Identification and Characterization of Host Shut-Off Proteins of Mycobacteriophages

Inaugural Dissertation

zur

Erlangung des Doktorgrades
Dr. nat. med.
der Medizinischen Fakultät
und
der Mathematisch-Naturwissenschafltichen Fakultät
der Universität zu Köln

vorgelegt von

Dr. med. Jan Rybniker
aus Karlsruhe

Köln, 2011

Berichterstatter: Professor Dr. rer. nat. Karin Schnetz

                            Privatdozent Dr. med. Georg Plum

Tag der letzten mündlichen Prüfung: 17.10.2011

Table of Contents

1. Abbreviations

dsDNA   **doublestranded desoxyribonucleic acid**

gp   **geneproduct**

Fe-S   **iron sulfur**

mRNA   **messenger ribonucleic acid**

ORF   **open reading frame**

PMF   **peptide mass fingerprint**

SdhA   **Serine dehydratase A**

UV-Vis   **ultraviolet-visible spectrophotometry**

2. Summary (English)

Mycobacteriophages are viruses that infect exclusively mycobacteria. In this work I screened mycobacteriophage genomes for host shut-off proteins, proteins that are capable of down-regulating the mycobacterial metabolism early in the infectious cycle. These proteins are of interest in the drug-target discovery process of important pathogens such as *Mycobacterium tuberculosis*. Several host shut-off proteins were identified and characterized with regard to functional and regulatory aspects. Geneproduct 49, the WhiB-like protein of mycobacteriophage TM4 (WhiBTM4) was shown to be growth inhibitory in the mycobacterial host upon induction of expression. Like its homologue in the host (WhiB2), this viral protein is capable of co-ordinating an iron-sulfur (Fe-S) cluster. The UV-visible absorption spectra obtained from freshly purified and reconstituted WhiBTM4 were consistent with the presence of an oxygen sensitive [2Fe-2S] cluster. The quantification of mRNA-levels during phage infection showed that *whiBTM4* is a highly transcribed early phage gene and a dominant negative regulator of WhiB2, an essential mycobacterial protein. Strikingly, both apo-WhiB2 of *M. tuberculosi*s and apo-WhiBTM4 were capable of binding to the conserved promoter region upstream of the *whiB2* gene indicating that WhiB2 regulates its own synthesis which is inhibited in the presence of WhiBTM4. Thus, within this work, substantial evidence could be provided, supporting the hypothesis of viral and bacterial WhiB proteins being important Fe-S containing transcriptional regulators with DNA-binding capability.

In mycobacteriophage L5 three open reading frames within an early operon were identified as toxic to the host *M. smegmatis* when expressed from an inducible expression vector. These ORFs coding for gp77, gp78 and gp79 presumably function as shut-off genes during early stages of phage replication. There is evidence, that the cell division is affected by one of the proteins (gp79). The transcription of the cytotoxic polypeptides is directed by a promoter situated in ORF83 and transcription control is achieved through the phage repressor gp71 which was shown by co-expression of this protein. The findings presented here can provide useful tools for the molecular genetics of mycobacteria. The mycobacteriophage L5 early protein gp77 was further characterized with regard to its possible function within the host. I provide data showing that this purified phage protein of unknown function specifically binds to protein MSMEG_3532 when incubated with protein lysates of *Mycobacterium smegmatis*. This interaction was confirmed by pull-down assays using purified MSMEG_3532 as bait which co-purified with gp77. The amino acid sequence of MSMEG_3532 is nearly identical

to that of threonine dehydratases, serine dehydratases and an L-threo-3-hydroxyaspartate dehydratase. An enzymatic assay confirmed this host protein as a pyridoxal-5'-phosphate-dependent L-serine dehydratase (SdhA) which converts L-serine to pyruvate. This is the first biochemical characterization of a serine dehydratase derived from mycobacteria. Though the addition of purified gp77 to the established *in vitro* assay had no influence on the enzymatic activity of MSMEG_3532, the specific interaction of phage protein and dehydratase *in vivo* may well have a role in altering the amino acid pool or the products of amino acid metabolism in favour of phage maturation.

3. Summary (German)

Mykobakteriophagen sind Viren die ausschließlich Mykobakterien infizieren. In dieser Arbeit wurden Genome von Mykobakteriophagen dahin gehend untersucht, ob sie für *host shut-off* proteine kodieren. Diese Proteine führen direkt nach der Injektion viraler DNS in den Wirt zur Herabregulation des Wirtsstoffwechsels und können für die Identifikation von *drug targets* wichtiger humanpathogener Erreger wie zum Beispiel *Mycobacterium tuberculosis* von Interesse sein. Mehrere *host shut-off* Proteine konnten in dieser Arbeit identifiziert und charakterisiert werden. Das WhiB-ähnliche Genprodukt 49 des Mykobakteriophagen TM4 (WhiBTM4) hatte nach konditioneller Expression wachstumsinhibierende Eigenschaften. Sowohl das aufgereinigte Protein WhiBTM4 wie auch ein homolges und essentielles Protein des Wirts (WhiB2) waren in der Lage einen Eisen-Schwefel-Cluster (Fe-S) zu bilden. Das UV-vis Absorptionsspektrum von frisch aufgereinigtem und anaerob rekonstituiertem WhiBTM4-Protein zeigte einen sauerstofflabilen [2Fe-2S] cluster. Die Quantifizierung der mRNA während der Phageninfektion ergab, dass *whiBTM4* früh im Infektionszyklus transkribiert wird und ein dominant negativer Regulator von WhiB2 ist. Interessanter weise waren sowohl apo-WhiB2 von *M. tuberculosis* wie auch apo-WhiBTM4 in der Lage die hoch konservierte *whib2* Promoter-Region zu binden. Man kann also schlussfolgern, dass das WhiB2-Protein ein selbst-regulierendes Protein ist, das in der Gegenwart von WhiBTM4 herabreguliert wird. Somit konnten in dieser Arbeit wichtige Hinweise identifiziert werden, die zeigen, dass es sich bei viralen und bakteriellen WhiB-Proteinen um Fe-S haltige Transkriptionsfaktoren mit DNS-Bindungseigenschaften handelt.

Drei *open reading frames* eines Operons des Mykobakteriophagen L5 wirkten nach konditioneller Expression ebenfalls wachstumsinhibierend auf den Wirt *Mycobacterium smegmatis*. Diese ORFs (gp77, gp78 und gp79) fungieren wahrscheinlich als *host shut-off*

proteine und es gibt Hinweise darauf, dass gp79 die Zellteilung des Wirts unterbricht. Die Transkription der zytotoxischen Polypeptide wird durch einen neu identifizierten Promoter im Gen 83 angeschaltet. Reguliert wird dieser Promoter durch das Repressor-Protein gp71, was durch Co-Expressions-Versuche gezeigt werden konnte. Diese Daten können für die Entwicklung neuer Werkzeuge der mykobakteriellen Molekulargenetik von Nutzen sein. Das Mykobakteriphagen-Protein gp77 wurde bezüglich seiner möglichen intrazellulären Funktion im Wirt untersucht. So konnte gezeigt werden, dass das aufgereinigte Protein das Wirtsprotein MSMEG_3532 *in vitro* bindet. Diese Interaktion konnte über umfangreiche *pull-down*-Experimente verifiziert werden. Die Aminosäuresequenz des Proteins ist nahezu identisch mit der Sequenz bereits charakterisierter Enzyme wie der Threonin Dehydratase, Serin Dehydratase and der L-Threo-3-Hydroxyaspartat Dehydratase. In einem enzymtischen Assay konnte gezeigt werden, dass es sich um eine Pyridoxal-5'-Phosphat-abhängige L-Serin Dehydratase (SdhA) handelt, die L-Serin zu Pyruvat umwandelt. Es handelt sich somit um die erste biochemische Charakterisierung eines solchen Enzyms bei Mykobakterien. Auch wenn im beschriebenen Ansatz die Interaktion von gp77 mit SdhA *in vitro* nicht zur Inhibition der enzymatischen Aktivität führte, könnte die Interaktion *in vivo* durchaus eine Rolle spielen, zum Beispiel in der Alteration des Aminosäurepools oder des Aminosäurestoffwechsels während der Bakteriophageninfektion.

4. List of publications

I.

Rybniker, J., Nowag, A., van Gumpel, E., Nissen, N., Robinson, N., Plum, G., and Hartmann, P. 2010. Insights into the function of the WhiB-like protein of mycobacteriophage TM4-a transcriptional inhibitor of WhiB2. Mol Microbiol 77: 642-657.

II.

Rybniker, J., Plum, G., Robinson, N., Small, P.L., and Hartmann, P. 2008. Identification of three cytotoxic early proteins of mycobacteriophage L5 leading to growth inhibition in *Mycobacterium smegmatis*. Microbiology 154:2304-14.

III.

Rybniker, J., Krumbach, K., van Gumpel, E., Plum, G., Eggeling, L., Hartmann, P. The cytotoxic early protein 77 of mycobacteriophage L5 interacts with MSMEG_3532, an L-serine dehydratase of *Mycobacterium smegmatis*. Journal of Basic Microbiology.

IV.

Robinson, N., Kolter, T., Wolke, M., **Rybniker, J.**, Hartmann, P. and Plum, G. 2008. Mycobacterial phenolic glycolipid inhibits phagosome maturation and subverts the pro-inflammatory cytokine response. Traffic 9:1936-47

5. Awards

I.

First winner "Young Investigator Award 2010", Jahreskongress der Deutschen Gesellschaft für Innere Medizin, Wiesbaden 2010

II.

Second winner "Young Investigator Award 2008", Jahreskongress der Deutschen Gesellschaft für Innere Medizin, Wiesbaden 2008

6. Contributions to the present investigation (chapter 8) and publications

Contribution to 8.1. and publication I

In this project I designed the concept and established the experimental procedures such as qRT-PCR, protein purification and DNA-binding assays. After having established the purification of WhiBTM4, I instructed Nicole Nissen on the purification of WhiB2 which was part of her diploma-thesis. During her diploma-thesis Nicole Nissen was also instructed by Edeltraud van Gumpel and Angela Nowag. I finally calculated the statistics and wrote the manuscript. The manuscript was proofread by PD Dr. Pia Hartmann, PD Dr. Georg Plum and Nirmal Robinson and finally approved by all Co-authors.

Contribution to 8.2. and publication II

Within this project I designed and performed all cloning steps leading to the construction of several plasmids used in this study. Furthermore I performed the β-galactosidase assays and primer extension analysis and was thereby instructed by Prof. Karin Schnetz. I wrote the manuscript which was proofread by PD Dr. Pia Hartmann, PD Dr. Georg Plum, Nirmal Robinson and Prof. P. Small.

Contribution to 8.3. and publication III

In this study I performed the cloning steps and protein purification steps as well as the protein pull-down assays. Enzymatic assays were set up with the help of Prof. Lothar Eggeling and Karin Krumbach (Instituf für Biotechnologie 1, Forschungszentrum Jülich). I wrote the manuscript which was proofread by PD Dr. Pia Hartmann and PD Dr. Georg Plum.

7. Introduction

The systemic isolation and characterization of mycobacteriophages, viruses that infect mycobacteria, has led to the development of many genetic tools for the dissection and understanding of their mycobacterial hosts (Bardarov *et al.*, 1997; Rybniker *et al.*, 2003; van Kessel *et al.*, 2008). To date over 250 mycobacteriophages have been isolated and the genomic sequences of more than 45 phages are available in the NCBI database. Taken together, these relatively large genomes code for more than 4000 putative proteins, many of which are without a match in the database (Hatfull *et al.*, 2006; Pedulla *et al.*, 2003). The functions of the vast majority of these mycobacteriophage open reading frames (ORFs) remain unknown and only a few proteins have been expressed and examined in detail. Even more striking is the genomic diversity of mycobacteriophages and exchange of genes among the phages themselves and among their hosts has led to a vivid and probably highly influential evolutionary relationship between virus and bacterial host.

In this work I focussed on host shut-off proteins of mycobacteriophages. Viral host shut-off is performed by regulatory proteins expressed early in the infectious cycle. These proteins lead to a rapid and sustained down-regulation of essential metabolic functions of the host in favour of viral replication. The host shut-off phenomenon is well known for many lytic bacteriophages as well as for viruses of the eukaryotic cell (Fenwick and Clark, 1982; Sharma *et al.*, 2004; Svenson and Karlstrom, 1976). It has recently been shown that the systemic exploration of staphylococcal bacteriophage derived shut-off proteins and their targets in *Staphylococcus aureus* can be exploited for the identification of new antibiotic chemical compounds (Liu *et al.*, 2004). In this publication the authors describe the systemic sequencing of newly isolated bacteriophages and the conditional expression of single ORFs in *S. aureus*. Bacteriophage proteins that lead to growth inhibition upon induction are further characterized with regard to their biochemical function. In this work, a similar approach was used for the identification of mycobacteriophage derived host shut-off proteins. Early regulatory genes of mycobacteriophage TM4 were amplified and cloned into the conditional mycobacterial expression vector pSD24 which carries the inducible acetamidase promoter (Daugelat *et al.*, 2003). Mycobacteriophage TM4 is a dsDNA-tailed phage which is fully sequenced and has a broad host range among the fast and slow growing mycobacteria including *M. tuberculosis*, *M. bovis* bacille Calmette-Guérin (BCG), *M. ulcerans* and *M. smegmatis* (Ford *et al.*, 1998; Rybniker *et al.*, 2006) (Fig. 1). Plasmids harbouring TM4 genes were transformed in *M. smegmatis* and clones were induced using acetamide.
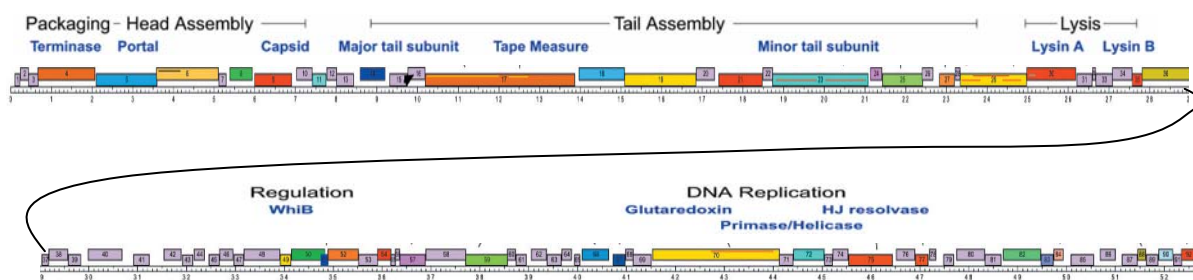
Figure 1. Genome organization of the mycobacteriophage TM4 52797 bp genome. Boxes above the genome represent TM4 ORFs (Ford *et al.*, 1998). Proteins of the early regulatory part (ORF40 to 92) were expressed in *M. smegmatis*.

Using this method, geneproduct 49 of ORF49 was identified as a growth inhibitory protein in *M. smegmatis*. An aminoacid sequence alignment of gp49 with the sequences of proteins currently available in the NCBI database revealed high homology with the WhiB protein family of actinomycetes. These proteins are putative transcription factors which have been identified in all actinomycetes sequenced so far, but not in other organisms. The majority of the WhiB-like proteins contain four perfectly conserved cysteines. WhiD, a protein of *Streptomyces coelicolor* which is required for late stages of sporulation, was the first WhiB protein shown to co-ordinate a Fe-S cluster with the help of these essential cysteines (Jakimowicz *et al.*, 2005). Fe-S clusters represent one of the simplest and most functionally versatile prosthetic groups (Beinert *et al.*, 1997). By undergoing oxidation-reduction reactions they play an important role in metabolic pathways and regulatory processes across all kingdoms of life. In these clusters Fe ions are linked to each other through sulphide bridges on a cysteine rich protein scaffold.

The determination and annotation of the *Mycobacterium tuberculosis* genome sequence revealed the presence of seven *whiB*-like genes (*whiB1-whiB7*) (Cole *et al.*, 1998). The proteins are characterized by the presence of four invariant cysteine residues and a C-terminal helix-turn-helix (HTH) motif with a GV/IWGG amino acid sequence signature in the putative β-turn. Though these motifs are conserved in all seven WhiB proteins, their cellular functions seem to differ substantially and are believed to involve pathogenesis, cell division, stress response as well as antibiotic resistance (Alam *et al.*, 2008; Gomez and Bishai, 2000; Morris *et al.*, 2005; Steyn *et al.*, 2002). Fe-S cluster co-ordinating properties have recently been shown for all *M. tuberculosis* derived WhiB proteins (Alam *et al.*, 2008). Though several reports on mycobacterial WhiB proteins give insight into their possible function as regulatory proteins, only WhiB3 has been examined in detail with regard to regulation and transcription (Singh *et al.*, 2007; Singh *et al.*, 2009). This protein was shown to bind DNA with high

affinity in its clusterless and oxidized apo-form (WhiB3-SS) and with lower affinity in the cluster carrying holo-form (WhiB3-[4Fe-4S]).

