

Expression, Purification and in Silico Characterization of Mycobacterium Smegmatis Alternative Sigma Factor SigB

Rakesh Kumar Singh

Banaras Hindu University

Lav Kumar Jaiswal

Banaras Hindu University

Tanmayee Nayak

Banaras Hindu University

Ravindra Singh Rawat

Vellore Institute of Technology: VIT University

Sanjit Kumar

Vellore Institute of Technology: VIT University

Ankush Gupta (✉ ankushgupta@bhu.ac.in)

Banaras Hindu University <https://orcid.org/0000-0001-7575-3187>

Research Article

Keywords: Alternative sigma factor, recombinant SigB, Mycobacterium, stress-responsive, ECF

Posted Date: September 15th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-856694/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Disease Markers on May 20th, 2022. See the published version at <https://doi.org/10.1155/2022/7475704>.

Abstract

Sigma factor B (SigB), an alternative sigma factor (ASF) is very similar to primary sigma factor SigA (σ^{70}) but dispensable for growth in both *Mycobacterium smegmatis* (Msmeg) and *Mycobacterium tuberculosis* (Mtb). It is involved in general stress responses including heat, oxidative, surface, starvation stress and macrophage infections. Despite having an extremely short half-life, SigB tends to operate downstream of at least three stress-responsive Extra Cytoplasmic Function (ECF) sigma factors (SigH, SigE, SigL) and SigF involved in multiple signalling pathways. There is very little information available regarding the regulation of SigB sigma factor and its interacting protein partners. Hence, we cloned the SigB gene into pET28a vector and optimized its expression in three different strains of *E. coli*, viz; (BL21(DE3), C41(DE3) and CodonPlus(DE3)). We also optimized several other parameters for the expression of recombinant SigB including IPTG concentration, temperature, and time-duration. We achieved the maximum expression of SigB at 25 °C in the soluble fraction of the cell which was purified by affinity chromatography using Ni-NTA and further confirmed by Western blotting. Further, structural characterization demonstrates the instability of SigB in comparison to SigA is carried out using homology modelling and structure function relationship. We have done protein-protein docking of RNA polymerase (RNAP) of *Mycobacterium smegmatis* (Msmeg) and sigB. This effort provides a platform for pulldown assay, structural and other studies with the recombinant protein to deduce the SigB interacting proteins, which might pave the way to study its signalling networks along with its regulation.

Introduction

Prokaryotic transcription is mediated by RNA polymerase (RNAP) which is made up of five core subunits ($\alpha_2\beta\beta'\omega$) that binds with the sigma factor (σ), also known as transcription initiation factor that provides specificity during transcription initiation. Based on the structure and function, σ factors are classified into two major groups, viz; (i) Sigma 70 (σ^{70}) i.e., primary sigma and (ii) Sigma 54 (σ^{54}) which is involved in nitrogen fixation and found mostly in plant growth promoting rhizobacteria. Further, σ^{70} is classified into four major groups based on the domain organization, viz; Group 1, 2, 3 and 4. Group 1 (SigA) has an extended N-terminal region and four highly conserved regions namely; region 1, 2, 3 and 4. The Group 2 (SigB) does not have the N-terminus extended region, but all the four conserved regions are present. The Group 3 (SigF) has only three regions viz; 2, 3 and 4 while the Group 4 (also known as the Extra Cytoplasmic Function (ECF) sigma factor) contains only regions 2 and 4. The regions 2 and 4 are extremely important in promoter recognition and binds at -10 and -35 regions of the promoter, respectively [1, 2].

σ^{70} being the principle σ factor, contributes in the expression of all housekeeping genes under normal conditions [3, 4]. However, under various stressful conditions, alternative sigma factors (ASF) redirect the transcription machinery towards the expression of a wide range of genes, which help the organism to respond to the internal or environmental stresses [5, 6]. Due to unequal division of tasks, σ^{70} have broad range promoter recognition efficiency but ASF deal with very high promoter recognition stringency. By

these mechanisms that regulate transcription initiation, the organism ensures its survival and adaptability [7–9]. SigB responds to different stresses like heat shock, surface, oxidative, starvation stresses and macrophage infections [3, 10, 11]. It is also expressed while transition to stationary phase during bacterial growth to cope with the environmental stresses [12–15].

In contrast to SigA, the half-life of Msmeg SigB is very short (less than 8 mins.) and in *E. coli* it is only 2 minutes [12, 16–18]. This indicates towards the nature and role of SigB as a protein with a short half-life and involved in regulatory functions. The transcription of SigB was shown to occur from two distinct types of promoters: one recognized by the stress-inducible sigma factors *viz*, SigE, SigH and SigL and the other recognized by SigF [19–22]. This infers that multiple sigma factor signalling networks converge at the SigB promoter. On the other hand, this observation is also supported by the deletion analysis of SigC, SigE, SigF, SigH, SigL sigma factors resulting in the reduced expression of SigB, demonstrated in mice model [19, 23–26]. The regulatory network and structure of the primary sigma factor SigA is already deduced. However, very little information about the regulatory network of SigB and other stress-responsive ECF sigma is available in the literature. Although, recombinant and purified SigB is essential for the study of protein interactions / elucidation of its regulatory network as well as crystallization studies; detailed and optimized expression and purification of this short-lived protein in higher quantities and highly soluble form is not available in the literature. Here, we optimize and demonstrate the expression and purification of highly soluble Msmeg SigB.

Material And Methods

Bacterial strains, chemicals and reagents

E. coli DH5 α (Zymo Research, USA) was used as cloning host and BL21(DE3), C41(DE3) (Millipore Sigma (Novagen)) and CodonPlus(DE3) (Agilent Technologies) cells were used expression host strains. The pET28a expression vector (Thermo Scientific, USA), Taq DNA polymerase (G-Biosciences, USA), restriction enzymes (New England Bio Labs Ltd. UK), T4 DNA ligase (New England Bio Labs Ltd., UK) and molecular biology grade reagents were obtained from Sigma Aldrich (USA). His-Tag monoclonal antibody and secondary Anti mouse IgG (H+L) (Peroxidase/ HRP conjugated) were from Puregene: Genetix Biotech Asia Pvt. Ltd. respectively.

Plasmid construction

The non-directional cloning strategy was employed for the cloning of *SigB* (MSMEG_2752) gene of *M. smegmatis* mc²155. The primer pair used to amplify the *SigB* gene from the genomic DNA of *M. smegmatis* mc²155 were, forward primer SigB_2752_EcoRI 5'-CGGAATTCATGGCAAATGCCACCAAGCC-3' and reverse primer SigB_2752_EcoRI 5'-CGGAATTCGGAGGCGTAGGAGCGGAGGCGG-3' respectively, where underline sequences are showing the EcoRI restriction site. The condition for PCR amplification were initial denaturation at 95°C for 5 minutes, 35 cycles (denaturation 95°C for 30 sec, annealing 60°C for 45 sec, elongation 72°C for 1 minute) and final elongation 72°C for 5 minute. The amplified PCR

product was ligated into the EcoRI digested and dephosphorylated pET28a expression vector. The ligation product was transformed into DH5α cells and the recombinant positive clones (His-SigB) were confirmed by EcoRI restriction digestion and sequencing.

