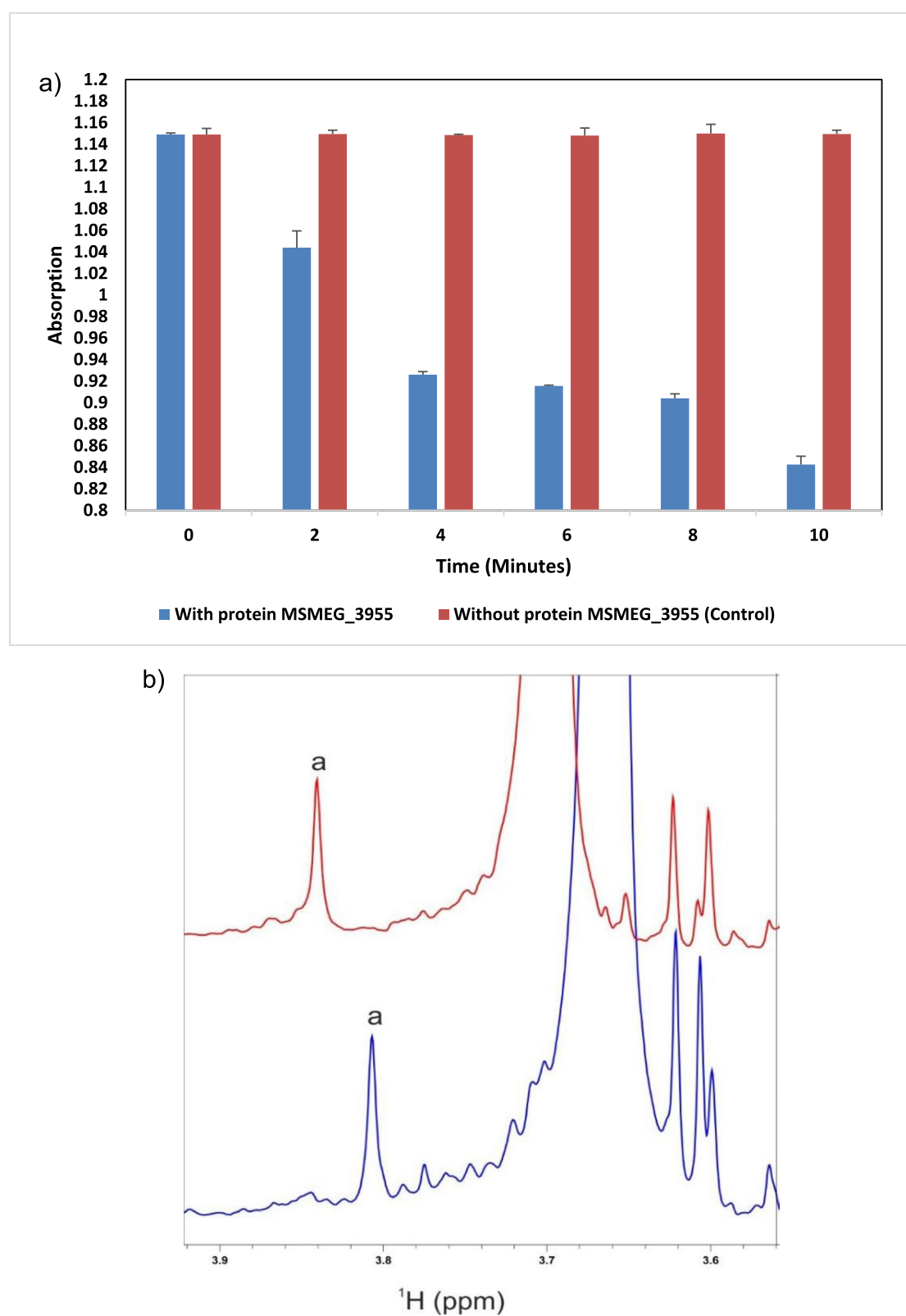
The curve represent data fits a first-order reaction and straight fitting of plot  $\ln[A]$  versus time will yield rate constant ( $k$ ) =  $-2 \times 10^{-3} s^{-1}$ . In the second set of experiment, NADPH absorbance was decreased as FMN bound to the enzyme itself was reduced. In the  $^1H$  NMR spectrum of FMN all the peaks were observed as reported in BMRB data for FMN (Entry number- bmse000257) shown in Fig. S2. The Flavin ring proton exhibit upfield chemical shift on reduction as reported earlier [35] labelled as "a" (Fig. 6b) and full spectra were shown in Fig. S3.