The 76 amino acid sequence of gp49 of mycobacteriophage TM4 (WhiBTM4) revealed highest sequence identity in 66 amino acids of the C-terminal region of WhiB2 (Fig. 2).

```
              10          20          30          40
              |           |           |           |
WhiBTM4  - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
WhiB2Mtb MSYEH - - - - - - - - - - - - LRGVMGGTPHTTTGSATASATAVLRPHL  33
WhiB2Ms  MSYESGDFDRVVRFDNRLLGSVSHAPHIDTGSTPTGAAG - -RPQL  43

              50          60          70          80
              |           |           |     ↓     |
WhiBTM4  - - - - - - - - - - - - - - MHMHMGGDPSAICAQTDPELWFPDKGQSTR  30
WhiB2Mtb SLVPEAPAPFEEPLPPEATDQWQDRALCAQTDPEAFFPEKGGSTR  78
WhiB2Ms  SLVPDS - - -FDVAPEAEE-DQWQERALCAQTDPEAFFPEKGGSTR  84

           ↓     ↓ 100       ↓        110         120         130
           |     |           |          |           |           |
WhiBTM4  DAKRMCMRCPLLDECRALALRDPHLV GVWGG LSAQERRRIRKGASA 76
WhiB2Mtb EAKKICMGCEVRHECLEYALAHDERFG IWGG LSERERRRLKRGII -123
WhiB2Ms  EAKRICQGCEVRDACLEYALAHDERFG IWGG LSERERRRLKRGII -129
```
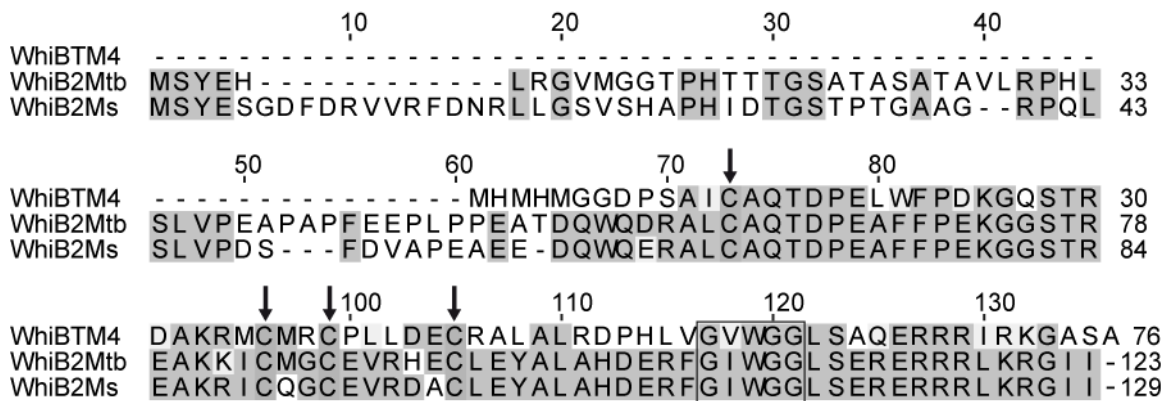
Figure 2. Alignment of WhiBTM4 with *M. smegmatis* (Ms) WhiB2 and *M. tuberculosis* (Mtb) WhiB2 amino acid sequences. The conserved cysteines are highlighted. The conserved β-turn of the C-terminal HTH motif is marked at position 117-121

The N-terminal 56 amino acids of WhiB2, which are of unknown function, are missing in the short bacteriophage protein. In *M. smegmatis* WhiB2 was shown to be an essential gene for cell division and a conditional WhiB2 mutant exhibited irreversible filamentous and branched growth with aberrant septum formation (Gomez and Bishai, 2000). This conditional mutant could be complemented by both WhiB2 of *M. smegmatis* (WhiB2Ms) and WhiB2 of *M. tuberculosis* (WhiB2Mtb), showing that these proteins are functionally equivalent (Raghunand and Bishai, 2006). The role of WhiB2 in mycobacterial cell division as well as its essential nature and uniqueness to mycobacteria makes this protein an interesting drug target.

I was able to show that over-expression of WhiBTM4 leads to a WhiB2 knock-out phenotype in *M. smegmatis* by down-regulating WhiB2 expression. Basal expression from the leaky mycobacterial shuttle vector leads to a phage-resistant phenotype of the host. This phenomenon is known as superinfection exclusion and can be observed in both lytic and lysogenic bacteriophages as well as in many eukaryotic viruses (Lu and Henning, 1994).

The spectroscopic analysis of the purified protein suggests that WhiBTM4 co-ordinates a [2Fe-2S] cluster which is oxygen sensitive and can be restored under anaerobic conditions. Date could be generated, showing that WhiB2 as well as WhiBTM4 are capable of binding regulatory DNA of the host, a key feature in the establishment of WhiB proteins as regulatory proteins. A mutational analysis revealed that both the Fe-S cluster co-ordinating cysteines as

well as amino acid residues of the C-terminal HTH motif are essential for the observed defect in septation.

An additional screen for mycobacteriophage derived host shut-off proteins was performed using Mycobacteriophage L5. This is a temperate phage isolated from *Mycobacterium smegmatis* and it was the first mycobacteriophage to be sequenced (Fig. 3)(Hatfull and Sarkis, 1993).
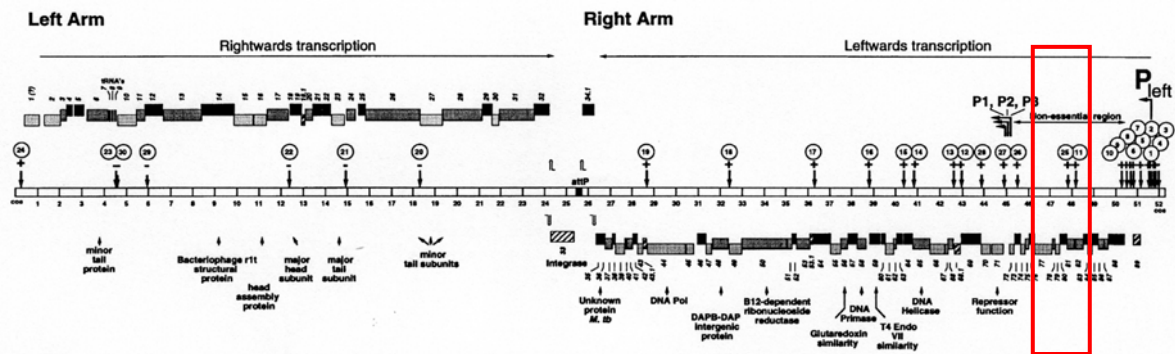

Figure 3. Annotation of the L5 genome, 52297 bp dsDNA, 88 genes. The red box shows the genetic element that can not be transformed into *M. smegmatis*.

L5 forms stable lysogens in *Mycobacterium smegmatis* and has a broad host range among the pathogenic mycobacteria (Rybniker *et al.*, 2006). Though L5 is a temperate phage the following evidence suggests that the phage switches off host protein synthesis during early lytic growth. First, a quantitative decrease of host proteins is seen minutes after induction of a thermo inducible L5 lysogen (Hatfull and Sarkis, 1993) and secondly, it has been reported that a genetic element spanning ORF83 to ORF77 does not transform in *M. smegmatis* (Fig. 3) (Donnelly-Wu *et al.*, 1993). These ORFs are situated on the right arm of the L5 genome which encodes primarily early regulatory proteins. Seven candidate genes were amplified from the L5 genetic region that does not transform in *M. smegmatis* and were inserted in a set of modified acetamidase inducible expression plasmids. Using this method I present data showing that the L5 gene products (gp) 77, 78 and 79 are toxic to its natural host.

The potential presence of host shut-off proteins in the temperate phage L5 implies an extremely tight silencing of the responsible genes by the L5 repressor. The repressor encoded by ORF71 is responsible for the temperate phenotype and the superinfection immunity of lysogens. Transcriptional silencing is achieved by binding of the repressor protein within or downstream of the four early promoter sequences identified so far (Brown *et al.*, 1997; Donnelly-Wu *et al.*, 1993). In this work I provide new insight on the regulation of the toxic

proteins by the description of a new promoter element and the transcription control through the L5 repressor gp71.

In an attempt to identify the functions of L5 toxic proteins I focussed on gp77. Intracellular minute amounts of this protein act bacteriostatically on the natural host *M. smegmatis*. The amino acid sequence of gp77 shows homology with hypothetical proteins of actinomycetes suggesting that this protein was acquired by the phage from a former host. Sequence analysis of gp77 identified homology to the conserved domain DUF2786. This domain which is approximately 40 amino acids in length is found in a family of bacterial proteins with unknown function (Marchler-Bauer *et al.*, 2009).

To identify host proteins that are capable of interacting with gp77 affinity purified gp77 was incubated with soluble proteins of *M. smegmatis*. In this assay gp77 specifically interacted with the host protein MSMEG_3532 *in vitro*. MSMEG_3532 was identified as a pyridoxal-5'-phosphate-dependent L-serine dehydratase (SdhA) by an enzymatic assay. In *Escherichia coli* L-serine dehydratase deficiency has recently been shown to impair growth through the inhibition of cell division (Zhang and Newman, 2008). Though the addition of gp77 to the enzymatic assay did not affect the activity of the amino acid dehydratase *in vitro*, the specific interaction of phage protein and dehydratase may play a role in providing amino acids or products of the amino acid metabolism required for phage maturation.

8. Present investigation – "Identification and Characterization of Host Shut-Off Proteins of Mycobacteriophages"

8.1 Insights into the function of the WhiB-like protein of mycobacteriophage TM4 – a transcriptional inhibitor of WhiB2

Overexpression of WhiBTM4 in *M. smegmatis* is toxic and leads to a WhiB2 knock-out phenotype. Conditional expression of WhiBTM4 in liquid media with the help of an acetamidase inducible promoter system in plasmid pSD24 led to the formation of elongated and highly branched cells (Fig. 4 E, G). Furthermore, these cells had lost their acid-fast properties in the Ziehl-Neelsen stain (Fig. 4B). Fluorescence microscopy with SYTO11, which specifically stains nucleic acids, revealed the presence of multiple nucleoids in the filamented cells (Fig. 4 F, H). Thus, over-expression of WhiBTM4 in liquid medium leads to a phenotype similar to the WhiB2Ms knock-out mutant with hindered septum formation and fragmentation as previously shown by Gomez and Bishai (Gomez and Bishai, 2000).

Induction of WhiBTM4-expression on agar plates containing 0.2% acetamide was not tolerated by *M. smegmatis*.
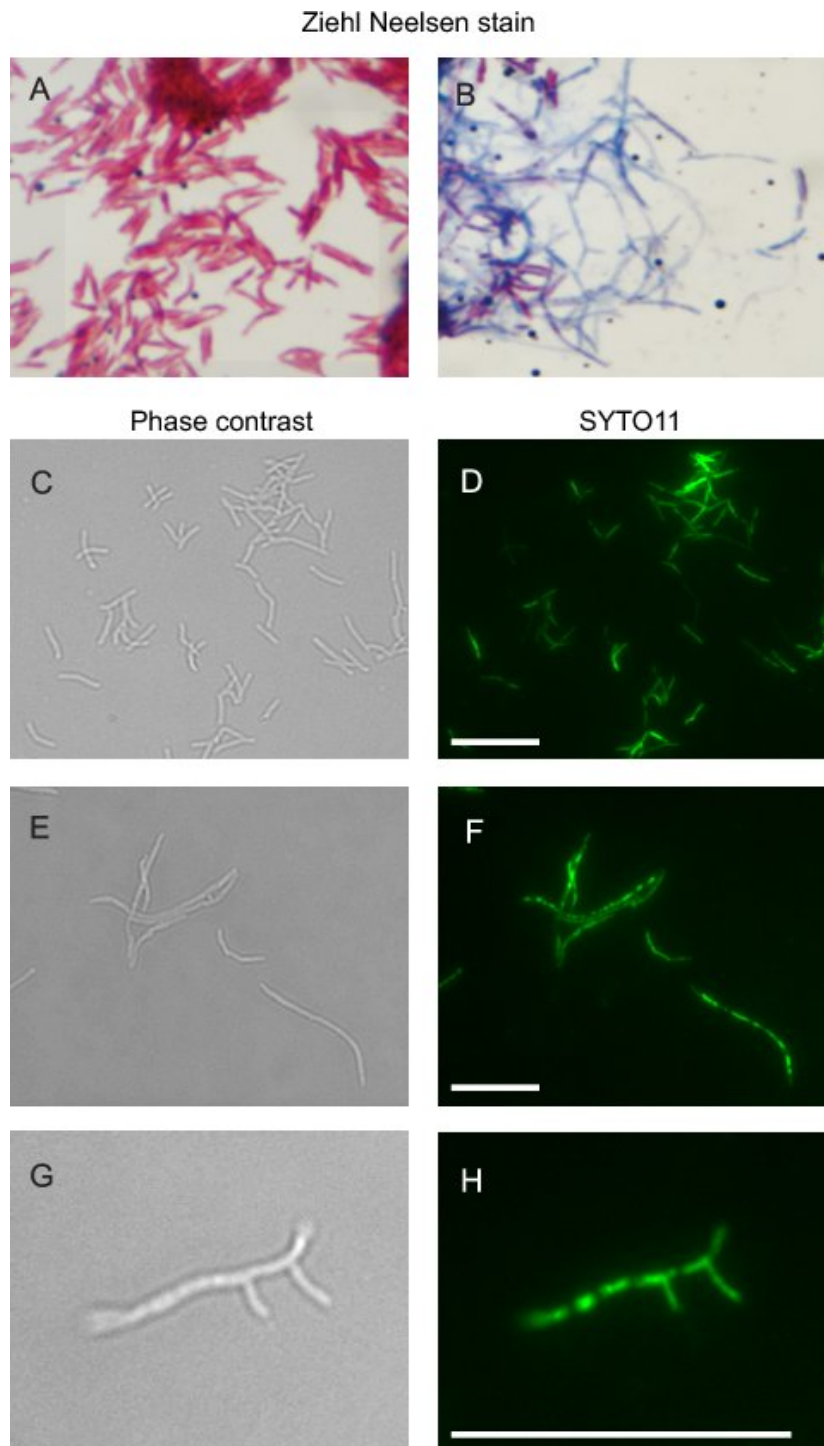
Ziehl Neelsen stain

Figure 4. Microscopic analysis. *M. smegmatis* cells containing pSD24WhiBTM4 were grown in the presence of 0.02 % acetamide (B, E-H) or in the absence of acetamide (A, C, D). Ziehl-Neelson stained cells (A, B) show a loss of acid fast properties of cells expressing WhiBTM4. After induction for 12 h, cells become highly filamentous and branched (E-H). SYTO11-stained induced clones show nucleoids throughout the elongated filament suggesting multiple chromosomes per cell (F, H). Magnification of light and fluorescence microscopy: x 1000. Scale bars: 20 μm.

WhiBTM4 is highly expressed during the early phase of phage replication and leads to a transcriptional down-regulation of WhiB2Ms. Expression of WhiBTM4 leads to morphological changes comparable to a WhiB2 knock-out phenotype. Thus the expression pattern of WhiBTM4 and WhiB2Ms during phage infection was determined. *M. smegmatis* cells were infected with TM4 and mycobacterial RNA was extracted at different time-points. mRNA levels of *whiBTM4* and three host genes (*whiB2Ms, hsp60* and *16S rRNA*) were determined using qRT-PCR. The phage gene was highly transcribed 20 minutes after phage infection and expression levels gradually declined at later time-points (Fig. 5A). This establishes WhiBTM4 as an early phage gene. *whiB2*-mRNA levels decreased in the presence of WhiBTM4 whereas 16S rRNA levels were unchanged until the late phase of infection. HSP60 was slightly down-regulated after 40 minutes of infection but levels increased during the late phase of phage replication. These findings are supported by similar data in *M. smegmatis* cells harbouring pSD24WhiBTM4. Here, induction of WhiBTM4 expression with acetamide for four hours led to a 90 % decrease in *whiB2*-specific RNA (Fig. 5B).
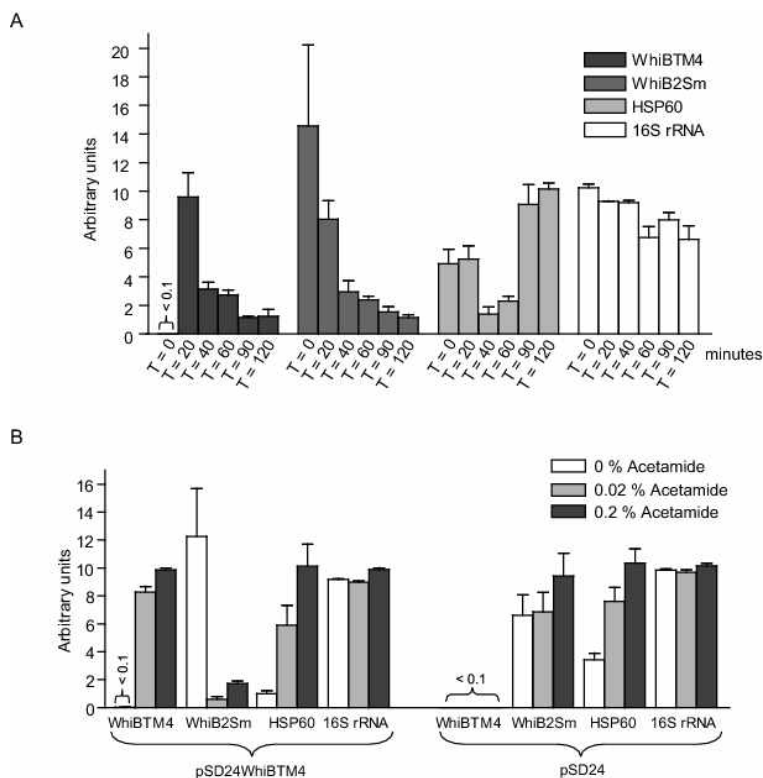

Figure 5. A. Quantitative RT-PCR of *M. smegmatis* infected with mycobacteriophage TM4. Infection of *M. smegmatis* with TM4 leads to detection of *whiBTM4* RNA after 20 minutes. Presence of WhiBTM4 (■) leads to a decline of *whiB2* (■) and *HSP60* (▨) specific RNA. Values for *HSP60* increase during later time-points whereas *whiB2* levels remain low (arbitrary unit).
B. Quantitative RT-PCR of *M. smegmatis* carrying pSD24WhiBTM4. Induction of pSD24WhiBTM4 with different concentrations of acetamide leads to a decrease of *whiB2Ms*

levels which is not seen in cells carrying pSD24 only. (no acetamide: □; 0.2 % acetamide: ■; 0.02 % acetamide: ▪). Error bars indicate the standard error of the mean, n = 3.