Expression of recombinant His-SigB protein

a. Expression host strain optimization

The positive clones containing His-SigB were transformed into three expression host strains of *E.coli viz*, BL21(DE3), C41(DE3) and CodonPlus(DE3) [27]. These expression strains possess λDE3 lysogen that carries the gene T7 RNA polymerase under the control of lacUV5 promoter. So, IPTG (Isopropyl β-D-1-thiogalactopyranoside) induction for protein expression was performed to obtain maximum expression. The selected clone was inoculated from a single colony and cultured over night in 5 ml Luria Bertani (LB) containing 50 µg/ml kanamycin at 37°C in a shaking incubator. Next day, 1% (v/v) of secondary culture was incubated at 37°C until the OD₆₀₀ of the culture reached ~0.4 O.D., the culture was induced with 0.1mM IPTG and grown for 3 hrs at 37°C at 150 rpm. For analysis of His-SigB protein expression, sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) was performed as described by [28]. For SDS-PAGE, 0.1 OD₆₀₀ equivalent cell lysate prepared in 4X sample loading buffer (10% SDS, 40% glycerol, 0.25M Tris–Cl (pH 6.8), 200 mM DTT, and 0.05% bromophenol blue) after heat denaturation at 100°C were loaded onto the gel. For visualization, the gel was stained by Coomassie brilliant blue (R-250) followed by destaining.

b. IPTG optimization

Commonly, IPTG was used to achieve high level of expression for recombinant proteins. To optimize the concentration of IPTG on *E. coli* strains and the optimum protein expression, we selected best expressing host strain for a range of IPTG concentrations *viz*, 0.05, 0.1, 0.25, 0.5mM and induced at approximately < 0.4 OD₆₀₀ at 37°C 150 rpm for 3 hrs. To determine the solubility of protein at various IPTG concentrations, the pellet/supernatant fractionation of the lysate was performed. The harvested cell cultures were washed and resuspended in 1X phosphate buffer saline (PBS, pH 8.0) and thereafter sonicated at 15% amplitude (ultrasonic homogenizer) for 2-3 min. The lysate was separated by centrifugation at 13,000 rpm for 20 min at 4°C containing soluble His-SigB in the supernatant as well as the pellet fraction resuspend in equal volume of the lysis buffer (1X PBS). The protein bands were visualized by denaturing SDS-PAGE after staining with Coomassie R-250 as mentioned above. Semi-quantitative analysis of the gels were performed with Quantity One software (BIO-RAD)

c. Temperature optimization

To optimize suitable temperature for protein expression, the best expressing host strain cultured at diff

d. Time point optimization

To check best time point for the expression of His-SigB in the best expressing host strain, it was induce

His-SigB purification and Immunoblotting

The recombinant His-SigB was purified by affinity using Ni-NTA beads as per manufacturer's instructions. Briefly, the cells were lysed by sonication at 15% amplitude in Buffer A (Tris.Cl 50 mM (pH 8.0), NaCl 300 mM, Imidazole 10 mM, Triton X-100 1 % (v/v), PMSF 2 mM, Lysozyme 1 mg/ml) and the soluble lysate was clarified by centrifugation at 13,000 rpm (rcf 14,926 g) for 15 min. at 4°C. The pre-equilibrated Ni-NTA beads (Qiagen) in Buffer A was incubated with the lysate at 4°C for 30 mins., washed thrice with Buffer B (Tris.Cl 50 mM (pH 8.0), NaCl 300 mM, Imidazole 50 mM, Triton X-100 0.5 % (v/v)) and eluted thrice with Buffer C (Tris.Cl 50 mM (pH 8.0), NaCl 300mM, Imidazole 400 mM, Triton X-100 0.5 % (v/v)).

The purified His-SigB was resolved by SDS-PAGE and transferred to the PVDF membrane followed by blocking in 5% (w/v) skimmed milk powder (Himedia laboratories) in 1XPBST (Phosphate saline buffer containing 0.1% Tween-20). The blot was washed thrice by 1X PBST and incubated with anti-His antibody (0.5 µg/µl, 1:1000 in 1% PBST; Puregene) for 1 hr., washed with PBST and further incubated with secondary antibody anti-Mouse IgG (H+L Peroxidase/HRP conjugated, 1:5000 Pure gene Genetix Biotech) in PBST. Finally the blot was washed thrice by PBST, developed in Clarity™ ECL Bio-RAD, and analysed by ChemiDoc Imaging System (Bio-RAD).

Homology modeling and Structural comparison of SigA and SigB

The protein sequences of *M. smegmatis* sigma factors SigA and SigB were retrieved from National Center for Biotechnology Information (NCBI) database [29]. To construct the homology models, the amino acid sequences of both the proteins were submitted to Phyre² server [30]. Phyre² server utilizes multiple templates to produce the accurate model of the protein. In case of *M. smegmatis* SigA, three templates were used, namely, 5TW1(RNA polymerase sigma factor) [31], 6C05(crystal structure of RNA polymerase sigma factor from *M. tuberculosis*) [32], 4YG2 (X-ray crystal structure of sigma70 holoenzyme from *E. coli*) [33] with percent identities of 99%, 97% and 55% respectively. Similarly, *M. smegmatis* SigB model was build using two templates 5TW1 (RNA polymerase sigma factor) [31]and 6C05 (crystal structure of RNA polymerase sigma factor from *M. tuberculosis*) [32] with percent identities of 65% and 64%, respectively. Both the models were validated on SAVES server using PROCHECK tool [34].

Protein-Protein docking of core RNA Polymerase (RNAP) and SigB

The protein-protein docking utilizes a multi-stage method for generating poses followed by their rankings. Here, we have used MOE platform [35] for the protein-protein docking experiment of core RNA Polymerase (RNAP) and SigB upon importing RNAP (PDB ID: 6EYD [36]) to MOE (<https://www.chemcomp.com/Products.htm>), workbench. Structure issues such as missing atoms, missing chains were automatically corrected using quick prep module. Optimal hydrogen positions and charges were optimized using protonated 3D module using the MOE software

recommended settings. We have used here core enzyme of 6YED [36] PDB without sigma factor A (SigA). Before commencement of docking, bad crystallographic contacts or other imperfect geometries were overcome by the energy minimization method using Amber EHT 10 force field recommended by MOE platform. [Energy Minimize](#) application itself can be used to adjust hydrogens and lone pairs and to calculate partial charges. Similarly, SigB protein was prepared before docking. SigB model was considered as ligand and PDB ID: 6YED [36] was considered as a receptor. After the docking run, the appropriate protein-protein fingerprints were generated for each pose which can be viewed using the PLIF visualization panel. Approximately, 100 coordinates were developed according to the binding energy. The best twenty dock complexes were considered for analysis.

