# Production of recombinant proteins in Mycobacterium smegmatis for structural and functional studies

**ARTICLE** *in* PROTEIN SCIENCE · JANUARY 2015

Impact Factor: 2.85 · DOI: 10.1002/pro.2584 · Source: PubMed



**2 AUTHORS:**

Ghader Bashiri
University of Auckland
**17** PUBLICATIONS **132** CITATIONS
SEE PROFILE

Edward N Baker
University of Auckland
**247** PUBLICATIONS **6,402** CITATIONS
SEE PROFILE

REVIEWS

# Production of recombinant proteins in *Mycobacterium smegmatis* for structural and functional studies

**Ghader Bashiri* and Edward N. Baker**

Structural Biology Laboratory, School of Biological Sciences and Maurice Wilkins Centre for Molecular Biodiscovery, The University of Auckland, Auckland 1010, New Zealand

Received 22 September 2014; Revised 7 October 2014; Accepted 8 October 2014
DOI: 10.1002/pro.2584
Published online 00 Month 2014 proteinscience.org

**Abstract: Protein production using recombinant DNA technology has a fundamental impact on our understanding of biology through providing proteins for structural and functional studies. *Escherichia coli* (*E. coli*) has been traditionally used as the default expression host to over-express and purify proteins from many different organisms. *E. coli* does, however, have known shortcomings for obtaining soluble, properly folded proteins suitable for downstream studies. These shortcomings are even more pronounced for the mycobacterial pathogen *Mycobacterium tuberculosis*, the bacterium that causes tuberculosis, with typically only one third of proteins expressed in *E. coli* produced as soluble proteins. *Mycobacterium smegmatis (M. smegmatis)* is a closely related and non-pathogenic species that has been successfully used as an expression host for production of proteins from various mycobacterial species. In this review, we describe the early attempts to produce mycobacterial proteins in alternative expression hosts and then focus on available expression systems in *M. smegmatis*. The advantages of using *M. smegmatis* as an expression host, its application in structural biology and some practical aspects of protein production are also discussed. *M. smegmatis* provides an effective expression platform for enhanced understanding of mycobacterial biology and pathogenesis and for developing novel and better therapeutics and diagnostics.**

**Keywords: *Mycobacterium smegmatis;* recombinant protein; expression system; protein production; *Mycobacterium tuberculosis*; *Mycobacterium leprae*; shuttle vectors**

*Abbreviations:* MAD, multi-wavelength anomalous diffraction; M-PFC, mycobacterial protein fragment complementation; *Mtb*, *Mycobacterium tuberculosis*; MTBSGC, *M. tuberculosis* Structural Genomics Consortium; PDB, Protein Data Bank; SAD, single-wavelength anomalous diffraction; SeMet, selenomethionine; TetR, tetracycline repressor.

*Correspondence to: Ghader Bashiri; Structural Biology Laboratory, School of Biological Sciences and Maurice Wilkins Centre for Molecular Biodiscovery, The University of Auckland, Auckland 1010, New Zealand. E-mail: g.bashiri@auckland.ac.nz

## Introduction

Use of recombinant DNA technologies to produce large quantities of proteins for structural and functional studies is today a routine practice. Considering the substantial role of recombinant protein production in almost all fields of life sciences, the choice of an appropriate expression method is a crucial variable in producing proteins in high quantity and quality.[1] In this regard, both improving existing

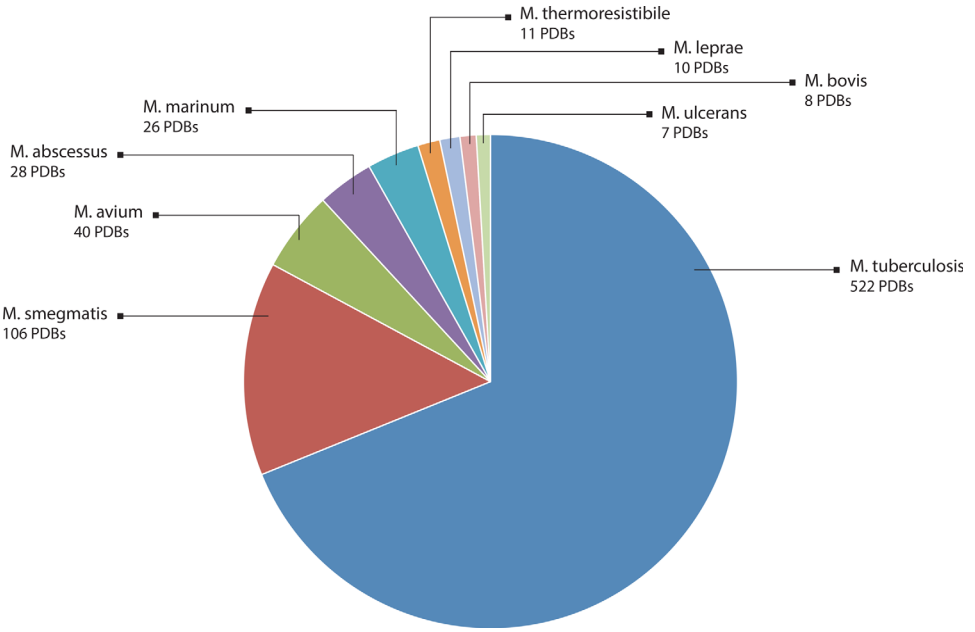


**Figure 1.** Protein structures present in the Protein Data Bank from different mycobacterial species. The numbers indicate discrete entries with <90% sequence identity. For clarity, the figure does not include species for which only two (*M. xenopi* and *M. fortuitum*) or one (*M. gastri*, *M. goodie*, *M. vanbaalenii*, *M. sp* JLS, and *M. sp* JC1) entry is found.

expression strategies and developing new expression systems are considered as the greatest challenge for structural genomics projects.[2]

*Escherichia coli* (*E. coli*) is the most prevalent microbial factory in biology.[3] The major bottleneck, however, in the production of recombinant proteins in *E. coli* is the formation of insoluble inclusion bodies.[3] This has been particularly discouraging in the expression of proteins from mycobacterial species including *Mycobacterium tuberculosis* (*Mtb*), the bacterium that causes tuberculosis. For this species, success in soluble expression, even with the best of *E. coli* strategies, has rarely exceeded 30%.[4] This is partly attributed to the different G+C content between *Mtb* (65.6%)[5] and *E. coli* (50.8%)[6] genomes, requiring different machinery for efficient transcription and translation of mycobacterial genes.[2] Several other factors that may play a role in the inability of *E. coli* cells to produce soluble and properly folded recombinant proteins, and strategies to overcome those limitations, are discussed elsewhere.[3]

In recognition of the limitations of *E. coli* host cells as a "one size fits all" strategy, various alternative expression systems for producing recombinant proteins have been developed.[7] The central idea is that problems in protein expression can be most effectively addressed by using host organisms more closely related to the organism from which the proteins are derived. There are, for instance, ~70 different expression hosts reported for protein production in the Protein Data Bank (PDB).[7] We presume the number of alternative expression hosts could be even larger since not all proteins are produced for structural studies.