Evidence for WhibTM4 co-ordinating an oxygen sensitive [2Fe-2S] cluster. When 6xHis-tagged WhiBTM4 was purified using affinity chromatography, both, the column and the protein eluate had a brown appearance. This suggests the presence of an iron-sulfur chromophore. Native protein samples were examined for their UV-visible absorption spectra which showed peaks at 322 and 423 nm (Fig. 6A). The highly concentrated cluster extracted under denaturing conditions showed two additional shoulders at 462 and 590 nm. This pattern is characteristic for mycobacterial proteins carrying a [2Fe-2S] cluster as well as for [2Fe-2S] ferredoxins from other organisms (Messerschmidt, 2001). The A_{423}/A_{280} ratio of freshly purified WhiBTM4 was 0.25. Iron-sulfur clusters are known to sense the presence of various oxidants and reductants making them intracellular redox sensing molecules. I was interested in similar properties of the mycobacteriophage TM4 derived Fe-S cluster and found that the intensity of the characteristic UV-visible absorption peaks decreased when the protein was exposed to atmospheric oxygen. The peaks as well as the intense brown colour of the protein were lost after three hours, indicating that the cluster is oxygen sensitive (Fig. 6B). This rapid cluster degradation observed in WhiBTM4 stands in contrast to recently published data on WhiB proteins of the host where cluster loss through atmospheric oxygen is achieved within days (Alam *et al.*, 2008). Recently it was shown that fully oxygenated WhiB4 (apo-WhiB4) of *M. tuberculosis* is capable of co-ordinating an [4Fe-4S] cluster under semi-anaerobic conditions in the presence of DTT (Alam *et al.*, 2007). During anaerobic incubation of colourless apo-WhiBTM4 with $FeCl_3$, Na_2S and DTT, a characteristic brown colour developed within three to four hours. The peak pattern of the UV-visible absorption spectra of this reconstituted WhiBTM4 was similar to that of freshly purified protein indicating that apo-WhiBTM4 can be reconstituted to a [2Fe-2S] cluster (Fig. 6C). During anaerobic incubation and dialysis, a substantial amount of protein had aggregated. However, the A_{423}/A_{280} ratio of reconstituted WhiBTM4 was higher than the ratio of the freshly purified cluster and with 0.39 this ratio indicates the presence of an intact [2Fe-2S] cluster (Rouhier *et al.*, 2007).

Similar experiments were performed with purified WhiB2Mtb. Results regarding cluster assembly, decay and reconstitution were comparable to WhiBTM4.
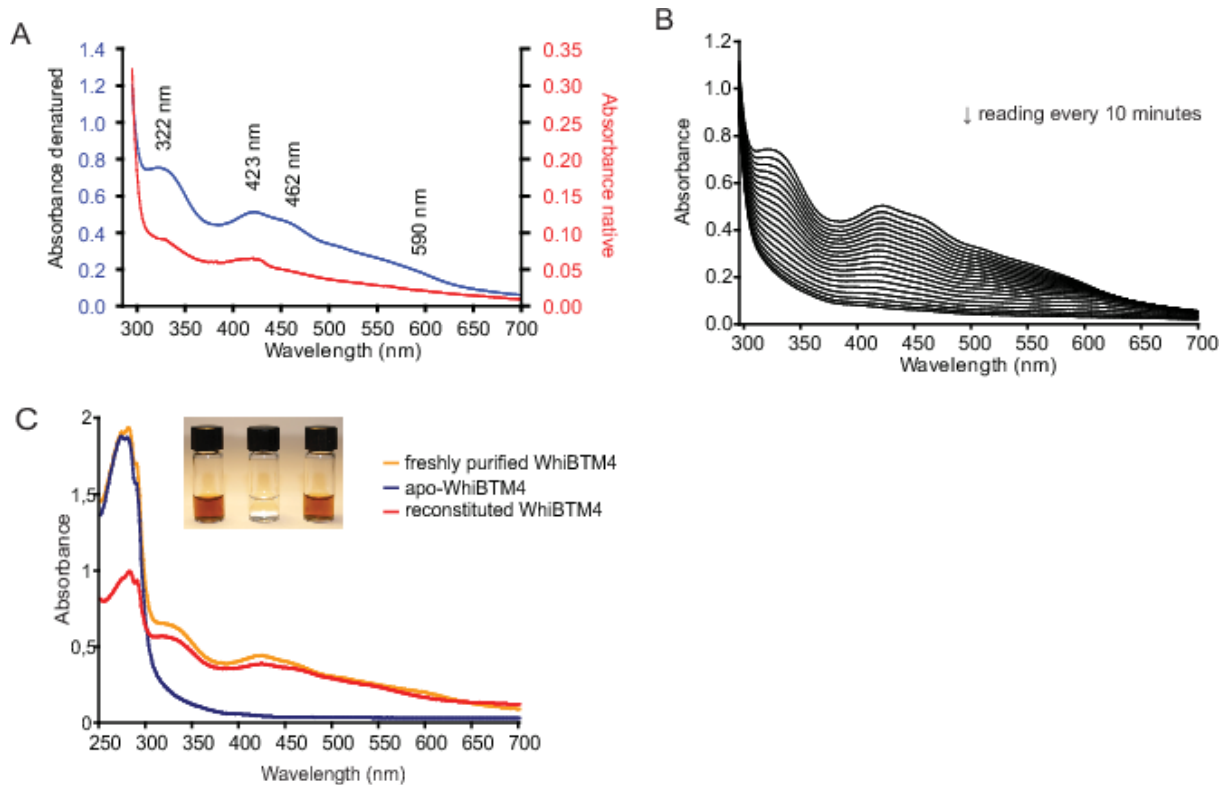
Figure 6. A. UV-visible absorption spectra of WhiBTM4 purified under native conditions (red line) and purified in the presence of 8M urea (blue line). Numbers indicate the peak at the specified wavelength. This peak pattern is characteristic for proteins carrying a [2Fe-2S] cluster.
B. Effect of air exposure on the UV-visible absorption spectrum of WhiBTM4. The sample was kept under atmospheric oxygen at 25°C and measured every 10 minutes. The WhiBTM4 cluster is sensitive to oxygen.
C. UV-visible absorption spectra of freshly purified WhiBTM4 (orange line), fully oxygenated apo-WhiBTM4 (blue line) and reconstituted WhiBTM4 carrying the [2Fe-2S] cluster (red line). The WhiBTM4 [2Fe-2S] cluster can be rebuilt under anaerobic conditions.

WhiB2Mtb and WhiBTM4 specifically bind WhiB2Mtb promoter DNA. Since the presence of WhiBTM4 leads to a down-regulation of WhiB2, I was interested in DNA binding properties of the proteins. In both *M. smegmatis* and *M. tuberculosis* a highly identical promoter sequence can be found 126 bp (*M. smegmatis*) or 226 bp (*M. tuberculosis*) upstream of the respective *whiB2*-gene-start codon (Fig. 7A). Apo-WhiB2Mtb and apo-WhiBTM4 were generated and complete loss of cluster was confirmed by UV-Vis spectroscopy. Proteins were pre-incubated with biotin-labeled WhiB2-promoter DNA and protein-DNA complexes were examined by electrophoretic mobility shift assays (EMSA). Fig. 7B/C shows that apo-WhiB2Mtb as well as apo-WhiBTM4 generated a DNA complex with retarded mobility. No shift was observed when the proteins were incubated with biotin-labeled Epstein-Barr virus promoter DNA or *M. tuberculosis* Ag85A coding DNA showing that binding is sequence specific (Fig. 7 D/E). Increasing the protein concentration and decreasing the concentration of poly-dI-dC DNA led to the formation of a faint shifted band

in these controls, however, in these experiments the *whiB2*-promoter protein/DNA complex was clearly more distinct than the controls (not shown). DNA-binding of both proteins could be competed by the addition of a 200 fold molar excess of unlabeled promoter DNA (competitor 1) (Fig. 7 A-C). This phenomenon was used to identify the protein binding site within the promoter sequence by competing DNA binding with small unlabeled oligo duplexes 15 to 16 bp in size (competitor 2-4, Fig. 7 A). The addition of a 200 fold molar excess of competitor 2, but not competitor 3 or 4 inhibited the formation of a DNA complex with retarded mobility showing that WhiB2Mtb and WhiBTM4 specifically bind to the 5' end of the WhiB2 promoter (Fig. 7 F/G).
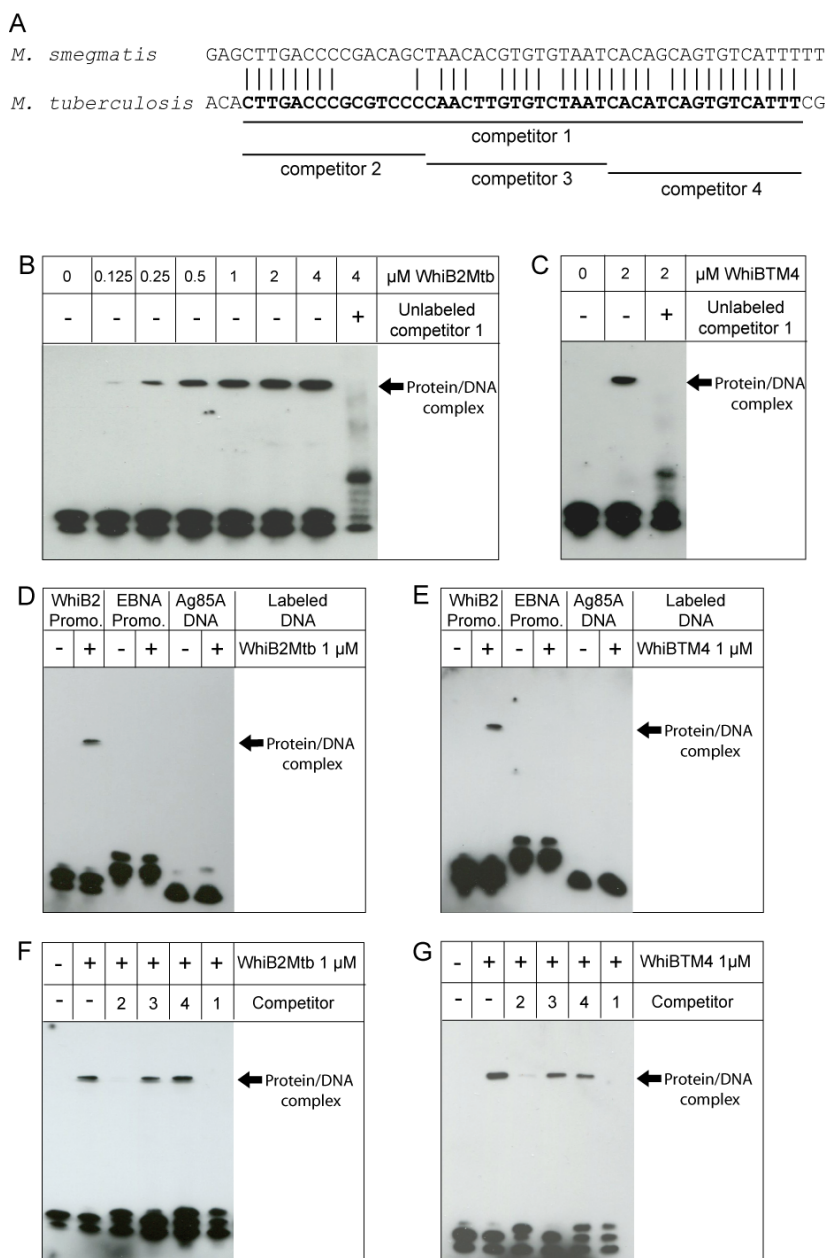

Figure 7. A. Promoter sequence upstream of the *whiBMtb/whiB2Ms* genes. Sequence in bold letters was biotinylated, annealed and used for EMSA experiments, the line segments below show the sequences of unlabeled oligo duplex competitors used in competition experiments.
B. Different concentrations of apo-WhiB2Mtb were incubated with 20 fmol of biotin labeled WhiB2Mtb promoter DNA. The mobility of promoter DNA is retarded in the presence of WhiB2 protein, the underlying Protein/DNA complex can be resolved in the presence of a 200 fold molar excess of unlabeled promoter DNA.
C. A similar shift of labeled DNA can be observed in the presence of apo-WhiBTM4. D/E. WhiB2Mtb and WhiBTM4 were incubated with 20 fmol of biotin labeled WhiB2Mtb promoter DNA, Epstein-Barr nuclear antigen promoter DNA (EBNA) and *M. tuberculosis* Ag85A coding sequence as controls. Only in the presence of WhiB2Mtb promoter DNA a shifted DNA complex can be observed.
F/G. DNA binding of the two proteins is inhibited in the presence a 200 fold molar excess of unlabeled competitor 2 but not through the addition of competitor 3 or 4 showing that the specific binding site is situated within the 5'-end of the promoter sequence. Individual experiments were performed at least three times to rule out pipetting errors.

Concluding remarks

This work identifies and characterizes the bacteriophage derived Fe-S protein WhiBTM4. This is the first description of a viral protein capable of co-ordinating a Fe-S cluster. Sequence comparison of WhiBTM4 with the seven WhiB proteins of mycobacteria revealed high sequence identity in 66 amino acids of the C-terminal region of WhiB2 of *M. tuberculosis* and *M. smegmatis*, a protein that is essential for proper septation and cell division. It was shown here that a phage derived WhiB protein and the essential host protein WhiB2 are capable of binding DNA. Both proteins specifically bound the promoter DNA upstream of the *whiB2* gene, strengthening the hypothesis that WhiB proteins are important transcriptional regulators During phage infection WhiBTM4 could represent a truncated, non-functional homologue of WhiB2 which acts as a transcriptional antagonist displacing WhiB2 from its promoter region and leading to down regulation of the host protein with subsequent alterations of the cell wall composition.

It is an intriguing finding that a viral protein expressed early in the infectious cycle is capable of sensing hypoxia, normoxia and the presence of reducing or oxidizing agents with the help of its Fe-S cluster. It is rather unlikely that WhiBTM4 undergoes the cycle of cluster assembly, loss and re-assembly as part of a redox sensing machinery during the brief presence of the protein in the early stage of phage infection. Redox sensing is not considered to be an essential part of the viral infectious cycle. Since it is the apo-forms of WhiB3, WhiB2 and WhiBTM4 that are capable of binding DNA, the Fe-S cluster in WhiBTM4 could rather be a by-product engendered by the presence of four highly conserved cysteines. Here, the cysteine mediated formation of intramolecular disulfide bonds is the prerequisite for DNA-binding and transcriptional regulation. In this scenario the formation of the observed Fe-S cluster is secondary, which may be mirrored in the relative fast cluster loss that was observed.

The analysis of phage genes that are homologous to, but not necessarily identical to host genes is a pioneering way to dissect the function of poorly understood genes of the host. Furthermore the data generated in this work may be useful in the exploitation of WhiB2 as a drug target in mycobacterial disease.

8.2. Identification of three cytotoxic early proteins of mycobacteriophage L5 leading to growth inhibition in *Mycobacterium smegmatis*

It was shown previously, that the L5 sequence spanning bp 48750 to 46034 which covers ORF83 to ORF77 does not produce transformants when introduced into *M. smegmatis* mc²155 (Donnelly-Wu *et al.*, 1993). I could confirm this finding by cloning these ORFs into pJR5, an acetamide inducible mycobacterial expression vector. This construct, pJR5::8377 (Fig. 8) could not be transformed in *M. smegmatis* mc²155. In a next step ORF83 was deleted from this construct leading to the plasmid pJR5::8277 which was transformable showing relatively small *M. smegmatis* colonies on 7H10 agar after 5 days of incubation at 37 °C. When subcultured on 7H10 agar containing 0.1 % acetamide there was no visible growth suggesting that the phenotype of clones carrying pJR5::8277 is due to one or more inhibitory ORFs within the phage sequence spanning ORF82 to ORF77. One could also assume that in pJR5::8277 these toxic ORFs are decoupled from a promoter sequence upstream of ORF82 (Fig 8). To define the exact ORF responsible for toxicity, six plasmids were designed individually carrying ORF82, 81, 80, 79, 78 or 77 under control of the acetamidase promoter (Fig. 8).
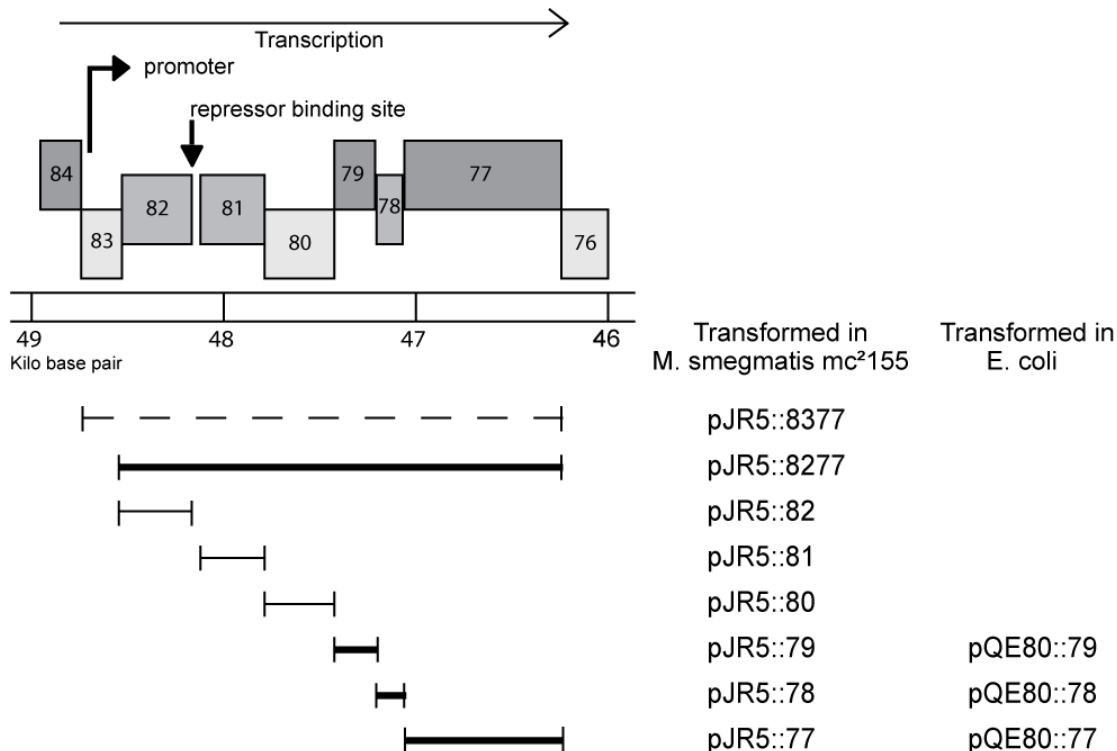

Figure 8. Recombinant plasmids created in this study. The upper part of this Figure illustrates L5 ORFs of the right arm which predominantly carries early genes. Bold lines indicate inserts with growth inhibitory properties when clones are grown in the presence of acetamide. pJR5::8377 does not produce transformants in *M. smegmatis* whereas pJR5::8277 does, leading to the assumption that ORF83 contains a putative promoter sequence.