## Discussion

The latency mechanism is found to be associated with the dormancy survival regulator transcription factors which together constitute the DosR regulon [4, 36]. There are multiple genes present in the DosR regulon

but the structure and function of most of their encoded proteins are unknown. Also, the DosR regulon associated proteins such as Acg family has evolved from the nitroreductase homodimer by gene duplication and fusion and then loss of one of the two active sites. This has resulted in a monomeric protein with a single active site but with an overall fold resembling the nitroreductase homodimer [19]. Therefore, it is believed that most of the DosR regulon encoded proteins may have the potential nitroreductase activity. In the present study we characterised a novel DosR regulon gene MSMEG\_3955 which was of *Msmeg* origin. MSMEG\_3955 was found to be a homotrimeric and showed FMN reductase activity. Flavin biosynthesis is important for the action of *Mtb* as it lacks the flavin uptake mechanism [37]. Riboflavin biosynthesis enzymes have been found to be conserved in *Mycobacterium leprae*, regarded as a minimal mycobacterial genome [38, 39]. However, regulation of flavin synthesis in *Mtb* is poorly explored. More than 3% of *Mtb* genes are flavoproteins and are particularly rich in FAD-dependent acyl-CoA dehydrogenases required for lipid degradation [40, 41]. The present study provides the first impression of FMN related activity of a DosR regulon encoded protein. The results so obtained can add to our understanding of the physiological state of the *Mtb* during latency and could be of utmost





interest in the fight against tuberculosis. In the present study it is observed that the reduced and non-reduced protein gave the similar size of bands in SDS-PAGE of ~36 kDa. On NATIVE-PAGE the band of recombinant protein corresponds to ~115 kDa. The elution profile of protein by size exclusion chromatography and DLS also verified the similar results. From these results we conclude that the protein of our interest is an oligomer and it exists in the homotrimeric form in the solution. Since  $\beta$ -mercaptoethanol breaks the disulphide linkage and reduce the protein to its monomeric form, thus the protein with its treatment got linearize and existed in its monomeric form and gave band of ~36 kDa but the conversion of trimeric form into monomeric form on SDS-PAGE run even in the absence of a reducing agent suggests the absence of disulphide bonds within the three protomers. Which was further confirmed with mutation of the cysteine residue at position 63 with alanine to see the change in oligomeric structure. Mutation did not give any significant change in the trimeric structure of the protein. Which suggest no role of cys-63 in structure stabilization by disulphide bond formation. Although the presence of disulphide bonds in between cysteine residues of protomers was expected to play role in stabilization of the trimeric structure.

The absorption spectra analysis and TLC concluded that the cofactor FMN is bound to the protein MSMEG\_3955. To quantify the number of FMN molecules attached to the protein, apoprotein was obtained by deflavination using different methods i.e. TCA and SDS treatment. Firstly, with the TCA treatment the flavin got detached from the protein and the apoprotein forms insoluble aggregates. We also tried to dissolve the apoprotein using 6 M and 8 M urea but it did not work. Further, the treatment of protein at different SDS concentrations of (0.05–3.0%) showed that deflavination occurred and the apoprotein got degraded at minimum concentration of 0.05% of SDS which showed that the FMN was bound to the protein non-covalently. These results concluded that protein was getting degraded after deflavination and requires FMN for its stability. Similarly, in a study conducted on protein WrbA, an oligomeric flavodoxin-like protein that binds one molecule of flavin mononucleotide (FMN) per monomer was found out to form tetramers, which are more thermoresistant than dimers or monomers, suggesting that multimerization underlies the FMN effect on WrbA thermostability [42]. Furthermore it was also showed a clear shift towards a higher molecular mass of Wild-type dodecin from *S. davaonensis* in the presence of FMN which indicate that the FMN promoted the formation of multimers (probably trimers) [43]. The protein AtHal3 is a homotrimer and three FMN

