5-2020

A Putative Cystathionine Beta-Synthase Homolog of Mycolicibacterium smegmatis is Involved in de novo Cysteine Biosynthesis

Saroj Kumar Mahato
University of Arkansas, Fayetteville

A Putative Cystathionine Beta-Synthase Homolog of *Mycolicibacterium smegmatis* is Involved in *de novo* Cysteine Biosynthesis

A thesis submitted in partial fulfillment
of the requirement for the degree of
Master of Science in Cell and Molecular Biology

by

Saroj Kumar Mahato
Purbanchal University
Bachelor of Science in Biotechnology, 2016

May 2020
University of Arkansas

This thesis is approved for recommendation to the Graduate Council.

Young Min Kwon, Ph.D.
Thesis Director

Suresh Thallapuranam, Ph.D.
Committee Member

Inés Pinto, Ph.D.
Committee Member

ABSTRACT

Mycobacteria include serious pathogens of humans and animals. *Mycolicibacterium smegmatis* is a non-pathogenic model that is widely used to study core mycobacterial metabolism. This thesis explores mycobacterial pathways of cysteine biosynthesis by generating and study of genetic mutants of *M. smegmatis*.

Published *in vitro* biochemical studies had revealed three independent routes to cysteine synthesis in mycobacteria involving separate homologs of cysteine synthase, namely CysK1, CysK2, and CysM. However, *in vivo* data were lacking. The *M. smegmatis* genome encodes only a CysM homolog and lacks orthologs for CysK1 or CysK2. The gene that codes for CysM is a part of an operon, *mec*$^{+}$*cysOM* whose products are involved in the cysteine biosynthesis pathway. The *M. smegmatis* genome also encodes a putative cystathionine beta-synthase (CBS) protein that has two domain – an N-terminal domain that shares a weak sequence similarity with CysK1 and a C-terminal domain that is specific to CBS enzymes. CBS is a metabolic enzyme that catalyzes the conversion of homocysteine to cystathionine in all three domains of life (Bacteria, Archaea, and Eukarya).

To dissect the roles of CysM and CBS proteins in cysteine biosynthesis *in vivo*, a series of unmarked knockout mutants and complementation strains of *M. smegmatis* were generated and analyzed phenotypically. Neither the Δ*mec*$^{+}$*cysOM* nor the Δ*cbs* mutants of *M. smegmatis* were auxotrophic for cysteine. However, a Δ*mec*$^{+}$*cysOM_cbs* double mutant of *M. smegmatis* was auxotrophic for cysteine. Genetic complementation of the double mutant using either *cbs* gene or *mec*$^{+}$*cysOM* operon rescued cysteine auxotrophy. Furthermore, the N-terminal CysK1-like domain of the putative CBS was sufficient to rescue cysteine auxotrophy. Thus, these *in vivo* data implicate a role for the putative CBS in cysteine biosynthesis and also suggest that the protein may have dual functions in mycobacteria.

Multidrug-resistant (MDR) strains of *M. tuberculosis*, the causative agent of Tuberculosis (TB), are becoming a global crisis. Mycobacterial sulfur metabolism has emerged as a vital target for developing novel drugs to treat MDR-TB. Our findings reveal a potentially new target in mycobacterial sulfur metabolism relevant to strategic development of novel TB drugs.

ACKNOWLEDGEMENTS

I would like to thank the Cell and Molecular Biology program, the Department of Biological Sciences, and the Graduate School at the University of Arkansas for providing me with the necessary resources to complete my research project. I am grateful to the Department of Biological Sciences for supporting me financially during my time here through research and teaching assistantships.

I am deeply grateful to Dr. Young Min Kwon for chairing my master's thesis committee. He has been a great help in my completion of the program.

I extend my sincere gratitude to my committee members, Dr. Suresh Thallapuranam and Dr. Inés Pinto for their time and support. The knowledge I have gained from interacting with them is greatly valuable to me.

I would also like to thank Dr. Ravi Barabote and Dr. Douglas Rhoads for their laboratory space and equipment without which this project would not have been possible. I am grateful to Dr. Barabote for directing me toward becoming an independent researcher.

Special thanks to my labmate, Jasleen Saini for her help, support, and motivation. I also learned several laboratory techniques from her which was vital to my research progress.

I would like to thank Dr. David McNabb for his kind support and motivation.

I would like to acknowledge the contribution of former and current department staff including Ms. Patty Drane, Ms. Becky Harris, Ms. Jan Warren, and Mr. Mike Belcher. Their work has been greatly helpful in the smooth progress of my graduate program and my research project.

DEDICATION

I would like to dedicate this thesis to my parents, Mr. Madusudan Mahato and Ms. Paramshila Devi Mahato who have made countless sacrifices for me to be able to go to school.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

I. INTRODUCTION

A. Mycobacteria

Mycobacteria (Family: Mycobacteriaceae) are scientifically classified as Domain: Bacteria; Phylum: Actinobacteria; Class: Actinobacteridae; Order: Actinomycetales; Suborder: Corynebacterineae. They are acid-fast, gram-positive bacilli. The prefix “myco-“ means fungus in Greek but does not refer to any genetic relationship with fungi; these bacteria grow in broth culture with fungus-like film (Kerr & Barrett, 1994). Mycobacteria have a unique bacterial capsule made up of mycolic acid covering the cell wall. This capsule helps them avoid phagocytosis by macrophages and also enhances their pathogenesis (Daffe & Etienne, 1999; Frehel et al., 1986; J. Liu et al., 1996).

Mycobacteria include serious and opportunistic pathogens of humans and animals. *Mycobacterium tuberculosis*, the causative agent of tuberculosis (TB) (S. V. Gordon & Parish, 2018), *M. leprae* the causative agent of Leprosy, also known as Hansen’s disease (Gillis, 2014) are serious human pathogens. *M. bovis*, the causative agent of tuberculosis in cattle (Grange & Yates, 1996).

1. Mycobacterium tuberculosis

M. tuberculosis is a serious pathogen of humans. It generally grows in the lungs and commonly causes symptoms like serious cough lasting for three or more weeks, pain in the chest, and coughing up blood and sputum. According to the Centers for Disease Control and Prevention (CDC), approximately 23% of the world’s population is infected by *M. tuberculosis*. Not all infected individuals develop TB disease. An estimated 10 million people developed the disease in 2017 and 1.57 million people died of TB related complications in the same year making it the deadliest infectious disease (MacNeil et al., 2019). Strikingly, over 42% of the TB death occur in South-East Asia. Most of the individuals infected with *M. tuberculosis* do not show any symptoms and carry dormant bacteria in their tissues for decades before developing the active

TB disease. During this latency period, the bacteria do not replicate (Wayne, 1994). So, these bacteria can survive nutrient deficiency in their site of infection. They can also synthesize mycothiol, a compound that can absorb oxidative stress generated by the host macrophages (Buchmeier et al., 2006). This helps them survive in the host organism for decades. When the immunity of the host weakens, for example in case of Acquired Immune Deficiency Syndrome (AIDS), the pathogen becomes active and causes the disease (Chaisson et al., 1987). In 2018, approximately 8.6% of the total TB cases were among the persons living with Human Immunodeficiency Virus (HIV) (MacNeil et al., 2020).

In the 1940s, streptomycin was introduced to treat tuberculosis (Hinshaw et al., 1945). When streptomycin-resistant TB started emerging, Isoniazid in combination with streptomycin proved to be effective (M. L. Cohn et al., 1959). Soon isoniazid-resistant strains started to appear (Selkon et al., 1964). Rifampin was found to be an effective drug against *M. tuberculosis* (Crowle et al., 1988). Then *M. tuberculosis* has also acquired resistance to rifampin (D. L. Cohn et al., 1997; Yuen et al., 1999). In recent years, we have been seeing the emergence of Multidrug-resistant TB (MDR-TB) that are resistant to both isoniazid and rifampin, and Extensively drug-resistant TB (XDR-TB) that are resistant to both isoniazid and rifampin plus any fluoroquinolone and at least one of the three injectable antibiotics: kanamycin, capreomycin, or amikacin (Seung et al., 2015). It is vital to keep up with the race against this pathogen by discovering new drug targets until alternative therapeutic approaches like bacteriophage therapy become established (Azimi et al., 2019).

2. Mycobacterium leprae

Leprosy, on the other hand, is one of the neglected tropical diseases. It manifests in humans as the infection of skins, peripheral nerves, respiratory tract, and eyes. This can lead to permanent sensory and motor impairments including blindness and paralysis. In 1985, 12 million people were estimated to have leprosy (WHO, 2002). Multidrug therapy (MDT), a regimen consisting of

three antibiotics: dapsone, rifampicin, and clofazimine, was recommended by the World Health Organization (WHO) in 1981. This proved to be a “game changer” which brought down the number of reported cases from 5.3 million in 1985 to 3.1 million by 1991 and to 597,000 by 2000 (Smith et al., 2017). However, the number of new cases reported every year has only declined very slowly from 265,661 in 2006 to 210,758 in 2015 (WHO, 2016). Early treatment can prevent permanent disabilities. The incubation period of the infection can last up to 20 years which makes the elimination of the disease difficult (Lastória & de Abreu, 2014). Another major hindrance to studying the biology of *M. leprae* is caused by the fact that it has not been cultured in an artificial growth medium.

3. Mycolicibacterium smegmatis

Mycolicibacterium smegmatis, formerly known as *Mycobacterium smegmatis*, is a non-pathogenic microorganism but has been reported to cause infections. *M.* smegmatis can be found in soil and water sources at high altitudes up to 3500 meters (King et al., 2017). These rapidly growing acid-fast bacteria were first described and reported by Lustgarten in 1884. Alvarez and Tavel later found this species in genital secretion (smegma) and consequently named the organism *Mycobacterium smegmatis* (R. E. Gordon & Smith, 1953). When the genus *Mycobacterium* was emended in 2018, *M. smegmatis* was classified into a new genus, *Mycolicibacterium* (R. S. Gupta et al., 2018). Quantitative structome analysis of *M. smegmatis* revealed that the species was morphologically more similar to *Escherichia coli* than *M. tuberculosis* which supported the introduction of the novel genus, *Mycolicibacterium* (Yamada et al., 2018).

It has commonly been used as a non-pathogenic model organism to study the metabolism in mycobacteria. *M. smegmatis* is a rapidly growing, Biosafety Level 1 organism which makes it a very good model organism for mycobacteria as *M tuberculosis* and *M. leprae* are slow-growing bacteria (R. E. Gordon & Smith, 1953; Singh & Reyrat, 2009). *M. smegmatis* is also easier to

genetically manipulate (Jacobs et al., 1991; Van Kessel & Hatfull, 2008) which makes it easier to study *in vivo* cysteine biosynthesis in mycobacteria (this study). One of the very important factors that make this organism a suitable model for this project is the fact that it can grow well in M9 minimal medium which provides researchers freedom to manipulate the medium and test its growth when certain compounds were the sole sources of sulfur that they might be assimilated into cysteine.

B. Cysteine

Cysteine (Figure 1E) is one of the two sulfur-containing proteogenic amino acids, the other being methionine. The molecular formula of cysteine is given by $COOHCH(NH_2)CH_2SH$ with a thiol group (–SH) serving as the functional group on the side chain of the amino acid. Two cysteine residues in a protein can covalently bond with each other to form a disulfide bond that aids the secondary and tertiary structures of proteins. Cysteine serves as a precursor for the biosynthesis of the second sulfur-containing proteogenic amino acid, methionine (Figure 1B). Cysteine also has antioxidant properties as the thiol group of the amino acid can undergo redox reactions. Cysteine can often be found in the active sites of enzymes participating as a nucleophile (Verma et al., 2016). It serves as a precursor for the biosynthesis of glutathione which protects the cells from oxidative stress in most of the gram-negative aerobic bacteria (Fahey et al., 1978) and eukaryotes (Fernandes et al., 2007). The mycobacterial equivalent of glutathione is called mycothiol (Figure 1D) (Newton et al., 2008). Cysteine forms a part of this compound (Spies & Steenkamp, 1994).

C. Cysteine Biosynthesis

Plants and bacteria can synthesize cysteine *de novo* and then, through the forward transsulfuration pathway, synthesize methionine. They can also synthesize cysteine using methionine as a precursor through the reverse transsulfuration pathway. In fungi and mammals including humans, cysteine is synthesized from methionine via the reverse transsulfuration

pathway (Mudd et al., 1965; Ono et al., 1999; Papet et al., 2019). Methionine is a nutritionally essential amino acid in mammals. Hence, we can derive that we do not have any pathway for assimilating inorganic sulfur into amino acids in humans. For this reason, components of the *de novo* biosynthesis of cysteine are viewed as a good field to look for drug targets against pathogenic bacteria. Figure 2 shows the predicted pathways for cysteine biosynthesis in mycobacteria.

1. Reverse Transsulfuration Pathway from Methionine

The reverse transsulfuration pathway starts with the conversion of methionine to S-adenosylmethionine (SAM). This is an ATP requiring reaction catalyzed by the enzyme methionine adenosyltransferase (Berger & Knodel, 2003). Methyltransferase, then, transfers the methyl group of SAM to a substrate to synthesize S-adenosylhomocysteine (SAH) (A. Gupta et al., 2001; Im et al., 2016) which is then converted to homocysteine by S-adenosylhomocysteine hydrolase (SAHH) (Singhal et al., 2013). Cystathionine β-synthase (CBS) can then combine homocysteine with serine to form cystathionine (Kery et al., 1994). Cystathionine can be broken down to α-ketoglutarate and cysteine by cystathionine γ-lyase (CGL) (Wheeler et al., 2005). This pathway for the biosynthesis of cysteine is not only found in mycobacteria but it is also common among all the three domains of life: Bacteria, Archaea, and Eukarya.

2. *De novo* Cysteine Biosynthesis Catalyzed by CysK1

Plants and many species of bacteria can assimilate inorganic sulfur into cysteine. For the *de novo* biosynthesis of cysteine, bacteria including mycobacteria take up sulfur from the environment in the form of sulfate (SO_4^{2-}) through ATP-binding channel (ABC) transporters (Wooff et al., 2002). Intracellular sulfate can then be activated to form adenosine 5'-phosphosulfate (APS) by the enzyme ATP sulfurylase. APS reductase reduced this activated sulfate to sulfite (SO_3^{2-}). Sulfite can further be reduced to sulfide (S^{2-}) (Pinto et al., 2007). CysK1, an *O*-acetyl-L-serine sulfhydrylase (OASS), can now transfer this sulfide to *O*-acetyl-L-

serine (OAS) to synthesize cysteine and release acetyl as a byproduct. Thus, CysK1 is a cysteine synthase. Homologs of the gene (*cysK1*) coding for this enzyme can be found in the genomes of plants and both gram-negative and gram-positive bacteria including *M. tuberculosis, M. bovis,* and *M. marinum*. Surprisingly, *M. smegmatis* genome appears to lack this gene.

3. *De novo* Cysteine Biosynthesis Catalyzed by CysK2

M. tuberculosis genome also codes for a second enzyme CysK2, an *O*-phospho-L-serine sulfhydrylase (OPSS), that can transfer the sulfide to an *O*-phospho-L-serine (OPS) to form cysteine, releasing a phosphate as a byproduct. This is a recently discovered pathway in *M. tuberculosis* (Steiner et al., 2014). Again, homologs of the gene (*cysK2)* coding for this enzyme can be found in the genomes of *M. bovis* and *M. marinum.* This gene is not found in *M. smegmatis* genome.

4. *De novo* Cysteine Biosynthesis Catalyzed by CysM

Actinobacteria have been reported to exhibit a yet another route to synthesize cysteine *de novo.* This pathway involves three proteins including a sulfur carrying protein, CysO. A sulfurtransferase transfers a sulfide group onto the C-terminus of the CysO protein (Burns et al., 2005). Cysteine synthase B or CysM enzyme condenses thiocarboxylated CysO with *O*-phospho-L-serine to form a CysO-cysteine adduct by displacing the phosphate group of OPS (Ågren et al., 2008; O'Leary et al., 2008). Cysteine is then cleaved from CysO by a peptidase enzyme known as mec^{+}. Genes coding for these three proteins, *mec*$^{+}$, *cysO*, and *cysM* are found in a cluster in *M. tuberculosis.* Homologs of these genes are found in operons in other mycobacteria too including *M. smegmatis* (Jurgenson et al., 2008).