Results

Expression host strain optimization for the expression of recombinant His-SigB

The pET28a vector belongs to the pET series of vectors that contain a T7 promoter with *lac* operator for the IPTG inducible expression of downstream recombinant genes. However, these genes can only be expressed when T7 RNA polymerase is made available in expression hosts containing a DE3 lysogen under the control of IPTG inducible lacUV5 promoter in their genomic DNA [27]. Hence, we utilized three expression host strains of *E. coli* containing the DE3 lysogen *viz*, BL21(DE3), C41(DE3) and CodonPlus(DE3) for the optimization of the expression of recombinant His-SigB. Empty pET28a vector and uninduced recombinant His-SigB clones were used as negative control for the expression host strains BL21(DE3), C41(DE3) and CodonPlus(DE3) in figure 1 (lanes: V.O. and U) respectively. These negative controls did not show any expression of the recombinant His-SigB as depicted in figure 1. However, maximum expression of the recombinant His-SigB (~42 kDa) is observed in the expression host CodonPlus(DE3) followed by C41(DE3) and BL21(DE3), respectively (lane I for each cell type in figure 1).

Optimization of IPTG concentration for recombinant His-SigB

CodonPlus(DE3) depicted the maximum expression of the recombinant His-SigB protein hence; it was decided to optimize the maximum protein expression at different IPTG concentrations (0.05-0.5 mM). The expression of His-SigB at IPTG concentrations; 0.05, 0.1 and 0.25 mM were comparable which was approximately 30% higher than the expression at 0.5 mM concentration (as depicted in Figure 2(a) and (c)). Amongst the three concentrations, best expression was found in 0.1 mM IPTG. When the pellet and supernatant fractions of the induced CodonPlus(DE3) cells were observed, comparable distribution of the soluble His-SigB protein was observed in the supernatant fractions at different IPTG concentrations as depicted in figure 2(b) and (d). However, total expression and relative solubility of the recombinant His-SigB was optimum at 0.1 mM IPTG concentration which was selected further studies.

Temperature optimization for the expression of recombinant His-SigB protein

At the optimized IPTG concentration of 0.1 mM, His-SigB was induced at 37°C as well as lower temperatures *viz*, 25°C and 16°C for increasing the relative solubility of His-SigB. Maximum expression of

His-SigB was obtained at 37°C and the expression decreased drastically to approximately 62% and 54% at temperatures; 25°C and 16°C, respectively (figure 3 (a) and (c)) as compared to that at 37°C. However, when the distribution of the soluble His-SigB protein was observed in the supernatant fractions at different temperatures, it was observed that almost 31%, 55% and 54% of the total induced proteins fractionated to supernatant at temperatures; 37°C, 25°C and 16°C respectively (figure 3(b) and (d)). Thus, higher total expression and maximum relative solubility of the recombinant His-SigB was observed at 25°C temperature.

Time course experiments, Ni-NTA affinity purification and Immunoblotting of purified His-SigB

The time course experiment was carried out with for determining the maximum expression of soluble His-SigB protein at the optimized conditions of 25°C with 0.1 mM IPTG for different time points *viz*, 0, 15, 30, 60, 120 and 180 minute. The maximum expression was observed at 180 minutes after the IPTG induction as depicted in figure 4 (a).

His-SigB protein was purified using Ni-NTA affinity chromatography. The affinity purified His-SigB protein was found to have minimal proteolytic degradation as depicted by the Commassie stained gel in figure 4 (b). In spite of the short half-life of His-SigB, approximately 2 mg - 2.2 mg/l culture of the protein was purified. To further rule out the possibility of proteolytic degradation of the recombinant protein, immunoblotting was performed using a monoclonal anti-6XHis antibody. It was observed that a highly purified His-SigB protein without any proteolytic degradation was obtained under the conditions optimized in this article as depicted in figure 4 (c).

Structural characterization of SigA and SigB

The Ramachandran plots for *M. smegmatis* SigA and SigB proteins demonstrate 98.1% and 98.7% residues falling in the allowed regions, respectively (Supplementary figure 1). The models of SigA and SigB however, also exhibited few structural differences. *M. smegmatis* SigA differs from SigB in possessing an N-terminal extended domain. It has been observed that the presence of N-terminal extended domain provides stability to the molecule and consequently, SigA is comparatively more stable than SigB.

The C-terminal domains of SigA and SigB are composed of four α -helices. The first and the fourth α -helices of the domain lie adjacent to one another in a parallel fashion and run in opposite directions. The second and fourth helices are shorter than the first and third helices. Furthermore, second and fourth helices are connected to the ends of third helix in perpendicular manner through short loops (figure 5 (a) and (b)). The region connecting the first and second helix could not be modeled precisely.

Molecular interactions of RNAP and SigB

Conformational changes of RNAP are required to initiate the transcription process. To investigate the interactions of SigB with RNAP, protein-protein docking experiment was performed. The protein-protein docking experiment revealed SigB bound to RNAP with binding energy of -70.39 (Sixteen dock complex).

The *M. smegmatis* RNAP is composed of five subunits including α (two copies of chain A and B), β (Chain C), β' (Chain D) and ω (Chain E) (figure 5 (c)). These five chains together formed the core enzyme and help to synthesize the RNA using DNA as template and ribonucleotides as the substrate. To start this process the core enzyme has to bind to a sigma factor. In RNAP of *M. smegmatis*, Chain C and D mainly interact with SigB (Figure 5 (d)). Arg846, Leu408 of chain C and Arg67, Thr38, Lys64, Thr150, Arg345, Glu250 and Glu275 of Chain D play a major role in the molecular interactions with SigB which are in contrast with the molecular interactions of SigA.