molecules are bound to it, one FMN present between the two adjacent protomers [44]. As FMN bound to the protein at lower temperature [42] and protein MSMEG\_3955 also expresses at 25 °C discussed in material and methods. Thus, the existence of trimeric structure of MSMEG\_3955 found in the present study could be that there may be three FMN bound to the protein at low temperature, one FMN in between the two adjacent protomers and might provide this protein a homotrimeric stable structure since, no disulphide linkage found within the three protomers to form a stable trimeric structure.

Rv3131 from *Mtb* a close homologue to MSMEG\_3955 is considered to be a putative NADPH nitroreductase. The sequence homology suggests that the gene MSMEG\_3955 belongs to nitroreductase family [3]. However, the present study did not show significant nitroreductase activity by purified recombinant protein MSMEG\_3955, using nitrofurantoin as substrate and both NADH and NADPH individually as an electron donor. There was no nitroreductase activity compared to the controls (data not shown). It was earlier reported that some flavoenzymes does not show nitroreductase activity [45]. The protein MSMEG\_5246 has unknown function and documented as putative nitroreductase, which also does not show nitroreductase activity [19]. The reason for no nitroreductase activity of protein MSMEG\_5246 may be assumed for the presence of a lid on the groove of protein's structure as described by the crystallographic structure as there is no access to the active site. Although, the protein MSMEG\_5246 also belong to Acg family [6] and exhibits about 37% identity with MSMEG\_3955 and gives top hit model for homologue modelling to generate the hypothetical structure of MSMEG\_3955 but do not show FMN reductase activity. However, the cofactor FMN was reduced aerobically by the protein MSMEG\_3955 showing FMN reductase activity as discussed in the present study by UV-Visible spectroscopy and  $^1\text{H}$  NMR spectra.

Thus, we are considering presence of FMN reductase activity of the protein MSMEG\_3955 is the evolutionary existence of its homotrimeric native form. This structural characterization of the protein MSMEG\_3955 is described for the first time in the present study. Since the protein is having FMN reductase activity and because of its existence as trimeric stable conformation, its enzymatic activity can be involved in biochemical pathways which ultimately help in adaptation of bacteria to survive in the dormancy and in the latent infection. The antibiotic treatment against TB requires a number of drugs depending upon its occurrence such as multiple drug resistant (MDR) and extremely drug resistant (XDR). These drugs work against different components of TB. Elucidating

a distinct activity of *Mtb* proteins can help in better immune surveillance and therefore can be used in the future control of TB.

## Conclusion

Our study conclude, the MSMEG 3955 protein is a flavo-protein that is bound by the FMN. In its natural state, it exists as a homotrimer. There are no disulphide linkages between the three protomers of homotrimer MSMEG 3955. This protein is a NADPH-dependent oxidoreductase that uses FMN as a cofactor.