## Expression Hosts for Mycobacterial Proteins

Publication of the *Mtb* H37Rv genome sequence in 1998[5] ignited a widespread interest in the biology and pathogenesis of mycobacteria, enabling functional and structural investigation of proteins from various mycobacterial species (Fig. 1). This led in turn to formation of the *Mtb* Structural Genomics Consortium (http://www.webtb.org) to "develop a foundation for tuberculosis diagnosis and treatment by determining the three-dimensional structures of proteins".[8] However, the success rate of obtaining soluble *Mtb* proteins from *E. coli* host cells is far from ideal, with only one third of the proteins being produced in soluble and properly folded form, suitable for structural and functional studies.[4] To address this problem, different expression hosts have been used for heterologous expression of proteins from *Mtb* and other pathogenic mycobacterial species.

### *Baculovirus/insect host cells*

The baculovirus/insect cell expression system is one of the most widely used systems for routine production of recombinant proteins.[9] This expression system was one of the first to be used in attempts to produce *Mtb* proteins in an alternative expression host[10] and has been used for production of a limited number of mycobacterial proteins.[11,12] This expression system has, however, been overtaken by other expression hosts with closer evolutionary links to mycobacteria.

### *Streptomyces host cells*

Mycobacteria and Streptomyces are members of the Actinobacteria family, and therefore share close evolutionary links. Initial observations that *Streptomyces*

**Table I.** *Examples of* Mtb *Proteins Expressed in* M. smegmatis *Expression Host*

| ORF | Function/ annotation | Molecular weight (kDa) | Theoretical pI | Solubility | | Oligomeric state | References |
|---|---|---|---|---|---|---|---|
| | | | | Soluble | Insoluble | | |
| Rv0065 | VapC1 | 14.3 | 4.8 | ✓ | | | 79 |
| Rv0132c | fHMAD | 38.4 | 6.4 | ✓ | | Dimer | 63,80 |
| Rv0407 | FGD1 | 36.9 | 5.1 | ✓ | | Dimer | 72 |
| Rv0413 | MutT3 | 23.4 | 4.8 | | ✓ | | 4 |
| Rv0475 | HBHA | 21.5 | 9.8 | ✓ | | | 59 |
| Rv0542c | MenE | 36.5 | 5.6 | | ✓ | | n/a |
| Rv0617 | VapC29 | 13.9 | 5.7 | ✓ | | | 79 |
| Rv0901 | CP | 18.9 | 10.1 | | | | 81 |
| Rv0993 | GalU | 32.3 | 4.6 | ✓ | | | 82 |
| Rv1096 | Possible glycosyl transferase | 31 | 7 | ✓ | | | 83 |
| Rv1160 | MutT2 | 15.1 | 6.1 | ✓ | | | 4 |
| Rv1186c | PruC | 57.5 | 5.9 | ✓[a] | | | 84 |
| Rv1187 | PruA | 58.8 | 5.7 | ✓ | | Dimer | 73 |
| Rv1188 | PruB | 36.2 | 8.3 | ✓[a] | | | 73 |
| Rv1212c | GlgA | 41.5 | 6.3 | ✓ | | | 82 |
| Rv1213 | GlgC | 43.8 | 4.9 | ✓ | | | 82 |
| Rv1601 | HisB | 22.7 | 6.8 | ✓ | | Monomer | 85 |
| Rv1665 | PKS11 | 37.6 | 5.6 | ✓ | | Dimer | 71 |
| Rv1886c | FbpB | 34.5 | 5.7 | ✓ | | Monomer | 86 |
| Rv1980c | MPT64 | 24.8 | 4.6 | ✓ | | | 41 |
| Rv1991c | MazF6 | 12.2 | 9.9 | ✓ | | | 44 |
| Rv2141c | DapE2 | 48.1 | 4.7 | ✓ | | | 4 |
| Rv2220 | GlnA1 | 53.5 | 4.8 | ✓ | | Dodecamer | 87,88 |
| Rv2245 | KasA | 43.2 | 4.9 | ✓ | | Dimer | 89 |
| Rv2420c | RsfS | 13.8 | 4.5 | ✓ | | Dimer | PDB code 4WCW |
| Rv2611c | Possible acetyltransferase | 35 | 9.1 | ✓ | | | 4 |
| Rv2903c | LepB | 31.8 | 5.7 | | ✓ | | n/a |
| Rv2924c | MutM | 31.9 | 10.3 | ✓ | | | 4 |
| Rv3127 | CP | 38.4 | 7.6 | | ✓ | | 4 |
| Rv3177 | Possible peroxidase | 30.8 | 8.9 | | ✓ | | 4 |
| Rv3248c | SahH | 54.3 | 4.8 | ✓ | | | 90 |
| Rv3490 | OtsA | 55.8 | 6.7 | ✓ | | | 82 |
| Rv3547 | Ddn | 17.3 | 9.9 | ✓[a] | | | 4 |
| Rv3565 | AspB | 41 | 4.6 | ✓ | | | 91 |
| Rv3713 | CobQ2 | 24.1 | 6.1 | ✓ | | | 4 |
| Rv3772 | HisC2 | 38 | 6.5 | ✓ | | | 92 |
| Rv3855 | EthR | 23.7 | 4.9 | ✓ | | Dimer | 93 |
| Rv3874 | CFP-10 | 10.8 | 4.3 | ✓ | | Dimer | 94 |

[a] Soluble aggregates.
CP, conserved protein.

*lividans* (*S. lividans*) could utilize the *Mycobacterium bovis* (*M. bovis*) BCG translational signals suggested its use as an expression host for mycobacterial genes.[13] *S. lividans* is also an efficient microorganism for expression and secretion of recombinant proteins[14] and has thus been successfully used to express and purify *Mtb* proteins from the culture medium.[15] *S. lividans* also supports glycosylation of *Mtb* proteins,[16] reinforcing its usefulness for heterologous production of proteins from pathogenic mycobacteria.

### *Mycobacterial host cells*

Proteins from pathogenic mycobacterial species have been previously expressed in non-pathogenic mycobacteria, including *Mycobacterium smegmatis* (*M. smegmatis*)[17] and *Mycobacterium vaccae* (*M. vaccae*).[18] The former, however, is now the species of choice, given its successful application as an efficient expression host (Table I). In this review, we discuss the various expression systems and strategies developed for protein expression and purification from *M. smegmatis* host cells.

### M. smegmatis *expression systems*

Since the first report of introduction of foreign DNA into mycobacteria in 1987,[19] *M. smegmatis* has been used as a model for pathogenic and slower-growing mycobacterial species. Isolation of an efficient plasmid-transformation mutant strain of *M. smegmatis* ($mc^2155$) from the wild-type ATCC 607 strain[20] established an effective way of delivering DNA into this strain. Expression of foreign genes in mycobacteria was later made possible by generating

shuttle vectors that used the *M. bovis* BCG *hsp60* promoter for constitutive expression[21] (*e.g.* pMV261 and pMV361). Inducible expression systems, on the other hand, provide a means of regulating gene expression and allow direct comparison of the induced and non-induced cells. Given the wide range of plasmids generated by the mycobacterial research community,[22] a variety of inducible expression systems have been used to express and purify recombinant proteins from *M. smegmatis* host cells.