21

Using these plasmids we were able to show that gp77, gp78 as well as gp79 possess growth inhibitory properties with gp77 being the most potent (Fig. 9 A, B).

Uninduced *M. smegmatis* clones containing pJR5::77 clearly show smaller colonies on 7H10 agar compared to the other plasmid carrying clones (Fig. 9 A). This is most likely due to the expression of minute amounts of gp77 from an incompletely repressed acetamidase promoter. Growing induced *M. smegmatis* clones carrying pJR5::79 to stationary phase in liquid medium leads to the development of non-homogeneous turbidity and the formation of small clumps despite the fact that Tween 80 was added to the medium (Fig. 9 B). When grown on solid medium containing 0.1 % acetamide these clones have a rough surface morphology compared to that of wildtype *M. smegmatis*. Fluorescence microscopy of induced pJR5::79 clones showed filamentation with most of the bacteria being more than three times longer than uninduced clones and *M. smegmatis* expressing LacZ indicating an impaired cell division (Fig. 9 C). Interestingly some pJR5::79 cells showed a branching phenotype (Fig. 9 C).
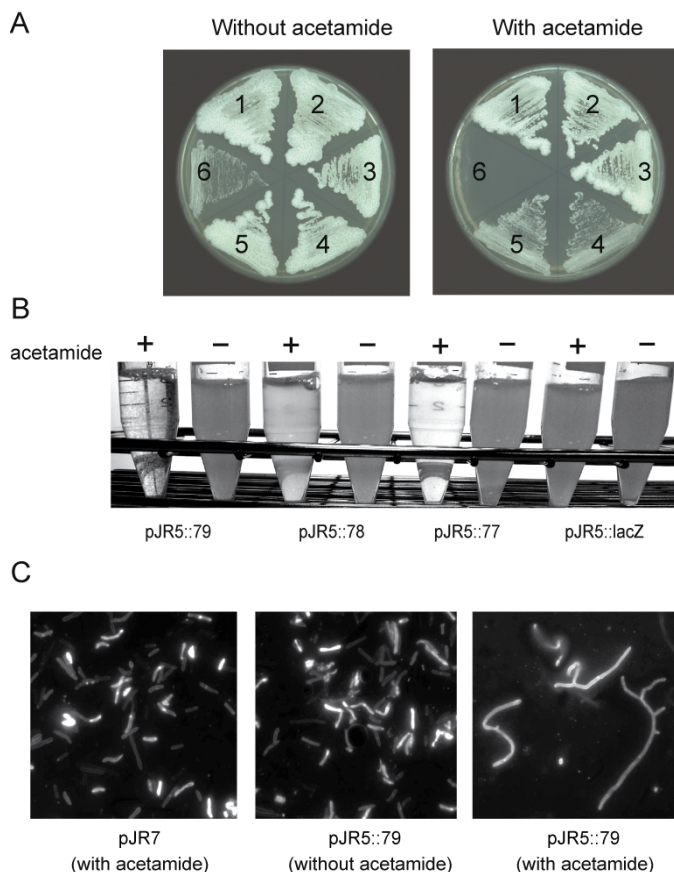

Figure 9. A. Recombinant plasmids carrying ORF82 to ORF77 (1-6) under control of the acetamidase promoter cloned into *M. smegmatis* mc²155 and subcultured with or without the presence of inducing agent. B. *M. smegmatis* clones grown to stationary phase in Middlebrook 7H9 medium plus Tween 80 with or without acetamide (0.05 %). Expression of gp77, gp78 and 79 is growth inhibitory, the presence of gp79 leads to clumping. C. Fluorescence microscopy of FITC stained pJR5::79 clones grown with or without acetamide (0.05 %), a clone expressing LacZ was used as a control. Expression of gp79 leads to elongation of the individual bacteria and a branching phenotype (magnification x 1000).

Regulatory aspects of the cytotoxic phage genes during lysogenic growth.

The potential presence of host shut-off proteins in the temperate phage L5 implies an extremely tight silencing of the responsible genes by the L5 repressor. Furthermore the data gathered so far indicated the presence of a promoter sequence within ORF83. To confirm these assumptions a set of plasmids was constructed allowing tests for the presence of a promoter and the function of the phage repressor gp71 on this promoter. These plasmids are shown in Fig. 10.
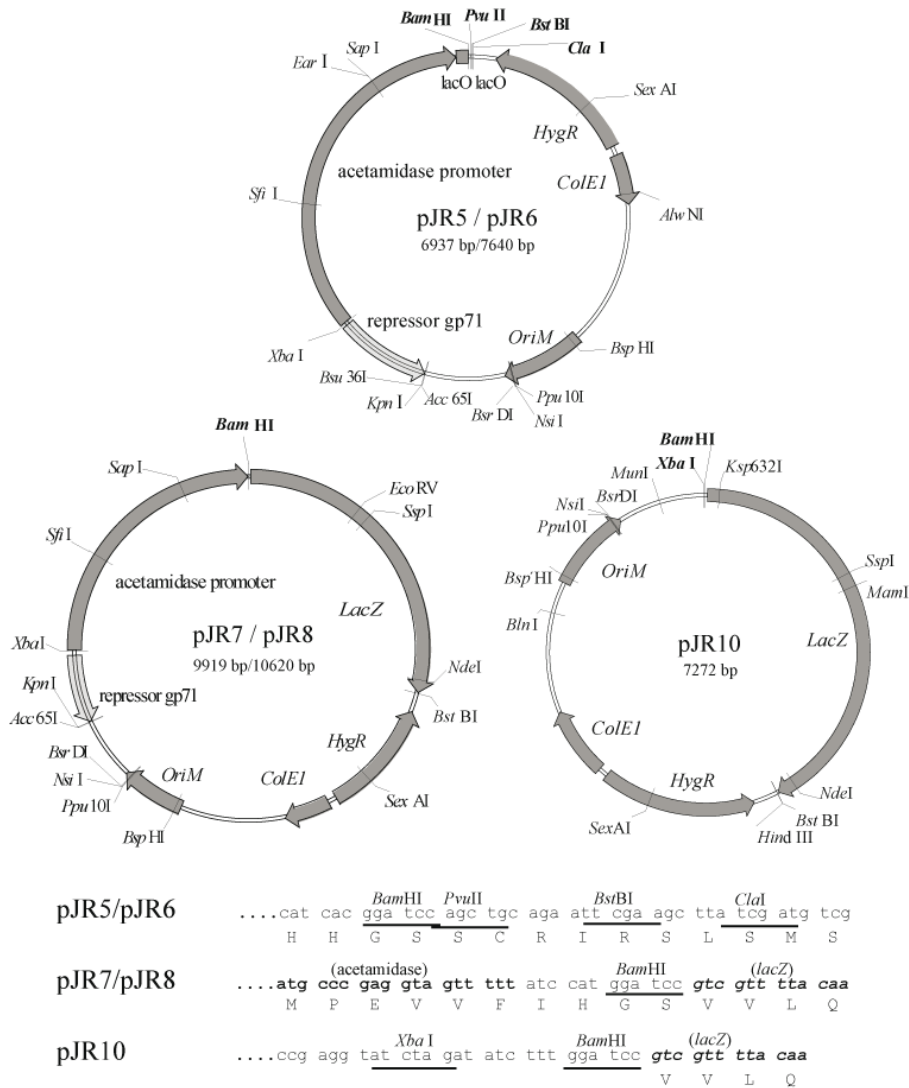

Figure 10. Maps of the modified acetamidase expression plasmid pJR5 and the new promoter probe plasmids pJR7/pJR8 and pJR10. Only pJR6 and pJR8 carry the L5-repressor gene gp71 used for co-expression studies. The lower part shows the multiple cloning sites.

Co-expression of the L5 repressor allows transformation of ORF83 to ORF77.

In contrast to pJR5::8377 which does not transform competent *M. smegmatis* cells, electroporation of pJR6::8377, which constitutively expresses the L5 repressor gp71 resulted

in approximately 2.4 x 10^4 transformants per µg of plasmid DNA. Interestingly the colonies obtained varied largely in colony size and morphology. Subcultures of these clones on media containing acetamide led to growth inhibition showing that the extremely strong acetamidase promoter may overcome transcriptional silencing of the L5 repressor (not shown).

ORF83 contains a putative promoter which is functional in *M. smegmatis* but not in *E. coli*. Transformability of pJR5::8277 but not of pJR5::8377 indicates the presence of a positive regulatory element within ORF83 which possibly induces the transcription of the three cytotoxic genes. To confirm this hypothesis two new mycobacterial promoter probe vectors were created using pSD24 as a backbone. The insertion of *lacZ* from plasmid pMC1871 downstream of the acetamidase promoter led to pJR7. This construct allows the insertion of candidate promoter sequences between the acetamidase promoter and the *lacZ* gene using the single *Bam*HI site (Fig. 10). pJR7 grown without acetamide shows a relatively low but measurable β-galactosidase activity whereas addition of acetamide to the medium results in an extremely high activity (Fig. 11). This is consistent with recently published data stating that the acetamidase promoter in pSD24 is about three times stronger than the mycobacterial hsp60 promoter (Daugelat *et al.*, 2003). Induced pJR7 was used as a positive control in subsequent experiments for the evaluation of the promoter activity of L5 genes. The presence of a functional promoter inserted between acetamidase promoter and *lacZ* gene results in an increased β-galactosidase activity compared to uninduced clones carrying pJR7.

To abolish any background activity of the acetamidase promoter the complete element was excised resulting in pJR10 (Fig. 10). This construct had no detectable β-galactosidase activity comparable to wild type *M. smegmatis* mc²155.

The insertion of ORF83 together with a small part of the 5'-end of ORF82 into the linearized promoter probe vectors (leading to a fusion protein of gp82 and LacZ) gave pJR7::83 and pJR10::83 (Fig. 11). *M. smegmatis* transformants harbouring these plasmids showed an increase of β-galactosidase activity compared to clones containing pJR10 and uninduced pJR7 (Fig. 11). These constructs mimic the situation seen in pJR5::8377 with the toxic ORFs being exchanged by the *lacZ* gene. Compared to fully induced pJR7, promoter activity was about three times lower in (uninduced) pJR7::83 and about four times lower in pJR10::83 suggesting that the promoter within ORF83 is of moderate strength (Fig. 11).

Uninduced pJR7 had measurable activity indicating incomplete suppression of the acetamidase promoter in pSD24 and related constructs (Fig. 11). This explains the slower growth of uninduced clones containing pJR5::77.

Interestingly *E. coli* DH5α containing pJR10::83 (which has shuttle vector properties) or pMC1871, an *E. coli* promoter probe vector (Table 1), containing ORF83 did not show any β-galactosidase activity in the ONPG assay (data not shown). This indicates that promoter activity within ORF83 is restricted to mycobacteria.

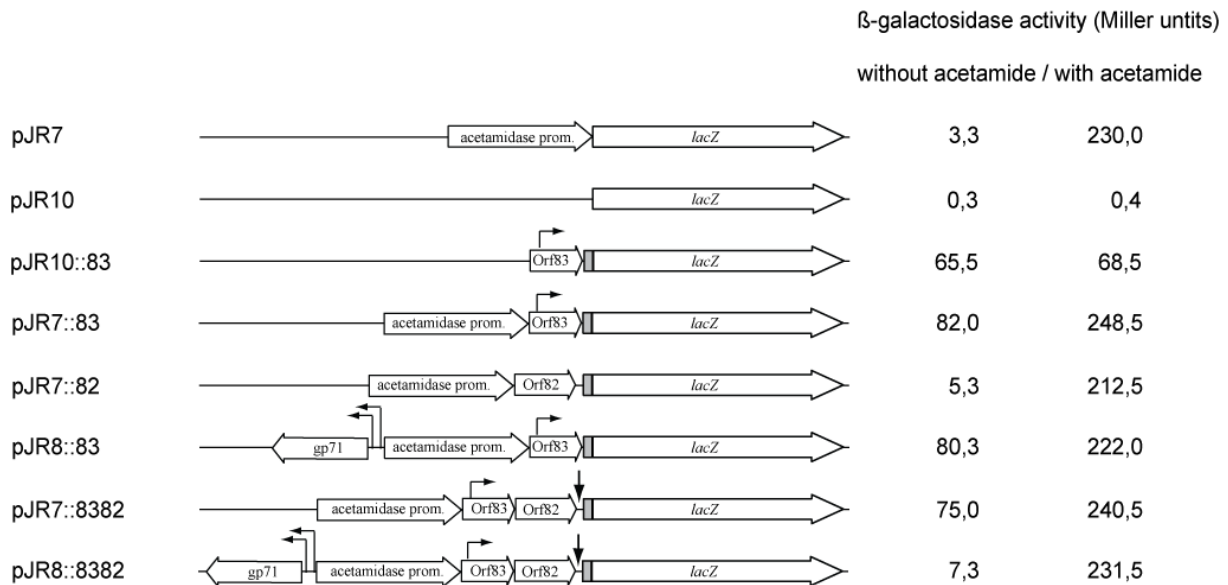|		ß-galactosidase activity (Miller untits)	
		without acetamide	with acetamide
pJR7		3,3	230,0
pJR10		0,3	0,4
pJR10::83		65,5	68,5
pJR7::83		82,0	248,5
pJR7::82		5,3	212,5
pJR8::83		80,3	222,0
pJR7::8382		75,0	240,5
pJR8::8382		7,3	231,5

Figure 11. β-galactosidase assay of different plasmid constructs carrying the *lacZ* gene. The new promoter probe vectors pJR7 and pJR10 grown with and without acetamide (0.1 %). Induced pJR7 functions as a positive control. Uninduced pJR7 shows some activity whereas deletion of the promoter element (pJR10) leads to no detectable β-galactosidase activity indicating leakiness of the acetamidase promoter in pSD24. Constructs carrying ORF83 upstream of the *lacZ* gene show promoter activity within this ORF. Only the presence of both, the repressor gp71 and the repressor binding site (↓) between ORF82 and ORF81 (pJR8::8382) leads to a 10 fold reduction of promoter activity which can be overcome by the strong acetamidase promoter.

Transcriptional regulation of the putative promoter is achieved through repressor binding between ORF82 and ORF81.

Lysogeny of mycobacteriophage L5 is achieved through binding of the repressor gp71 to asymmetric DNA sites with the consensus sequence GGTGG(C/A)TGTCAAG which are situated either within some promoter sites or downstream of promoters (Brown *et al.*, 1997). The binding sites within promoters represent true operators which negatively regulate transcription initiation whereas the downstream ones act as stoperators inhibiting transcription elongation leading to a down-regulation of gene expression (Brown *et al.*, 1997). One such stoperator is situated downstream of ORF83 between the stop codon of ORF82 and the start codon of ORF81 (Fig. 8). To show that gp71 and its binding site have regulatory influence on the promoter element in ORF83 the repressor gene 71 was cloned into the promoter probe plasmid pJR7 leading to pJR8 (Fig. 10). Insertion of ORF83 upstream of the *lacZ* gene of

pJR8 led to similar β-galactosidase activity as seen for pJR7::83 in the ONPG assay whereas integration of ORF83 together with the intergenic region of 82 and 81 (pJR8::8382) led to a 10-fold reduction of β-galactosidase activity (Fig. 11). Addition of acetamide to the medium abolished gp71 activity. A single stoperator site downstream of the extremely strong acetamidase promoter seems to have only little negative regulatory effect on promoter activity and thus may not be detectable in this assay.

Detection of the transcriptional start site within ORF83. Primer extension analysis was performed on *M. smegmatis* RNA containing either pJR10::83 or pJR10 as a negative control. As shown in Fig. 12 a transcriptional start point was detected mapping to position 37 of ORF83. A putative core promoter sequence with -35 (GTGACC) and -10 (TCAAAT) hexamers is situated upstream of this position. The -35 site is homologous with the -35 site of promoter P4 of mycobacteriophage L1, a closely related homoimmune phage of L5 (Chattopadhyay *et al.*, 2003). The experiment was performed with three individual sets of total RNA which showed identical results.
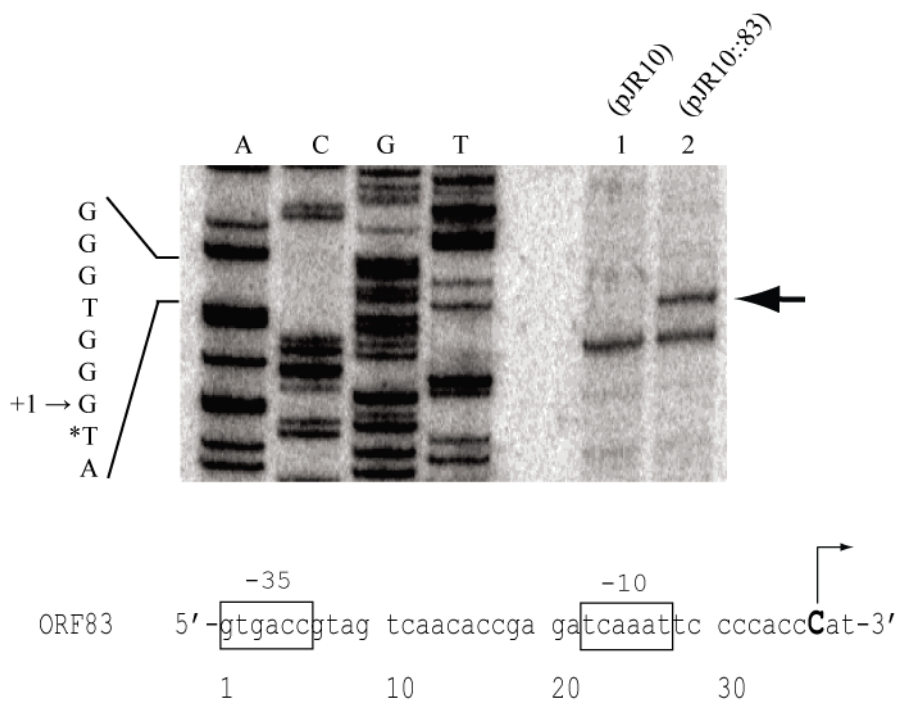

Figure 12. Primer extension analysis. Lanes A, C, G, T, sequencing ladder generated with pJR10::83 plasmid DNA. Lane 1 and 2, primer extension products obtained with RNA isolated from *M. smegmatis* clones harbouring pJR10 (promoterless control) and pJR10::83, respectively. A single band was detected matching position 37 (*) of the sequencing ladder and position 37 of ORF83 (one base upstream due to usage of [α-32P]-dCTP for the sequencing ladder and the end labelled primer for the primer extension experiment). The two lower bands observed in lane 1 and 2 are presumably nonspecific cDNA products. The lower part of the figure shows the DNA sequence of ORF83 with the transcriptional start site at position 37 and putative, hexameric -35 and -10 sites indicated by boxes.