## Abbreviations

Acr: Acr co-regulated gene; BLAST: Basic Local Alignment Search Tool; BSA: Bovine serum albumin; CD: Circular Dichroism; cDNA: Complementary DNA; CTAB: Hexadecyltrimethyl ammonium bromide; DLS: Dynamic Light Scattering; DosR: Dormancy survival regulator; FMN: Flavin mononucleotide; GPC: Gel permeation chromatography; IPTG: Isopropyl  $\beta$ -D-1-thiogalactopyranoside; LB: Luria bertani; MDR: Multiple drug resistant; NADH: Nicotinamide adenine dinucleotide (Reduced form); NADPH: Nicotinamide adenine dinucleotide phosphate (Reduced form); NCBI: National Center for Biotechnology Information; Ni-NTA: Nickel-Nitrilotriacetic Acid; NMR: Nuclear magnetic resonance; OADC: Oleic acid dextrose catalase; PAGE: Polyacrylamide Gel Electrophoresis; PCR: Polymerase Chain Reaction; PDB: Protein Data Bank; PMSF: Phenylmethylsulfonyl fluoride; SDM: Site directed mutagenesis; SDS: Sodium dodecyl sulfate; SEC: Size exclusion chromatography; TCA: Trichloroacetic acid; TLC: Thin layer chromatography; UV: Ultraviolet; XDR: Extremely drug resistant.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12866-021-02330-y>.

**Additional file 1 : Figure S1.** Sequence Alignment of *MSMEG\_3955* from *Mycobacterium smegmatis* MC2-155 with *Rv3131* from *Mycobacterium tuberculosis* H37Rv showing Identities: 202/323(63%), Positives: 249/323(77%), Gaps: 0/323(0%). **Figure S2.** The <sup>1</sup>H NMR spectra of FMN oxidised form recorded at 500 MHz, pH-7.6. **Figure S3.** The <sup>1</sup>H NMR spectra of FMN in oxidized (red) and reduced (blue) form recorded at 500 MHz, pH-7.6. **Figure S4.** Validation of generated monomeric MSMEG\_3955. a) PROCHECK gave results as 89.9% in most favoured regions, 8.3% in additional allowed regions, 1.4% in generously allowed regions and 0.4% in disallowed regions. b) Verify 3D gave 96.68% showing the generated trimeric protein model of good quality. c) ERRAT gave overall quality factor 90.7 expressed as the percentage of the protein for which the calculated error value falls below 95% rejection limit. **Figure S5.** Absorbance-time curve for the reaction of Protein + NADPH + FMN.

## Acknowledgements

We are thankful to Ashutosh Vashist for his valuable suggestions in manuscript writing. The Department of Science and Technology, India WOS-A funding to SMT.

## Authors' contributions

NK: Experimental design, experimentation, data analysis and wrote the main manuscript. SMT: Data analysis. RJ: Data analysis. AK\*: Conception of idea, experimental design, data analysis and manuscript writing. All authors reviewed the manuscript. The author(s) read and approved the final manuscript.

## Funding

NK was funded by Rajiv Gandhi National Fellowship (UGC- Govt. of India).

## Availability of data and materials

The dataset used in the current study was submitted to NCBI (GenBank) and the accession number is MZ198152.

## Declarations

### Ethics approval and consent to participate

Not applicable.

### Consent for publication

Not applicable.

### Competing interests

The authors declare that they have no competing interests.

## Author details

<sup>1</sup>Department of Molecular Biology and Biochemistry, Guru Nanak Dev University, Amritsar, Punjab 143005, India. <sup>2</sup>Department of Botanical and Environmental Sciences, Guru Nanak Dev University, Amritsar, Punjab 143005, India. <sup>3</sup>Current Address: Department of Biotechnology & Microbiology, Kannur University, Dr. E.K. Janaki Ammal Campus, Palayad, Thalassery, Kannur, Kerala, India.