Discussion

Although, SigB is dispensable for growth in both Msmeg and Mtb [15,37] nevertheless, SigB acts as general stress responsive transcription factor. The importance of SigB protein is elucidated by the fact that it occurs downstream in the signalling pathways of several stress responsive ECF and other sigma factors [13, 14, 22]. The expression of recombinant SigB has been mentioned in reference to the *in vitro* transcription assay experiment in BL21(DE3) strain which is deficient of lon and omp-t proteases and is therefore suitable for expression of non-toxic genes but the SigB protein is very poorly expressed in this strain as demonstrated from our results (Fig. 1) [14]. This lead us to express SigB in other alternative strains like C41(DE3) and CodonPlus(DE3). C41(DE3), derived from BL21(DE3), contains an uncharacterized mutation that prevents cell death due to expression of many toxic recombinant proteins. Hence, it is often utilized for expression of toxic and membrane proteins from all classes of organisms [38]. However, in our case it lead to only a marginal increase in the expression of recombinant SigB (Fig. 1). CodonPlus(DE3) improves protein expression by preventing codon bias and by providing additional copies of specific tRNA genes *viz*, argU, ileY, leuW and proL genes that recognize AGA/AGG, AUA, CUA and CCC codons, respectively that are rare in *E. coli* and occur commonly in GC-rich genomes like Mycobacterium [39, 40]. SigB of Msmeg in our case contains at least 8 Prolines out of which 2 are encoded by CCC codons which might be an important reason for the lower expression in other strains utilized.

During the optimization with different IPTG concentrations, it was observed that higher relative expression of soluble SigB is obtained at lower IPTG concentrations like 0.1 mM which might be due to decreased formation of inclusion bodies due to lower rate of induction (Fig. 2). Greater IPTG often leads to faster protein expression resulting in increased formation of inclusion bodies. Similarly, when three different temperatures were utilized for SigB expression, there was a drastic decrease in expression from 38% – 46% at 25°C and 16°C respectively, as compared to that at 37°C. However, the relative fraction of total protein expressed as soluble fraction in the supernatant was much greater at lower temperatures like 25°C and 16°C. The intermediate temperature of 25°C was best suited for our experimental work (Fig. 3). The IPTG induction of recombinant proteins like human RNase L, human growth hormone (hGH) and several other proteins at lower temperatures in bacteria favours relatively higher yield of the expressed protein in the soluble fraction of the cell [41–47]. The experimental conditions optimized above for recombinant SigB expression provided a high yield of soluble and proteolytically undegraded protein

which can be visualized by the immunoblotting of purified SigB using a monoclonal anti-6XHis antibody since, no degradation products could be visualized at higher concentrations of the purified protein (Fig. 5).

However, the highly unstable nature of this protein can be further improved by using various stabilizers and (or) another cofactor so that future experiments like protein-protein interactions and crystallization of SigB can be carried out smoothly. *M. smegmatis* RNA polymerase SigA and Sig B showed 60% sequence identity. Sigma factor A differs from sigma factor B in possessing N-terminal domain, which may provide stability. Our docking studies demonstrated that SigB binds to RNAP where chain C and chain D of RNAP play a significant role in molecular interactions. *In silico* comparative binding studies of SigA and SigB demonstrated that they both bind to similar regions of RNAP. In future, our aim will be to co-crystallize SigB-RNAP together to better understand the mechanism of their molecular interactions.

Conclusion

SigB has significant similarities with house-keeping sigma factor SigA with regards to the domain organization however; it has drastically short half-life as compared to SigA. Despite, its short half-life, it is also involved in multiple stress response regulatory pathways of *Mycobacterium* in coordination with ECF and other sigma factors. Higher quantities of soluble and purified recombinant SigB is essential for the study of its structure, function and regulation but there is dearth of literature that depicts it. In this report, we demonstrate an efficient method for the production of soluble and purified recombinant His-SigB protein that can be used as a tool to unravel the structure, function and regulation of its signalling cascade.

Declarations

Acknowledgments

We acknowledge the financial assistance in the form of Senior Research Fellowships (SRF) to RKS and LKJ from the Indian Council of Medical Research (ICMR), and to TN from University Grants Commission (UGC), New Delhi and to the Inter Disciplinary School of Life Sciences (ISLS), BHU, Varanasi for providing the instrumentation facility.

Declarations

Not applicable

Conflicts of interest

The authors declare no conflict of interest

References

1. Murakami KS, Masuda S, Darst SA (2002) Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 Å resolution. *Science* 296:1280–1284
2. Vassylyev DG, Sekine SI, Laptenko O et al (2002) Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution. *Nature* 417:712–719
3. Manganello R, Dubnau E, Tyagi S, Kramer FR, Smith I (1999) Differential expression of 10 sigma factor genes in *Mycobacterium tuberculosis*. *Mol Microbiol* 31(2):715–724
4. Wu S, Howard ST, Lakey DL et al (2004) The principal sigma factor sigA mediates enhanced growth of *Mycobacterium tuberculosis* in vivo. *Mol Microbiol* 51(6):1551–1562
5. Paget MSB, Helmann JD (2003) The $\sigma 70$ family of sigma factors. *Genome Biol* 4(1):1–6
6. Browning DF, Busby SJW (2004) The regulation of bacterial transcription initiation. *Nat Rev Microbiol* 2(1):1–9
7. Gruber TM, Gross CA (2003) Multiple Sigma Subunits and the Partitioning of Bacterial Transcription Space. *Annu Rev Microbiol* 57:441–466
8. Koo BM, Rhodius VA, Nonaka G, DeHaseth PL, Gross CA (2009) Reduced capacity of alternative σ s to melt promoters ensures stringent promoter recognition. *Genes Dev* 23(20):2426–2436
9. Feklistov A, Darst SA (2009) Promoter recognition by bacterial alternative σ factors: The price of high selectivity? *Genes Dev* 23(20):2371–2375
10. Betts JC, Lukey PT, Robb LC, McAdam RA, Duncan K (2002) Evaluation of a nutrient starvation model of *Mycobacterium tuberculosis* persistence by gene and protein expression profiling. *Mol Microbiol* 43(3):717–731
11. Schnappinger D, Ehrh S, Voskuil MI et al (2003) Transcriptional adaptation of *Mycobacterium tuberculosis* within macrophages: Insights into the phagosomal environment. *J Exp Med* 198(5):693–704
12. Hu Y, Coates ARM (1999) Transcription of two sigma 70 homologue genes, sigA and sigB, in stationary-phase *Mycobacterium tuberculosis*. *J Bacteriol* 181(1):469–476
13. Sachdeva P, Misra R, Tyagi AK, Singh Y (2010) The sigma factors of *Mycobacterium tuberculosis*: Regulation of the regulators. *FEBS J* 277(3):605–626
14. Lee JH, Karakousis PC, Bishai WR (2008) Roles of SigB and SigF in the *Mycobacterium tuberculosis* sigma factor network. *J Bacteriol* 190(2):699–707
15. Fontán PA, Voskuil MI, Gomez M et al (2009) The *Mycobacterium tuberculosis* sigma factor σB is required for full response to cell envelope stress and hypoxia in vitro, but it is dispensable for in vivo growth. *J Bacteriol* 191(18):5628–5633
16. Rustad TR, Minch KJ, Brabant W et al (2013) Global analysis of mRNA stability in *Mycobacterium tuberculosis*. *Nucleic Acids Res* 41(1):509–517
17. Bernstein JA, Khodursky AB, Lin PH, Lin-Chao S, Cohen SN (2002) Global analysis of mRNA decay and abundance in *Escherichia coli* at single-gene resolution using two-color fluorescent DNA microarrays. *Proc Natl Acad Sci U S A* 99(15):9697–9702