### Acetamidase promoter-based expression system

This system makes use of the inducible acetamidase promoter of *M. smegmatis*.[23] Protein expression is induced by the addition of acetamide to cell cultures,[24] although the induction mechanism is not fully understood. The vector pJAM2 was developed by Triccas *et al*. in 1998[25] through cloning a 1.5 kb fragment of the acetamidase promoter of *M. smegmatis* into the pJEM12 shuttle vector.[26] This was the first mycobacterial expression vector possessing two important characteristics; the ability to regulate high-level expression of recombinant genes (acetamide induction) and a simple and efficient method to purify the recombinant proteins (through a $His_6$-tag). The pJAM2 vector is reported to be unstable for expression of proteins in *Mtb*, due to deletion of the promoter region.[27]

A similar system was later developed to study the proteins present in the region of deletion 1 (RD1) of *Mtb*.[28] The pSD vector series generated in this study were constructed by inserting a 2.6 kb fragment of the *M. smegmatis* acetamidase promoter into the promoterless plasmid pMV206. Recombinant proteins could be expressed with $His_6$- or HA-tags for downstream analyses.[28]

### T7 promoter-based vectors

The T7 expression system for *E. coli*[29] provides the most common strategy used for heterologous protein production. In this system, the RNA polymerase from T7 bacteriophage transcribes the genes cloned downstream of the T7 promoter.[29] Exploiting the efficiency of this system in *E. coli* and its adaptability for use with other expression systems,[30] T7 promoter-based vectors[31] (*e.g.* pYUB1049 and pYUB1062 shuttle vectors) have been developed for protein expression in *M. smegmatis* host cells. These vectors have the advantage that expression and solubility trials can be carried out in parallel in *E. coli* and *M. smegmatis*.

The pYUB1049 vector has been used as the basis for three further vectors; pDESTsmg,[4] pYUB28b and pYUBDuet.[32] pDESTsmg is a destination vector for the Gateway® cloning system with the capability of expressing recombinant proteins with an N-terminal $His_6$-tag.[4] The pYUB28b vector harbors the multiple cloning site from the pET28b plasmid (Novagen) and can express proteins with N– and C–terminal $His_6$–tags.[32] The pYUBDuet vector is a co-expression vector for simultaneous expression of two genes from two separate T7 promoters.[32] The multiple cloning site for this vector comes from pETDuet-1 vector (Novagen), expressing one of the proteins with an N-terminal $His_6$-tag and the other one with a C-terminal S-tag.

### Tetracycline-inducible expression system

This system has been widely used in both prokaryotic and eukaryotic organisms, and is capable of regulating gene expression both *in vitro* and *in vivo*.[33] Expression in this system is regulated through a constitutively expressed tetracycline repressor (TetR) protein, where TetR binding to the operator prevents transcription in the absence of inducer (*e.g.* anhydrotetracycline). Inducer binding to TetR, however, relieves the repression and transcription takes place.[34] Considering the ability of the tetracycline-inducible expression system to tightly regulate gene expression in mycobacteria,[35,36] it has been used in various contexts for conditional expression of essential genes in pathogenic mycobacteria.[33,37]

A series of tetracycline-inducible Gateway®-compatible vectors have been reported that use tandem affinity purification tags ($His_6$-Strep-Strep) at either N- or C-termini of recombinant proteins.[38] Another vector toolbox (pST series) allows tetracycline-inducible expression of proteins with N-terminal ($His_6$-tag) and C-terminal (FLAG) tags.[39] The authors of this work have also generated a suicide plasmid for creating gene replacements, a plasmid for chromosomal integration, a plasmid for gene expression under hypoxic conditions and a vector for co-expression of two proteins under control of two independent promoters.[39]

### L-lactamase promoter-based expression system

Expression of recombinant proteins in this system is under the control of the β-lactamase promoter of *Mycobacterium fortuitum*.[40] This expression system has been mainly used to express and purify a variety of mycobacterial antigens from *M. smegmatis*.[41–43] The protein yield of the *Mycobacterium leprae* 35 kDa antigen was reported to be similar to that observed with the acetamidase-based expression system[25] but purification is based on antibody affinity chromatography.

### Other expression systems

Alternative inducible expression systems for mycobacterial proteins include arabinose-inducible,[44] pristinamycin-inducible,[45] nitrile-responsive, and thiostrepton-responsive systems.[46] These expression systems provide a means of regulating gene expression in a wider range of mycobacterial hosts. The latter two expression systems have been reviewed elsewhere.[46]

### *Co-expression systems*

A number of methods for investigating protein-protein interactions in mycobacterial host cells have been reported, based on bacterial two-hybrid systems.[47] These methods include the mycobacterial protein fragment complementation (M-PFC)[48] and Split-Trp[49] systems. These do not, however, allow purification of the native protein complex due to the presence of fused reporter proteins.

In another approach, a tunable co-expression vector (pTetCoex) has been developed, based on the tetracycline-inducible expression system, to study protein-protein interactions in *M. smegmatis* cells.[50] This vector allows co-expression to be optimized by keeping the expression of one protein constant, while changing the expression level of its partner. Expression is monitored by antibody-based methods.[50]

As described earlier, pYUBDuet[32] and pST-2K[39] are co-expression vectors that can be used to overexpress and purify stable protein complexes from *M. smegmatis* host cells. A recent report describes expression of a recombinant protein in *M. smegmatis* host cells, to be used as a bait for identification and purification of partner proteins that show weak or transient interactions, using *in vivo* formaldehyde crosslinking.[38] This method could potentially also be used with the aforementioned co-expression vectors, where the components of protein complex are overexpressed together.

## M. smegmatis *strains*

Since the isolation of the *M. smegmatis* $mc^2155$ strain,[20] it has been used as the workhorse of mycobacterial research, especially for expressing genes from mycobacterial species. It has been speculated that the *M. smegmatis* $mc^2155$ strain has acquired changes in plasmid replication, enabling a high propensity for transformation.[20] Observed modifications in the composition and structure of the cell envelope may also contribute to explain this useful property.[51] The exact mechanism by which this strain has acquired its high transformability, however, is not yet understood.[51]

The *M. smegmatis* $mc^24517$ strain is an expression strain that has been generated to make use of T7 promoter-based expression system.[31] In this strain, a copy of T7 RNA polymerase is integrated into the chromosome under control of the acetamidase promoter from *M. smegmatis*.[31] T7 RNA polymerase production is induced with acetamide and in turn transcribes the recombinant genes for subsequent translation. Expression can also be induced using lactose or IPTG in this system.[52] This strain is used to express proteins using T7 promoter-based vectors described earlier.

When recombinant proteins containing a $His_6$-tag are expressed from $mc^2155$ or $mc^24517$ strains, a major contaminant protein (the Hsp60 chaperone GroEL1) often co-purifies with the proteins.[4,53] This protein has a histidine-rich C-terminal tail, enabling it to bind to the purification beads. A mutant strain of *M. smegmatis* $mc^2155$, so called GroEL1ΔC, has been generated to improve purification efficiency by removing the coding sequence for the histidine residues.[54]

## M. smegmatis *advantages*

The rationale behind using mycobacterial host cells for production of mycobacterial proteins is that these micro-organisms are more closely related and could therefore provide the specific requirements for proper folding, soluble expression and production of biologically and immunologically active proteins.