Concluding remarks:

In this study I have shown that the temperate mycobacteriophage L5 produces toxic proteins leading to growth inhibition of the natural host *M. smegmatis*. The ORFs encoding the toxic gene products are located in the right arm of the phage genome which primarily codes for regulatory proteins such as DNA helicase, primase and the repressor (gp71). Tight regulation of these early gene products is a prerequisite to achieve successful L5-lysogeny. The co-expression results shown here are in favour of the explanation that transcriptional silencing of the downstream genes during lysogeny is exerted by the phage repressor gp71. These results emphasize the major role of gp71 for globally silencing prophage gene expression, a crucial mechanism for a temperate virus.

The distinct interaction of the three proteins with the host remains unclear. The predicted 28.2 kDa protein encoded by ORF77, which possesses the strongest activity in *M. smegmatis*, does not contain an excess quantity of positively charged amino acids and we were not able to identify any DNA-binding motifs such as helix-turn-helix structures enabling DNA-binding activity. Membrane association of gp77 is unlikely since there is a lack of hydrophobic regions, signal peptides or predicted transmembrane helices. Using p-BLAST to compare the amino acid sequence of gp77 with the available database revealed sequence homology not only to other mycobacteriophages but also to hypothetical proteins of actinomycetes such as *Frankia* spp., *Streptomyces* spp. and *Mycobacteria* spp.. This indicates that L5 or a L5-related phage may have acquired gp77 from a former host by illegitimate recombination. The toxic protein may have been retained conferring a selective advantage upon the host.

The N-terminal segment of gp79 (amino acid 1 to 41) shares sequence similarity with the signal peptide of the D-alanyl-D-alanine carboxypeptidase of *Bacillus licheniformis*. This enzyme removes C-terminal D-alanyl residues from sugar-peptide cell wall precursors and is also a penicillin-binding protein (PBP). It is synthesized with a hydrophobic signal peptide to target it to the cell wall. In most bacterial species including mycobacteria PBPs are essential for the biogenesis of the peptidoglycan layer of the cell wall. Knockout of PBP-A in *M. smegmatis* results in hindered growth and defective cell septation and division (Dasgupta *et al.*, 2006). This study presents evidence that gp79 interferes with the cell membrane or cell wall synthesis of *M. smegmatis*. First, colonies of clones carrying pJR5::79 grown for several days on 7H10 agar containing 0.1 % acetamide have a rough surface compared to the smooth phenotype of wild type *M. smegmatis* and clones expressing other L5 proteins. Second, the same clones grown in 7H9 broth containing acetamide form relatively large clumps even in the presence of Tween 80, also in contrast to wild type *M. smegmatis*. Third, microscopy of
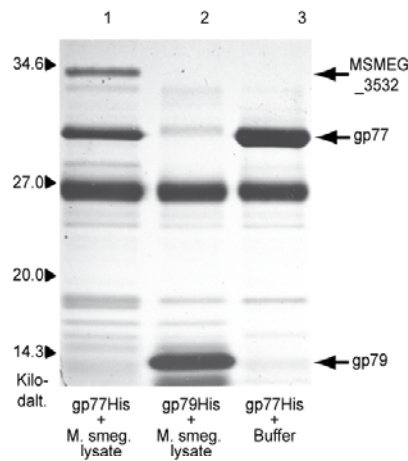
induced clones shows elongated rods compared to the uninduced control and in some bacterial cells a branching phenomenon can be observed indicating that gp79 interferes with the cell septation. The reason for L5 interfering with the cell wall of its host in the early stage of phage propagation is not clear, the genes responsible for lysis of the cell wall after completion of the lytic cycle are situated in the far left arm of the phage genome.

8.3 The cytotoxic early protein 77 of mycobacteriophage L5 interacts with MSMEG_3532, an L-serine dehydratase of *Mycobacterium smegmatis*

gp77 interacts with MSMEG_3532 a host protein of *M. smegmatis*.

To identify host proteins that may interact with gp77 of mycobacteriophage L5 6xHis-tagged gp77 was purified on Ni-NTA magnetic beads. Loaded beads were incubated with soluble protein lysates of *M. smegmatis*, washed extensively and eluted in buffer containing a high concentration of imidazole. Eluates were examined by SDS-PAGE gel electrophoresis and bands present only in the gp77 (bait) plus *M. smegmatis* lysate (prey) lane were excised. In three independent experiments a specific mycobacterial protein co-purified with gp77. This protein was not present in the controls (Fig. 13A). PMF analysis of the excised band identified the protein as MSMEG_3532. To verify the interaction MSMEG_3532 was expressed in *E. coli* BL21(DE3) and *E. coli* lysates were incubated with gp77 coupled on magnetic beads. Again MSMEG_3532 co-purified with the phage protein (Fig. 13B). As a final proof the experiment was also performed with 6xHis-tagged MSMEG_3532 bound to Ni-NTA magnetic beads as bait and untagged gp77 as prey. Here gp77 co-purified with MSMEG_3532 (Fig. 13B).
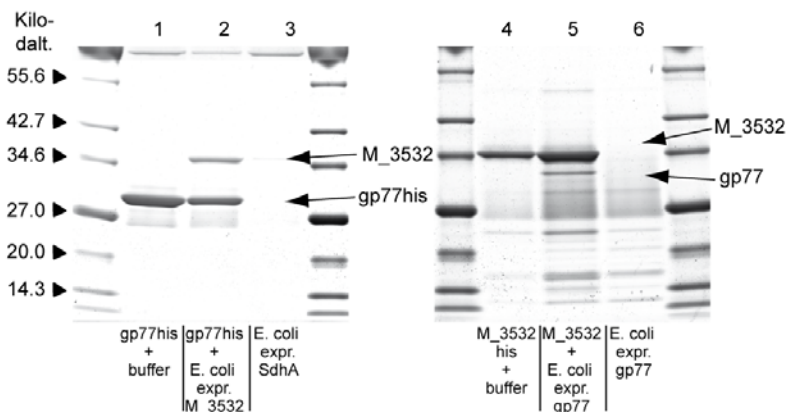
Figure 13. A. SDS-PAGE of eluates from pull down assay of purified gp77 incubated with *M. smegmatis* lysates. gp77 (lane 1) and gp79 (control; lane 2) of mycobacteriophage L5 were bound to Ni-NTA magnetic beads and incubated with *M. smegmatis* lysates for 30 min at 4°C. Beads were washed and bound proteins eluted. Lane 3: gp77 bound to beads and incubated with lysis buffer instead of *M. smegmatis* lysate (second control). MSMEG_3532 co-purifies with gp77. The additional single band at approximately 18 kD in lane 1 is an unspecific binding product derived from *E. coli* (as determined by PMF).
B. SDS-PAGE of pull down assays for confirmation. All proteins were expressed in *E. coli* BL21. gp77his was bound to magnetic beads (lanes 1 and 2) and incubated with lysate of *E. coli* expressing MSMEG_3532 from pET29b (lane 2), in lane 3 this lysate was incubated with beads only (control). MSMEG_3532 expressed in *E. coli* binds to and co-purifies with gp77. MSMEG_3532 was bound to magnetic beads (lanes 4 and 5) and incubated with lysate of *E. coli* expressing gp77 from pET29b (lane 5), in lane 6 this lysate was incubated with beads only (control). gp77 binds to and co-purifies with MSMEG_3532. The differences in size of proteins expressed from pQE80 or pET29b are due to the relatively large N-terminal S-tag/thrombin cleavage site-fusion of pET29b

Identification and characterisation of MSMEG_3532. The protein encoded by MSMEG_3532 exhibits high identities to pyridoxal-5'-phosphate-dependent threonine dehydratases, serine dehydratases and an L-threo-3-hydroxyaspartate dehydratase. These lyases are specific for their substrate as well as to the enantiomer they attack. In order to identify the function of MSMEG_3532, His-tagged protein was isolated and an enzymatic assay performed using a mixture of D-serine, L-serine, D-threonine and L-threonine. After 20 min reaction time, 2-oxoacids were derivatized with 1.2-diamino-4.5-dimethoxybenzol and products analysed by HPLC together with pyruvate and 2-oxobutyrate as standard (Hara, 1985). Only pyruvate was detectable, whereas no 2-oxobutyrate was formed (Fig. 14 A). Subsequent assays analyzing the time-dependent conversion of L-serine but not D-serine to pyruvate clearly identified MSMEG_3532 as an L-serine dehydratase (SdhA) despite its annotation as a threonine dehydratase (Fig. 14 B).
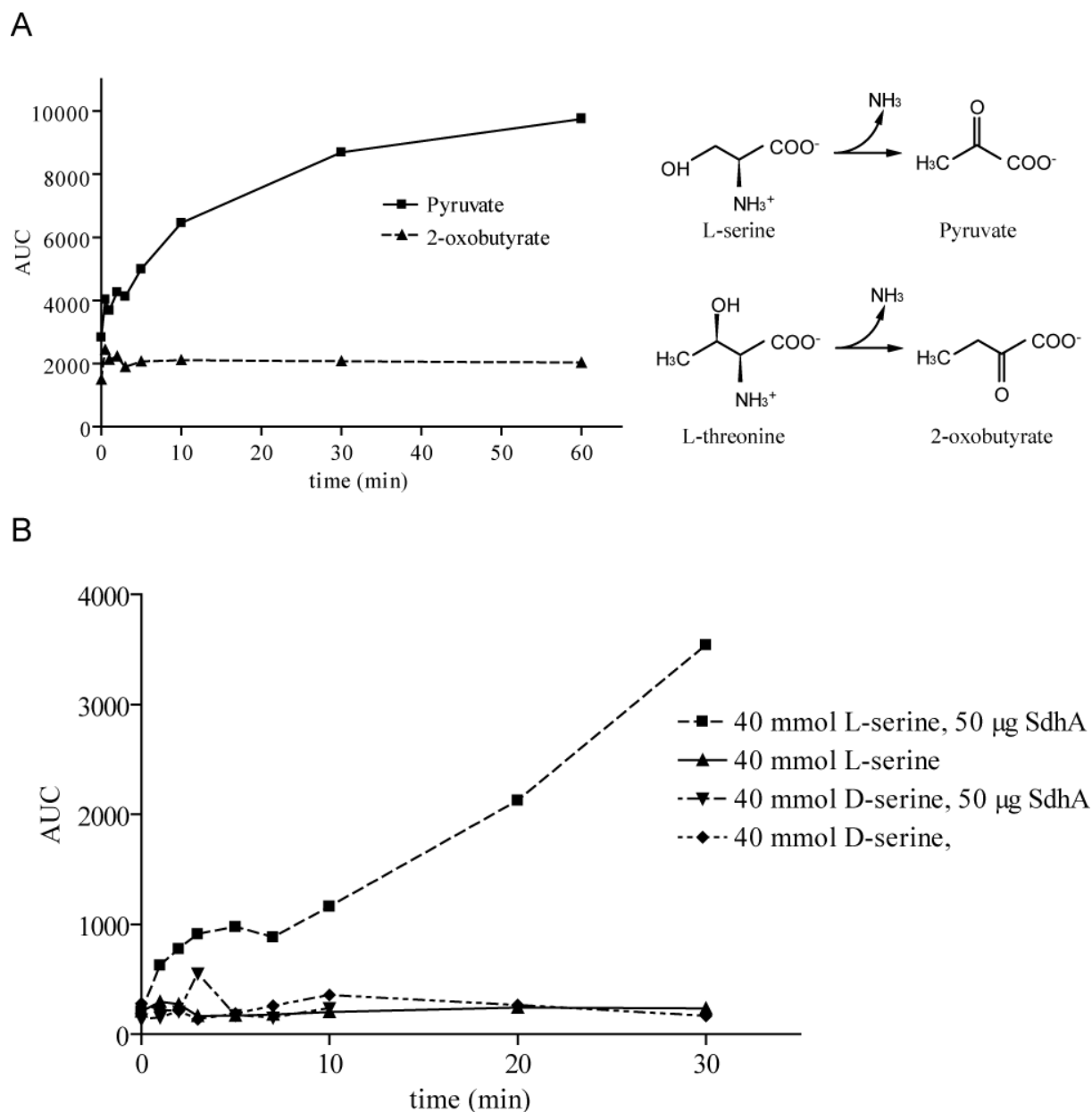
Figure 14. A. Enzyme assay performed using a mixture of D-serine, L-serine, D-threonine and L-threonine added to recombinant SdhA of *M. smegmatis*. In this assay pyruvate but no 2-oxobutyrate could be detected making SdhA a serine-dehydratase. AUC: area under the curve. Amounts of pyruvate and 2-oxobutyrate are shown as area under the concentration curve as determined by the HPLC software.

B. Pyruvate formation with the recombinant serine dehydratase, SdhA, of *M. smegmatis*. Samples from assays containing L-serine plus protein (■), and a control without protein (▲), as well as samples containing D-serine plus protein (▼) and D-serine without protein (♦) were taken at different time intervals and the pyruvate formed was quantified. Only the addition of L-serine to SdhA led to a measurable increase of pyruvate making SdhA an L-serine dehydratase. AUC: area under the curve. Amounts of pyruvate are shown as area under the concentration curve as determined by the HPLC software.

Addition of gp77 to SdhA has no effect on enzymatic activity *in vitro*. In an attempt to assay whether the interaction of gp77 with SdhA has an influence on L-serine dehydratase activity, both proteins were incubated together in 0.1 M potassium phosphate buffer pH 8.2,

containing 0.1 mM pyridoxal-5'-phosphate and 10 % glycerol. Three assays were set up where at a constant concentration of 15.6 µg ml^{-1} gp77, 17.8µg SdhA protein was added, corresponding to a molar ratio of 1:1. The same reaction was performed with a SdhA/gp77 molar ratio of 1:2 or 1:0·2. The assays were incubated for 30 min at RT, and subsequently L-serine dehydratase activity was determined. As a result no difference in enzyme activity was seen between the three assays compared to the SdhA-only control (Fig. 15).
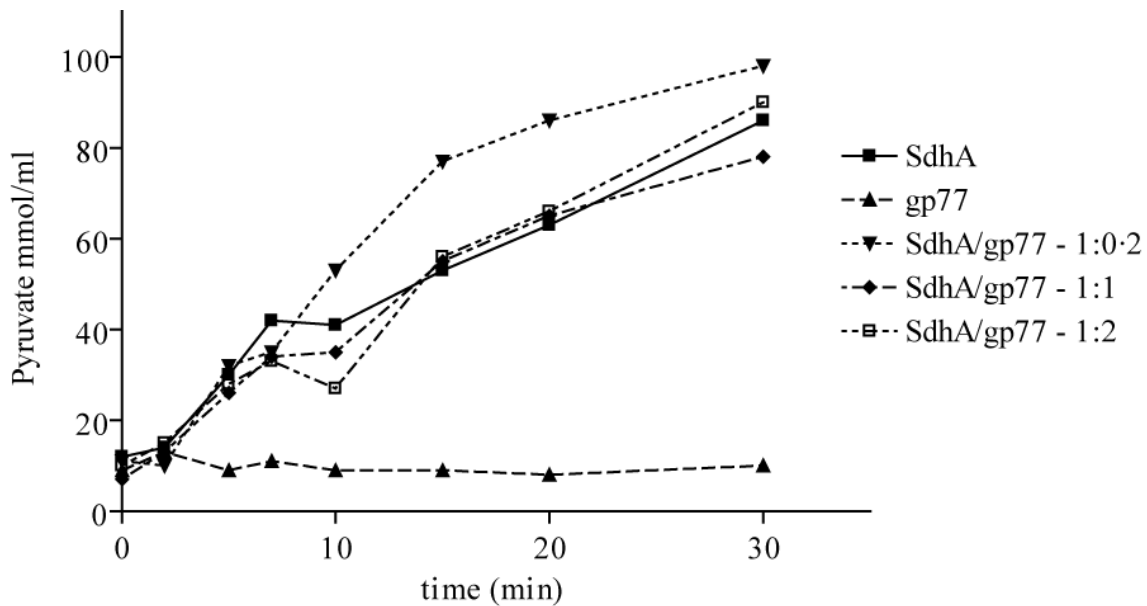

Figure 15. Effect of gp77 on SdhA *in vitro*. Different amounts of gp77 were added to SdhA and pyruvate was quantified after the addition of L-serine. The addition of the phage protein had no statistical significant effect on the function of SdhA.

Concluding remarks

In this study a protein of *M. smegmatis* was identified which interacts with the early cytotoxic protein gp77 of mycobacteriophage L5. This host protein (MSMEG_3532) is a pyridoxal-5'-phosphate-dependent L-serine dehydratase, with the gene accordingly termed *sdhA*. This catabolic enzyme converts L-serine to pyruvate through dehydration and deamination steps. Consumption of L-serine provides pyruvate as a substrate for the tricarboxylic acid cycle and the acetate switch in bacteria (Wolfe, 2005). The deletion of the three *E. coli* genes involved in the L-serine to pyruvate conversion leads to a specific growth defect. It has been assumed that the accumulation of serine starves the cell for S-adenosylmethionine, an important methyl donor in bacterial replication (Zhang and Newman, 2008). In the annotated genome of *M. smegmatis* a peptide transporter is located adjacent to the *sdhA* gene. This supports the hypothesis that SdhA might enable the organism to utilize L-serine as a component of extracellular peptides. Uptake and hydrolysis of the latter would provide convenient access to

L-serine as carbon and energy source. In both *M. smegmatis* as well as in *M. flavescens sdhA* is erroneously annotated as threonine dehydratase, *ilvA*. The serine dehydratase exhibits about 33-36% identities over its entire length to a second threonine dehydratase that can be found in many mycobacterial species. In *M. smegmatis* and *M. flavescens* this *ilvA* gene is encoded by MSMEG_3183 and Mflv_3634, respectively.