II. METHODS

A. Bacterial strains and growth conditions

Escherichia coli DH5α cells were grown in Luria Bertani (LB) broth or agar media. *M. smegmatis* mc^2155 and the generated mutant strains of *M. smegmatis* were routinely grown on Middlebrook 7H10 agar supplemented with 0.5% (v/v) glycerol and in Middlebrook 7H9 supplemented with 0.2% (v/v) glycerol or tryptic soy agar (TSA), or Tryptic soy broth (TSB) media. Liquid mycobacterial growth media were supplemented with 0.05% (v/v) Tween 80. Media were supplemented with 100 µg/mL hygromycin B, and/or 25 µg/mL kanamycin sulfate when needed. *E. coli* DH5α was grown at 37 ^{O}C and 250 rpm (liquid). *M. smegmatis* strains were grown at 37 ^{O}C and 150 rpm (liquid). Media and other chemicals were purchased from Becton, Dickinson and Company (BD) and MilliporeSigma respectively.

B. DNA manipulation

All primers used in this study are listed in Table 3. Knockout mutants of *M. smegmatis* were generated through an allelic exchange strategy as described previously with some modifications on the Xer-cise technique (Figure 3) (Cascioferro et al., 2010; Van Kessel & Hatfull, 2007). A hygromycin-resistance gene-containing 1.1 kb region of pYUB28b (hyg^R) (Bashiri et al., 2010) was amplified using primers containing *M. tuberculosis* putative *dif* sequence. To knockout the *mec*$^+$*cysOM* operon, the upstream and the downstream regions (~600 base pairs each) of the operon were amplified from *M. smegmatis* chromosome with the *dif* sequence introduced at the ends that were to be fused with the hygromycin-resistance gene amplicon. *Dif* sites are 28 bp regions with identical sequences that can recombine to excise the region of DNA between them. Mycobacterial genomes contain *dif*-recombinase that can catalyze this recombination. Here, I utilized this feature of *dif* sites to excise the selective marker to generate unmarked knockout mutants. A fusion PCR using the forward primer of the upstream region and the reverse primer of the downstream region resulted in the "upstream-*dif*-hyg^R-*dif*-downstream" cassette which

was then electroporated into *M. smegmatis* cells containing pJV53 (Van Kessel & Hatfull, 2007). PJV53 plasmid encodes mycobacteriophage genes gp60-61 that facilitate site-specific recombination. Transformants were selected on TSA plates supplemented with hygromycin and kanamycin. Unmarked knockout mutants were obtained through a passage on non-selective plate followed by replica plating on selective (supplemented with hygromycin) and non-selective media. Knockout mutants were cured of the pJV53 plasmids through a passage on non-selective medium and then replica plating on selective (supplemented with kanamycin) and non-selective media. The same approach was followed to generate unmarked Δ*cbs* mutant. The Δ*cbs* mutation was also introduced in pJV53 containing *M. smegmatis* Δ*mec*$^{+}$*cysOM* mutant to result in Δ*mec*$^{+}$*cysOM_cbs* double mutant. The knockout mutations were confirmed using PCR (Figure 7).

For complementation, the *mec*$^{+}$*cysOM* operon amplified from the *M. smegmatis* chromosome using the primers containing the NdeI and NheI restriction sites at the 5' termini of the forward and reverse primers, respectively. The region of pJV53 encoding the two phage recombinases, gp60 and gp61 was excised using NdeI and NheI restriction enzymes (Figure 4A). The PCR amplicon was restriction-digested using the same restriction enzymes and then ligated into the pJV53 vector to result in pJmeccysOM (Figure 5A). The same approach was followed to construct the plasmid pJcbs (Figure 4B) for the cloning of the *cbs* gene. To construct pJcbs(1:311) (Figure 5B), the N-terminal domain coding region, the first 933 bases of the DNA was amplified. A stop codon was introduced in the amplicon by adding a complementary sequence of a stop codon in the reverse primer (P17). This amplicon was restriction-digested and ligated into the vector using the method described for the construction of the plasmid pJmeccysOM. Plasmids thus constructed were electroporated into kanamycin-sensitive cells of the unmarked knockout mutants. The complemented mutants were confirmed using PCR. All plasmids used in this study are listed in Table 1. The resulting strains are included in Table 2.

pYUB28b was a gift by Ted Baker (Addgene plasmid #37277) (Bashiri et al., 2010). pJV53 was a gift by Graham Hatfull (Addgene plasmid #26904) (Van Kessel & Hatfull, 2007). Enzymes were purchased from New England Biolabs (NEB).

C. Sulfur Utilization Assay

Sulfur utilization experiments were carried out in sulfur-free M9 (BD) medium (47.76 mM disodium phosphate, 22.04 mM monopotassium phosphate, 8.56 mM sodium chloride, 18.69 mM ammonium chloride, 0.2% (v/v) glycerol, 1 mM $MgCl_2$ and 0.1 mM $CaCl_2$) supplemented with a sulfur-containing compound: 2 mM $MgSO_4$, 1 mM Na_2S, 60 µM methionine, 60 µM homocysteine, 60 µM cystathionine, or 30 µM cysteine. Strains of *M. smegmatis* were grown in triplicates in M9 medium containing cysteine to log phase. The cells were washed 3x with H_2O and then resuspended in M9 medium containing $MgSO_4$. The cultures were then diluted to an OD_{600} of 0.020. OD_{600} was measured for all the cultures after every 6 hours for 48 hours unless otherwise specified.

D. Bioinformatics Analysis

The protein sequences were obtained from NCBI (Reference sequence: NC_008596.1 for *M. smegmatis* and Reference sequence: NC_000962.3 for *M. tuberculosis*). Multiple sequence alignment of the protein sequences was carried out using Muscle on Lasergene MegAlign Pro software from DNASTAR. The alignment image was created using ESPript (Robert & Gouet, 2014). A phylogenetic tree was drawn using Neighbor-joining method on MEGA X (Kumar et al., 2018; Saitou & Nei, 1987; Zuckerkandl & Pauling, 1965).

III. RESULTS

A. Confirmation of mutants

The mutants generated in this study were confirmed using PCR. Colonies were used as templates for direct PCR to amplify a length of the chromosome that included the allelic exchange region and a 98 bp stretch downstream of the region to confirm the allelic exchange at the target site. A PCR product of approximately 1.3 kb in Lane 2 amplified from Hyg^S colonies of *M. smegmatis* Δ*mec*$^+$*cysOM* compared to that of approximately 2.7 kb in Lane 1 (Figure 7A), amplified from the wild type strain indicated the deletion of the operon.

For the confirmation of the deletion of *cbs* gene, the stretch of DNA including allelic exchange region and 76 bp upstream was amplified from the chromosome. A PCR product of approximately 1.4 kb in Lane 8, amplified from Δ*cbs* strain was observed while the wild type strain gave a product of 2.5 kb in Lane 6 (Figure 7B). The shorter bands for the PCR amplification of the operon region and the gene region compared to those from the wild type strains in Figure 7A and Figure 7B indicated the deletion of both regions.

For complemented mutants, using a primer specific to the insert (P13) and the other specific to the vector (P18), a PCR product of approximately 2.3 kb was obtained with Δ*mec*$^+$*cysOM::mec*$^+$*cysOM* (Lane 3) and *M. smegmatis* Δ*mec*$^+$*cysOM_cbs::mec*$^+$*cysOM* (Lane 4) in Figure 8A. Using a primer specific to the insert (P15) and a primer specific to the vector (P18), a PCR product of approximately 1.9 kb was obtained with Δ*cbs::cbs* (Lane 8) and Δ*mec*$^+$*cysOM_cbs::cbs* (Lane 9) while a PCR product of approximately 1.3 kb was obtained with Δ*mec*$^+$*cysOM_cbs::cbs*(1:311) (Lane 10) in Figure 8B.

B. Sulfur Utilization Assay

1. $MgSO_4$

Figure 9 shows the growth curve of various strains of *M. smegmatis* in M9 with Magnesium sulfate as the sole sulfur source. All strains start out at an OD_{600} of ~0.015. The wild type strain

has the best growth rate among all the strains throughout the 48 hours of observation. The Δ*mec*$^{+}$*cysOM* single knockout mutant grows as well as the wild type strain until the OD_{600} of ~0.5 but falls behind afterwards. The Δ*cbs* single mutant grows slower than the wild type strain and the Δ*mec*$^{+}$*cysOM* strain until the OD_{600} of ~1.0 and then grows faster than the Δ*mec*$^{+}$*cysOM* strain. The Δ*mec*$^{+}$*cysOM_cbs* double mutant does not appear to grow at all in this medium. The complemented single knockout mutants show similar growth to the non-complemented knockout strains throughout the observation period. Complementation with either *cbs* or *mec*$^{+}$*cysOM* in the double mutant appears to restore the ability to grow in this medium. The Δ*mec*$^{+}$*cysOM_cbs::cbs*(1:311) does grow in this medium but the growth rate is slower than the wild type strain and any other strains that grow in this medium.

2. Na_2S

Figure 10 shows the growth rate comparison of the various strains of *M. smegmatis* in M9 with Sodium sulfide as the sole sulfur source. Just like in M9 medium with $MgSO_4$ as the sole sulfur source, the wild type strain shows the best growth rate among all the strains in the minimal medium with sodium sulfide as the sole sulfur source. The single knockout mutants and the complemented single knockout mutants appear to grow as well as the wild type strain until the OD_{600} of ~0.8 and then start to grow slower. The Δ*mec*$^{+}$*cysOM_cbs* double mutant does not grow at all in this medium. The double knockout mutant complemented with either *mec*$^{+}$*cysOM* or *cbs* grow as well as their corresponding single mutants, Δ*cbs* or Δ*mec*$^{+}$*cysOM*, respectively. The Δ*mec*$^{+}$*cysOM_cbs::cbs*(1:311) strain does grow in this medium but has a very slow growth rate.

3. Methionine

Figure 11 shows the comparison of various strains of *M. smegmatis* in M9 minimal medium with methionine as the sole sulfur source. In this medium, the Δ*mec*$^{+}$*cysOM* mutant appears to grow better than or as well as the wild type strain while the Δ*cbs* mutant grows much slower than the

wild type strain. Complemented strains of the single knockouts show similar growth rate as that of the wild type strain. The Δ*mec*$^{+}$*cysOM_cbs* double mutant does not appear to grow at all in this medium. Growth of the double mutant is restored when complemented with *cbs*. The *mec*$^{+}$*cysOM_cbs::mec*$^{+}$*cysOM* shows a growth pattern similar to the Δ*cbs* mutant. Interestingly, Δ*mec*$^{+}$*cysOM_cbs::cbs*(1:311) strain grows much better than the Δ*cbs* mutant in this medium.

4. Homocysteine

Growth curves of the different strains of *M. smegmatis* in M9 with homocysteine as the sole sulfur source are shown in Figure 12. Just like in the minimal medium with methionine as the sole sulfur source, the Δ*mec*$^{+}$*cysOM* mutant appears to grow better than or as well as the wild type strain while the Δ*cbs* mutant grows at a much slower rate than the wild type strain in M9 with homocysteine as the sole sulfur source. The single knockout mutants complemented with the missing gene/operon grow as well as the wild type strain. The Δ*mec*$^{+}$*cysOM_cbs* double mutant does not grow at all in this medium. Growth of the double mutant is restored when complemented with *cbs*. The *mec*$^{+}$*cysOM_cbs::mec*$^{+}$*cysOM* shows a growth pattern similar to the Δ*cbs* mutant. Similar to what was observed in M9 with methionine, Δ*mec*$^{+}$*cysOM_cbs::cbs*(1:311) strain grows much better the Δ*cbs* mutant in this medium.

5. Cystathionine

Figure 13 shows the growth curves of the various strains of *M. smegmatis* in M9 minimal medium with cystathionine as the only source of sulfur. In the minimal medium, the wild type strain as well as all the mutants were able to grow including the Δ*mec*$^{+}$*cysOM_cbs* double mutant. Mutant strains do however fall behind the wild type strain after an OD_{600} of ~0.6. The growth rates among the knockout and complemented knockout mutants do not show any striking difference.

6. Cysteine

Growth curves of *M. smegmatis* wild type and mutant strains are shown in Figure 14. Just like in the medium with cystathionine as the sole sulfur source, all the mutants and the wild type strain are able to grow in this medium. As expected, the Δ*mec*$^{+}$*cysOM_cbs* double mutant is able to grow when cysteine is provided in the medium. The wild type strain grew much faster than the knockout and complemented knockout mutant strains with one exception of the Δ*mec*$^{+}$*cysOM_cbs::cbs*(1:311) strain which grew as well as the wild type strain in this medium.

C. Bioinformatic Analysis

NCBI conserved domain search for *M. smegmatis* CBS protein against the conserved domain database returned revealed that the protein has two conserved domains. The N-terminal domain extending from Ile^{1} to Lys^{304} matched with a cysteine synthase domain. Protein BLAST (NCBI) revealed 60% positives between the N-terminal domain of *M. smegmatis* CBS (1:304) and *M. tuberculosis* CysK1. The protein sequences of CysM and CBS of *M. smegmatis* were aligned with those of CysK1, CysK2 and CysM of *M. tuberculosis* whose structural and biochemical information were described previously (Ågren et al., 2008; Schnell et al., 2007; Steiner et al., 2014). A phylogenetic tree was constructed for the alignment using Neighbor-joining method (Figure 15). CBS proteins from the two species were in a cluster with the CysK1 while, CysK2 appeared to be the least similar protein among the analyzed cysteine synthases and cystathionine beta-synthase proteins. Multiple sequence alignment (Figure 16) showed *M. smegmatis* CBS has K^{44}, ^{181}GTGGT^{185}, S^{269}, and N^{74} aligned with the same residues in CysK1 that help in covalent boding with the Pyridoxal phosphate cofactor, hydrogen bonding with the phosphate group, hydrogen bonding with the N-1 atom of the pyrimidine ring, and hydrogen bonding with the 3'-hydroxyl group of the cofactor. The *M. smegmatis* CBS protein also contains ^{71}TSGNT^{75}, D^{296}, and G^{225} which aligned with the residues in the active site of the CysK1 holoenzyme.

IV. DISCUSSION

Growth curves in sulfur-free M9 minimal medium supplemented with one of the sulfur-containing compounds ($MgSO_4$, Na_2S, methionine, homocysteine, cystathionine and cysteine) were obtained for all the knockout mutants and complemented knockout mutants generated in this study along with the wild type strain of *M. smegmatis* (Figures 9 – 14 respectively). Since the mutations were made in the genes involved in cysteine biosynthesis, all the strains should be able to grow in the presence of cysteine. It was vital to make sure that all the strains showed growth in M9 medium with cysteine. In this medium, the growth rates of all the mutants were close to that of the wild type strain during the log phase (OD_{600} ~0.6) which indicates that there was not any significant growth defect in any of the mutants.

To analyze the growth of strains in media with different sulfur sources, it is important to consider where these sulfur sources are placed in the biosynthetic pathway of cysteine. In the reverse transsulfuration pathway, cystathionine is the product of cystathionine β-synthase. The enzyme that catalyzes the conversion of cystathionine to cysteine, cystathionine γ-lyase, was not mutated in any of the strains. Therefore, all the strains must also be able to grow in M9 medium with cystathionine as the sole source of sulfur. As expected, all the strains grew in this medium.