18. Zgurskaya HI, Keyhan M, Matin A (1997) The σ (s) level in starving *Escherichia coli* cells increases solely as a result of its increased stability, despite decreased synthesis. *Mol Microbiol* 24(3):643–651
19. Dainese E, Rodrigue S, Delogu G et al (2006) Posttranslational regulation of *Mycobacterium tuberculosis* extracytoplasmic-function sigma factor σ L and roles in virulence and in global regulation of gene expression. *Infect Immun* 74(4):2457–2461
20. Riccardo M, Martin IV, Gary KS, Issar S (2001) The *Mycobacterium tuberculosis* ECF sigma factor: role in global gene expression and survival in macrophages. *Mol Microbiol* 41(2):423–437
21. Raman S, Song T, Puyang X, Bardarov S, Jacobs J, Husson RN (2001) The alternative sigma factor σ H regulates major components of oxidative and heat stress responses in *Mycobacterium tuberculosis*. *J Bacteriol* 183(20):6119–6125
22. Chauhan R, Ravi J, Datta P et al (2016) Reconstruction and topological characterization of the sigma factor regulatory network of *Mycobacterium tuberculosis*. *Nat Commun* (7):1–12
23. Ando M, Yoshimatsu T, Ko C, Converse PJ, Bishai WR (2003) Deletion of *Mycobacterium tuberculosis* Sigma Factor E Results in Delayed Time to Death with Bacterial Persistence in the Lungs of Aerosol-Infected Mice. *Infect Immun* 71(12):7170–7172
24. Geiman DE, Kaushal D, Ko C et al (2004) Attenuation of Late-Stage Disease in Mice Infected by the *Mycobacterium tuberculosis* Mutant Lacking the SigF Alternate Sigma Factor and Identification of SigF-Dependent Genes by Microarray Analysis. *Infect Immun* 72(3):1733–1745
25. Hahn MY, Raman S, Anaya M, Husson RN (2005) The *Mycobacterium tuberculosis* extracytoplasmic-function sigma factor SigL regulates polyketide synthases and secreted or membrane proteins and is required for virulence. *J Bacteriol* 187(20):7062–7071
26. Kaushal D, Schroeder BG, Tyagi S et al (2002) Reduced immunopathology and mortality despite tissue persistence in a *Mycobacterium tuberculosis* mutant lacking alternative σ factor, SigH. *Proc Natl Acad Sci U S A* 99(12):8330–8335
27. Sambrook J, Fritsch E, Maniatis T (1989) *Molecular Cloning: A Laboratory Manual*. Cold spring Harbor laboratory
28. Tsai WC, Wu TC, Chiang BL, Wen HW (2017) Cloning, expression, and purification of recombinant major mango allergen Man i 1 in *Escherichia coli*. *Protein Exp Purif* 130:35–43
29. Agarwala R, Barrett T, Beck J et al (2018) Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* 46(D1):D8–D13
30. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE (2015) The Phyre2 web portal for protein modelling, prediction and analysis. *Nat Protoc* 10(6):845–858
31. Hubin EA, Fay A, Xu C et al (2017) Structure and function of the mycobacterial transcription initiation complex with the essential regulator RbpA. *Elife* 6:1–40
32. Boyaci H, Chen J, Lilic M et al (2018) Fidaxomicin jams *Mycobacterium tuberculosis* RNA polymerase motions needed for initiation via RBPA contacts. *Elife* 7:1–19

33. Murakami KS (2013) X-ray crystal structure of *Escherichia coli* RNA polymerase $\sigma 70$ holoenzyme. *J Biol Chem* 288(13):9126–9134
34. Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Crystallogr* 26(2):283–291
35. Molecular Operating Environment (2013) Vol 910
36. Kouba T, Pospíšil J, Hnilicová J, Šanderová H, Barvík I, Krásný L (2019) The core and holoenzyme forms of RNA polymerase from *Mycobacterium smegmatis*. *J Bacteriol* 201(4):e00583–e00518
37. Mukherjee R, Gomez M, Jayaraman N, Smith I, Chatterji D (2005) Hyperglycosylation of glycopeptidolipid of *Mycobacterium smegmatis* under nutrient starvation: Structural studies. *Microbiology* 151(7):2385–2392
38. Dumon-Seignovert L, Cariot G, Vuillard L (2004) The toxicity of recombinant proteins in *Escherichia coli*: A comparison of overexpression in BL21(DE3), C41(DE3), and C43(DE3). *Protein Expr Purif* 37(1):203–206
39. Carstens C-P, Waesche A (1999) Codon bias-adjusted BL21 derivatives for protein expression. *Strategies* 12(2):49–51
40. Hu X, Shi Q, Yang T, Jackowski G (1996) Specific replacement of consecutive AGG codons results in high-level expression of human cardiac troponin T in *Escherichia coli*. *Protein Expr Purif* 7(3):289–293
41. Gupta A, Rath PC (2012) Expression, purification and characterization of the interferon-inducible, antiviral and tumour-suppressor protein, human RNase L. *J Biosci* 37(1):103–113
42. Kim MJ, Park HS, Seo KH, Yang HJ, Kim SK, Choi JH (2013) Complete Solubilization and Purification of Recombinant Human Growth Hormone Produced in *Escherichia coli*. *PLoS One* 8(2):1–8
43. Li M, Cui W, Mo C, Wang J, Zhao Z, Cai M (2014) Cloning, expression, purification, antiserum preparation and its characteristics of the truncated UL6 protein of Herpes Simplex Virus 1. *Mol Biol Rep* 41(9):5997–6002
44. Santosh KM, Nitish K, Gautam K, Tara K, Krishna P (2018) Recombinant human interferon regulatory factor-1 (IRF-1) protein expression and solubilisation study in *Escherichia coli*. *Mol Biol Rep* 45(5):1367–1374
45. Kaur J, Kumar A, Kaur J (2018) Strategies for optimization of heterologous protein expression in *E. coli* Roadblocks and reinforcements. *Int J Biol Macromol* 106:803–822
46. Zou X, Xu Z, Wang Y et al (2019) Expression, purification, and antiserum production of the truncated UL31 protein of herpes simplex virus 1. *Iran J Biotechnol* 17(1):71–76
47. Sørensen HP, Mortensen KK (2005) Soluble expression of recombinant proteins in the cytoplasm of *Escherichia coli*. *Microb Cell Fact* 4(1):1–8