### *Molecular chaperones*

The co-purification of *M. smegmatis* GroEL1 protein with recombinant proteins,[4,53] described above, implies that the presence of specific mycobacterial chaperones may facilitate the efficient folding and soluble expression of *Mtb* proteins. It is interesting to note that the *Mtb* homologs of GroEL are different from the *E. coli* counterparts, in that they are ATP independent and do not form the classical assemblies observed in *E. coli*.[55]

### *Post-translational modifications*

Post-translational modifications of proteins are more common in eukaryotes, although new technologies have revealed the presence of a diverse set of post-translational modifications in prokaryotes.[56] A wide range of post-translational modifications have been reported in the *Mtb* proteome,[57] and a clear advantage of mycobacterial host cells is their capability for post-translational modifications of the recombinant proteins.[28,58–61] The ability of mycobacterial expression systems to provide physiologically relevant post-translational modifications may in part contribute to their efficient production of biologically and immunologically active proteins.

### *Metabolites/ligands*

Mycobacteria occupy a distant location on the evolutionary tree compared with *E. coli*, consistent with their unique environmental niches and physiological needs. This could, in principle, be associated with differences in biochemical pathways, enzymatic functions and metabolites in mycobacteria. For instance, a metabolome-based reaction network analysis revealed that reactions involving glutamate are central to the metabolism of *Mtb* and *M. leprae*, whereas pyruvate fills this role in the *E. coli* metabolome.[62]

A great advantage of using *M. smegmatis* as an expression host is the presence of ligands that might not be present in *E. coli*. In this context, co-purification of a rare flavin derivative, coenzyme

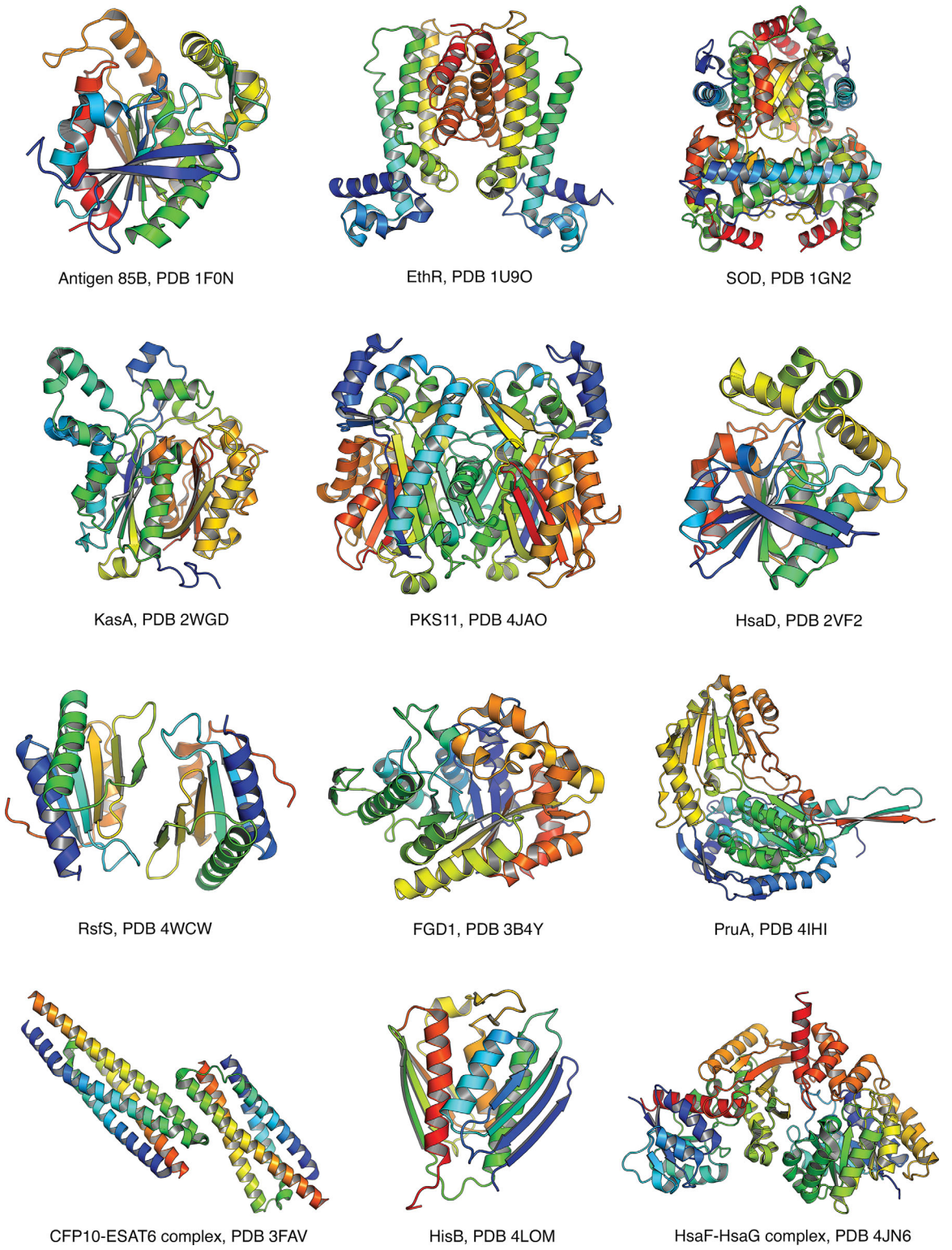


**Figure 2.** Crystal structure of *Mtb* proteins expressed using alternative expression hosts. All proteins were expressed in *M. smegmatis* expression host, except for 1GN2 (*M. vaccae*), 4JN6 (*Rhodococcus jostii*), and 2VF2 (*Pseudomonas putida*).

$F_{420}$, has been demonstrated upon over-expression and purification of an *Mtb* protein from *M. smegmatis*.[63] This coenzyme is sporadically present in microorganisms and not present in *E. coli*, likely explaining the unsuccessful attempts to express this as a soluble protein in *E. coli*.[63]

### *Applications in structural biology*

In macromolecular crystallography, single- and multi-wavelength anomalous diffraction (SAD/MAD) methods provide a powerful approach for experimental phasing of crystal structures.[64] Selenomethionine(SeMet)-substituted proteins provide a valuable

method for the solution of protein crystal structures using SAD/MAD methods.[65] These proteins are prepared by replacing methionine residues with SeMet during protein expression. SeMet incorporation protocols have already been described for *E. coli*,[66] baculovirus,[67] mammalian,[68] and yeast[69] expression systems, as well as *in vitro* expression system.[70]

Both T7 promoter-based[52] and acetamide-based[71] expressions systems in *M. smegmatis* have been successfully reported to prepare SeMet-substituted proteins for experimental phasing. This method has led to determination of the crystal structures of three proteins from *Mtb* by the use of MAD phasing; $F_{420}$-dependent glucose-6-phosphate dehydrogenase (FGD1),[72] polyketide synthase 11 (PKS11),[71] and $\Delta^1$-pyrroline-5-carboxylic dehydrogenase (PruA).[73]

A search of the PDB for entries with a reported alternative expression host for mycobacterial proteins reveals the presence of 40 structures, belonging to 12 unique proteins from *Mtb* (Fig. 2). Nine of these 12 proteins were produced in *M. smegmatis* and one in each of *M. vaccae*,[18] *Rhodococcus jostii,*[74] and *Pseudomonas putida*[75] as the expression host (Fig. 2). Although the present list may not be exhaustive, it clearly demonstrates the successful application of *M. smegmatis* as a host for obtaining proteins in the quantity and high quality required for X-ray crystallography.