The gene MSMEG_5270 (*cbs*) of *M. smegmatis* has been annotated to encode cystathionine β-synthase (CBS). A protein BLAST with the sequence of CBS of another actinobacterium, *Streptomyces venezuelae*, shows a 69% identity and 79% similarity. Chang & Vining, in 2002, described the protein in *S. venezuelae* to catalyze the conversion of homocysteine to cystathionine. When homocysteine is the sole source of sulfur in the M9 medium, the Δ*mec*$^+$*cysOM* strain of *M. smegmatis* can utilize the reverse transsulfuration system to synthesize cystathionine from homocysteine and then cysteine from cystathionine with cystathionine β-synthase and cystathionine γ-lyase to catalyze the reactions respectively. However, the deletion of both *mec*$^+$*cysOM* operon and *cbs* gene did not allow the cells to grow

in this medium as expected. Interestingly, the Δ*cbs* mutant was able to grow in this medium which may be explained by a possible degradation of homocysteine to release inorganic sulfur that can be assimilated into cysteine via the CysM catalyzed biosynthesis (Reisch et al., 2011). Another possible explanation would be the catalysis of the homocysteine to cystathionine reaction by the intact CysM enzyme. There is however no experimental evidence to support this hypothesis. CysM protein sequence shows 38% identity and 53% similarity with that of CBS, while also missing the C-terminal domain of the CBS enzyme. The fates of homocysteine have not been explored extensively in mycobacteria. It would be interesting to track the route sulfur takes in the Δ*cbs* mutant by feeding it with ^{34}S-enriched homocysteine and unenriched cysteine to see whether there is a reverse transsulfuration activity taking place in this strain.

Methionine is upstream of homocysteine as the sulfur donor in the biosynthesis of cysteine via the reverse transsulfuration pathway. Since the genes encoding for the enzymes involved in the conversion of methionine to homocysteine were intact in all the strains, consistent with the expectation, they showed similar phenotypes in M9 medium with methionine as the sole sulfur source as they did with homocysteine as such.

Since the *M. smegmatis* genome encodes neither CysK1 nor CysK2 cysteine synthases, the only cysteine synthase known to be present in *M. smegmatis* is CysM encoded by the gene *cysM* which is a part of an operon shared with *mec*$^+$ and *cysO*. The wild type strain, like *M. tuberculosis* (Ågren et al., 2008), is expected to assimilate sulfide (S^{2-}) into cysteine via the CysM catalyzed reaction while the strain with the *mec*$^+$*cysOM* operon knocked out would not be able to grow in an M9 medium with sulfide as the sole sulfur source. Surprisingly, the Δ*mec*$^+$*cysOM* strain was able to grow in this medium indicating that there must be at least one alternative pathway for the *de novo* biosynthesis of cysteine in this organism. The Δ*mec*$^+$*cysOM_cbs* did not grow at all in this medium indicating that CBS is involved in the *de novo* biosynthesis of cysteine in *M. smegmatis.* There are two potential ways this enzyme may

be involved in cysteine biosynthesis. The first potential route followed by *M. smegmatis* might be through its ability to possibly synthesize homocysteine by direct sulfhydrylation of *O*-acetyl-L-homoserine with inorganic sulfide as the sulfur donor. This route of homocysteine biosynthesis has been reported in methanogenic archaea (Allen et al., 2015; Y. Liu et al., 2010), *Saccharomyces cerevisiae* (Thomas & Surdin-Kerjan, 1997), *Leptospira meyeri* (Belfaiza et al., 1998), and also in *Corynebacterium glutamicum* (Lee & Hwang, 2003). It is important to note that *C. glutamicum* is from the same taxonomic order of Actinomycetales, as is *M. smegmatis*. The enzyme catalyzing this reaction is *O*-acetylhomoserine sulfhydrylase (OAHS). The protein sequence of the enzyme of *M. smegmatis* shows a 69% identity and 73% similarity with the OAHS of *C. glutamicum*. Thus, it is possible that this enzyme assimilates inorganic sulfide to synthesize homocysteine which is then converted to cystathionine by CBS and then via the CGL-catalyzed reaction, it is converted to cysteine making the Δ*mec*$^{+}$*cysOM* strain of *M. smegmatis* able to grow in M9 with sulfide as the sole sulfur source. The second potential route of *de novo* cysteine biosynthesis involving CBS in the Δ*mec*$^{+}$*cysOM* strain may be followed as CBS being able to catalyze the assimilation of sulfide to *O*-acetyl-L-serine to generate cysteine. CBS has two domains: The N-terminal domain shows a 44% identity and 60% similarity with the *O*-acetylserine sulfhydrylase (CysK1) encoded by the gene *cysK1* of *M. tuberculosis*, and the C-terminal domain is highly conserved among the cystathionine beta-synthases of organisms from all the three domains of life: Bacteria, Archaea, and Eukarya. The C-terminal domain has been reported to act as a regulatory domain. This domain was found to be non-essential in the biosynthesis of cysteine as the Δ*mec*$^{+}$*cysOM_cbs* complemented with the N-terminal domain of CBS was able to grow in M9 with sulfide as the sole sulfur source. Protein sequence alignment revealed, as discussed in the results, that the *M. smegmatis* CBS protein sequence has amino acid residues aligned with the ones in *M. tuberculosis* CysK1 that are involved in binding with the pyridoxal phosphate cofactor. Same is the case with the residues involved in the catalytic activity of the CysK1 enzyme. Also, the *M. smegmatis* genome encodes the enzymes serine *O*-

acetyltransferase (product of the gene *cysE*) which catalyzes the biosynthesis of *O*-acetyl-L-serine. The presence of an *O*-acetylserine sulfhydrylase would explain the significance of this enzyme *O*-acetyltransferase and its product in *M. smegmatis*. For these reasons, the hypothesis that *M. smegmatis* CBS may be acting as a cysteine synthase (*O*-acetylserine sulfhydrylase) cannot be rejected without further investigation.

One way to determine which route CBS takes to synthesize cysteine *de novo* in *M. smegmatis* would be to knock out the gene MSMEG_2394 coding for cystathionine γ-lyase (CGL) in the Δ*mec*$^{+}$*cysOM* background and then test this strain's ability to grow in M9 medium with sulfide as the sole sulfur source. Since a functional CGL catalyzes the conversion of cystathionine to cysteine, this strain will be unable to synthesize cysteine via both CysM-catalyzed biosynthesis and reverse transsulfuration pathway. If the double mutant strain is able to grow in this medium, it will support the hypothesis that CBS may be acting as a cysteine synthase using sulfide and *O*-acetylserine as its substrates. If the strain fails to grow in this medium, it will support that CBS may only be catalyzing the conversion of homocysteine to cystathionine and that *M. smegmatis* must have a pathway to directly assimilate sulfide into homocysteine catalyzed by an *O*-acetylhomoserine sulfhydrylase.

Another way to determine the route taken by CBS in *de novo* biosynthesis of cysteine would be to use ^{34}S-labeled sulfide in the growth of Δ*mec*$^{+}$*cysOM* strain of *M. smegmatis.* This strain would be fed with labeled sulfide and unlabeled homocysteine. If CBS has an *O*-acteylserine sulfhydrylase activity, we should see a higher percentage of labeled cysteine vs labeled cystathionine. However, if CBS is merely converting homocysteine to cystathionine, then, cells would use up unlabeled homocysteine for the biosynthesis cysteine while more homocysteine is synthesized by the *O*-acetylhomoserine sulfhydrylase. This would result in a higher percentage of homocysteine being labeled than cysteine.

Sulfate is reduced to sulfide via the activities of APS sulfurylase, APS reductase, and Sulfide reductase enzymes in *M. smegmatis*. The genes coding for these enzymes were not mutated in any of the strains used in this study. Hence, using sulfate as the sulfur donor in M9 medium should yield similar growth curves for any strain that they exhibited in M9 with sulfide as the sole sulfur source. Thus, the results for all the strains of *M. smegmatis* were consistent with this expectation.

Based on the observation and above discussion, CBS appears to be involved in *de novo* biosynthesis of cysteine in *M. smegmatis*, either in the transsulfuration pathway where homocysteine is synthesized by assimilation of sulfide onto *O*-acetylhomoserine or through the assimilation of sulfide onto *O*-acetylserine to synthesize cysteine directly. So far, only three *de novo* cysteine biosynthesis pathways have been described in *M. tuberculosis*: CysK1, CysK2, and CysM being the major enzymes in those pathways. However, these enzymes have been described largely based on biochemical studies. Not much has been explored *in vivo*. My data supports that CysM is involved in cysteine biosynthesis in *M. smegmatis*. This project also introduces a new pathway for cysteine biosynthesis in *M. smegmatis* which employs CBS as an enzyme catalyzing one of the steps. Considering that *M. smegmatis* CBS sequence shows an 86% identity and 93% similarity with *M. tuberculosis* CBS sequence, it is highly likely that the enzyme has the same function in *M. tuberculosis.* This means that there may be a yet another route for the *de novo* cysteine biosynthesis in *M. tuberculosis* different from the three previously described pathways. Certainly, experimental evidence is necessary to verify this hypothesis in *M. tuberculosis* to make any conclusions. If verified to be true, this would introduce a novel target for the development of anti-tuberculosis drugs.

V. REFERENCES

Ågren, D., Schnell, R., Oehlmann, W., Singh, M., & Schneider, G. (2008). Cysteine synthase (CysM) of Mycobacterium tuberculosis is an O-phosphoserine sulfhydrylase: Evidence for an alternative cysteine biosynthesis pathway in Mycobacteria. *Journal of Biological Chemistry*, *283*(46), 31567–31574. https://doi.org/10.1074/jbc.M804877200

Allen, K. D., Miller, D. V., Rauch, B. J., Perona, J. J., & White, R. H. (2015). Homocysteine is biosynthesized from aspartate semialdehyde and hydrogen sulfide in methanogenic archaea. *Biochemistry*, *54*(20), 3129–3132. https://doi.org/10.1021/acs.biochem.5b00118

Azimi, T., Mosadegh, M., Nasiri, M. J., Sabour, S., Karimaei, S., & Nasser, A. (2019). Phage therapy as a renewed therapeutic approach to mycobacterial infections: A comprehensive review. *Infection and Drug Resistance*, *12*, 2943–2959. https://doi.org/10.2147/IDR.S218638

Bashiri, G., Rehan, A. M., Greenwood, D. R., Dickson, J. M. J., & Baker, E. N. (2010). Metabolic engineering of cofactor F 420 production in mycobacterium smegmatis. *PLoS ONE*, *5*(12), 1–10. https://doi.org/10.1371/journal.pone.0015803

Belfaiza, J., Martel, A., Margarita, D., & Saint Girons, I. (1998). Direct sulfhydrylation for methionine biosynthesis in Leptospira meyeri. *Journal of Bacteriology*, *180*(2), 250–255. https://doi.org/10.1128/jb.180.2.250-255.1998

Berger, B. J., & Knodel, M. H. (2003). Characterisation of methionine adenosyltransferase from Mycobacterium smegmatis and M. tuberculosis. *BMC Microbiology*, *3*, 1–13. https://doi.org/10.1186/1471-2180-3-12

Buchmeier, N. A., Newton, G. L., & Fahey, R. C. (2006). A mycothiol synthase mutant of Mycobacterium tuberculosis has an altered thiol-disulfide content and limited tolerance to stress. *Journal of Bacteriology*, *188*(17), 6245–6252. https://doi.org/10.1128/JB.00393-06

Burns, K. E., Baumgart, S., Dorrestein, P. C., Zhai, H., McLafferty, F. W., & Begley, T. P. (2005). Reconstitution of a new cysteine biosynthetic pathway in mycobacterium tuberculosis. *Journal of the American Chemical Society*, *127*(33), 11602–11603. https://doi.org/10.1021/ja053476x

Cascioferro, A., Boldrin, F., Serafini, A., Provvedi, R., Palù, G., & Manganelli, R. (2010). Xer site-specific recombination, an efficient tool to introduce unmarked deletions into mycobacteria. *Applied and Environmental Microbiology*, *76*(15), 5312–5316. https://doi.org/10.1128/AEM.00382-10

Chaisson, R. E., Schecter, G. F., Theuer, C., Rutherford, G. W., Echenberg, D. F., Hopewell, P. C., & others. (1987). Tuberculosis in patients with the acquired immunodeficiency syndrome. *Am Rev Respir Dis*, *136*, 570–574.

Chang, Z., & Vining, L. C. (2002). Biosynthesis of sulfur-containing amino acids in Streptomyces venezuelae ISP5230: Roles for cystathionine β-synthase and transsulfuration. *Microbiology*, *148*(7), 2135–2147. https://doi.org/10.1099/00221287-148-7-2135

Cohn, D. L., Bustreo, F., & Raviglione, M. C. (1997). Drug-Resistant Tuberculosis: Review of the Worldwide Situation and the WHO/IUATLD Global Surveillance Project. *Clinical Infectious Diseases*, *24*(Supplement_1), S121–S130. https://doi.org/10.1093/clinids/24.supplement_1.s121

Cohn, M. L., Middlebrook, G., & Russell, W. F. (1959). Combined drug treatment of tuberculosis. I. Prevention of emergence of mutant populations of tubercle bacilli resistant to both streptomycin and isoniazid in vitro. *The Journal of Clinical Investigation*, *38*(8), 1349–1355.

Crowle, A. J., Elkins, N., & May, M. H. (1988). Effectiveness of ofloxacin against Mycobacterium tuberculosis and Mycobacterium avium, and rifampin against M. tuberculosis in cultured human macrophages. *The American Review of Respiratory Disease*, *137*(5), 1141–1146. https://doi.org/10.1164/ajrccm/137.5.1141

Daffe, M., & Etienne, G. (1999). The capsule of Mycobacterium tuberculosis and its implications for pathogenicity. *Tubercle and Lung Disease : The Official Journal of the International Union against Tuberculosis and Lung Disease*, *79*(3), 153–169. https://doi.org/10.1054/tuld.1998.0200

Fahey, R. C., Brown, W. C., Adams, W. B., & Worsham, M. B. (1978). Occurrence of glutathione in bacteria. *Journal of Bacteriology*, *133*(3), 1126–1129. https://doi.org/10.1128/jb.133.3.1126-1129.1978

Fernandes, P. N., Mannarino, S. C., Silva, C. G., Pereira, M. D., Panek, A. D., & Eleutherio, E. C. A. (2007). Oxidative stress response in eukaryotes: Effect of glutathione, superoxide dismutase and catalase on adaptation to peroxide and menadione stresses in Saccharomyces cerevisiae. *Redox Report*, *12*(5), 236–244. https://doi.org/10.1179/135100007X200344

Frehel, C., Ryter, A., Rastogi, N., & David, H. (1986). The electron transparent zone in phagocytized Mycobacterium avium and other Mycobacteria Formation, persistence and role in bacterial survival. *Annales de l'Institut Pasteur / Microbiologie*, *137*(1), 239–257. https://doi.org/10.1016/S0769-2609(86)80115-6

Gillis, T. P. (2014). Mycobacterium leprae. In *Molecular Medical Microbiology: Second Edition*. https://doi.org/10.1016/B978-0-12-397169-2.00093-7

Gordon, R. E., & Smith, M. M. (1953). Rapidly growing, acid fast bacteria. I. Species' descriptions of Mycobacterium phlei Lehmann and Neumann and Mycobacterium smegmatis (Trevisan) Lehmann and Neumann. *Journal of Bacteriology*, *66*(1), 41–48.

Gordon, S. V., & Parish, T. (2018). Microbe profile: Mycobacterium tuberculosis: Humanity's deadly microbial foe. *Microbiology (United Kingdom)*, *164*(4), 437–439. https://doi.org/10.1099/mic.0.000601

Grange, J. M., & Yates, M. D. (1996). Guidelines for speciation within the Mycobacterium tuberculosis complex. *World Health Organization*, *2*, 1–23. https://doi.org/10.1111/j.1440-1754.2011.02214.x

Gupta, A., Kumar, P. H., Dineshkumar, T. K., Varshney, U., & Subramanya, H. S. (2001). Crystal structure of Rv2118c: An AdoMet-dependent methyltransferase from Mycobacterium tuberculosis H37Rv. *Journal of Molecular Biology*. https://doi.org/10.1006/jmbi.2001.4935

Gupta, R. S., Lo, B., & Son, J. (2018). Phylogenomics and Comparative Genomic Studies Robustly Support Division of the Genus Mycobacterium into an Emended Genus Mycobacterium and Four Novel Genera. *Frontiers in Microbiology*, *9*(67), 1–41. https://doi.org/10.3389/fmicb.2018.00067

Hinshaw, Hc., Feldman, W. H., & others. (1945). Streptomycin in Treatment of Clinical Tuberculosis: a Preliminary Report. *Proceedings of Staff Meetings of the Mayo Clinic*, *20*(18), 313–318.