Little is known about the amino acid requirements of mycobacteriophages during lytic growth and the present knowledge is mostly based on investigations that were carried out in the 1950s. For bacteriophages infecting gram negative bacteria, it is well known that the quality of the host's amino acid pool has major influence on phage replication and some data indicate that bacteriophages actively modify this pool according to their needs. In *E. coli*, phage development is inhibited in the absence of amino acids such as leucine, valine, isoleucine and threonine (Gots and Hunt, 1953). Addition of pyruvate to the growth medium of lysogenic *E. coli* K12 led to increased survival of bacteria after induction of the phage. Furthermore *E. coli*-cells grown on L-serine rich media accumulate pyruvate after infection with phage T7 indicating that the host's pyruvate metabolism is altered by bacteriophages (Borek *et al.*, 1956). There are also data showing that the activity of catabolic enzymes such as the serine dehydratase is decreased upon bacteriophage infection of *E. coli* (Pardee and Kunkee, 1952).

It is an interesting finding that a mycobacteriophage protein such as gp77 specifically interacts with the catabolic enzyme SdhA. Currently one can not determine how gp77 modifies SdhA during lytic phage replication since the addition of the phage protein had no influence on the conversion of serine to pyruvate in our *in vitro* assay. The reason for this may be found in general limitations of simplified *in vitro* assays where important co-factors or accessory proteins for the alteration of enzymatic processes by secondary binding proteins may be missing. Also binding of gp77 may inhibit the dimerization of SdhA *in vivo* which could alter enzymatic activity − a mechanism that can not be detected *in vitro*. Additionally the role of SdhA in interacting with gp77 could be to provide L-serine or amino groups for further metabolic processes involving gp77 without influencing enzymatic activity of SdhA.

Copies of gp77 are present in some other mycobacteriophages. In addition, the amino acid sequence is homologous to mostly hypothetical proteins of actinomycetes and the short conserved domain DUF2786 of unknown function which can be found in both, gram negative and gram positive bacteria (Marchler-Bauer *et al.*, 2009). A careful sequence analysis showed partial identity to cytosolic catabolic enzymes such as a putative amino acid decarboxylase of *Streptomyces ambofaciens* or enzymes involved in peptidoglycan catabolism such as the N-acetylmuramyl-L-alanine amidase of *Bacillus cereus*. In prokaryotes the interaction of

enzymes involved in the processing of amino acids with other cytosolic proteins is a well known phenomenon (Arifuzzaman *et al.*, 2006; Parrish *et al.*, 2007). Presumably the phage gene 77 was acquired from a former host of L5 by homologous recombination and modified in such a way that binding to SdhA leads to a favourable situation for the phage, for example, by increasing the amino acid pool for the generation of new phage particles or modifying energy recruiting cell cycles in favour of the phage production machinery.

9. References

Alam, M.S., Garg, S.K., and Agrawal, P. (2007) Molecular function of WhiB4/Rv3681c of Mycobacterium tuberculosis H37Rv: a [4Fe-4S] cluster co-ordinating protein disulphide reductase. Mol Microbiol 63: 1414-1431.

Alam, M.S., Garg, S.K., and Agrawal, P. (2008) Studies on structural and functional divergence among seven WhiB proteins of Mycobacterium tuberculosis H37Rv. Febs J.

Arifuzzaman, M., Maeda, M., Itoh, A., Nishikata, K., Takita, C., Saito, R., Ara, T., Nakahigashi, K., Huang, H.C., Hirai, A., Tsuzuki, K., Nakamura, S., Altaf-Ul-Amin, M., Oshima, T., Baba, T., Yamamoto, N., Kawamura, T., Ioka-Nakamichi, T., Kitagawa, M., Tomita, M., Kanaya, S., Wada, C., and Mori, H. (2006) Large-scale identification of protein-protein interaction of Escherichia coli K-12. Genome Res 16: 686-691.

Bardarov, S., Kriakov, J., Carriere, C., Yu, S., Vaamonde, C., McAdam, R.A., Bloom, B.R., Hatfull, G.F., and Jacobs, W.R., Jr. (1997) Conditionally replicating mycobacteriophages: a system for transposon delivery to Mycobacterium tuberculosis. Proc Natl Acad Sci U S A 94: 10961-10966.

Beinert, H., Holm, R.H., and Munck, E. (1997) Iron-sulfur clusters: nature's modular, multipurpose structures. Science 277: 653-659.

Borek, E., Rockenbach, J., and Ryan, A. (1956) The Effect of Lipoic Acid on Recovery in Induced Lysogenic Organisms. Proc Natl Acad Sci U S A 42: 708-710.

Brown, K.L., Sarkis, G.J., Wadsworth, C., and Hatfull, G.F. (1997) Transcriptional silencing by the mycobacteriophage L5 repressor. Embo J 16: 5914-5921.

Chattopadhyay, C., Sau, S., and Mandal, N.C. (2003) Cloning and characterization of the promoters of temperate mycobacteriophage L1. J Biochem Mol Biol 36: 586-592.

Cole, S.T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S.V., Eiglmeier, K., Gas, S., Barry, C.E., 3rd, 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, M.A., Rajandream, M.A., Rogers, J., Rutter, S., Seeger, K., Skelton, J., Squares, R., Squares, S., Sulston, J.E., Taylor, K., Whitehead, S., and Barrell, B.G. (1998) Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. Nature 393: 537-544.

Dasgupta, A., Datta, P., Kundu, M., and Basu, J. (2006) The serine/threonine kinase PknB of Mycobacterium tuberculosis phosphorylates PBPA, a penicillin-binding protein required for cell division. Microbiology 152: 493-504.

Daugelat, S., Kowall, J., Mattow, J., Bumann, D., Winter, R., Hurwitz, R., and Kaufmann, S.H. (2003) The RD1 proteins of Mycobacterium tuberculosis: expression in Mycobacterium smegmatis and biochemical characterization. Microbes Infect 5: 1082-1095.

Donnelly-Wu, M.K., Jacobs, W.R., Jr., and Hatfull, G.F. (1993) Superinfection immunity of mycobacteriophage L5: applications for genetic transformation of mycobacteria. Mol Microbiol 7: 407-417.

Fenwick, M.L., and Clark, J. (1982) Early and delayed shut-off of host protein synthesis in cells infected with herpes simplex virus. J Gen Virol 61 (Pt l): 121-125.

Ford, M.E., Stenstrom, C., Hendrix, R.W., and Hatfull, G.F. (1998) Mycobacteriophage TM4: genome structure and gene expression. Tuber Lung Dis 79: 63-73.

Gomez, J.E., and Bishai, W.R. (2000) whmD is an essential mycobacterial gene required for proper septation and cell division. Proc Natl Acad Sci U S A 97: 8554-8559.

Gots, J.S., and Hunt, G.R., Jr. (1953) Amino acid requirements for the maturation of bacteriophage in lysogenic Escherichia coli. J Bacteriol 66: 353-361.

Hara, S.T., Y., Iwata,T., Yamaguchi,M., Nakamura, M. (1985) Fluorometric determination of alpha-keto acids with 4,5-dimethoxy-1,2-diaminobenzene and its application to high-performance liquid chromatography. Anal Chim Acta 172: 167-173.

Hatfull, G.F., and Sarkis, G.J. (1993) DNA sequence, structure and gene expression of mycobacteriophage L5: a phage system for mycobacterial genetics. Mol Microbiol 7: 395-405.

Hatfull, G.F., Pedulla, M.L., Jacobs-Sera, D., Cichon, P.M., Foley, A., Ford, M.E., Gonda, R.M., Houtz, J.M., Hryckowian, A.J., Kelchner, V.A., Namburi, S., Pajcini, K.V., Popovich, M.G., Schleicher, D.T., Simanek, B.Z., Smith, A.L., Zdanowicz, G.M., Kumar, V., Peebles, C.L., Jacobs, W.R., Jr., Lawrence, J.G., and Hendrix, R.W. (2006) Exploring the mycobacteriophage metaproteome: phage genomics as an educational platform. PLoS Genet 2: e92.

Jakimowicz, P., Cheesman, M.R., Bishai, W.R., Chater, K.F., Thomson, A.J., and Buttner, M.J. (2005) Evidence that the Streptomyces developmental protein WhiD, a member of the WhiB family, binds a [4Fe-4S] cluster. J Biol Chem 280: 8309-8315.

Liu, J., Dehbi, M., Moeck, G., Arhin, F., Bauda, P., Bergeron, D., Callejo, M., Ferretti, V., Ha, N., Kwan, T., McCarty, J., Srikumar, R., Williams, D., Wu, J.J., Gros, P., Pelletier, J., and DuBow, M. (2004) Antimicrobial drug discovery through bacteriophage genomics. Nat Biotechnol 22: 185-191.

Lu, M.J., and Henning, U. (1994) Superinfection exclusion by T-even-type coliphages. Trends Microbiol 2: 137-139.

Marchler-Bauer, A., Anderson, J.B., Chitsaz, F., Derbyshire, M.K., DeWeese-Scott, C., Fong, J.H., Geer, L.Y., Geer, R.C., Gonzales, N.R., Gwadz, M., He, S., Hurwitz, D.I., Jackson, J.D., Ke, Z., Lanczycki, C.J., Liebert, C.A., Liu, C., Lu, F., Lu, S., Marchler, G.H., Mullokandov, M., Song, J.S., Tasneem, A., Thanki, N., Yamashita, R.A., Zhang, D., Zhang, N., and Bryant, S.H. (2009) CDD: specific functional annotation with the Conserved Domain Database. Nucleic Acids Res 37: D205-210.

Messerschmidt, A. (2001) Handbook of metalloproteins. Chichester ; New York: Wiley.

Morris, R.P., Nguyen, L., Gatfield, J., Visconti, K., Nguyen, K., Schnappinger, D., Ehrt, S., Liu, Y., Heifets, L., Pieters, J., Schoolnik, G., and Thompson, C.J. (2005) Ancestral antibiotic resistance in Mycobacterium tuberculosis. Proc Natl Acad Sci U S A 102: 12200-12205.

Pardee, A.B., and Kunkee, R.E. (1952) Enzyme activity and bacteriophage infection. II. Activities before and after virus infection. J Biol Chem 199: 9-24.

Parrish, J.R., Yu, J., Liu, G., Hines, J.A., Chan, J.E., Mangiola, B.A., Zhang, H., Pacifico, S., Fotouhi, F., DiRita, V.J., Ideker, T., Andrews, P., and Finley, R.L., Jr. (2007) A proteome-wide protein interaction map for Campylobacter jejuni. Genome Biol 8: R130.

Pedulla, M.L., Ford, M.E., Houtz, J.M., Karthikeyan, T., Wadsworth, C., Lewis, J.A., Jacobs-Sera, D., Falbo, J., Gross, J., Pannunzio, N.R., Brucker, W., Kumar, V., Kandasamy, J., Keenan, L., Bardarov, S., Kriakov, J., Lawrence, J.G., Jacobs, W.R., Jr., Hendrix, R.W., and Hatfull, G.F. (2003) Origins of highly mosaic mycobacteriophage genomes. Cell 113: 171-182.

Raghunand, T.R., and Bishai, W.R. (2006) Mycobacterium smegmatis whmD and its homologue Mycobacterium tuberculosis whiB2 are functionally equivalent. Microbiology 152: 2735-2747.

Rouhier, N., Unno, H., Bandyopadhyay, S., Masip, L., Kim, S.K., Hirasawa, M., Gualberto, J.M., Lattard, V., Kusunoki, M., Knaff, D.B., Georgiou, G., Hase, T., Johnson, M.K., and Jacquot, J.P. (2007) Functional, structural, and spectroscopic characterization of a glutathione-ligated [2Fe-2S] cluster in poplar glutaredoxin C1. Proc Natl Acad Sci U S A 104: 7379-7384.

Rybniker, J., Wolke, M., Haefs, C., and Plum, G. (2003) Transposition of Tn5367 in Mycobacterium marinum, using a conditionally recombinant mycobacteriophage. J Bacteriol 185: 1745-1748.

Rybniker, J., Kramme, S., and Small, P.L. (2006) Host range of 14 mycobacteriophages in Mycobacterium ulcerans and seven other mycobacteria including Mycobacterium tuberculosis--application for identification and susceptibility testing. J Med Microbiol 55: 37-42.

Sharma, R., Raychaudhuri, S., and Dasgupta, A. (2004) Nuclear entry of poliovirus protease-polymerase precursor 3CD: implications for host cell transcription shut-off. Virology 320: 195-205.

Singh, A., Guidry, L., Narasimhulu, K.V., Mai, D., Trombley, J., Redding, K.E., Giles, G.I., Lancaster, J.R., Jr., and Steyn, A.J. (2007) Mycobacterium tuberculosis WhiB3 responds to O2 and nitric oxide via its [4Fe-4S] cluster and is essential for nutrient starvation survival. Proc Natl Acad Sci U S A 104: 11562-11567.

Singh, A., Crossman, D.K., Mai, D., Guidry, L., Voskuil, M.I., Renfrow, M.B., and Steyn, A.J. (2009) Mycobacterium tuberculosis WhiB3 maintains redox homeostasis by regulating virulence lipid anabolism to modulate macrophage response. PLoS Pathog 5: e1000545.

Steyn, A.J., Collins, D.M., Hondalus, M.K., Jacobs, W.R., Jr., Kawakami, R.P., and Bloom, B.R. (2002) Mycobacterium tuberculosis WhiB3 interacts with RpoV to affect host survival but is dispensable for in vivo growth. Proc Natl Acad Sci U S A 99: 3147-3152.

Svenson, S.B., and Karlstrom, O.H. (1976) Bacteriophage T4-induced shut-off of host-specific translation. J Virol 17: 326-334.

van Kessel, J.C., Marinelli, L.J., and Hatfull, G.F. (2008) Recombineering mycobacteria and their phages. Nat Rev Microbiol 6: 851-857.

Wolfe, A.J. (2005) The acetate switch. Microbiol Mol Biol Rev 69: 12-50.

Zhang, X., and Newman, E. (2008) Deficiency in l-serine deaminase results in abnormal growth and cell division of Escherichia coli K-12. Mol Microbiol 69: 870-881.

10. Acknowledgments

11. Appendix: Publications I-III

Insights into the function of the WhiB-like protein of mycobacteriophage TM4 – a transcriptional inhibitor of WhiB2

Jan Rybniker,[1]* Angela Nowag,[1]
Edeltraud van Gumpel,[1] Nicole Nissen,[1]
Nirmal Robinson,[2] Georg Plum[2] and Pia Hartmann[1]
[1]*1st Department of Internal Medicine, Division of Infectious Diseases, 50924 Cologne, Germany.*
[2]*Institute for Medical Microbiology, Immunology and Hygiene, University of Cologne, 50924 Cologne, Germany.*

Summary

WhiB-like proteins of actinomycetes are known to co-ordinate iron-sulfur (Fe-S) clusters and are believed to have regulatory functions in many essential bacterial processes. The systematic determination of the genome sequences of mycobacteriophages has revealed the presence of several *whiB*-like genes in these viruses. Here we focussed on the WhiB-like protein of mycobacteriophage TM4, WhiBTM4. We provide evidence that this viral protein is capable of co-ordinating a Fe-S cluster. The UV-visible absorption spectra obtained from freshly purified and reconstituted WhiBTM4 were consistent with the presence of an oxygen sensitive [2Fe-2S] cluster. Expression of WhiBTM4 in the mycobacterial host led to hindered septation resembling a WhiB2 knockout phenotype whereas basal expression of WhiBTM4 led to superinfection exclusion. The quantification of mRNA-levels during phage infection showed that *whiBTM4* is a highly transcribed early phage gene and a dominant negative regulator of WhiB2. Strikingly, both apo-WhiB2 of *Mycobacterium tuberculosis* and apo-WhiBTM4 were capable of binding to the conserved promoter region upstream of the *whiB2* gene indicating that WhiB2 regulates its own synthesis which is inhibited in the presence of WhiBTM4. Thus, we provide substantial evidence supporting the hypothesis of viral and bacterial WhiB proteins being important Fe-S containing transcriptional regulators with DNA-binding capability.

Introduction

Fe-S clusters represent one of the simplest and most functionally versatile prosthetic groups (Beinert *et al.*, 1997). By undergoing oxidation-reduction reactions they play an important role in metabolic pathways and regulatory processes across all kingdoms of life. In these clusters Fe ions are linked to each other through sulphide bridges on a cysteine rich protein scaffold. The family of WhiB proteins of actinomycetes are putative transcription factors which have been identified in all actinomycetes sequenced so far, but not in other organisms. The majority of the WhiB-like proteins contain four perfectly conserved cysteines. WhiD, a protein of *Streptomyces coelicolor* which is required for late stages of sporulation, was the first WhiB protein shown to co-ordinate a Fe-S cluster with the help of these essential cysteines (Jakimowicz *et al.*, 2005).

The determination and annotation of the *Mycobacterium tuberculosis* genome sequence revealed the presence of seven *whiB*-like genes (*whiB1*-*whiB7*) (Cole *et al.*, 1998; Soliveri *et al.*, 2000). The proteins are characterized by the presence of four invariant cysteine residues and a C-terminal helix-turn-helix (HTH) motif with a GV/IWGG amino acid sequence signature in the putative β-turn. Though these motifs are conserved in all seven WhiB proteins, their cellular functions seem to differ substantially and are believed to involve, pathogenesis, cell division, stress response as well as antibiotic resistance (Gomez and Bishai, 2000; Steyn *et al.*, 2002; Morris *et al.*, 2005; Alam *et al.*, 2007). Fe-S cluster co-ordinating properties have recently been shown for all *M. tuberculosis* derived WhiB proteins (Alam *et al.*, 2007; 2009; Singh *et al.*, 2007). Though several reports on mycobacterial WhiB proteins give insight into their possible function as regulatory proteins, only WhiB3 has been examined in detail with regard to regulation and transcription (Singh *et al.*, 2009). This protein was shown to bind DNA with high affinity in its clusterless and oxidized apo-form (WhiB3-SS) and with lower affinity in the cluster carrying holo-form (WhiB3-[4Fe-4S]).