Figures

Expression host strain optimization for the expression of recombinant His-SigB

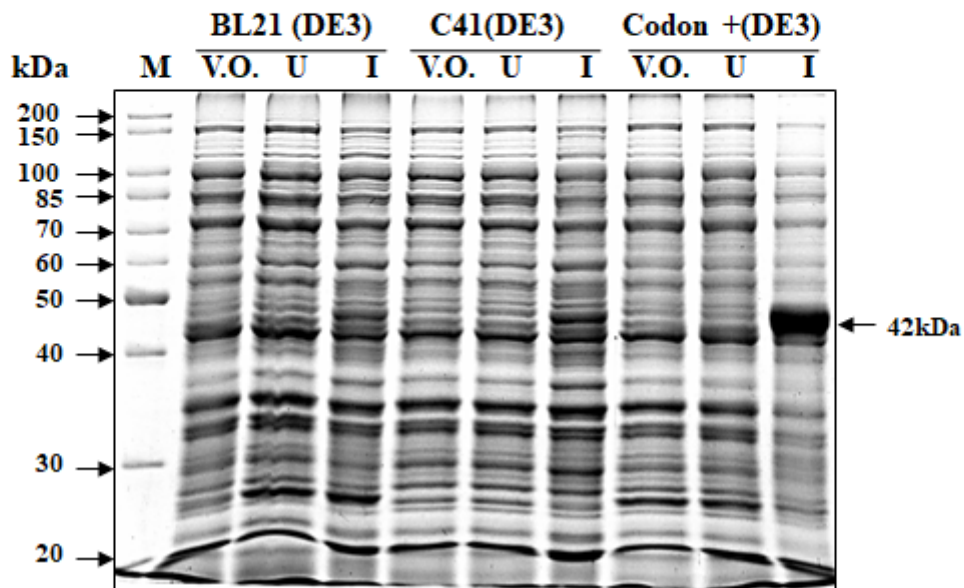
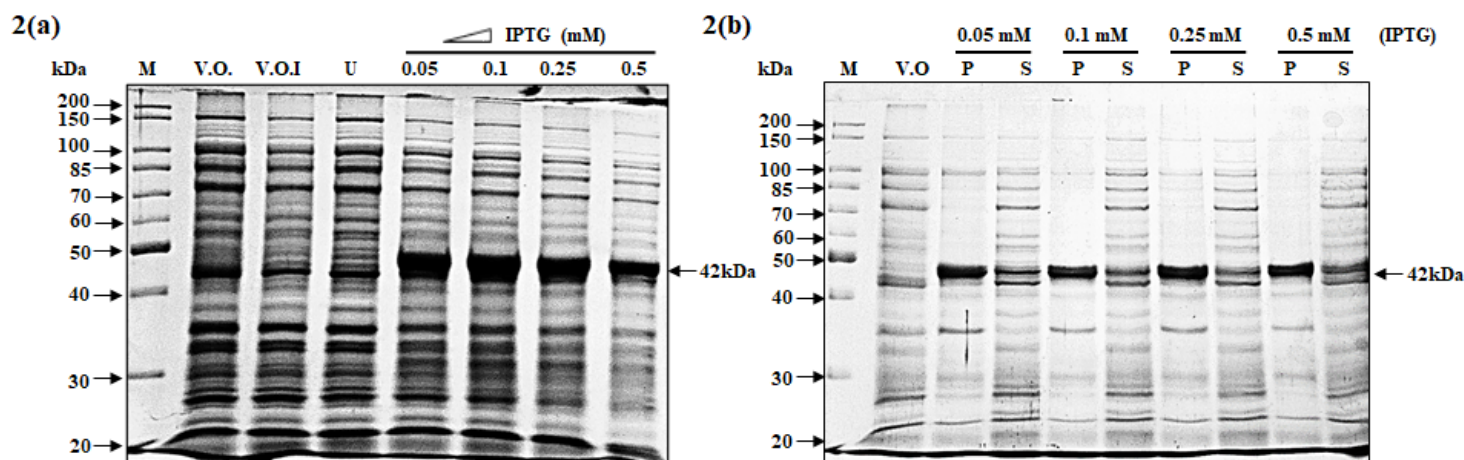


Figure 1

Expression host strain optimization for the expression of recombinant His-SigB. 10% SDS PAGE depicting recombinant His-SigB protein expression profile from three expression host strains namely; BL21(DE3), C41(DE3) and CodonPlus(DE3). Highest protein expression is observed in the host strain CodonPlus(DE3). M= Molecular size marker, V.O.= pET28a vector only control, U= Uninduced His-SigB clone, I=Induced His-SigB clone.

IPTG concentration optimization for the expression of recombinant His-SigB



IPTG concentration optimization for the expression of recombinant His-SigB

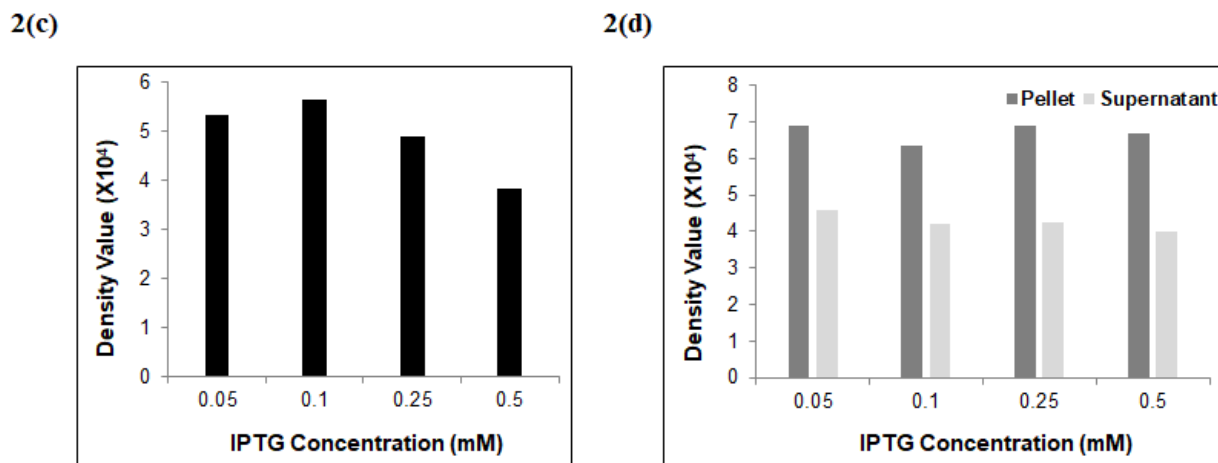
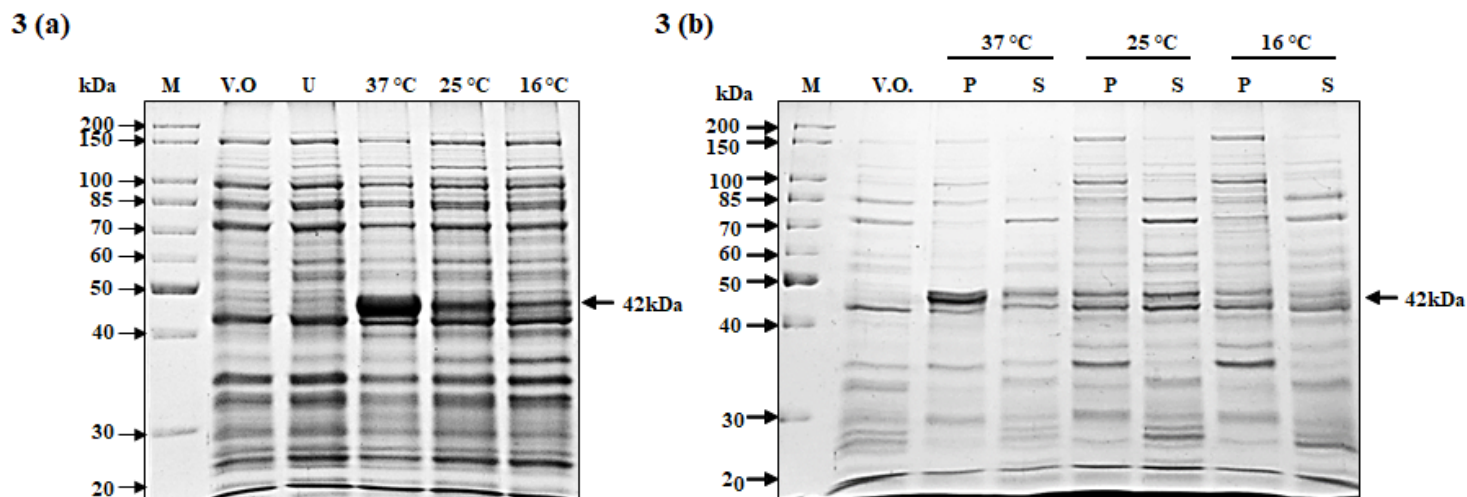