Im, H. N., Kim, H. S., An, D. R., Jang, J. Y., Kim, J., Yoon, H. J., Yang, J. K., & Suh, S. W. (2016). Crystal structure of Rv2258c from Mycobacterium tuberculosis H37Rv, an S-adenosyl-l-methionine-dependent methyltransferase. *Journal of Structural Biology*, *193*(3), 172–180. https://doi.org/10.1016/j.jsb.2016.01.002

Jacobs, W. R., Kalpana, G. V., Cirillo, J. D., Pascopella, L., Snapper, S. B., Udani, R. A., Jones, W., Barletta, R. G., & Bloom, B. R. (1991). [25] Genetic systems for mycobacteria. *Methods in Enzymology*, *204*, 537–555. https://doi.org/10.1016/0076-6879(91)04027-L

Jurgenson, C. T., Burns, K. E., Begley, T. P., & Ealick, S. E. (2008). Crystal structure of a sulfur carrier protein complex found in the cysteine biosynthetic pathway of Mycobacterium tuberculosis. *Biochemistry*, *47*(39), 10354–10364. https://doi.org/10.1021/bi800915j

Kerr, J. H., & Barrett, T. L. (1994). Atypical Mycobacterial Diseases. *Military Dermatology*, 391–422.

Kery, V., Bukovska, G., & Kraus, J. P. (1994). Transsulfuration depends on heme in addition to pyridoxal 5'-phosphate. Cystathionine β-synthase is a heme protein. *Journal of Biological Chemistry*, *269*(41), 25283–25288.

King, H. C., Khera-Butler, T., James, P., Oakley, B. B., Erenso, G., Aseffa, A., Knight, R., Wellington, E. M., & Courtenay, O. (2017). Environmental reservoirs of pathogenic mycobacteria across the Ethiopian biogeographical landscape. *PLoS ONE*, *12*(3), 1–15. https://doi.org/10.1371/journal.pone.0173811

Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution*, *35*(6), 1547–1549. https://doi.org/10.1093/molbev/msy096

Lastória, J. C., & de Abreu, M. A. M. M. (2014). Leprosy: Review of the epidemiological, clinical, and etiopathogenic aspects - Part 1. *Anais Brasileiros de Dermatologia*, *89*(2), 205–218. https://doi.org/10.1590/abd1806-4841.20142450

Lee, H. S., & Hwang, B. J. (2003). Methionine biosynthesis and its regulation in Corynebacterium glutamicum: Parallel pathways of transsulfuration and direct sulfhydrylation. *Applied Microbiology and Biotechnology*, *62*(5–6), 459–467. https://doi.org/10.1007/s00253-003-1306-7

Liu, J., Barry, C. E., Besra, G. S., & Nikaido, H. (1996). Mycolic acid structure determines the fluidity of the mycobacterial cell wall. *Journal of Biological Chemistry*, *271*(47), 29545–29551. https://doi.org/10.1074/jbc.271.47.29545

Liu, Y., Sieprawska-Lupa, M., Whitman, W. B., & White, R. H. (2010). Cysteine is not the sulfur source for iron-sulfur cluster and methionine biosynthesis in the methanogenic archaeon Methanococcus maripaludis. *Journal of Biological Chemistry*, *285*(42), 31923–31929. https://doi.org/10.1074/jbc.M110.152447

MacNeil, A., Glaziou, P., Sismanidis, C., Date, A., Maloney, S., & Floyd, K. (2020). Global Epidemiology of Tuberculosis and Progress Toward Meeting Global Targets - Worldwide, 2018. *MMWR. Morbidity and Mortality Weekly Report*, *69*(11), 281–285. https://doi.org/10.15585/mmwr.mm6911a2

MacNeil, A., Glaziou, P., Sismanidis, C., Maloney, S., & Floyd, K. (2019). Global Epidemiology of Tuberculosis and Progress Toward Achieving Global Targets — 2017. *MMWR Morb Mortal Wkly Rep*, *68*(11), 263-3–266. https://doi.org/10.15585/mmwr.mm6811a3

Mudd, S. H., Finkelstein, J. D., Irreverre, F., & Laster, L. (1965). Transsulfuration in mammals. Microassays and tissue distributions of three enzymes of the pathway. *Journal of Biological Chemistry*, *240*(11), 4382–4392.

Newton, G. L., Buchmeier, N., & Fahey, R. C. (2008). Biosynthesis and Functions of Mycothiol, the Unique Protective Thiol of Actinobacteria. *Microbiology and Molecular Biology Reviews*, *72*(3), 471–494. https://doi.org/10.1128/mmbr.00008-08

O'Leary, S. E., Jurgenson, C. T., Ealick, S. E., & Begley, T. P. (2008). O-phospho-L-serine and the thiocarboxylated sulfur carrier protein CysO-COSH are substrates for CysM, a cysteine synthase from Mycobacterium tuberculosis. *Biochemistry*, *47*(44), 11606–11615. https://doi.org/10.1021/bi8013664

Ono, B. I., Hazu, T., Yoshida, S., Kawato, T., Shinoda, S., Brzvwczy, J., & Paszewski, A. (1999). Cysteine biosynthesis in Saccharomyces cerevisiae: A new outlook on pathway and regulation. *Yeast*, *15*(13), 1365–1375. https://doi.org/10.1002/(SICI)1097-0061(19990930)15:13<1365::AID-YEA468>3.0.CO;2-U

Papet, I., Rémond, D., Dardevet, D., Mosoni, L., Polakof, S., Peyron, M.-A., & Savary-Auzeloux, I. (2019). Sulfur Amino Acids and Skeletal Muscle. In *Nutrition and Skeletal Muscle*. https://doi.org/10.1016/b978-0-12-810422-4.00020-8

Pinto, R., Harrison, J. S., Hsu, T., Jacobs, W. R., & Leyh, T. S. (2007). Sulfite reduction in mycobacteria. *Journal of Bacteriology*, *189*(18), 6714–6722. https://doi.org/10.1128/JB.00487-07

Reisch, C. R., Moran, M. A., & Whitman, W. B. (2011). Bacterial catabolism of dimethylsulfoniopropionate (DMSP). *Frontiers in Microbiology*, *2*(AUG), 1–12. https://doi.org/10.3389/fmicb.2011.00172

Robert, X., & Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Research*, *42*(W1), 320–324. https://doi.org/10.1093/nar/gku316

Saitou, N., & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, *4*(4), 406–425. https://doi.org/10.1093/oxfordjournals.molbev.a040454

Schnell, R., Oehlmann, W., Singh, M., & Schneider, G. (2007). Structural insights into catalysis and inhibition of O-acetylserine sulfhydrylase from Mycobacterium tuberculosis: Crystal structures of the enzyme α-aminoacrylate intermediate and an enzyme-inhibitor complex. *Journal of Biological Chemistry*, *282*(32), 23473–23481. https://doi.org/10.1074/jbc.M703518200

Selkon, J. B., Devdatta, S., Kulkarni, K. G., Mitchison, D. A., Narayana, A. S. L., Nair, C. N., & Ramachandran, K. (1964). the Emergence of Isoniazid-Resistant Cultures in Patients With Pulmonary Tuberculosis During Treatment With Isoniazid Alone or Isoniazid Plus Pas. *Bulletin of the World Health Organization*, *31*(2), 273–294.

Seung, K. J., Keshavjee, S., & Rich, M. L. (2015). Multidrug-resistant tuberculosis and extensively drug-resistant tuberculosis. *Cold Spring Harbor Perspectives in Medicine*, *5*(9), 1–20. https://doi.org/10.1101/cshperspect.a017863

Singh, A. K., & Reyrat, J.-M. (2009). Laboratory Maintenance of Mycobacterium smegmatis. *Current Protocols in Microbiology*, *14*(1), 10C.1.1-10C.1.12. https://doi.org/10.1002/9780471729259.mc10c01s14

Singhal, A., Arora, G., Sajid, A., Maji, A., Bhat, A., Virmani, R., Upadhyay, S., Nandicoori, V. K., Sengupta, S., & Singh, Y. (2013). Regulation of homocysteine metabolism by Mycobacterium tuberculosis S-adenosylhomocysteine hydrolase. *Scientific Reports*, *3*. https://doi.org/10.1038/srep02264

Smith, C. S., Aerts, A., Saunderson, P., Kawuma, J., Kita, E., & Virmond, M. (2017). Multidrug therapy for leprosy: a game changer on the path to elimination. *The Lancet Infectious Diseases*, *17*(9), e293–e297. https://doi.org/10.1016/S1473-3099(17)30418-8

Spies, H. S. C., & Steenkamp, D. J. (1994). Thiols of Intracellular Pathogens: Identification of Ovothiol A in Leishmania Donovani and Structural Analysis of a Novel Thiol from Mycobacterium Bovis. *European Journal of Biochemistry*, *224*(1), 203–213. https://doi.org/10.1111/j.1432-1033.1994.tb20013.x

Steiner, E. M., Bo¨th, D., Lo¨ssl, P., Vilaplana, F., Schnell, R., & Schneider, G. (2014). CysK2 from Mycobacterium tuberculosis is an O-phospho-L-serine-dependent S-sulfocysteine synthase. *Journal of Bacteriology*, *196*(19), 3410–3420. https://doi.org/10.1128/JB.01851-14

Thomas, D., & Surdin-Kerjan, Y. (1997). Metabolism of sulfur amino acids in Saccharomyces cerevisiae. *Microbiology and Molecular Biology Reviews : MMBR*, *61*(4), 503–532. https://doi.org/10.1128/.61.4.503-532.1997

Van Kessel, J. C., & Hatfull, G. F. (2007). Recombineering in Mycobacterium tuberculosis. *Nature Methods*, *4*(2), 147–152. https://doi.org/10.1038/nmeth996

Van Kessel, J. C., & Hatfull, G. F. (2008). Efficient point mutagenesis in mycobacteria using single-stranded DNA recombineering: Characterization of antimycobacterial drug targets. *Molecular Microbiology*, *67*(5), 1094–1107. https://doi.org/10.1111/j.1365-2958.2008.06109.x

Verma, S., Dixit, R., & Pandey, K. C. (2016). Cysteine proteases: Modes of activation and future prospects as pharmacological targets. *Frontiers in Pharmacology*, *7*(APR), 1–12. https://doi.org/10.3389/fphar.2016.00107

Wayne, L. G. (1994). Dormancy of Mycobacterium tuberculosis and latency of disease. In *European Journal of Clinical Microbiology & Infectious Diseases*. https://doi.org/10.1007/BF02111491

Wheeler, P. R., Coldham, N. G., Keating, L., Gordon, S. V., Wooff, E. E., Parish, T., & Hewinson, R. G. (2005). Functional demonstration of reverse transsulfuration in the Mycobacterium tuberculosis complex reveals that methionine is the preferred sulfur source for pathogenic mycobacteria. *Journal of Biological Chemistry*, *280*(9), 8069–8078. https://doi.org/10.1074/jbc.M412540200

WHO. (2002). Global situation. *Wkly Epidemiol. Rec*, *77*, 1–8.

WHO. (2016). *Global leprosy update, 2015: time for action, accountability and inclusion*. *91*(35), 405–420.

Wooff, E., Michell, S. L., Gordon, S. V., Chambers, M. A., Bardarov, S., Jacobs, W. R., Hewinson, R. G., & Wheeler, P. R. (2002). Functional genomics reveals the sole sulphate transporter of the Mycobacterium tuberculosis complex and its relevance to the acquisition of sulphur in vivo. *Molecular Microbiology*, *43*(3), 653–663. https://doi.org/10.1046/j.1365-2958.2002.02771.x

Yamada, H., Yamaguchi, M., Igarashi, Y., Chikamatsu, K., Aono, A., Murase, Y., Morishige, Y., Takaki, A., Chibana, H., & Mitarai, S. (2018). Mycolicibacterium smegmatis, Basonym Mycobacterium smegmatis, Expresses Morphological Phenotypes Much More Similar to Escherichia coli Than Mycobacterium tuberculosis in Quantitative Structome Analysis and CryoTEM Examination. *Frontiers in Microbiology*, *9*(SEP). https://doi.org/10.3389/fmicb.2018.01992

Yuen, L. K. W., Leslie, D., & Coloe, P. J. (1999). Bacteriological and molecular analysis of rifampin-resistant Mycobacterium tuberculosis strains isolated in Australia. *Journal of Clinical Microbiology*, *37*(12), 3844–3850. https://doi.org/10.1128/jcm.37.12.3844-3850.1999

Zuckerkandl, E., & Pauling, L. (1965). Evolutionary Divergence and Convergence in Proteins. *Evolving Genes and Proteins*, 97–166. https://doi.org/10.1016/b978-1-4832-2734-4.50017-6

VI. FIGURES

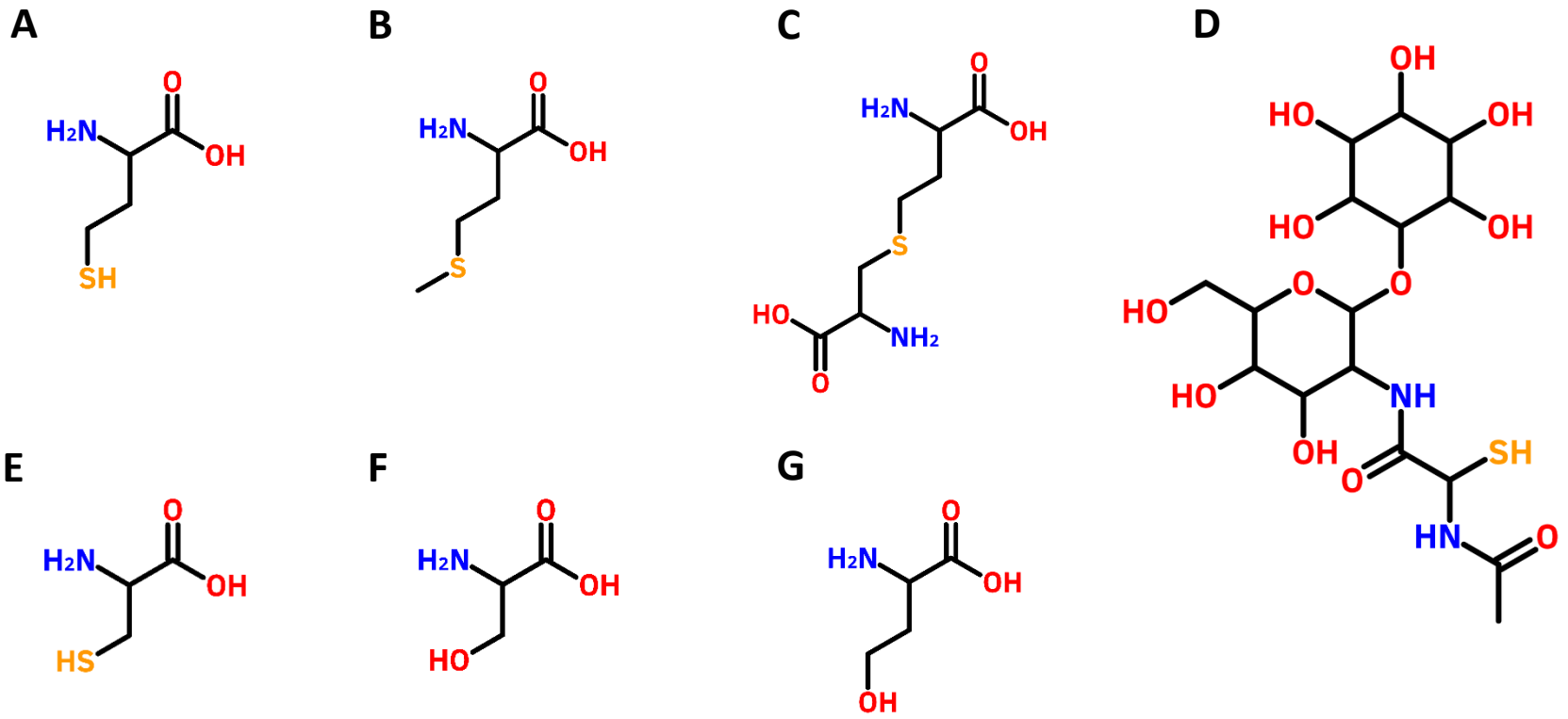

Figure 1: Structural formula of compounds relevant to this study.
(A) Homocysteine, (B) Methionine, (C) Cystathionine, (D) Mycothiol, (E) Cysteine, (F) Serine, and (G) Homoserine