The systematic sequencing and comparison of bacteriophage genomes revealed the presence of WhiB-like
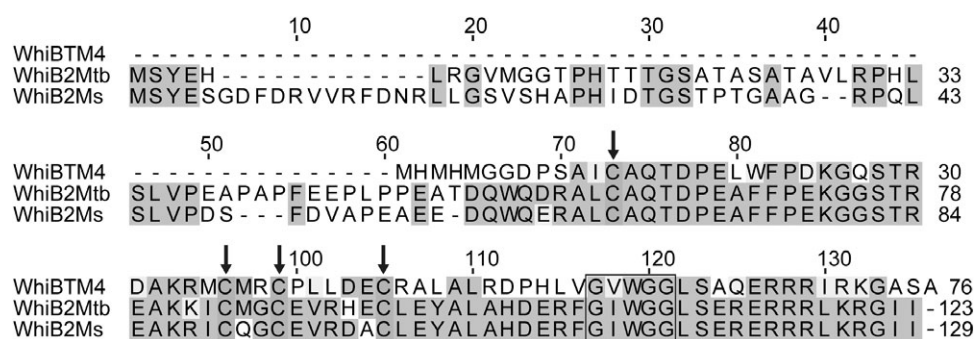
Fig. 1. Alignment of WhiBTM4 with *M. smegmatis* (Ms) WhiB2 and *M. tuberculosis* (Mtb) WhiB2 amino acid sequences. The conserved cysteines are highlighted. The conserved β-turn of the C-terminal HTH motif is marked at position 117–121.

proteins in several mycobacteriophages and streptomyces phages (Pedulla *et al.*, 2003; Van Dessel *et al.*, 2005; Morris *et al.*, 2008). Mycobacteriophages have proven to be useful tools for the exploration of mycobacterial genetics. Furthermore, the recent sequencing of more than 30 mycobacteriophage genomes revealed an extraordinary high genetic diversity among these phages (Pedulla *et al.*, 2003; Hatfull *et al.*, 2006). The variety and high number of phage genes of bacterial origin reflect the frequent recombination events which were probably evolutionarily important for both virus and host. However, the interaction of bacteriophage and mycobacterium during the phage replication cycle is not yet fully understood and only a few phage proteins have been examined in detail.

In this study we have focussed on mycobacteriophage TM4. This dsDNA-tailed phage is fully sequenced and has a broad host range among the fast and slow growing mycobacteria including *M. tuberculosis*, *Mycobacterium bovis* bacille Calmette-Guérin (BCG), *Mycobacterium ulcerans* and *Mycobacterium smegmatis* (Ford *et al.*, 1998; Rybniker *et al.*, 2006). The annotation of the TM4 genome showed a single putative WhiB-like protein situated in a genomic region which encodes primarily for regulation- and DNA metabolism-associated proteins. We termed this protein WhiBTM4. The alignment of the 76 amino acid sequence of WhiBTM4 with the seven mycobacterial WhiB proteins identified so far revealed high sequence identity in 66 amino acids of the C-terminal region of WhiB2 (Fig. 1). The N-terminal 56 amino acids of WhiB2, which are of unknown function, are missing in the short bacteriophage protein. In *M. smegmatis* WhiB2 was shown to be an essential gene for cell division and a conditional WhiB2 mutant exhibited irreversible filamentous and branched growth with aberrant septum formation (Gomez and Bishai, 2000). This conditional mutant could be complemented by both WhiB2 of *M. smegmatis* (WhiB2Ms) and WhiB2 of *M. tuberculosis* (WhiB2Mtb), showing that these proteins are functionally equivalent (WhiB2Ms has also been termed WhmD in some publications, but to maintain consistency in this publication we use the term WhiB2Ms) (Raghunand and Bishai, 2006a). The role of WhiB2 in mycobacterial cell division as well as its essential nature and uniqueness to mycobacteria makes this protein an interesting drug target.

Here we show that overexpression of WhiBTM4 leads to a WhiB2 knockout phenotype in *M. smegmatis* by downregulating WhiB2 expression. Basal expression from the leaky mycobacterial shuttle vector leads to a phage-resistant phenotype of the host. This phenomenon is known as superinfection exclusion and can be observed in both lytic and lysogenic bacteriophages as well as in many eukaryotic viruses (Lu and Henning, 1994; Tscherne *et al.*, 2007).
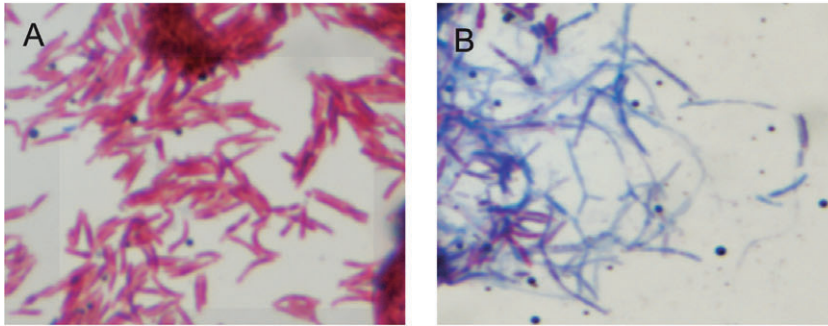
The spectroscopic analysis of the purified protein suggests that WhiBTM4 co-ordinates a [2Fe-2S] cluster which is oxygen sensitive and can be restored under anaerobic conditions. We present data showing that WhiB2 as well as WhiBTM4 are capable of binding regulatory DNA of the host, a key feature in the establishment of WhiB proteins as regulatory proteins. A mutational analysis revealed that both the Fe-S cluster co-ordinating cysteines as well as amino acid residues of the C-terminal HTH motif are essential for the observed defect in septation. We also provide evidence that this C-terminal HTH motif is involved in DNA-binding of WhiBTM4. Furthermore, we show that in contrast to its homologue in the host, WhiBTM4 is a non-essential phage protein under laboratory conditions.
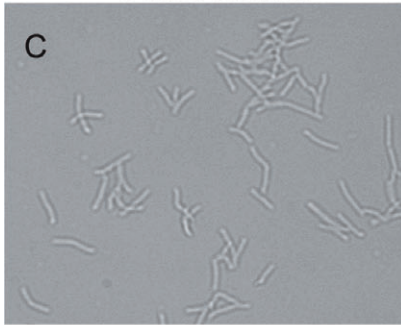
Results

Overexpression of WhiBTM4 in M. smegmatis *leads to a WhiB2 knockout phenotype*

Conditional expression of WhiBTM4 in liquid media with the acetamidase inducible promoter in pSD24 led to the formation of elongated and highly branched cells (Fig. 2E and G). Furthermore, these cells had lost their acid-fast properties in the Ziehl-Neelsen stain (Fig. 2B). Fluorescence microscopy with SYTO11, which specifically stains nucleic acids, revealed the presence of multiple nucleoids
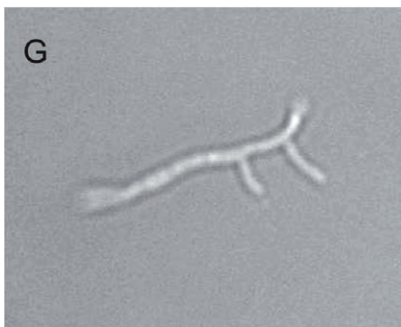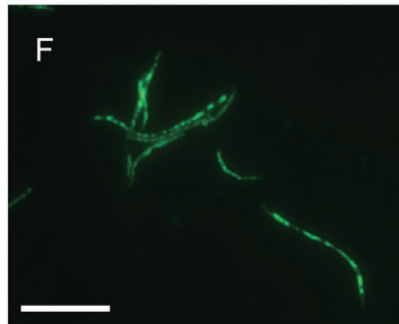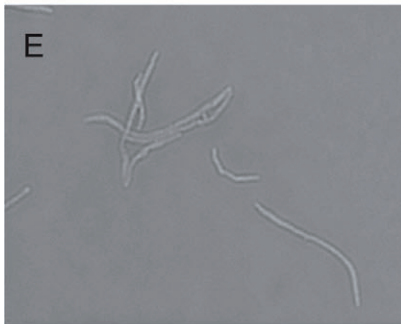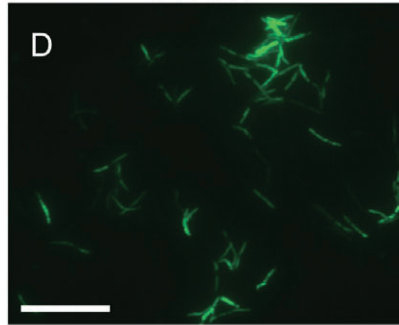
Ziehl Neelsen stain

Fig. 2. Microscopic analysis. *M. smegmatis* cells containing pSD24WhiBTM4 were grown in the presence of 0.02% acetamide (B, E–H) or in the absence of acetamide (A, C, D). Ziehl-Neelson stained cells (A, B) show a loss of acid fast properties of cells expressing WhiBTM4. After induction for 12 h, cells become highly filamentous and branched (E–H). SYTO11-stained induced clones show nucleoids throughout the elongated filament suggesting multiple chromosomes per cell (F, H). Magnification of light and fluorescence microscopy: ×1000. Scale bars: 20 μm.

in the filamented cells (Fig. 2F and H). Thus, overexpression of WhiBTM4 in liquid medium leads to a phenotype similar to the WhiB2Ms knockout mutant with hindered septum formation and fragmentation as previously shown by Gomez and Bishai (2000). Induction of WhiBTM4 expression on agar plates containing 0.2% acetamide was not tolerated by *M. smegmatis* (not shown).

WhiBTM4 is highly expressed during the early phase of phage replication and leads to a transcriptional downregulation of WhiB2Ms

The fact that expression of WhiBTM4 leads to morphological changes comparable to a WhiB2 knockout phenotype directed our interest towards the expression pattern
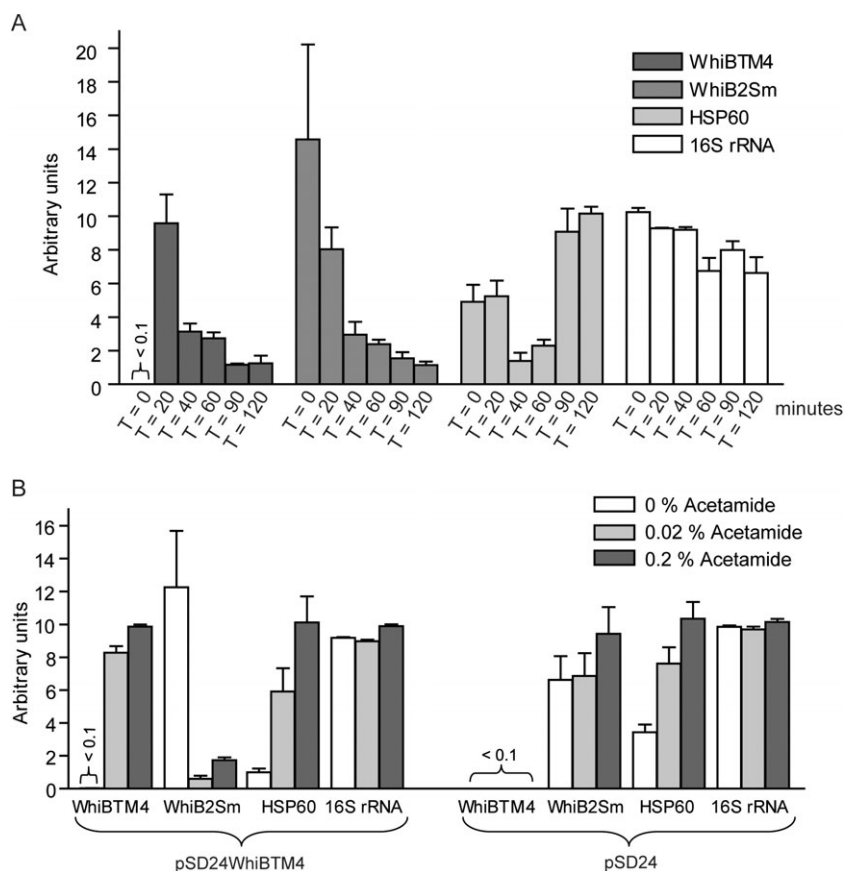
Fig. 3. A. Quantitative RT-PCR of *M. smegmatis* infected with mycobacteriophage TM4. All data were obtained as triplicates. For better visualization and comparability for each gene at the time point of highest expression the middle value of the triplicate was defined as 10 (arbitrary unit). All other values for this particular gene were calculated as the ratio of 10. Infection of *M. smegmatis* with TM4 leads to detection of *whiBTM4* RNA after 20 min. Presence of WhiBTM4 (■) leads to a decline of *whiB2* (■) and *HSP60* (■) specific RNA. Values for *HSP60* increase during later time points whereas *whiB2* levels remain low. B. Quantitative RT-PCR of *M. smegmatis* carrying pSD24WhiBTM4. Induction of pSD24WhiBTM4 with different concentrations of acetamide leads to a decrease of *whiB2Ms* levels which is not seen in cells carrying pSD24 only. Here *HSP60* RNA seems to increase independently of the presence of WhiBTM4 (no acetamide: □; 0.2% acetamide: ■; 0.02% acetamide: ■). Error bars indicate the standard error of the mean, $n = 3$. Three individual cultures were used to extract RNA which was then used for RT-PCR experiments. The actual qRT-PCR values determined in the experiments are given in Table S1.

of WhiBTM4 and WhiB2Ms during phage infection. *Mycobacterium smegmatis* cells were infected with TM4 and mycobacterial RNA was extracted at different time points. mRNA levels of *whiBTM4* and three host genes (*whiB2Ms*, *hsp60* and *16S rRNA*) were determined using qRT-PCR. The phage gene was highly transcribed 20 min after phage infection and expression levels gradually declined at later time points (Fig. 3A). This establishes WhiBTM4 as an early phage gene. *whiB2*-mRNA levels decreased in the presence of WhiBTM4 whereas 16S rRNA levels were unchanged until the late phase of infection. HSP60 was slightly downregulated after 40 min of infection but levels increased during the late phase of phage replication. These findings are supported by similar data in *M. smegmatis* cells harbouring pSD24WhiBTM4. Here, induction of WhiBTM4 expression with acetamide for 4 h led to a 90% decrease in *whiB2*-specific RNA (Fig. 3B).

Purification of WhiBTM4 and WhiB2Mtb

As the results obtained so far suggested that WhiBTM4 is a transcriptional inhibitor of WhiB2, we were interested to study the biophysical properties of the protein. Overexpression of 6x His tagged WhiBTM4 in *Escherichia coli*

BL21 with subsequent loading of cell lysate to a Ni-NTA column led to a deep brown colouring of the column. When purified and eluted under native conditions, total yield of protein was low and most of the brown substance remained on the column, indicating the formation of protein aggregates. However, the soluble fraction had a light brown colour and the eluates could be further examined by SDS-PAGE and UV-visible absorption spectroscopy. Purification and elution in the presence of 8 M urea led to a rapid clearance of the column and elution of a dark brown fluid. On SDS-PAGE the eluates of both natively purified protein and protein isolated in the presence of 8 M urea showed the presence of highly pure WhiBTM4 with prominent bands at approximately 9 kDa (Fig. 4A). Protein samples loaded without reducing agents such as dithiothreitol (DTT) revealed a second band at 18 kDa which disappeared in the presence of 1 mM DTT. All visible bands were examined by peptide mass fingerprint (PMF). The peptide pattern was consistent with his-tagged WhiBTM4. The second larger band at 18 kDa, which resolves in the presence of reducing agents, indicates that WhiBTM4 is capable of forming protein dimers. Furthermore, in the presence of DTT the prominent bands at 9 kDa migrated slower through the gel, most likely through the reduction of intramolecular
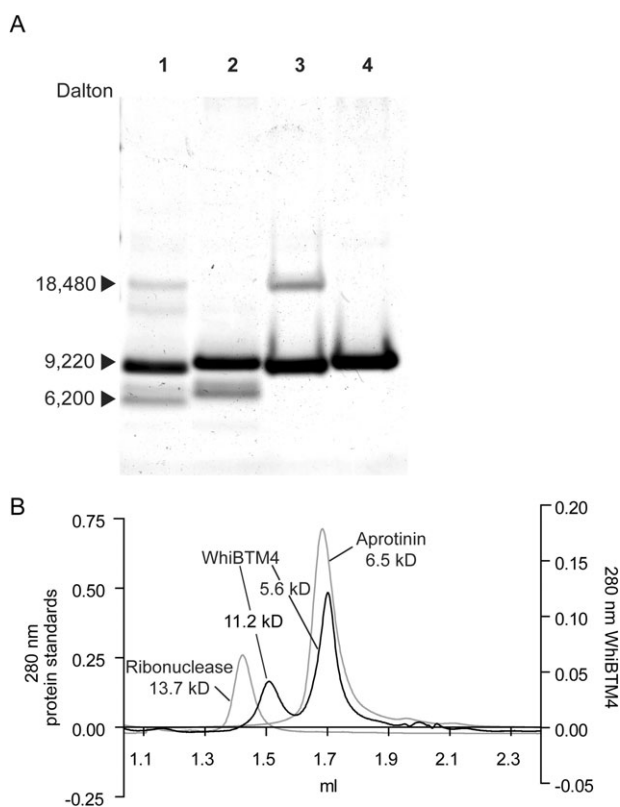
A

B

Fig. 4. A. Coomassie blue stained SDS-PAGE. A total of 5 μg of WhiBTM4 were loaded in each lane. Lane 1 and 2, native WhiBTM4; lane 3 and 4 WhiBTM4 purified under denaturing conditions. Samples of lane 2 and 4 were treated with 1 mM DTT leading to a loss of the second higher band at approximately 18 kDa and slower migration of the protein. The bands at approximately 6 kDa in lane 1 and 2 are most likely truncated protein samples due to proteolytic digest [note the weak additional band between the 9 and 18 kDa bands (lane 1), indicating the presence of a dimer formed by the truncated form of the protein]. Molecular mass is given as determined by MALDI-TOF. All prominent bands were excised and examined with PMF.
B. Gelfiltration of native WhiBTM4. Two peaks can be detected showing that in native buffer the protein is present as a dimer and monomer. The monomer has a calculated mass of 5.6 kD showing that the folded protein is highly compact most likely through the formation of disulfide bonds between the four cysteines.

disulfide bonds formed between the four cysteine residues of the protein leading to a less compact protein formation (Fig. 4A). This phenomenon has also been observed in other WhiB proteins (Raghunand and Bishai, 2006b). Gelfiltration analysis of native WhiBTM4 mirrored

these results; two peaks were visible at 5.6 and 11.2 kD showing that WhiBTM4 is capable of forming a protein dimer (Fig. 4B). The monomer eluted after the 6.5 kD aprotinin standard showing that the protein forms a highly compact tertiary structure under native conditions. Dimerization has also been shown for other WhiB proteins; the role of these dimers is not yet understood (Alam *et al.*, 2007; Crack *et al.*, 2009).