### *Practical aspects*

Given the different doubling time of *M. smegmatis* cells (~3 h) compared with that of *E. coli* (~20 min) and the waxy nature of the mycobacterial cell wall, growth conditions, and protocols should be accordingly adapted for *M. smegmatis* cultures.[76] Once transformed, it takes ~3 days for colonies to appear on plates, and starting from single colonies it would take 2–3 days for 5–10 mL starter cultures to grow. Expression cultures are inoculated at a dilution of 1:100 using starter cultures and expression is usually carried out for 3–4 days. The cultures must also be supplemented with detergents (*e.g.* Tween80 or tyloxapol) to prevent cell clumping.

The protein yield obtained from *M. smegmatis* cultures varies significantly, as for other expression systems. In the author's lab, ~10 mg protein per liter of cell culture is achieved for some proteins[52] while a few others are expressed as insoluble proteins. Our early results on using *M. smegmatis* as a rescue expression host indicated that >60% of proteins that formed inclusion bodies when expressed in *E. coli* are soluble when expressed in *M. smegmatis*.[4] Since then, we have successfully used *M. smegmatis* to express and purify over 25 mycobacterial proteins, leading to crystallization and determination of the crystal structures for eight of these proteins. These results support the use of *M. smegmatis* host cells as a routine strategy for protein production rather than just a rescue/salvage one (Table I).

Expression of recombinant proteins in *M. smegmatis* sometimes results in slow growth, very low or no expression or insoluble expression. These symptoms may be due to toxicity of these particular proteins, imposing heavy physiological burden on to the cells. A number of methods could be used to resolve this issue, including using plasmids with lower copy numbers,[22] use of weaker promoters with lower expression level[77] or using a promoter that is more tightly regulated.[39] Alternatively, other expression hosts that are described in this manuscript could be of use. In addition, some proteins that are part of hetero-oligomeric protein complexes might not be expressed as individual proteins, in which case co-expression with the partner protein could rectify the expression problem.[50]

## Concluding Remarks

Genetic manipulation of mycobacterial species has gone through an extensive progression over the last 30 years, leading to a better understanding of the function, structure and regulation of complex biological mechanisms in mycobacteria. Determination of the genome sequences for two major pathogens, *Mtb*[5] and *M. leprae*,[78] has since opened up a new window to investigate the biology and pathogenesis of these micro-organisms through a wide range of methods. The ultimate aim of the studies is to clear the way for novel and better therapeutics and diagnostics through studying the structure and function of mycobacterial proteins. Using *M. smegmatis* as an expression host can enable researchers to tackle proteins that were not amenable before and thereby produce biologically and immunologically active proteins for research and therapeutic applications.

## Acknowledgments

We apologize to our colleagues whose results are not covered in this review due to space limitations and also the focus of review. We thank all our colleagues in the Structural Biology Laboratory that provided the results of their research prior to publication. Research in the Baker lab on TB is supported by grants through the Health Research Council of New Zealand.

## References

1. Brondyk WH (2009) Selecting an appropriate method for expressing a recombinant protein. Methods Enzymol 463:131–147.
2. Bellinzoni M, Riccardi G (2003) Techniques and applications: the heterologous expression of *Mycobacterium tuberculosis* genes is an uphill road. Trends Microbiol 11:351–358.
3. Baneyx F (1999) Recombinant protein expression in *Escherichia coli*. Curr Opin Biotechnol 10:411–421.
4. Goldstone RM, Moreland NJ, Bashiri G, Baker EN, Shaun Lott J (2008) A new Gateway vector and

expression protocol for fast and efficient recombinant protein expression in *Mycobacterium smegmatis*. Protein Expr Purif 57:81–87.

5. Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, Gordon SV, Eiglmeier K, Gas S, Barry CE Jr, Tekaia F, Badcock K, Basham D, Brown D, Chillingworth T, Connor R, Davies R, Devlin K, Feltwell T, Gentles S, Hamlin N, Holroyd S, Hornsby T, Jagels K, Krogh A, McLean J, Moule S, Murphy L, Oliver K, Osborne J, Quail MA, Rajandream M-A, Rogers J, Rutter S, Seeger K, Skelton J, Squares S, Squares R, Sulston JE, Taylor K, Whitehead S, Barrell BG (1998) Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. Nature 393:537–544.
6. Blattner FR, Plunkett G, Bloch CA, Perna NT, Burland V, Riley M, Collado-Vides J, Glasner JD, Rode CK, Mayhew GF, Gregor J, Davis NW, Kirkpatrick HA, Goeden MA, Rose DJ, Mau B, Shao Y (1997) The complete genome sequence of *Escherichia coli* K-12. Science 277:1453–1462.
7. Fernandez FJ, Vega MC (2013) Technologies to keep an eye on: alternative hosts for protein production in structural biology. Curr Opin Struct Biol 23:365–373.
8. Terwilliger TC, Park MS, Waldo GS, Berendzen J, Hung LW, Kim CY, Smith CV, Sacchettini JC, Bellinzoni M, Bossi R, De Rossi E, Mattevi A, Milano A, Riccardi G, Rizzi M, Roberts MM, Coker AR, Fossati G, Mascagni P, Coates AR, Wood SP, Goulding CW, Apostol MI, Anderson DH, Gill HS, Eisenberg DS, Taneja B, Mande S, Pohl E, Lamzin V, Tucker P, Wilmanns M, Colovos C, Meyer-Klaucke W, Munro AW, McLean KJ, Marshall KR, Leys D, Yang JK, Yoon HJ, Lee BI, Lee MG, Kwak JE, Han BW, Lee JY, Baek SH, Suh SW, Komen MM, Arcus VL, Baker EN, Lott JS, Jacobs WJ, Alber T, Rupp B (2003) The TB structural genomics consortium: a resource for *Mycobacterium tuberculosis* biology. Tuberculosis 83:223–249.
9. Kost TA, Condreay JP, Jarvis DL (2005) Baculovirus as versatile vectors for protein expression in insect and mammalian cells. Nat Biotechnol 23:567–575.
10. Atkins D, al-Ghusein H, Prehaud C, Coates AR (1994) Overproduction and purification of *Mycobacterium tuberculosis* chaperonin 10. Gene 150:145–148.
11. Bellinzoni M, De Rossi E, Branzoni M, Milano A, Peverali FA, Rizzi M, Riccardi G (2002) Heterologous expression, purification, and enzymatic activity of *Mycobacterium tuberculosis* NAD(+) synthetase. Protein Expr Purif 25:547–557.
12. Bigi F, Taboga O, Romano MI, Alito A, Fisanotti JC, Cataldi AA (1999) Expression of the *Mycobacterium bovis* P36 gene in *Mycobacterium smegmatis* and the baculovirus/insect cell system. Braz J Med Biol Res 32:29–37.
13. Kieser T, Moss MT, Dale JW, Hopwood DA (1986) Cloning and expression of *Mycobacterium bovis* BCG DNA in *Streptomyces lividans*. J Bacteriol 168:72–80.
14. Lammertyn E, Anne J (1998) Modifications of Streptomyces signal peptides and their effects on protein production and secretion. FEMS Microbiol Lett 160:1–10.
15. Vallin C, Ramos A, Pimienta E, Rodriguez C, Hernandez T, Hernandez I, Del Sol R, Rosabal G, Van Mellaert L, Anne J (2006) Streptomyces as host for recombinant production of *Mycobacterium tuberculosis* proteins. Tuberculosis 86:198–202.
16. Lara M, Servin-Gonzalez L, Singh M, Moreno C, Cohen I, Nimtz M, Espitia C (2004) Expression, secretion, and glycosylation of the 45- and 47-kDa glycoprotein of *Mycobacterium tuberculosis* in *Streptomyces lividans*. Appl Environ Microbiol 70:679–685.
17. Triccas JA, Ryan AA (2009) Heterologous expression of genes in mycobacteria. Methods Mol Biol 465:243–253.
18. Cooper JB, McIntyre K, Badasso MO, Wood SP, Zhang Y, Garbe TR, Young D (1995) X-ray structure analysis of the iron-dependent superoxide dismutase from *Mycobacterium tuberculosis* at 2.0 Angstroms resolution reveals novel dimer-dimer interactions. J Mol Biol 246: 531–544.
19. Jacobs WR Jr, Tuckman M, Bloom BR (1987) Introduction of foreign DNA into mycobacteria using a shuttle phasmid. Nature 327:532–535.
20. Snapper SB, Melton RE, Mustafa S, Kieser T, Jacobs JWR (1990) Isolation and characterization of efficient plasmid transformation mutants of *Mycobacterium smegmatis*. Mol Microbiol 4:1911–1919.
21. Stover CK, de la Cruz VF, Fuerst TR, Burlein JE, Benson LA, Bennett LT, Bansal GP, Young JF, Lee MH, Hatfull GF, Snapper SB, Barletta RG, Jacobs Jr WR, Bloom BR (1991) New use of BCG for recombinant vaccines. Nature 351:456–460.
22. Movahedzadeh F, Bitter W (2009) Ins and outs of mycobacterial plasmids. Methods Mol Biol 465:217–228.
23. Mahenthiralingam E, Draper P, Davis EO, Colston MJ (1993) Cloning and sequencing of the gene which encodes the highly inducible acetamidase of *Mycobacterium smegmatis*. J Gen Microbiol 139:575–583.
24. Draper P (1967) The aliphatic acylamide amidohydrolase of *Mycobacterium smegmatis*: its inducible nature and relation to acyl-transfer to hydroxylamine. J Gen Microbiol 46:111–123.
25. Triccas JA, Parish T, Britton WJ, Gicquel B (1998) An inducible expression system permitting the efficient purification of a recombinant antigen from *Mycobacterium smegmatis*. FEMS Microbiol Lett 167:151–156.
26. Timm J, Lim EM, Gicquel B (1994) *Escherichia coli*-mycobacteria shuttle vectors for operon and gene fusions to *lacZ*: the pJEM series. J Bacteriol 176:6749–6753.
27. Brown AC, Parish T (2006) Instability of the acetamide-inducible expression vector pJAM2 in *Mycobacterium tuberculosis*. Plasmid 55:81–86.
28. Daugelat S, Kowall J, Mattow J, Bumann D, Winter R, Hurwitz R, Kaufmann SH (2003) The RD1 proteins of *Mycobacterium tuberculosis*: expression in *Mycobacterium smegmatis* and biochemical characterization. Microbes Infect 5:1082–1095.
29. Studier FW, Moffatt BA (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. J Mol Biol 189:113–130.
30. Lussier FX, Denis F, Shareck F (2010) Adaptation of the highly productive T7 expression system to *Streptomyces lividans*. Appl Environ Microbiol 76:967–970.
31. Wang F, Jain P, Gulten G, Liu Z, Feng Y, Ganesula K, Motiwala AS, Ioerger TR, Alland D, Vilcheze C, Jacobs WR, Jr., Sacchettini JC (2010) *Mycobacterium tuberculosis* dihydrofolate reductase is not a target relevant to the antitubercular activity of isoniazid. Antimicrob Agents Chemother 54:3776–3782.
32. Bashiri G, Rehan AM, Greenwood DR, Dickson JMJ, Baker EN (2010) Metabolic engineering of cofactor $F_{420}$ production in *Mycobacterium smegmatis*. PLoS ONE 5: e15803.
33. Carroll P, Muttucumaru DG, Parish T (2005) Use of a tetracycline-inducible system for conditional expression in *Mycobacterium tuberculosis* and *Mycobacterium smegmatis*. Appl Environ Microbiol 71:3077–3084.
34. Berens C, Hillen W (2003) Gene regulation by tetracyclines. Eur J Biochem 270:3109–3121.