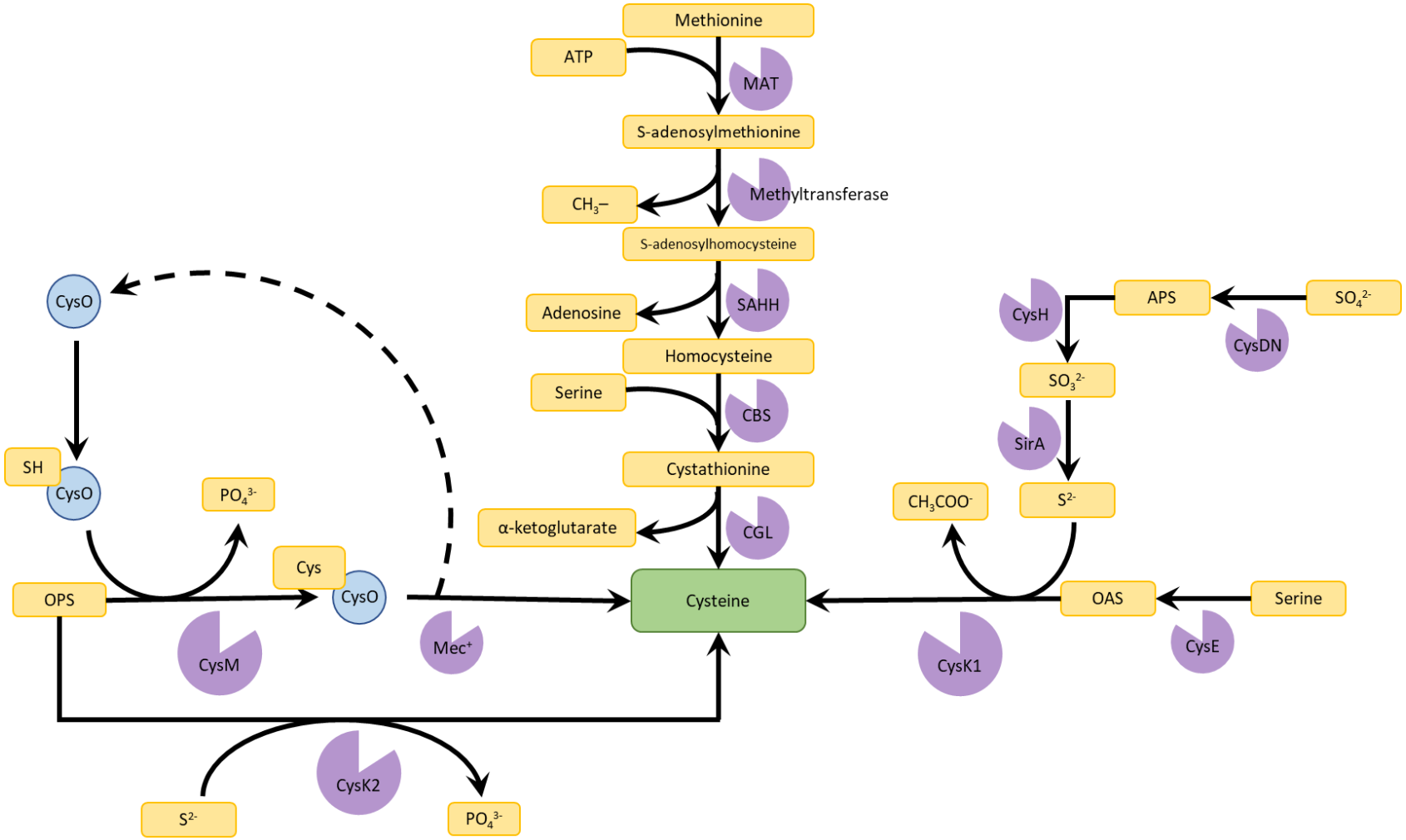

Figure 2: Predicted pathways for biosynthesis of cysteine in mycobacteria.
Rounded rectangles represent precursors, intermediates, or products in the biosynthetic pathways of cysteine. CysO is a sulfur carrier protein. Circular sectors in the diagram represent enzymes.

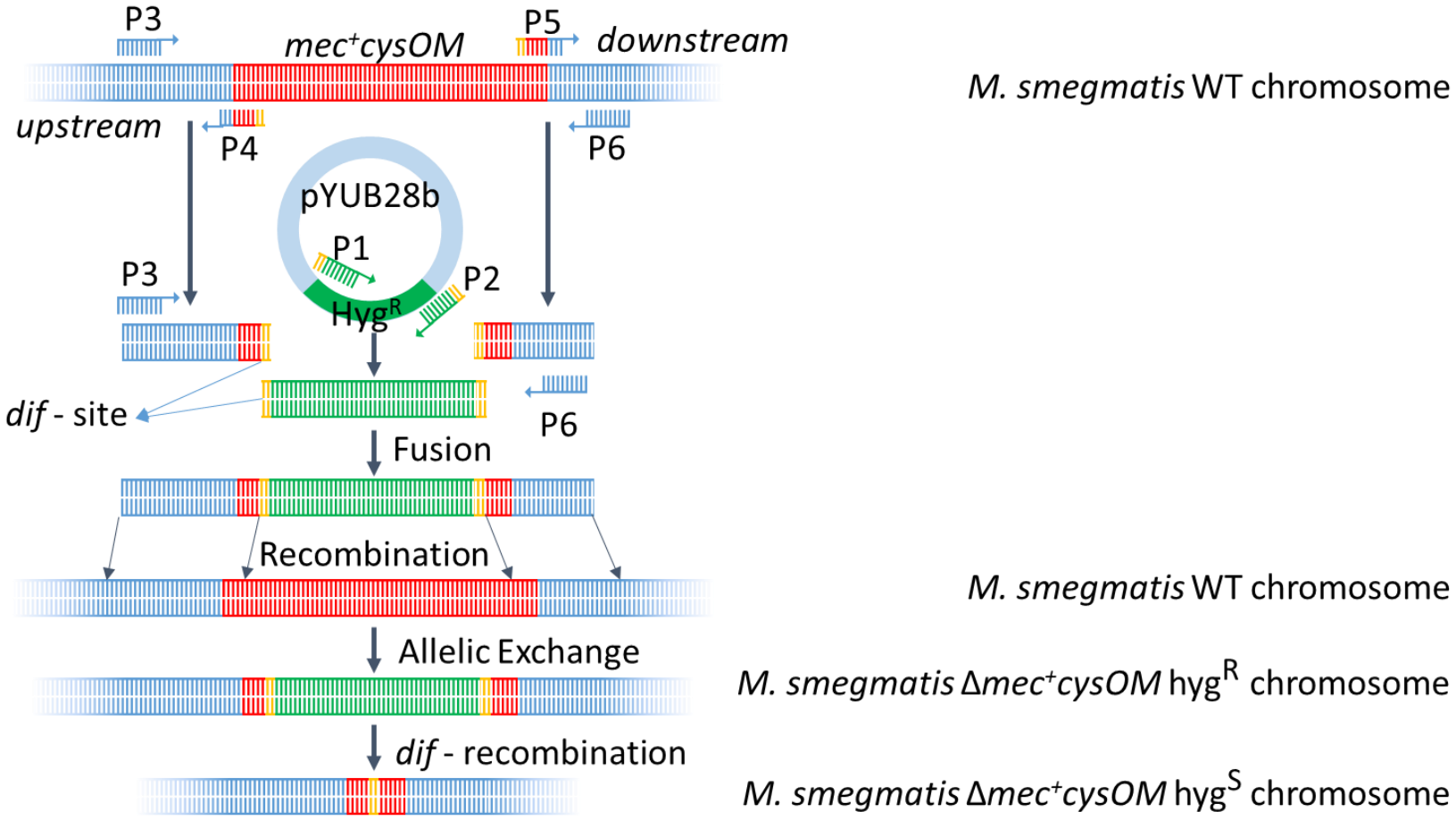

Figure 3: Representation of the steps followed in order to generate unmarked knockout mutant *M. smegmatis* Δ*mec*⁺*cysOM*

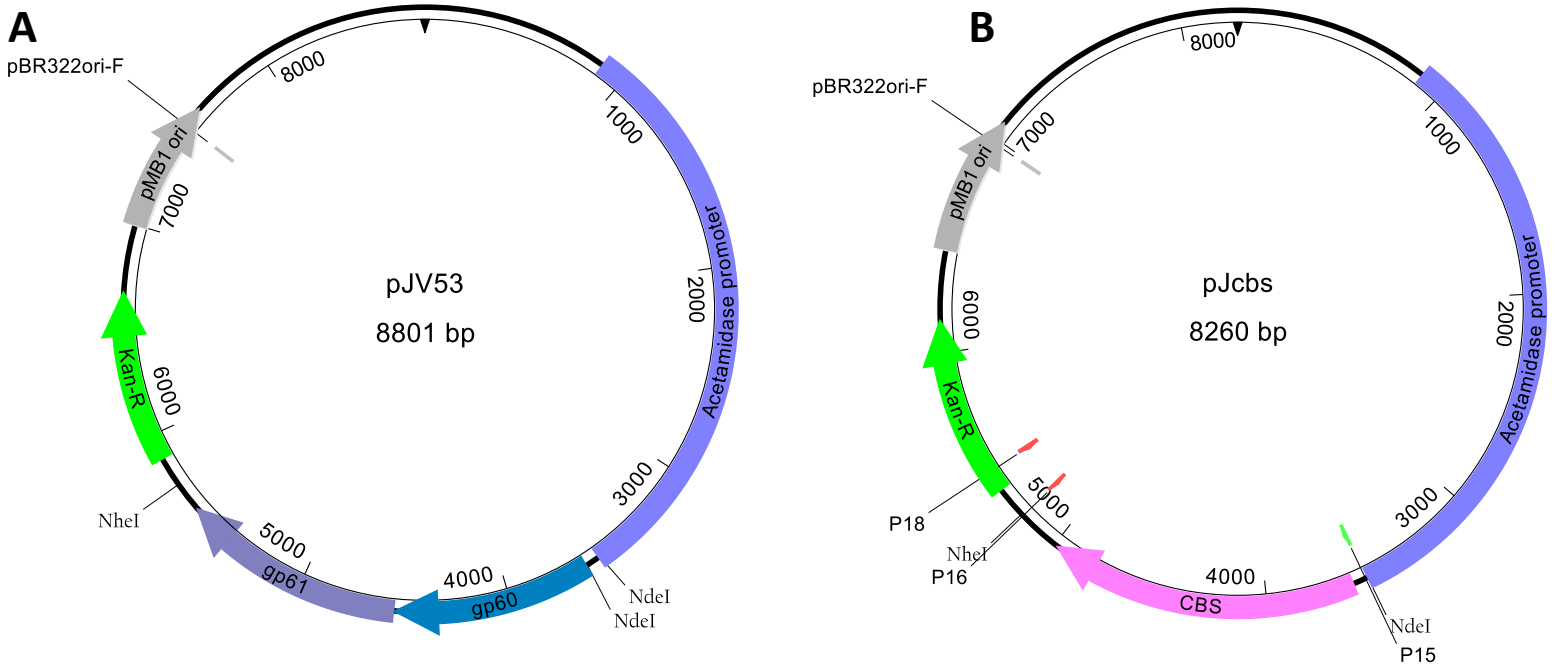

Figure 4: Plasmids pJV53 and pJcbs
Plasmids including the primer binding sites used in this study. The images were created using Lasergene SeqBuilder Pro software from DNASTAR. (A) pJV53 with the che9c genes, gp60 and gp61, (B) pJmeccysOM constructed by replacing gp60-gp61 from pJV53 with *mec*$^{+}$*cysOM* operon

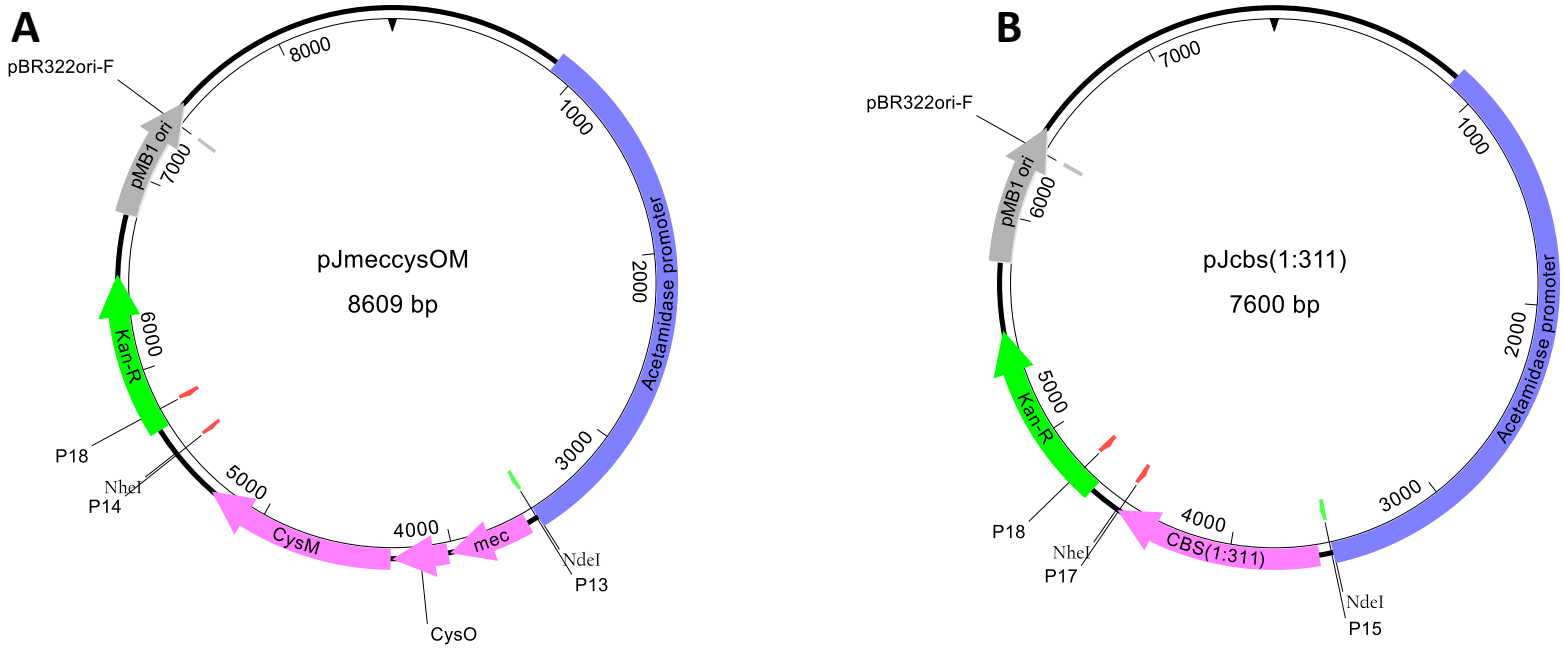

Figure 5: Plasmids pJmeccysOM and pJcbs(1:311)
Plasmids including the primer binding sites used in this study. The images were created using Lasergene SeqBuilder Pro software from DNASTAR. (A) pJcbs constructed by replacing gp60-gp61 from pJV53 with *cbs* gene, (B) pJcbs(1:311) constructed by replacing gp60-gp61 with the portion of cbs gene encoding the first 311 amino acids