Figure 2

IPTG concentration optimization for the expression of recombinant His-SigB. (a) 10% SDS PAGE depicting recombinant His-SigB protein expression profile in CodonPlus(DE3) host strain. Highest protein expression is observed in 0.1 mM IPTG. (b) Solubility of the recombinant His-SigB in different IPTG concentrations at 37°C. 10% SDS PAGE depicting recombinant His-SigB protein distribution in cell pellet and supernatant in CodonPlus(DE3) induced with IPTG concentration from 0.05- 0.5 mM IPTG. Highest protein solubility observed in the supernatant fraction at 0.1mM IPTG. M= Molecular size marker, V.O.= pET28a vector only control, V.O.I= pET28a vector only induced, U= Uninduced His-SigB clone, P= Pellet and S= Supernatant.

Temperature optimization for the expression of recombinant His-SigB protein



Temperature optimization for the expression of recombinant His-SigB protein

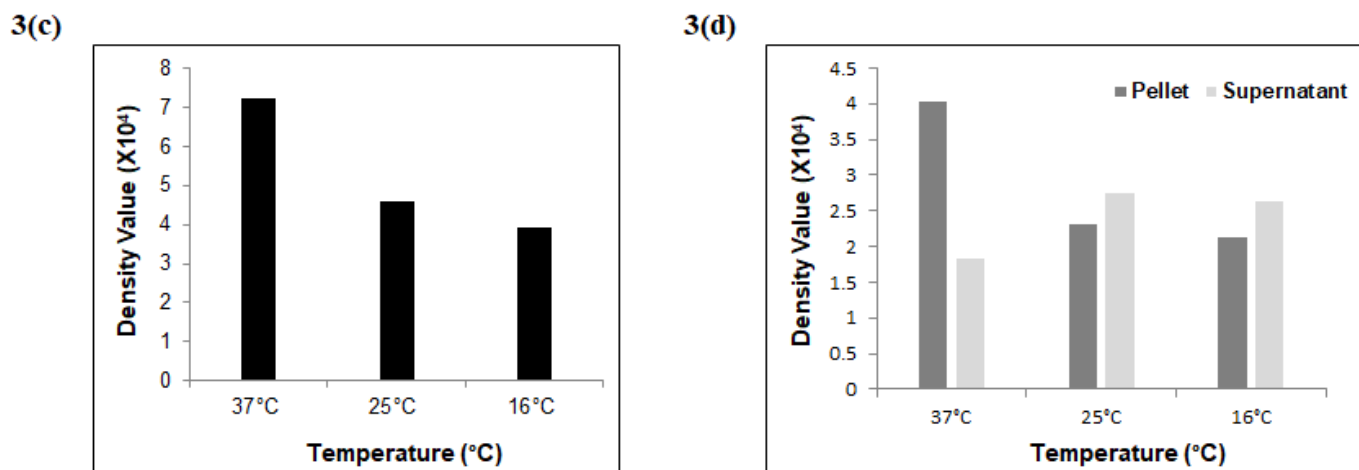


Figure 3

Temperature optimization for the expression of recombinant His-SigB protein. (a) 10% SDS PAGE depicting recombinant His-SigB protein expression profile in CodonPlus(DE3) cells at temperatures; 37°C, 25°C and 16°C with 0.1 mM IPTG. Highest protein expression was observed at 37°C. (b) 10% SDS PAGE depicting distribution of recombinant His-SigB protein in cell pellet and supernatant fractions at different temperatures. Highest protein solubility was observed at 25°C supernatant fraction. M= Marker, V.O.= pET28a Vector only control, U= uninduced His-SigB clone, P= Pellet and S= Supernatant. (c) Gel based semi-quantitative analysis from Fig. 3(a) depicts the highest expression of His-SigB at 37°C with 0.1 mM IPTG concentration. (d) Semi-quantitative analysis from Fig. 3(b) depicts higher relative protein solubility

observed in the supernatant fraction at 25°C with 0.1mM IPTG. Semi-quantitative analysis of the gels were performed with Quantity One software (BIO-RAD).