35. Blokpoel MC, Murphy HN, O'Toole R, Wiles S, Runn ES, Stewart GR, Young DB, Robertson BD (2005) Tetracycline-inducible gene regulation in mycobacteria. Nucleic Acids Res 33:e22.
36. Ehrt S, Guo XV, Hickey CM, Ryou M, Monteleone M, Riley LW, Schnappinger D (2005) Controlling gene expression in mycobacteria with anhydrotetracycline and Tet repressor. Nucleic Acids Res 33:e21.
37. Kim JH, O'Brien KM, Sharma R, Boshoff HI, Rehren G, Chakraborty S, Wallach JB, Monteleone M, Wilson DJ, Aldrich CC, Barry CE, Rhee KY, Ehrt S, Schnappinger D (2013) A genetic strategy to identify targets for the development of drugs that prevent bacterial persistence. Proc Natl Acad Sci U S A 110:19095–19100.
38. Lougheed KE, Bennett MH, Williams HD (2014) An *in vivo* crosslinking system for identifying mycobacterial protein-protein interactions. J Microbiol Methods 105C: 67–71.
39. Parikh A, Kumar D, Chawla Y, Kurthkoti K, Khan S, Varshney U, Nandicoori VK (2013) Development of a new generation of vectors for gene expression, gene replacement, and protein-protein interaction studies in mycobacteria. Appl Environ Microbiol 79:1718–1729.
40. Timm J, Perilli MG, Duez C, Trias J, Orefici G, Fattorini L, Amicosante G, Oratore A, Joris B, Frere JM, Pugsley AP, Gicquel B (1994) Transcription and expression analysis, using *lacZ* and *phoA* gene fusions, of *Mycobacterium fortuitum* beta-lactamase genes cloned from a natural isolate and a high-level beta-lactamase producer. Mol Microbiol 12:491–504.
41. Roche PW, Winter N, Triccas JA, Feng CG, Britton WJ (1996) Expression of *Mycobacterium tuberculosis* MPT64 in recombinant *Myco. smegmatis*: purification, immunogenicity and application to skin tests for tuberculosis. Clin Exp Immunol 103:226–232.
42. Triccas JA, Roche PW, Winter N, Feng CG, Butlin CR, Britton WJ (1996) A 35-kilodalton protein is a major target of the human immune response to *Mycobacterium leprae*. Infect Immun 64:5171–5177.
43. Triccas JA, Winter N, Roche PW, Gilpin A, Kendrick KE, Britton WJ (1998) Molecular and immunological analyses of the *Mycobacterium avium* homolog of the immunodominant *Mycobacterium leprae* 35-kilodalton protein. Infect Immun 66:2684–2690.
44. Carroll P, Brown AC, Hartridge AR, Parish T (2007) Expression of *Mycobacterium tuberculosis* Rv1991c using an arabinose-inducible promoter demonstrates its role as a toxin. FEMS Microbiol Lett 274:73–82.
45. Forti F, Crosta A, Ghisotti D (2009) Pristinamycin-inducible gene regulation in mycobacteria. J Biotechnol 140:270–277.
46. Sassetti CM (2009) Inducible expression systems for mycobacteria. Methods Mol Biol 465:255–264.
47. Karimova G, Pidoux J, Ullmann A, Ladant D (1998) A bacterial two-hybrid system based on a reconstituted signal transduction pathway. Proc Natl Acad Sci U S A 95:5752–5756.
48. Singh A, Mai D, Kumar A, Steyn AJ (2006) Dissecting virulence pathways of *Mycobacterium tuberculosis* through protein-protein association. Proc Natl Acad Sci U S A 103:11346–11351.
49. O'Hare H, Juillerat A, Dianiskova P, Johnsson K (2008) A split-protein sensor for studying protein-protein interaction in mycobacteria. J Microbiol Methods 73:79–84.
50. Chang Y, Mead D, Dhodda V, Brumm P, Fox BG (2009) One-plasmid tunable coexpression for mycobacterial protein-protein interaction studies. Protein Sci 18: 2316–2325.
51. Etienne G, Laval F, Villeneuve C, Dinadayala P, Abouwarda A, Zerbib D, Galamba A, Daffe M (2005) The cell envelope structure and properties of *Mycobacterium smegmatis* mc2155: is there a clue for the unique transformability of the strain? Microbiology 151:2075–2086.
52. Bashiri G, Squire CJ, Baker EN, Moreland NJ (2007) Expression, purification and crystallization of native and selenomethionine labeled *Mycobacterium tuberculosis* FGD1 (Rv0407) using a *Mycobacterium smegmatis* expression system. Protein Expr Purif 54:38–44.
53. Poulsen C, Akhter Y, Jeon AH, Schmitt-Ulms G, Meyer HE, Stefanski A, Stuhler K, Wilmanns M, Song YH (2010) Proteome-wide identification of mycobacterial pupylation targets. Mol Syst Biol 6:386.
54. Noens EE, Williams C, Anandhakrishnan M, Poulsen C, Ehebauer MT, Wilmanns M (2011) Improved mycobacterial protein production using a *Mycobacterium smegmatis* groEL1DeltaC expression strain. BMC Biotechnol 11:27.
55. Qamra R, Mande SC, Coates AR, Henderson B (2005) The unusual chaperonins of *Mycobacterium tuberculosis*. Tuberculosis 85:385–394.
56. Cain JA, Solis N, Cordwell SJ (2014) Beyond gene expression: the impact of protein post-translational modifications in bacteria. J Proteomics 97:265–286.
57. van Els C, Corbière V, Smits K, van Gaans-van den Brink J, Poelen M, Mascart F, Meiring H, Locht C (2014) Toward understanding the essence of post-translational modifications for the *Mycobacterium tuberculosis* immunoproteome. Front Immunol 5:[PAGE #S].
58. Herrmann JL, O'Gaora P, Gallagher A, Thole JE, Young DB (1996) Bacterial glycoproteins: a link between glycosylation and proteolytic cleavage of a 19 kDa antigen from *Mycobacterium tuberculosis*. EMBO J 15:3547–3554.
59. Delogu G, Bua A, Pusceddu C, Parra M, Fadda G, Prennan MJ, Zanetti S (2004) Expression and purification of recombinant methylated HBHA in *Mycobacterium smegmatis*. FEMS Microbiol Lett 239:33–39.
60. Garbe T, Harris D, Vordermeier M, Lathigra R, Ivanyi J, Young D (1993) Expression of the *Mycobacterium tuberculosis* 19-kilodalton antigen in *Mycobacterium smegmatis*: immunological analysis and evidence of glycosylation. Infect Immun 61:260–267.
61. Herrmann JL, Delahay R, Gallagher A, Robertson B, Young D (2000) Analysis of post-translational modification of mycobacterial proteins using a cassette expression system. FEBS Lett 473:358–362.
62. Verkhedkar KD, Raman K, Chandra NR, Vishveshwara S (2007) Metabolome based reaction graphs of *M. tuberculosis* and *M. leprae*: a comparative network analysis. PLoS One 2:e881.
63. Bashiri G, Perkowski EF, Turner AP, Feltcher ME, Braunstein M, Baker EN (2012) Tat-dependent translocation of an $F_{420}$-binding protein of *Mycobacterium tuberculosis*. PLoS One 7:e45003.
64. Hendrickson WA (1991) Determination of macromolecular structures from anomalous diffraction of synchrotron radiation. Science 254:51–58.
65. Hendrickson WA, Horton JR, LeMaster DM (1990) Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure. EMBO J 9:1665–1672.
66. Doublie S (1997) Preparation of selenomethionyl proteins for phase determination. Methods Enzymol 276:523–530.