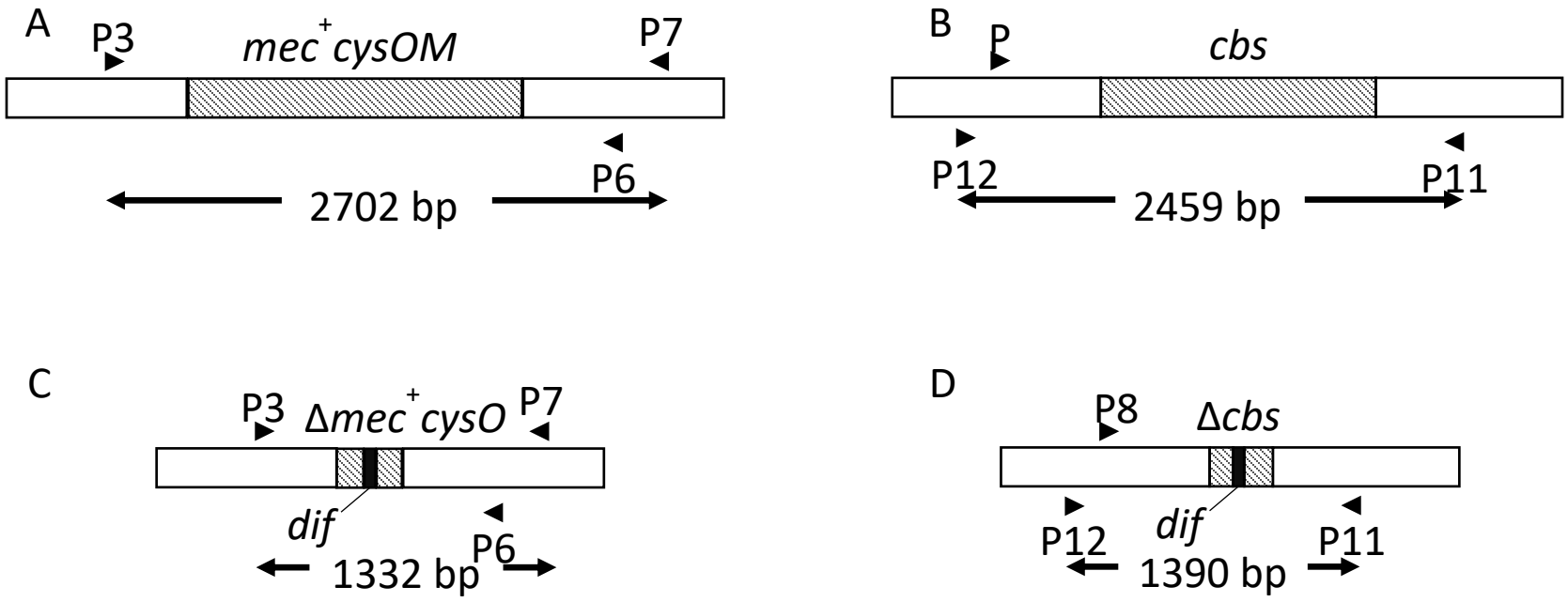

Figure 6: Organization of operon/gene in *M. smegmatis* with primer binding sites

The directions of the primers are represented by the arrow heads (A) intact *mec*$^{+}$*cysOM* operon, (B) intact *cbs* gene, (C) deleted *mec*$^{+}$*cysOM* operon, (D) deleted *cbs* gene

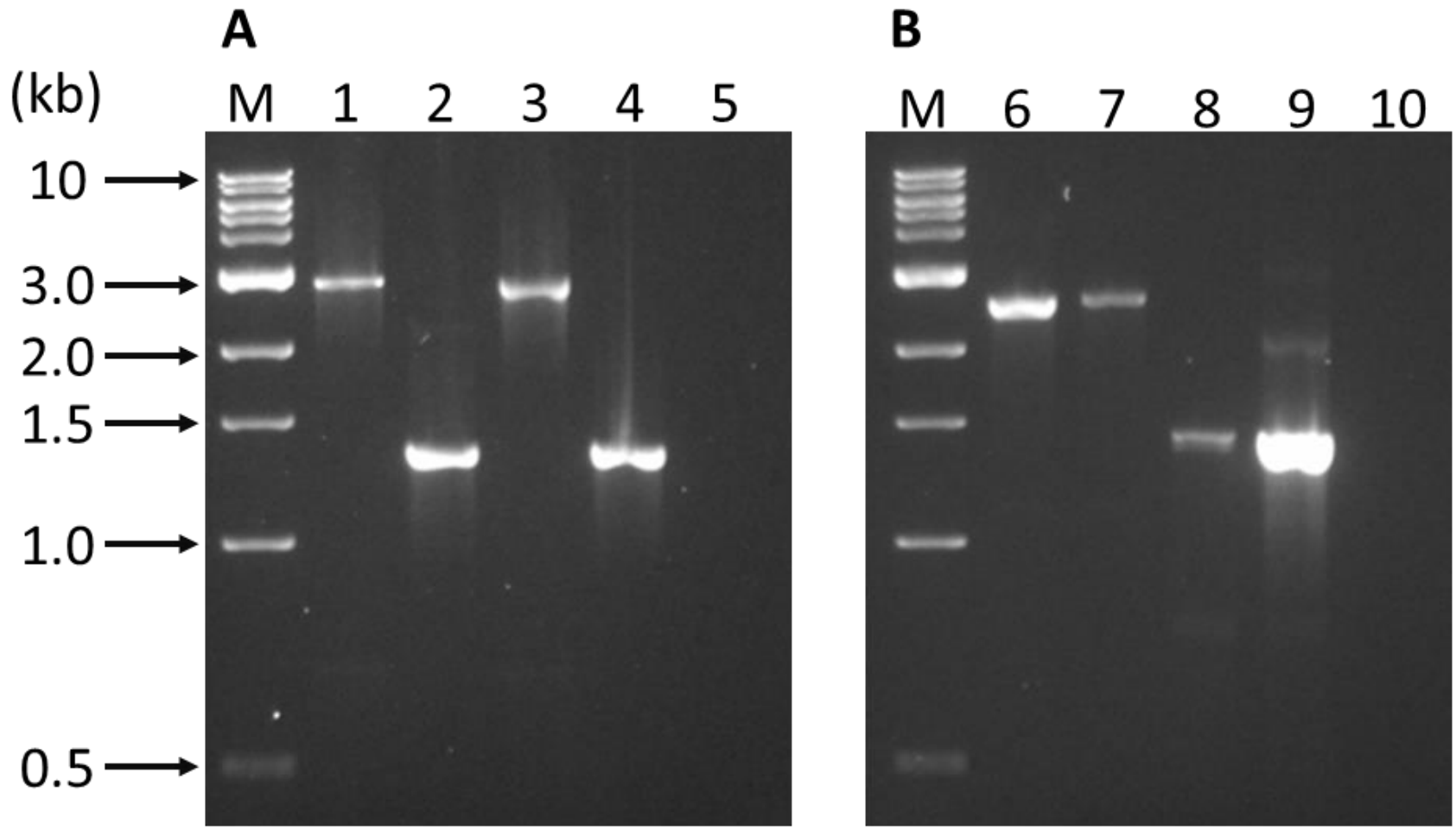

Figure 7: Confirmation of knockout mutants
Gel electrophoresis of PCR products using primers P3 and P7 (A) and P11 and P12 (B). DNA samples from the following strains of *M. smegmatis* were used: Wild type (Lanes 1 & 6), Δ*mec*$^{+}$*cysOM* (Lanes 2 & 7), Δ*cbs* (Lanes 3 & 8), Δ*mec*$^{+}$*cysOM_cbs* (Lanes 4 & 9). Lanes 5 and 10 were negative controls.

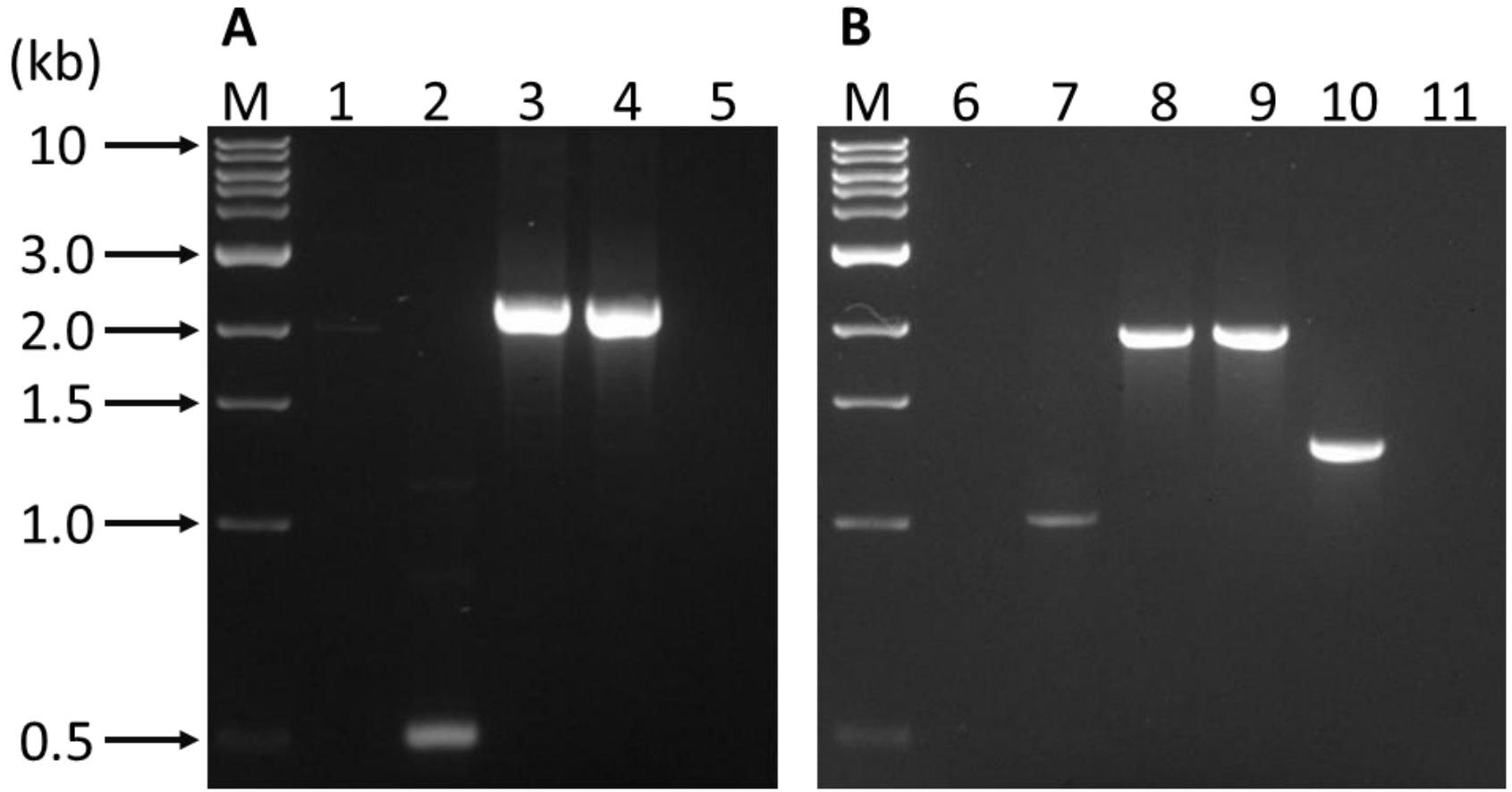

Figure 8: Confirmation of the complemented knockout mutants
Gel electrophoresis of PCR products using primers P13 & P18 (A) and P15 & P18 (B). pJV53 DNA was used as template in Lanes 2 & 7. Colonies of the following strains of *M. smegmatis* were used as samples: Wild type (Lanes 1 & 6), Δ*mec*$^{+}$*cysOM::mec*$^{+}$*cysOM* (Lane 3), *M. smegmatis* Δ*mec*$^{+}$*cysOM_cbs::mec*$^{+}$*cysOM* (Lane 4), Δ*cbs::cbs* (Lane 8), Δ*mec*$^{+}$*cysOM_cbs::cbs* (Lane 9), Δ*mec*$^{+}$*cysOM_cbs::cbs*(1:311) (Lane 10). Lanes 5 & 11 were negative controls.

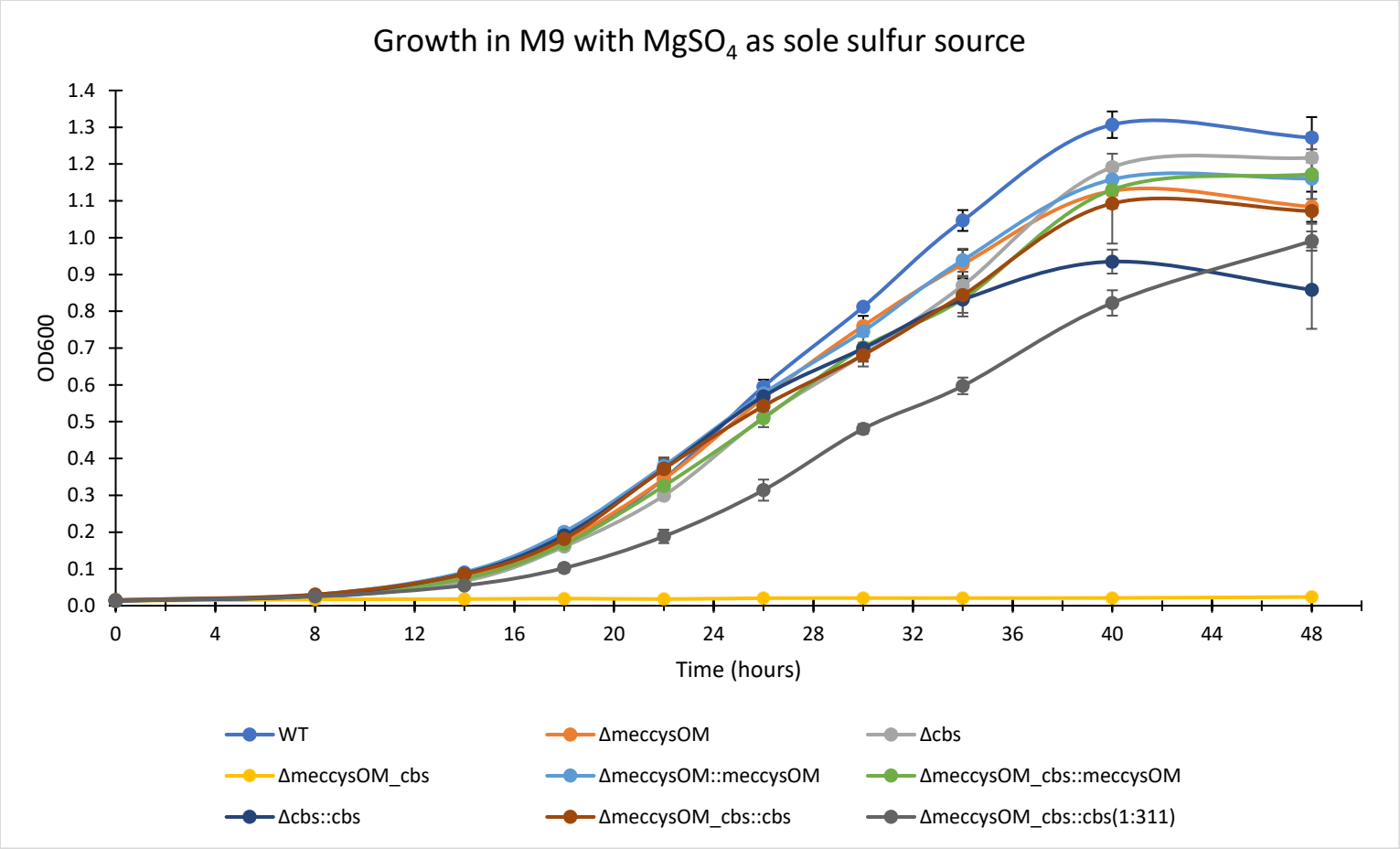

Figure 9: Growth curves of the mutants of *M. smegmatis* generated in this study along with the wildtype in sulfur-free M9 medium supplemented with $MgSO_4$