Received: 6 January 2021 Accepted: 27 September 2021

Published online: 19 November 2021

## References

- Wayne LG. Dormancy of mycobacterium tuberculosis and latency of disease. *Eur J Clin Microbiol Infect Dis*. 1994;13:908–14. <https://doi.org/10.1007/BF02111491>.
- Wayne LG, Sramek HA. Metronidazole is bactericidal to dormant cells of mycobacterium tuberculosis. *Antimicrob Agents Chemother*. 1994;38:2054–8. <https://doi.org/10.1128/AAC.38.9.2054>.
- Purkayastha A, McCue LA, McDonough KA. Identification of a mycobacterium tuberculosis putative classical nitroreductase gene whose expression is coregulated with that of the *acr* gene within macrophages, in standing versus shaking cultures, and under low oxygen conditions. *Infect Immun*. 2002;70:1518–29. <https://doi.org/10.1128/IAI.70.3.1518-1529.2002>.
- Park HD, Guinn KM, Harrell MI, Liao R, Voskuil MI, Tompa M, et al. *Rv3133c/dosR* is a transcription factor that mediates the hypoxic response of mycobacterium tuberculosis. *Mol Microbiol*. 2003;48:833–43. <https://doi.org/10.1046/j.1365-2958.2003.03474.x>.
- Voskuil MI, Schnappinger D, Visconti KC, Harrell MI, Dolganov GM, Sherman DR, et al. Inhibition of respiration by nitric oxide induces a mycobacterium tuberculosis dormancy program. *J Exp Med*. 2003;198:705–13.
- Hu Y, Coates ARM. Mycobacterium tuberculosis *acr* gene is required for growth and virulence in vivo. *PLoS One*. 2011;6:20958. <https://doi.org/10.1371/journal.pone.0020958>.
- Sassetti CM, Boyd DH, Rubin EJ. Genes required for mycobacterial growth defined by high density mutagenesis. *Mol Microbiol*. 2003;48:77–84. <https://doi.org/10.1046/j.1365-2958.2003.03425.x>.
- Tavares AFN, Nobre LS, Melo AMP, Saraiva LM. A novel nitroreductase of *Staphylococcus aureus* with S-nitrosoglutathione reductase activity. *J Bacteriol*. 2009;191:3403–6.
- Chauhan S, Tyagi JS. Powerful induction of divergent *tgsl-Rv3131* genes in mycobacterium tuberculosis is mediated by DevR interaction with a high-affinity site and an adjacent cryptic low-affinity site. *J Bacteriol*. 2009;191:6075–81.
- Hefti MH, Vervoort J, Van Berkel WJH. Dechlorination and reconstitution of flavoproteins: tackling fold and function. *Eur J Biochem*. 2003;270:4227–42.
- Nivière V, Vanoni MA, Zanetti G, Fontecave M. Reaction of the NAD(P) H:flavin oxidoreductase from *Escherichia coli* with NADPH and riboflavin: identification of intermediates. *Biochemistry*. 1998;37:11879–87.
- Galán B, Díaz E, Prieto MA, García JL. Functional analysis of the small component of the 4-hydroxyphenylacetate 3-monooxygenase of *Escherichia*