Time point experiment, Ni-NTA affinity purification and Western blotting of the recombinant His-SigB

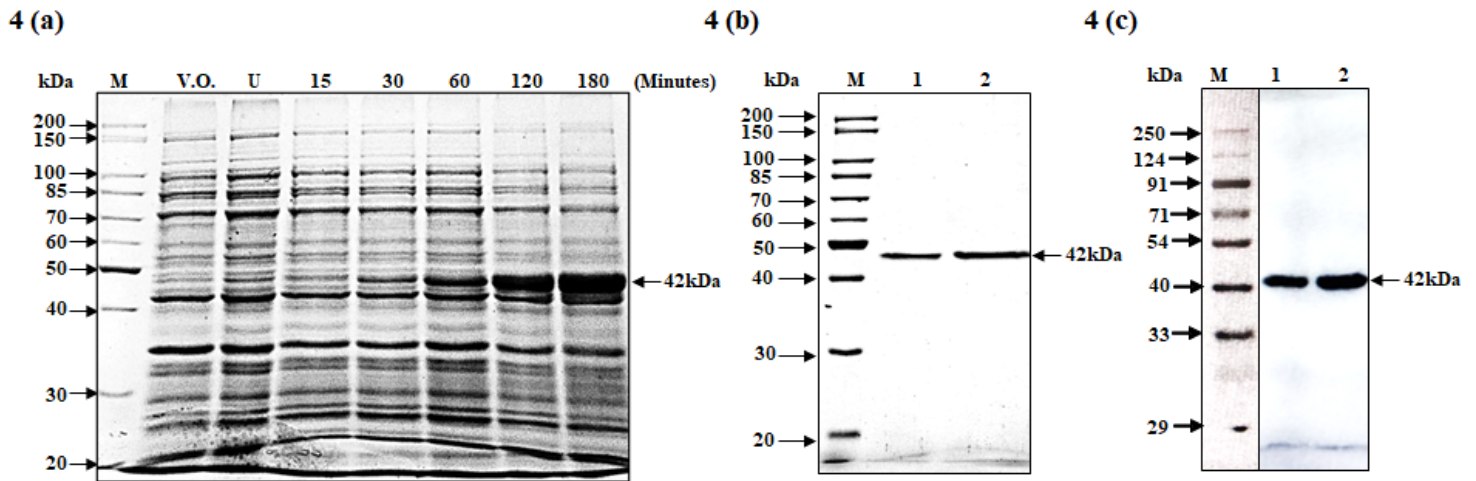


Figure 4

(a): Expression of recombinant His-SigB protein at different time points. 10% SDS PAGE depicting recombinant His-SigB protein expression profile collected at time points 15, 30, 60, 120 and 180 minutes . Highest protein expression was observed at 180 minutes after induction with 0.1mM IPTG. (b) Ni-NTA affinity purification of His-SigB: lanes denoting M= molecular size marker, 1= 5 µg and 2= 7.5 µg His-SigB purified protein stained with Commassie R-250. (c) Western blotting with purified recombinant His-SigB. Mouse Anti-His monoclonal antibody (Genetix) was used to probe purified His-SigB. M= Marker, V.O.= pET28a Vector only control, U= uninduced His-SigB clone, lane 1= 5 µg, lane 2= 10 µg purified His-SigB protein.

Homology modelling and docking analysis of SigA and SigB

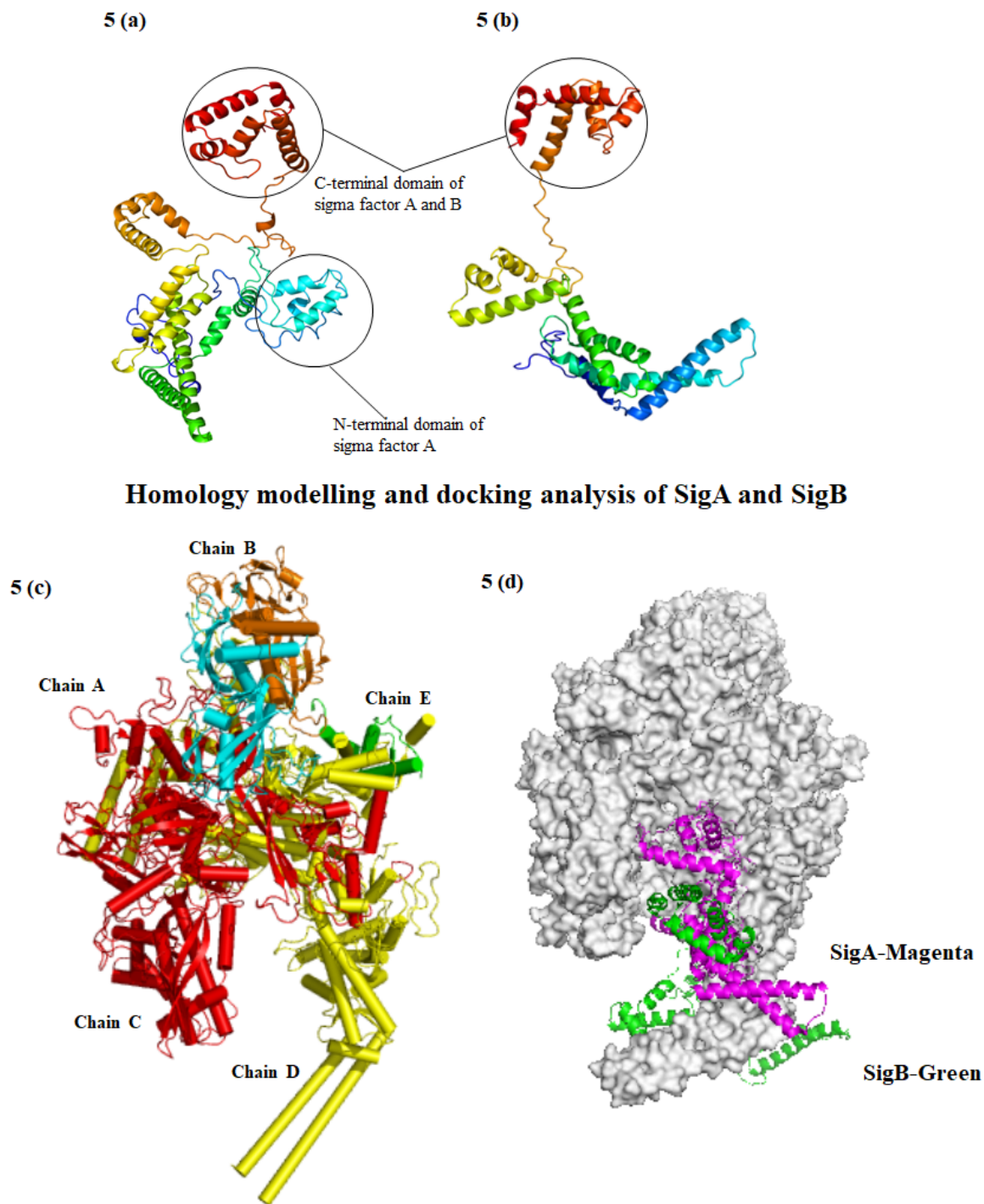


Figure 5

Homology models of a) *M. smegmatis* sigma factor SigA, b) *M. smegmatis* sigma factor SigB, c) structure of the core RNA Polymerase of *M. smegmatis* and d) comparative surface topology of SigA and SigB interactions with core RNA polymerase of *M. smegmatis*.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupFigure.ppt](#)