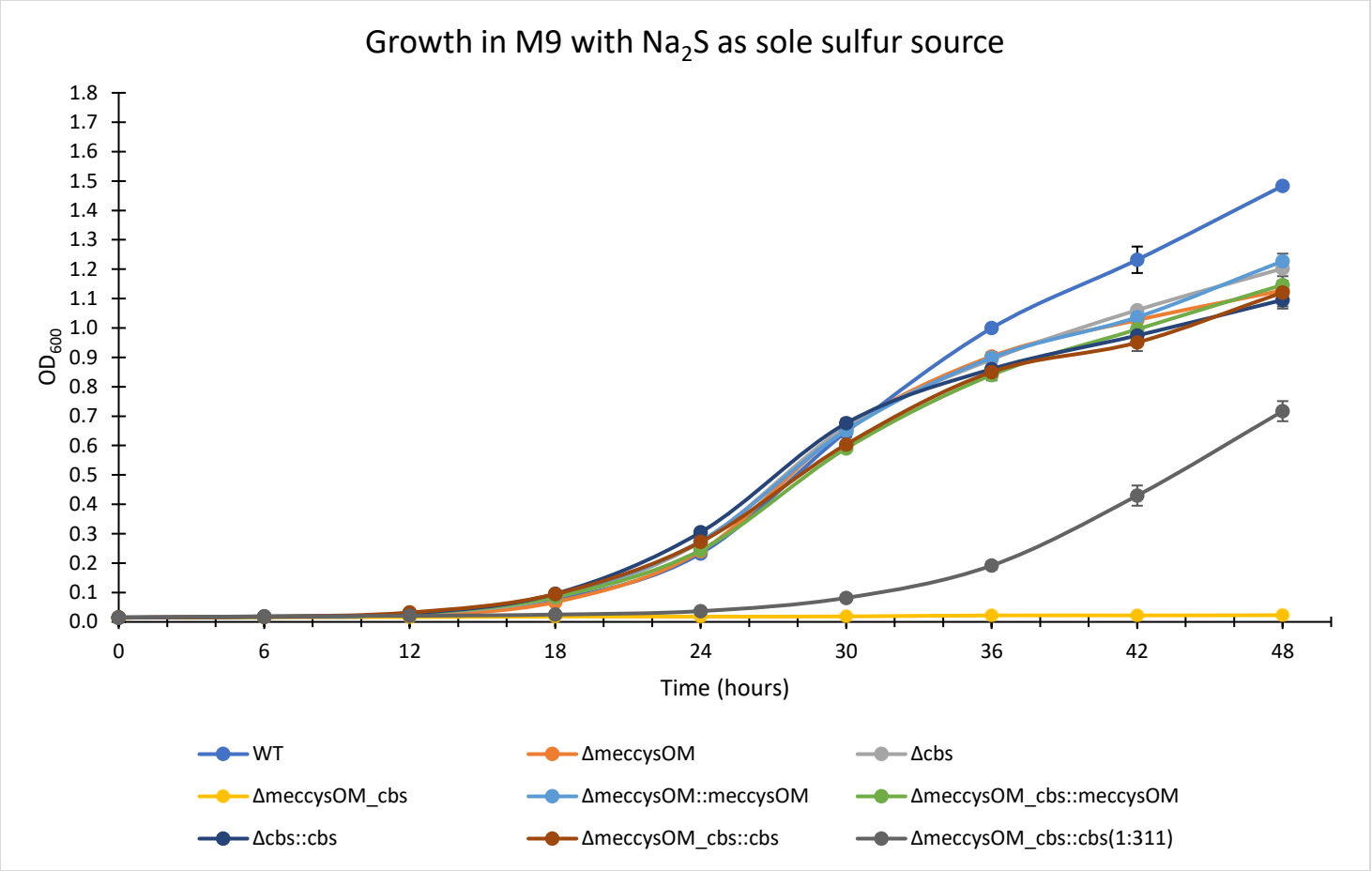

Figure 10: Growth curves of the mutants of *M. smegmatis* generated in this study along with the wildtype in sulfur-free M9 medium supplemented with Na_2S

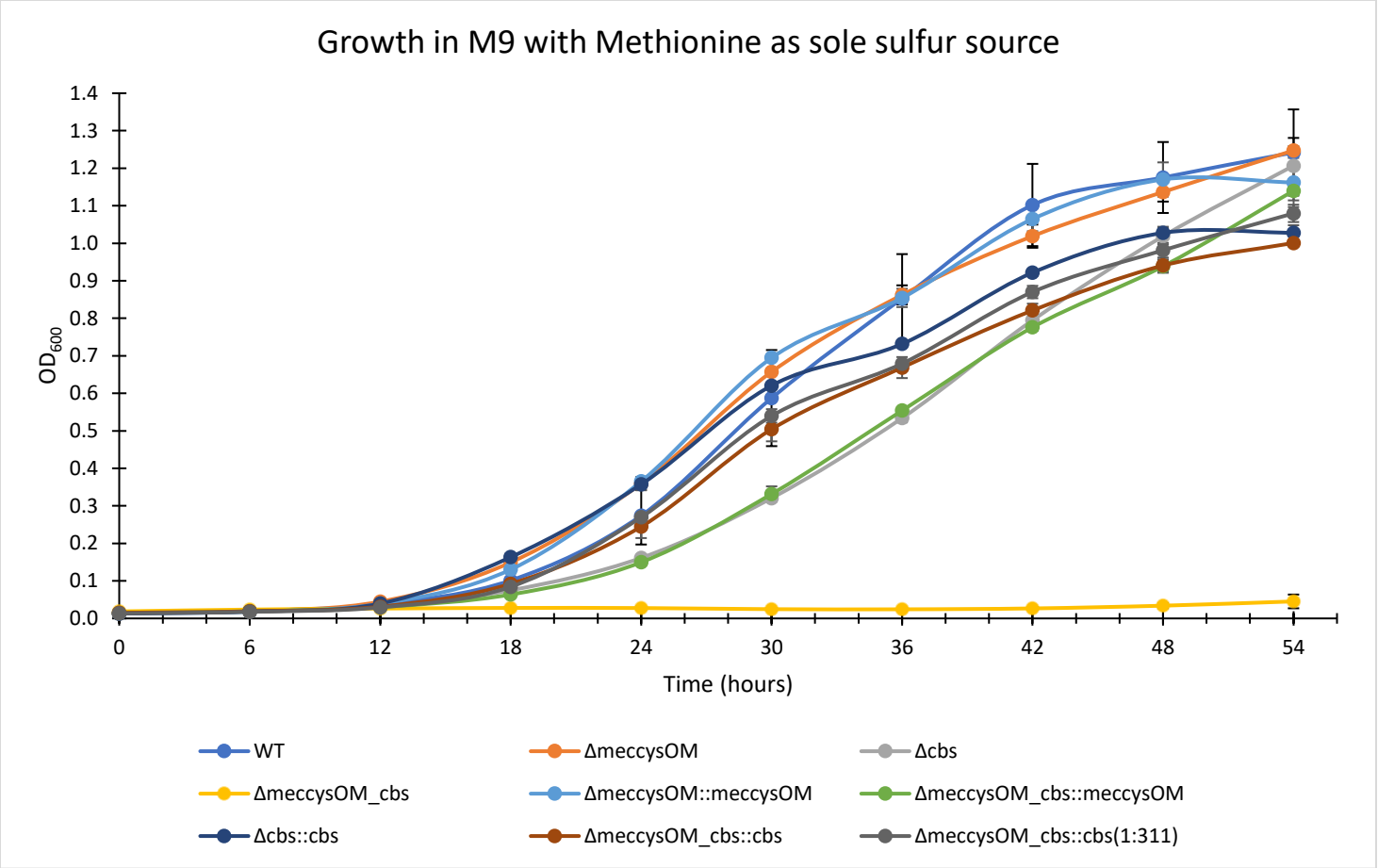

Figure 11: Growth curves of the mutants of *M. smegmatis* generated in this study along with the wildtype in sulfur-free M9 medium supplemented with Methionine

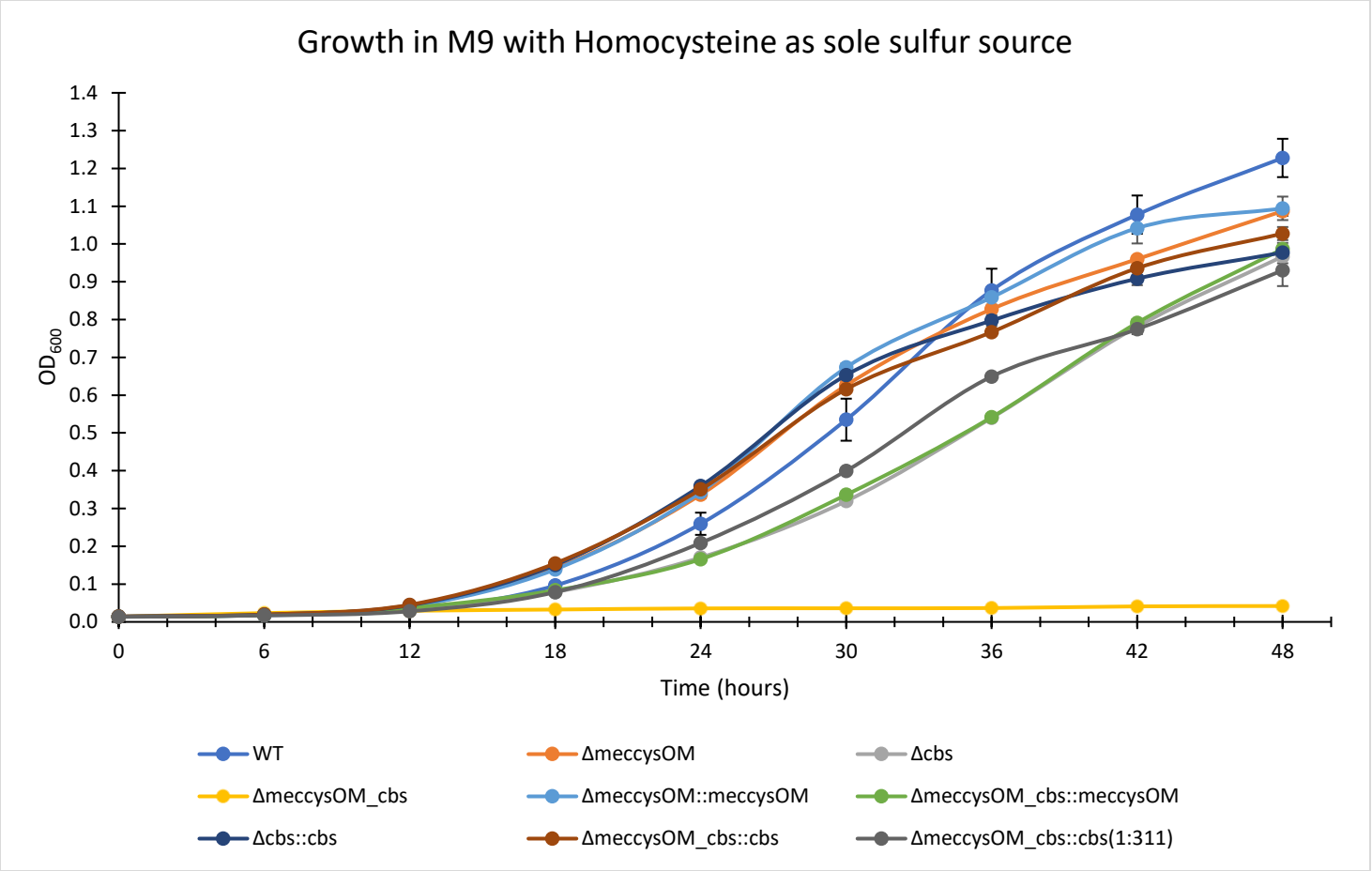

Figure 12: Growth curves of the mutants of *M. smegmatis* generated in this study along with the wildtype in sulfur-free M9 medium supplemented with Homocysteine

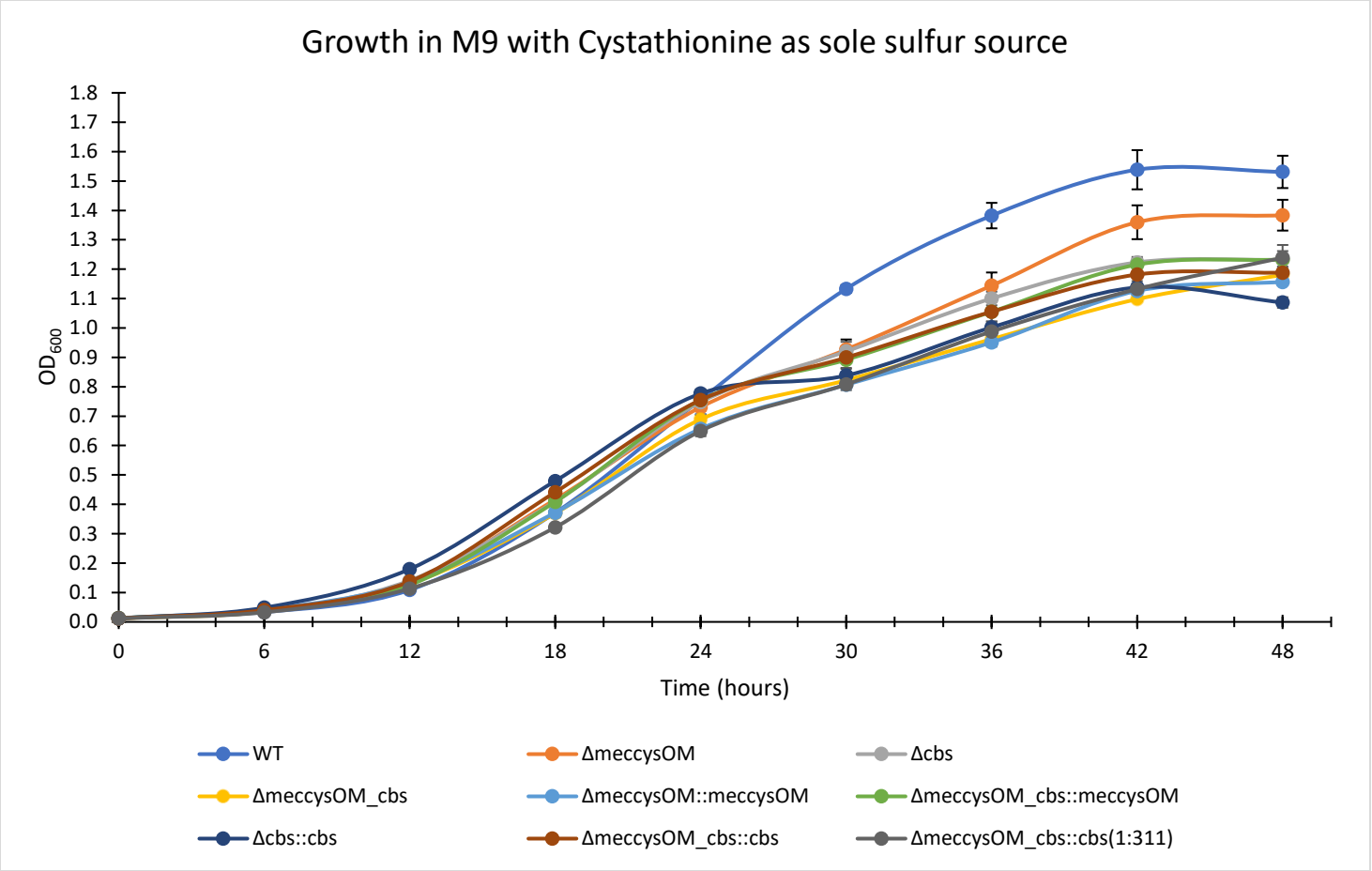

Figure 13: Growth curves of the mutants of *M. smegmatis* generated in this study along with the wildtype in sulfur-free M9 medium supplemented with Cystathionine