- coli W: a prototype of a new flavin:NAD(P)H reductase subfamily. *J Bacteriol.* 2000;182:627–36.
13. Otto K, Hofstetter K, Röthlisberger M, Witholt B, Schmid A. Biochemical characterization of StyAB from *Pseudomonas* sp. strain VLB120 as a two-co1. Otto K, Hofstetter K, Röthlisberger M, Witholt B, Schmid A. Biochemical characterization of StyAB from *Pseudomonas* sp. strain VLB120 as a two-component flavin-diffusibl. *J Bacteriol.* 2004;186:5292–302. <https://doi.org/10.1128/JB.186.16.5292-5302.2004>.
  14. Roldán MD, Pérez-Reinado E, Castillo F, Moreno-Viván C. Reduction of polynitroaromatic compounds: the bacterial nitroreductases. *FEMS Microbiol Rev.* 2008;32:474–500.
  15. Dick T, Lee BH, Murugasu-Oei B. Oxygen depletion induced dormancy in mycobacterium smegmatis. *FEMS Microbiol Lett.* 1998;163:159–64. <https://doi.org/10.1111/j.1574-6968.1998.tb13040.x>.
  16. Hill EB, Wayne LG, Gross WM. Purification of mycobacterial deoxyribonucleic acid. *J Bacteriol.* 1972;112:1033–9.
  17. Pour NR. Determination of protein concentration using Bradford microplate protein quantification assay. *Int Electron J Med.* 2015;4:11–7.
  18. Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, et al. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res.* 2014;42. <https://doi.org/10.1093/nar/gku340>.
  19. Chauviac F-X, Bommer M, Yan J, Parkin G, Daviter T, Lowden P, et al. Crystal structure of reduced MsAcp, a putative nitroreductase from mycobacterium smegmatis and a close homologue of mycobacterium tuberculosis Acp. *J Biol Chem.* 2012;287:44372–83. <https://doi.org/10.1074/jbc.M112.406264>.
  20. DeLano WL. Pymol: an open-source molecular graphics tool. *CCP4 Newsletter Protein Crystallography.* 2002;40:82–92.
  21. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Crystallogr.* 1993;26:283–91. <https://doi.org/10.1107/S0021889892009944>.
  22. Bowie JU, Ltcy R, Eisenberg D. A method to identify protein sequences that fold into a known three-dimensional structure, vol. 253; 1990.
  23. Colovos C, Yeates TO. Verification of protein structures: patterns of non-bonded atomic interactions. *Protein Sci.* 1993;2:1511–9. <https://doi.org/10.1002/pro.5560020916>.
  24. Zenno S, Koike H, Kumar AN, Jayaraman R, Tanokura M, Saigo K. Biochemical characterization of NfsA, the *Escherichia coli* major nitroreductase exhibiting a high amino acid sequence homology to Frp, a *Vibrio harveyi* flavin oxidoreductase. *J Bacteriol.* 1996;178:4508–14. <https://doi.org/10.1128/jb.178.15.4508-4514.1996>.
  25. Zenno S, Koike H, Tanokura M, Saigo K. Gene cloning, purification, and characterization of NfsB, a minor oxygen-insensitive nitroreductase from *Escherichia coli*, similar in biochemical properties to FRase I, the major flavin reductase in *Vibrio fischeri*. *J Biochem.* 1996;120:736–44. <https://doi.org/10.1093/oxfordjournals.jbchem.a021473>.
  26. Holt PJ, Efremov RG, Nakamaru-Ogiso E, Sazanov LA. Reversible FMN dissociation from *Escherichia coli* respiratory complex I. *Biochim Biophys Acta Bioenerg.* 1857;2016:1777–85.
  27. Grandori R, Khalifah P, Boice JA, Fairman R, Giovanielli K, Carey J. Biochemical characterization of WrbA, founding member of a new family of multimeric flavodoxin-like proteins. *J Biol Chem.* 1998;273:20960–6. <https://doi.org/10.1074/jbc.273.33.20960>.
  28. Ho SN, Hunt HD, Horton RM, Pullen JK, Pease LR. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene.* 1989;77:51–9. [https://doi.org/10.1016/0378-1119\(89\)90358-2](https://doi.org/10.1016/0378-1119(89)90358-2).
  29. Agarwal R, Bonanno JB, Burley SK, Swaminathan S. Structure determination of an FMN reductase from *Pseudomonas aeruginosa* PA01 using sulfur anomalous signal. *Acta Crystallogr D Biol Crystallogr.* 2006;62:383–91. <https://doi.org/10.1107/S0907444906001600>.