67. Bellizzi JJ, Widom J, Kemp CW, Clardy J (1999) Producing selenomethionine-labeled proteins with a baculovirus expression vector system. Structure 7:R263–R267.
68. Barton WA, Tzvetkova-Robev D, Erdjument-Bromage H, Tempst P, Nikolov DB (2006) Highly efficient selenomethionine labeling of recombinant proteins produced in mammalian cells. Protein Sci 15:2008–2013.
69. Malkowski MG, Quartley E, Friedman AE, Babulski J, Kon Y, Wolfley J, Said M, Luft JR, Phizicky EM, DeTitta GT, Grayhack EJ (2007) Blocking S-adenosylmethionine synthesis in yeast allows selenomethionine incorporation and multiwavelength anomalous dispersion phasing. Proc Natl Acad Sci U S A 104:6678–6683.
70. Kigawa T, Yamaguchi-Nunokawa E, Kodama K, Matsuda T, Yabuki T, Matsuda N, Ishitani R, Nureki O, Yokoyama S (2001) Selenomethionine incorporation into a protein by cell-free synthesis. J Struct Funct Genomics 2:29–35.
71. Gokulan K, O'Leary SE, Russell WK, Russell DH, Lalgondar M, Begley TP, Ioerger TR, Sacchettini JC (2013) Crystal structure of *Mycobacterium tuberculosis* polyketide synthase 11 (PKS11) reveals intermediates in the synthesis of methyl-branched alkylpyrones. J Biol Chem 288:16484–16494.
72. Bashiri G, Squire CJ, Moreland NJ, Baker EN (2008) Crystal structures of $F_{420}$-dependent glucose-6-phosphate dehydrogenase FGD1 involved in the activation of the anti-tuberculosis drug candidate PA-824 reveal the basis of coenzyme and substrate binding. J Biol Chem 283:17531–17541.
73. Lagautriere T, Bashiri G, Paterson NG, Berney M, Cook GM, Baker EN (2014) Characterization of the proline-utilization pathway in *Mycobacterium tuberculosis* through structural and functional studies. Acta Cryst D70:968–980.
74. Carere J, McKenna SE, Kimber MS, Seah SY (2013) Characterization of an aldolase-dehydrogenase complex from the cholesterol degradation pathway of *Mycobacterium tuberculosis*. Biochemistry 52:3502–3511.
75. Lack N, Lowe ED, Liu J, Eltis LD, Noble ME, Sim E, Westwood IM (2008) Structure of HsaD, a steroid-degrading hydrolase, from *Mycobacterium tuberculosis*. Acta Cryst F64:2–7.
76. Singh AK, Reyrat JM (2009) Laboratory maintenance of *Mycobacterium smegmatis*. Curr Protoc Microbiol Chapter 10:Unit10C 1.
77. Spratt JM, Britton WJ, Triccas JA (2003) Identification of strong promoter elements of *Mycobacterium smegmatis* and their utility for foreign gene expression in mycobacteria. FEMS Microbiol Lett 224:139–142.
78. Cole ST, Eiglmeier K, Parkhill J, James KD, Thomson NR, Wheeler PR, Honore N, Garnier T, Churcher C, Harris D, Mungall K, Basham D, Brown D, Chillingworth T, Connor R, Davies RM, Devlin K, Duthoy S, Feltwell T, Fraser A, Hamlin N, Holroyd S, Hornsby T, Jagels K, Lacroix C, Maclean J, Moule S, Murphy L, Oliver K, Quail MA, Rajandream MA, Rutherford KM, Rutter S, Seeger K, Simon S, Simmonds M, Skelton J, Squares R, Squares S, Stevens K, Taylor K, Whitehead S, Woodward JR, Barrell BG (2001) Massive gene decay in the leprosy bacillus. Nature 409:1007–1011.
79. Ahidjo BA, Kuhnert D, McKenzie JL, Machowski EE, Gordhan BG, Arcus V, Abrahams GL, Mizrahi V (2011) VapC toxins from *Mycobacterium tuberculosis* are ribonucleases that differentially inhibit growth and are neutralized by cognate VapB antitoxins. PLoS One 6: e21738.
80. Purwantini E, Mukhopadhyay B (2013) Rv0132c of *Mycobacterium tuberculosis* encodes a coenzyme $F_{420}$-dependent hydroxymycolic acid dehydrogenase. PLoS One 8:e81985.
81. Zhang L, Zhong Q, Bao L, Zhang Y, Gao L, Huang B, Zhang HD (2009) Rv0901 from *Mycobacterium tuberculosis*, a possible novel virulent gene proved through the recombinant *Mycobacterium smegmatis*. Jpn J Infect Dis 62:26–31.
82. Asencion Diez MD, Demonte AM, Syson K, Arias DG, Gorelik A, Guerrero SA, Bornemann S, Iglesias AA (2014) Allosteric regulation of the partitioning of glucose-1-phosphate between glycogen and trehalose biosynthesis in *Mycobacterium tuberculosis*. Biochim Biophys Acta [VOL:PAGE #S].
83. Yang S, Zhang F, Kang J, Zhang W, Deng G, Xin Y, Ma Y (2014) *Mycobacterium tuberculosis* Rv1096 protein: gene cloning, protein expression, and peptidoglycan deacetylase activity. BMC Microbiol 14:174.
84. Berney M, Weimar MR, Heikal A, Cook GM (2012) Regulation of proline metabolism in mycobacteria and its role in carbon metabolism under hypoxia. Mol Microbiol 84:664–681.
85. Ahangar MS, Vyas R, Nasir N, Biswal BK (2013) Structures of native, substrate-bound and inhibited forms of *Mycobacterium tuberculosis* imidazoleglycerol-phosphate dehydratase. Acta Cryst D69:2461–2467.
86. Anderson DH, Harth G, Horwitz MA, Eisenberg D (2001) An interfacial mechanism and a class of inhibitors inferred from two crystal structures of the *Mycobacterium tuberculosis* 30 kDa major secretory protein (Antigen 85B), a mycolyl transferase. J Mol Biol 307:671–681.
87. Harth G, Horwitz MA (1997) Expression and efficient export of enzymatically active *Mycobacterium tuberculosis* glutamine synthetase in *Mycobacterium smegmatis* and evidence that the information for export is contained within the protein. J Biol Chem 272:22728–22735.
88. Krajewski WW, Jones TA, Mowbray SL (2005) Structure of *Mycobacterium tuberculosis* glutamine synthetase in complex with a transition-state mimic provides functional insights. Proc Natl Acad Sci U S A 102: 10499–10504.
89. Luckner SR, Machutta CA, Tonge PJ, Kisker C (2009) Crystal structures of *Mycobacterium tuberculosis* KasA show mode of action within cell wall biosynthesis and its inhibition by thiolactomycin. Structure 17:1004–1013.
90. Singhal A, Arora G, Sajid A, Maji A, Bhat A, Virmani R, Upadhyay S, Nandicoori VK, Sengupta S, Singh Y (2013) Regulation of homocysteine metabolism by *Mycobacterium tuberculosis* S-adenosylhomocysteine hydrolase. Sci Rep 3:2264.
91. Saroj DC, Singh KH, Anant A, Biswal BK (2014) Overexpression, purification, crystallization and structure determination of AspB, a putative aspartate aminotransferase from *Mycobacterium tuberculosis*. Acta Cryst F70:928–932.
92. Nasir N, Vyas R, Chugh C, Ahangar MS, Biswal BK (2012) Molecular cloning, overexpression, purification, crystallization and preliminary X-ray diffraction studies of histidinol phosphate aminotransferase (HisC2) from *Mycobacterium tuberculosis*. Acta Cryst F68:32–36.
93. Frenois F, Engohang-Ndong J, Locht C, Baulard AR, Villeret V (2004) Structure of EthR in a ligand bound conformation reveals therapeutic perspectives against tuberculosis. Mol Cell 16:301–307.
94. Poulsen C, Panjikar S, Holton SJ, Wilmanns M, Song YH (2014) WXG100 protein superfamily consists of three subfamilies and exhibits an alpha-helical C-terminal conserved residue pattern. PLoS One 9:e89313.