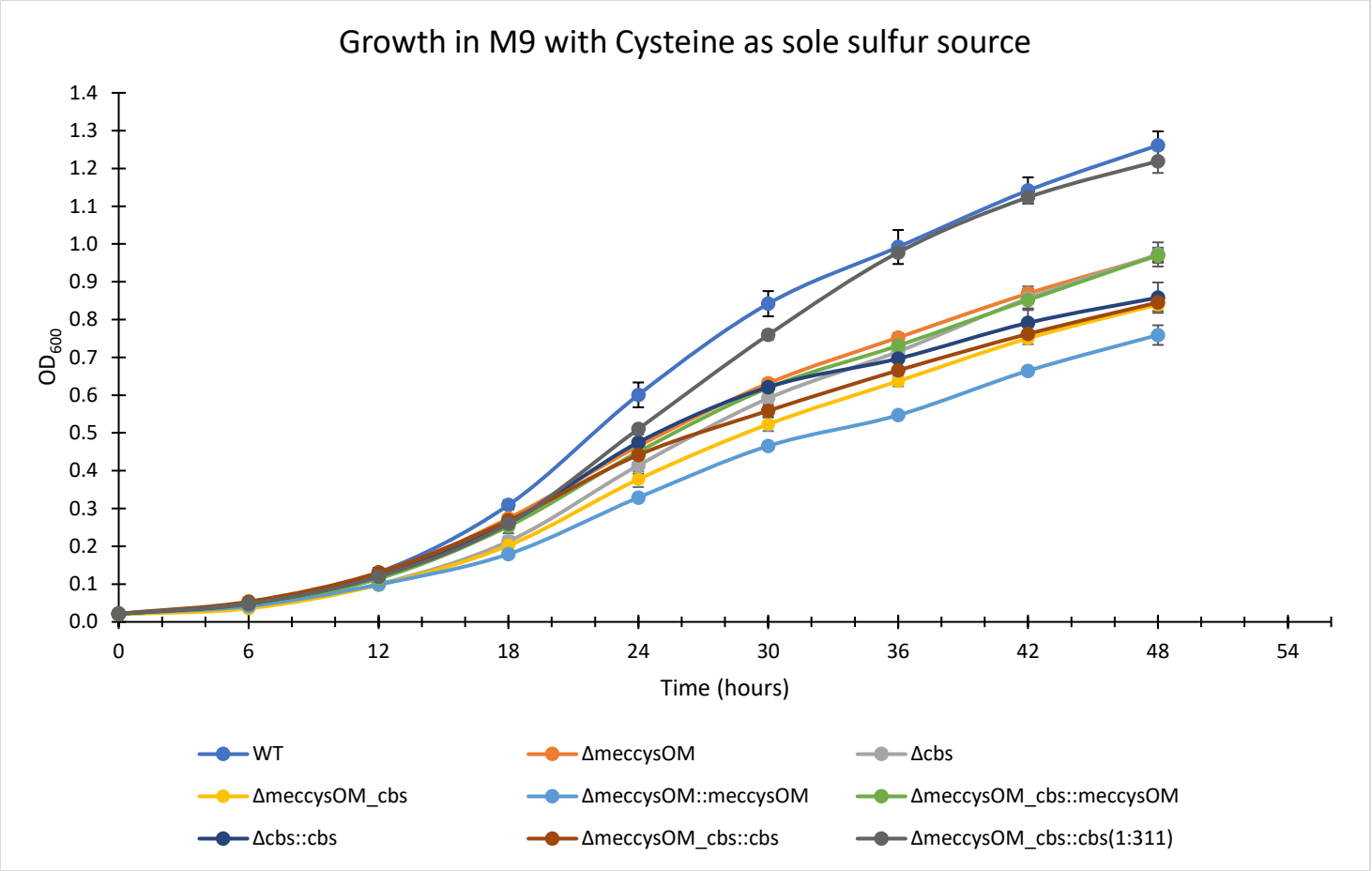

Figure 14: Growth curves of the mutants of *M. smegmatis* generated in this study along with the wildtype in sulfur-free M9 medium supplemented with Cysteine

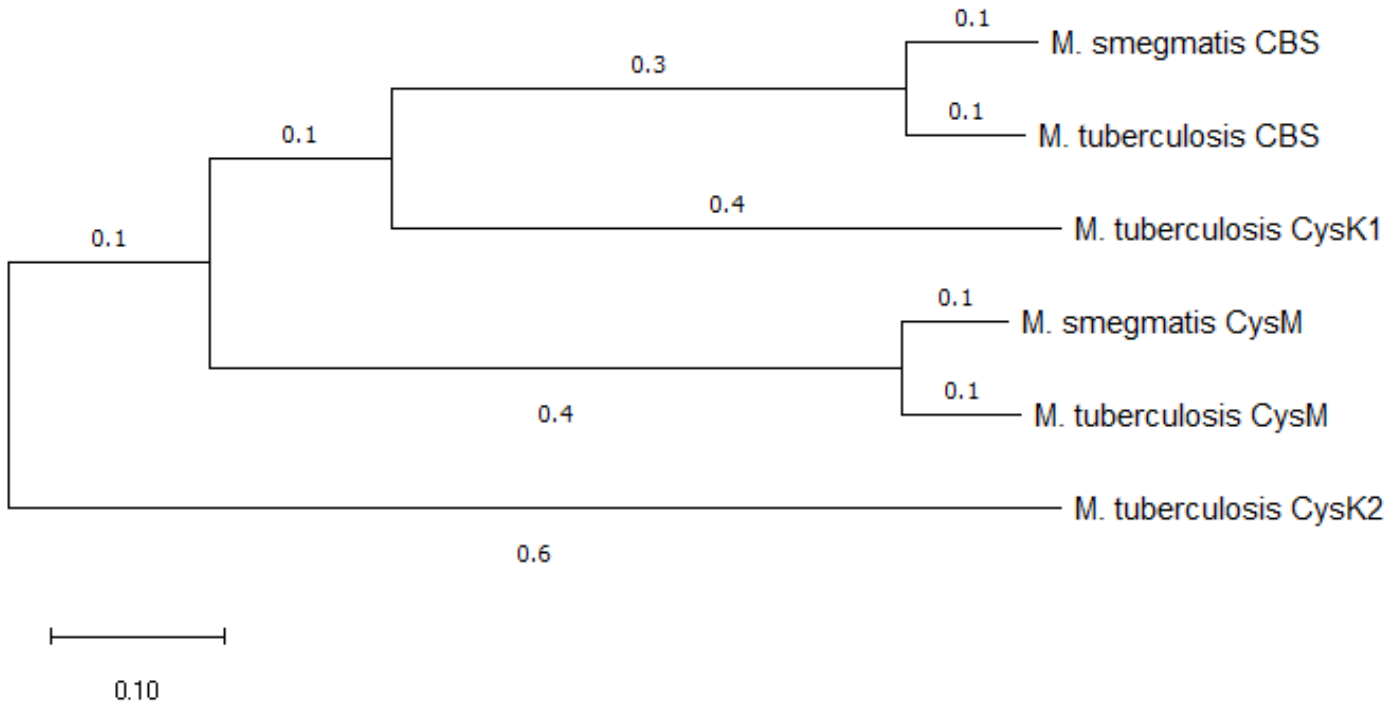

Figure 15: Phylogenetic relationship among Cysteine synthases and Cystathionine beta-synthases of *M. tuberculosis* and *M. smegmatis*

The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 2.19397737 is shown. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Poisson correction method and are in the units of the number of amino acid substitutions per site. This analysis involved 6 amino acid sequences. All ambiguous positions were removed for each sequence pair (pairwise deletion option). There were a total of 496 positions in the final dataset. Evolutionary analyses were conducted in MEGA X.

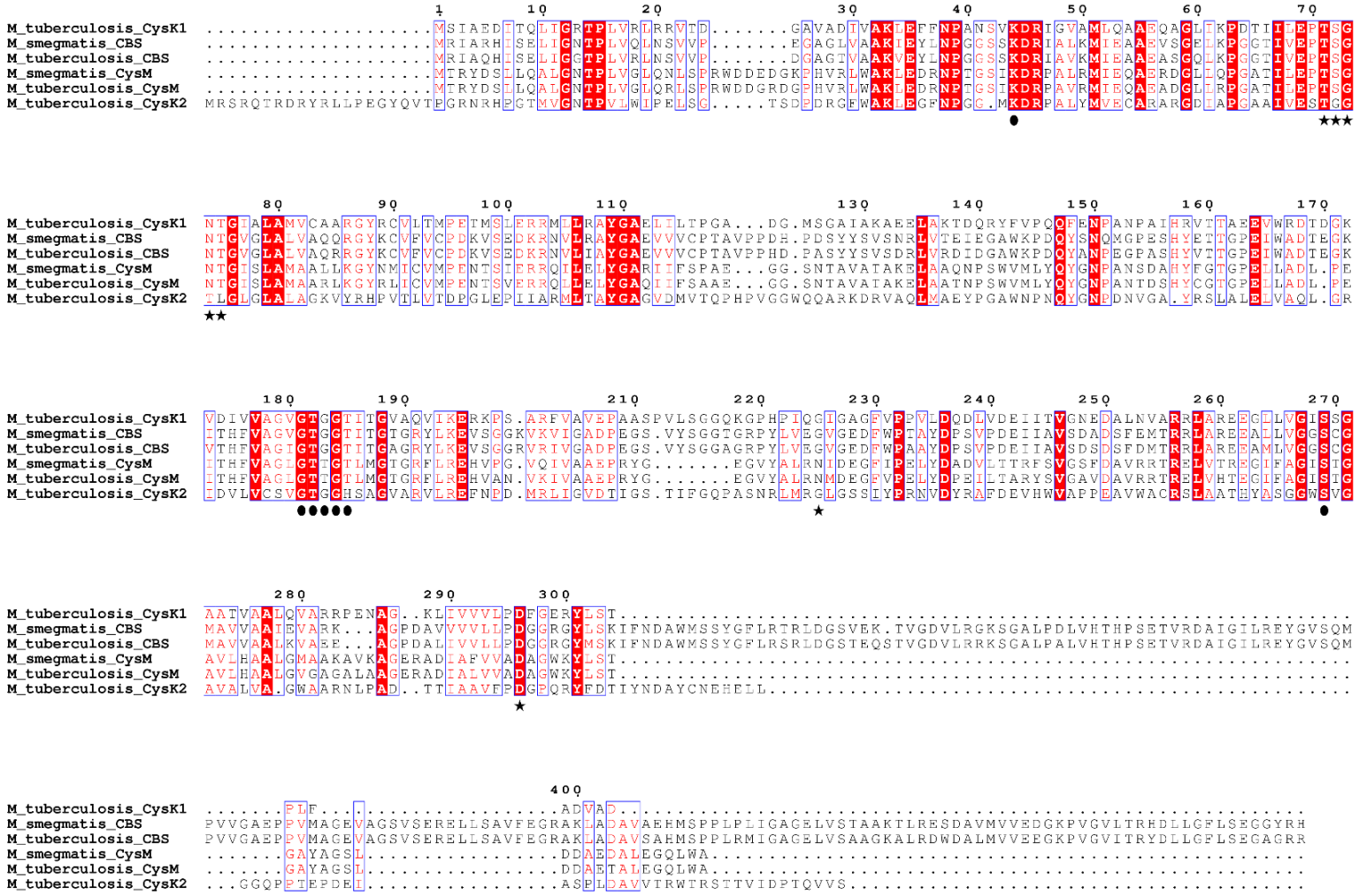

Figure 16: Multiple sequence alignment of cysteine synthases and cystathionine beta-synthases of *M. smegmatis* and *M. tuberculosis*

The columns are numbered by the position of residues of *M. smegmatis* CBS. The position of the residues in *M. tuberculosis* that contribute to binding of PLP cofactor (●) and those that are involved in the formation of the active site (★) are indicated. The image was created using ESPript.

VII. TABLES

Table 1: List of plasmids used in this study

Plasmid	**Description**
pYUB28b	For the amplification of hygromycin resistance gene
pJV53	For che9c gp60-gp61 recombinases and as the vector for complementation
pJmeccysOM	pJV53 replacing che9c gp60-gp61 with *M. smegmatis mec*$^{+}$*cysOM* operon
pJcbs	pJV53 replacing che9c gp60-gp61 with *M. smegmatis cbs gene*
pJcbs(1:311)	pJV53 replacing che9c gp60-gp61 with the sequence encoding the N-terminal domain of *M. smegmatis cbs gene* (the first 311 amino acid followed by a stop codon)

Table 2: List of bacterial strains used in this study

Strain	**Description**
E. coli DH5α	Carrier of plasmids pYUB28b or pJV53
M. smegmatis mc^2155	Wild type
M. smegmatis + pJV53	Transformed with pJV53 plasmid
M. smegmatis Δ*mec*$^+$*cysOM*	*mec*$^+$, *cysO*, *cysM* genes knocked out
M. smegmatis Δ*cbs*	*cbs* gene knocked out
M. smegmatis Δ*mec*$^+$*cysOM_cbs*	*mec*$^+$*cysOM* operon and *cbs* gene knocked out (double mutant)
M. smegmatis Δ*mec*$^+$*cysOM::mec*$^+$*cysOM*	*mec*$^+$*cysOM* knockout complemented with pJmeccysOM
M. smegmatis Δ*mec*$^+$*cysOM_cbs::mec*$^+$*cysOM*	Double mutant complemented with pJmeccysOM
M. smegmatis Δ*cbs::cbs*	*cbs* mutant complemented with pJcbs
M. smegmatis Δ*mec*$^+$*cysOM_cbs::cbs*	Double mutant complemented with pJcbs
M. smegmatis Δ*mec*$^+$*cysOM_cbs::cbs*(1:311)	Double mutant complemented with pJcbs(1:311)

Table 3: List of primers used in this study

Application	Primer Name	Sequence
Amplification of *hyg*^{R} gene	P1	AAGCTTCTCGAGTAAGCCGATAAGCGACATTATGTCAAGTCCCGGGTCTAGACCCGTCATCGTCAAC
	P2	GGTACCAAGCTTCTCGAGACTTGACATAATGTCGCTTATCGGCTTAATCGATCTAGATCACCGGCGCCGGGGG
Construction of Δ*mec*$^{+}$*cysOM* allele and confirmation of Δ*mec*$^{+}$*cysOM* mutation	P3	TACGCCGATCTGCGCACCGAACTCAGCCCGC
	P4	TCTAGACCCGGGACTTGACATAATGTCGCTTATCGGCTTACTCGAGAAGCTTGTTGGTCATCGCGATGAACCGCTC
	P5	TCTAGATCGATTAAGCCGATAAGCGACATTATGTCAAGTCTCGAGAAGCTTGGTACCCGGTGAGCGTGCCGACATCGCG
	P6	AGGATGCCGACGACGATCTGCCAGGCGTTGCG
	P7	CGATCGCGCCGCACAGATGTCCCTGCCACG
Construction of Δ*CBS* allele and confirmation of Δ*CBS* mutation	P8	TACGTGTGGCTGCGTCGCGCGGGCTACGAAC
	P9	TCTAGACCCGGGACTTGACATAATGTCGCTTATCGGCTTACTCGAGAAGCTTGCCGGGGTTGAGGTACTCGATTTTCGC
	P10	TCTAGATCGATTAAGCCGATAAGCGACATTATGTCAAGTCTCGAGAAGCTTGGTACCCGGCGCCGGTGAACTGGTCAGC
	P11	CGAAACCGCCCGGTTGCGCAGGGGGATAGC
	P12	CCCCGTGCTGGTGGTCGATTACCGCATGGTGCC
Cloning and confirmation of transformation	P13	GCAGCATATGATGTTGCCGGGAGCGTTCGAGT
	P14	GCATGCTAGCATCAGGTGTTCCCAGTTGGCGT
	P15	GCGGCATATGATCGGCGCCTACATACGTGAGG
	P16	ATAAGCTAGCATCGCACGCGCCGTCGTTAGCT
	P17	ATCGGCTAGCTCACATCCACGCGTCGTTGAAAATC
	P18	CACCTGATTGCCCGACATTATCGCGAGCCCATT