  30. Minton AP. Recent applications of light scattering measurement in the biological and biopharmaceutical sciences. *Anal Biochem.* 2016;501:4–22. <https://doi.org/10.1016/j.ab.2016.02.007>.
  31. Arzenšek D, Podgornik R, Kuzman D. Dynamic light scattering and application to proteins in solutions; 2010.
  32. Nobbmann U. Protein sizing by light scattering, molecular weight and polydispersity. *Int J Biol Macromol.* 1996;19:213–21. <http://www.ncbi.nlm.nih.gov/pubmed/8910062>.
  33. Heinig M, Frishman D. STRIDE: a web server for secondary structure assignment from known atomic coordinates of proteins. *Nucleic Acids Res.* 2004;32(WEB SERVER ISS):W500–2. <https://doi.org/10.1093/nar/gkh429>.
  34. Yang JT, Wu CSC, Martinez HM. Calculation of protein conformation from circular dichroism. *Methods Enzymol.* 1986;130(C):208–69. [https://doi.org/10.1016/0076-6879\(86\)30013-2](https://doi.org/10.1016/0076-6879(86)30013-2).
  35. Williamson G, Edmondson DE. [39] 1H NMR spectral analysis of the ribityl side chain of riboflavin and its ring-substituted analogs. *Methods Enzymol.* 1986;122 C:240–8. [https://doi.org/10.1016/0076-6879\(86\)22177-1](https://doi.org/10.1016/0076-6879(86)22177-1).
  36. Sherman DR, Voskuil M, Schnappinger D, Liao R, Harrell MI, Schoolnik GK. Regulation of the mycobacterium tuberculosis hypoxic response gene encoding  $\alpha$ -crystallin. *Proc Natl Acad Sci USA.* 2001;98:7534–9. <https://doi.org/10.1073/pnas.121172498>.
  37. Dahl SG, Sylte I, Ravna AW. Structures and models of transporter proteins. *J Pharmacol Exp Ther.* 2004;309:853–60. <https://doi.org/10.1124/jpet.103.059972>.
  38. Philipp WJ, Poulet S, Eiglmeier K, Pascopella L, Balasubramanian V, Heym B, et al. An integrated map of the genome of the tubercle bacillus, mycobacterium tuberculosis H37Rv, and comparison with mycobacterium leprae. *Proc Natl Acad Sci U S A.* 1996;93:3132–7. <https://doi.org/10.1073/pnas.93.7.3132>.
  39. Cole ST, Eiglmeier K, Parkhill J, James KD, Thomson NR, Wheeler PR, et al. Massive gene decay in the leprosy bacillus. *Nature.* 2001;409:1007–11. <https://doi.org/10.1038/35059006>.
  40. MacHeroux P, Kappes B, Ealick SE. Flavogenomics - a genomic and structural view of flavin-dependent proteins. *FEBS J.* 2011;278:2625–34.
  41. Garton NJ, Waddell SJ, Sherratt AL, Lee SM, Smith RJ, Senner C, et al. Cytological and transcript analyses reveal fat and lazy persister-like bacilli in tuberculous sputum. *PLoS Med.* 2008;5:0634–45. <https://doi.org/10.1371/journal.pmed.0050075>.
  42. Natalello A, Doglia SM, Carey J, Grandori R. Role of flavin mononucleotide in the thermostability and oligomerization of *Escherichia coli* stress-defense protein WrbA. *Biochemistry.* 2007;46:543–53. <https://doi.org/10.1021/bi061769c>.
  43. Bourdeaux F, Ludwig P, Paithankar K, Sander B, Essen LO, Grininger M, et al. Comparative biochemical and structural analysis of the flavin-binding dodecins from *Streptomyces davaonensis* and *Streptomyces coelicolor* reveals striking differences with regarto multimerization. *Microbiology (United Kingdom).* 2019;165:1095–106. <https://doi.org/10.1099/mic.0.000835>.
  44. Albert A, Martínez-Ripoll M, Espinosa-Ruiz A, Yenush L, Culiñez-Macià FA, Serrano R. The X-ray structure of the FMN-binding protein atHal3 provides the structural basis for the activity of a regulatory subunit involved in signal transduction. *Structure.* 2000;8:961–9.
  45. Anderson PJ, Cole LJ, McKay DB, Entsch B. A flavoprotein encoded in *Selemonas ruminantium* is characterized after expression in *Escherichia coli*. *Protein Expr Purif.* 2002;24:429–38. <https://doi.org/10.1006/prep.2001.1581>.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.