Samples loaded onto SDS gels were also examined by MALDI-TOF for an accurate analysis of protein mass (labels of Fig. 4A). Here the major peak was detected at 9220 Dalton, which corresponds to the calculated protein mass of the 6xHIS-tagged WhiBTM4.

For spectroscopic analysis and DNA-binding experiments we also purified WhiB2Mtb which is a functional homologue of WhiB2Ms. Interestingly, the purification of WhiB2Mtb could readily be performed in native buffer without the formation of protein aggregates.

Evidence for WhibTM4 co-ordinating an oxygen sensitive [2Fe-2S] cluster

The brown appearance of purified WhiBTM4 suggests the presence of an iron-sulfur chromophore. Samples were examined for their UV-visible absorption spectra which showed peaks at 322 and 423 nm (Fig. 5A). The highly concentrated cluster extracted with buffers containing 8 M urea showed two additional shoulders at 462 and 590 nm. This pattern is characteristic for mycobacterial proteins carrying a [2Fe-2S] cluster as well as for [2Fe-2S] ferredoxins from other organisms (Messerschmidt, 2001; Alam *et al.*, 2007). The A_{423}/A_{280} ratio of freshly purified WhiBTM4 was 0.25. With 0.34 atoms of iron per WhiBTM4 monomer the amount of protein containing the cluster was substoichiometric (Table 1). We note that the presence of 8 M urea had no effect on cluster stability indicating that the WhiBTM4 cluster is resistant to chaotropic agents.

Iron-sulfur clusters are known to sense the presence of various oxidants and reductants making them intracellular redox sensing molecules. We were interested in similar properties of the mycobacteriophage TM4 derived Fe-S cluster. We found that the intensity of the characteristic UV-visible absorption peaks decreased when the protein was exposed to atmospheric oxygen. The peaks as well as the intense brown colour of the

Table 1. A_{423}/A_{280} ratio and total iron content of freshly purified and reconstituted WhiBTM4. Shown are the means of three individual experiments and the standard deviation in parenthesis.

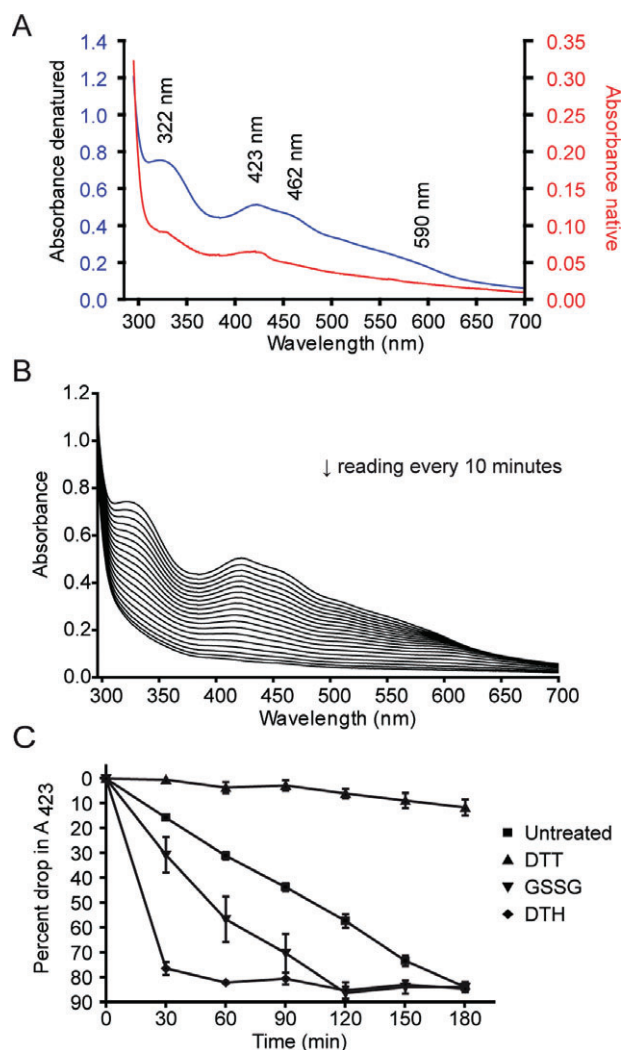| Sample	A_{423}/A_{280}	Atoms of iron per WhiBTM4 monomer	Atoms of iron per WhiBTM4 dimer
Freshly purified WhiBTM4	0.25 (+/–0.05)	0.34 (+/–0.09)	0.69
Reconstituted WhiBTM4	0.39 (+/–0.04)	1.13 (+/–0.12)	2.26

Fig. 5. A. UV-visible absorption spectra of WhiBTM4 purified under native conditions (red line) and purified in the presence of 8 M urea (blue line). Numbers indicate the peak at the specified wavelength. This peak pattern is characteristic for proteins carrying a [2Fe-2S] cluster.
B. Effect of air exposure on the UV-visible absorption spectrum of WhiBTM4. The sample was kept under atmospheric oxygen at 25°C and measured every 10 min. The WhiBTM4 cluster is sensitive to oxygen.
C. Kinetics of [2Fe-2S] cluster loss determined by the decrease of absorbance at A_{423} upon treatment with different oxidizing and reducing agents. Samples were incubated with 10 mM of DTT (dithiothreitol) (▲), GSSG (oxidized glutathione) (▼) or DTH (dithionite) (♦) and the absorbance recorded every 30 min (untreated control: ■). WhiBTM4 senses the presence of reducing and oxidizing agents *in vitro*. Experiments were performed in triplicates.
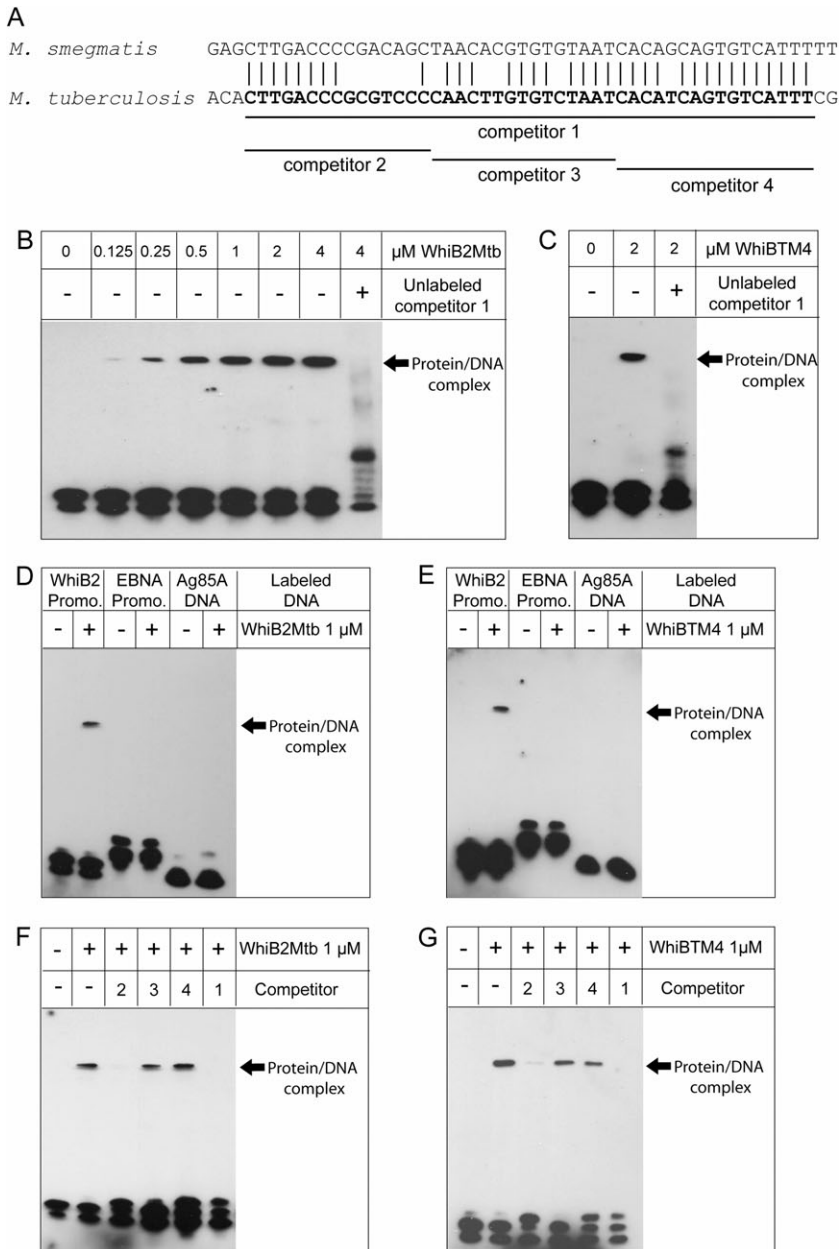
protein were lost after 3 h, indicating that the cluster is oxygen sensitive (Fig. 5B). This rapid cluster degradation observed in WhiBTM4 stands in contrast to recently published data on WhiB proteins of the host where cluster loss through atmospheric oxygen is achieved within days (Alam *et al.*, 2009). In WhiBTM4 cluster dis-

assembly could be arrested in the presence of 10 mM DTT which protects Fe-S clusters as an antioxidant and inhibits the formation of intramolecular disulphide bonds (Fig. 5C). On the contrary, the decay of the WhiBTM4 Fe-S cluster could be enhanced by the addition of 10 mM oxidized glutathione (GSSG) indicated by a faster drop of the peak at A_{423} (Fig. 5C). The characteristic light absorptions of [2Fe-2S] clusters are generated by sulfur-Fe^{3+} charge transfer transitions (Messerschmidt, 2001). Thus, a brown coloured $[2Fe-2S]^{2+}$ cluster can be reduced with the reducing agent dithionite to a $[2Fe-2S]^+$ cluster which results in visible bleaching of protein and a characteristic change in the UV-visible absorption spectrum (rapid loss of the peak at A_{423} and appearance of a peak at A_{322}). This UV-visible change of charge rather than disassembly of cluster could also be observed in WhiBTM4 (Fig. 5C).

Recently it was shown that fully oxygenated WhiB4 (apo-WhiB4) of *M. tuberculosis* is capable of co-ordinating an [4Fe-4S] cluster under semi-anaerobic conditions in the presence of DTT (Alam *et al.*, 2007). During anaerobic incubation of colourless apo-WhiBTM4 with $FeCl_3$, Na_2S and DTT, a characteristic brown colour developed within three to 4 hours. The peak pattern of the UV-visible absorption spectra of this reconstituted WhiBTM4 was similar to that of freshly purified protein indicating that apo-WhiBTM4 can be reconstituted to a [2Fe-2S] cluster (Fig. S1). During anaerobic incubation and dialysis, a substantial amount of protein had aggregated. However, the A_{423}/A_{280} ratio of reconstituted WhiBTM4 was higher than the ratio of the freshly purified cluster and with 0.39 this ratio indicates the presence of an intact [2Fe-2S] cluster (Rouhier *et al.*, 2007). This was mirrored by the high number of iron atoms per monomer/dimer of WhiBTM4 which was 1.13 and 2.26, respectively, as determined by atom absorption spectroscopy (Table 1).

Cluster degradation and anaerobic reconstitution of WhiB2Mtb

In contrast to data on host derived WhiB proteins (Alam *et al.*, 2009), cluster degradation through atmospheric oxygen in WhiBTM4 is a matter of minutes rather than hours. We therefore studied the kinetics of the host derived WhiB2-cluster when exposed to oxygen. WhiB2 of *M. tuberculosis* was isolated under native conditions. In our hands, freshly purified WhiB2Mtb showed a WhiBTM4-like spectroscopic behaviour with a relatively rapid cluster loss under atmospheric oxygen (Fig. S2B). Furthermore, reconstitution of this protein in an anaerobic atmosphere led to the formation of a [2Fe-2S]-like UV-visible absorption spectrum indicating that, like apo-WhiBTM4, apo-WhiB2Mtb can be reconstituted to a [2Fe-2S] cluster protein (Fig. S2C).

A

```
M. smegmatis    GAGCTTGACCCCGACAGCTAACACGTGTGTAATCACAGCAGTGTCATTTTT
                   |||||||        |  |||    ||||  ||||||||| ||||||||||||
M. tuberculosis ACACTTGACCCGCGTCCCCAACTTGTGTCTAATCACATCAGTGTCATTTCG
```

Fig. 6. Electrophoretic mobility shift assay (EMSA).

A. Promoter sequence upstream of the *whiBMtb/whiB2Ms* genes. Sequence in bold letters was biotinylated, annealed and used for EMSA experiments, the line segments below show the sequences of unlabeled oligo duplex competitors used in competition experiments.

B. Different concentrations of apo-WhiB2Mtb were incubated with 20 fmol of biotin labelled WhiB2Mtb promoter DNA. The mobility of promoter DNA is retarded in the presence of WhiB2 protein, the underlying Protein/DNA complex can be resolved in the presence of a 200-fold molar excess of unlabeled promoter DNA.

C. A similar shift of labelled DNA can be observed in the presence of apo-WhiBTM4.

D/E. WhiB2Mtb and WhiBTM4 were incubated with 20 fmol of biotin labelled WhiB2Mtb promoter DNA, Epstein-Barr nuclear antigen promoter DNA (EBNA) and *M. tuberculosis* Ag85A coding sequence as controls. Only in the presence of WhiB2Mtb promoter DNA a shifted DNA complex can be observed.

F/G. DNA binding of the two proteins is inhibited in the presence a 200-fold molar excess of unlabeled competitor 2 but not through the addition of competitor 3 or 4 showing that the specific binding site is situated within the 5′-end of the promoter sequence. Individual experiments were performed at least three times to rule out pipetting errors.

WhiB2Mtb and WhiBTM4 specifically bind WhiB2Mtb promoter DNA

Since the presence of WhiBTM4 leads to a downregulation of WhiB2, we were interested in DNA binding properties of the proteins. In both *M. smegmatis* and *M. tuberculosis* a highly identical promoter sequence can be found 126 bp (*M. smegmatis*) or 226 bp (*M. tuberculosis*) upstream of the respective *whiB2*-gene-start codon (Fig. 6A). Apo-WhiB2Mtb and apo-WhiBTM4 were generated and complete loss of cluster was confirmed by UV-Vis spectroscopy. Proteins were pre-incubated with biotin-labelled WhiB2-promoter DNA and protein-DNA complexes were examined by electrophoretic mobility shift assays (EMSA). Figure 6B and C shows that apo-WhiB2Mtb as well as apo-WhiB2TM4 generated a DNA complex with retarded mobility. No shift was observed when the proteins were incubated with biotin-labelled Epstein-Barr virus promoter DNA or *M. tuberculosis* Ag85A coding DNA showing that binding is sequence specific (Fig. 6D and E). We note that increasing the protein concentration and decreasing the concentration of poly-dI-dC DNA leads to the formation of a faint shifted band in these controls; however, in these experiments the *whiB2*-promoter protein/DNA complex is clearly more distinct than the controls (not shown). DNA-binding of both
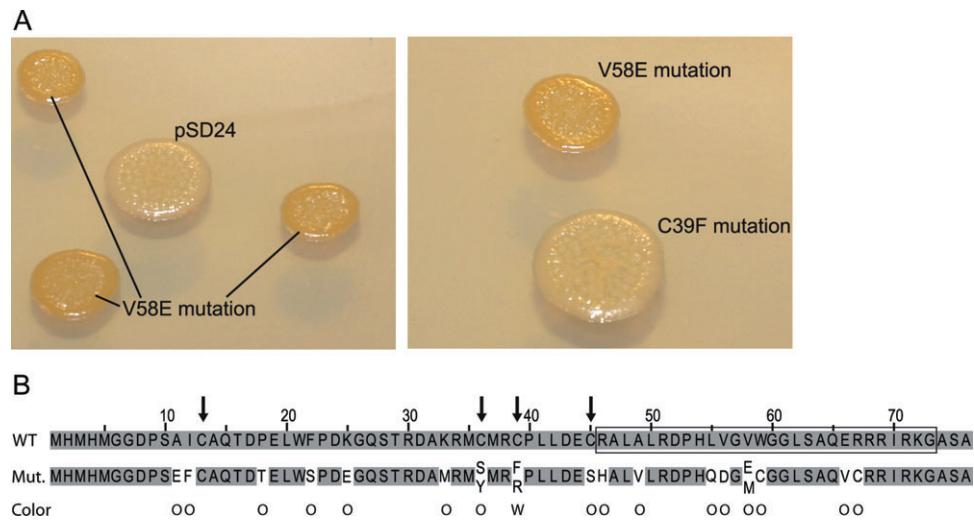
Fig. 7. A. Macroscopic appearance of *M. smegmatis* colonies expressing mutated WhiBTM4 species plated on 7H10 agar plates containing 0.2% acetamide. The red/orange phenotype of the V58E mutation is shown as an example. Cells expressing WhiBTM4 with a C39F mutation have the wild-type appearance of cells carrying pSD24 only.
B. Single amino acid mutations allowing WhiBTM4 expression in *M. smegmatis* and their colour (O = orange, W = white). The C-terminal HTH motif and conserved cysteines are highlighted.

proteins could be competed by the addition of a 200-fold molar excess of unlabeled promoter DNA (competitor 1) (Fig. 6A–C). This phenomenon was used to identify the protein binding site within the promoter sequence by competing DNA binding with small unlabeled oligo duplexes 15–16 bp in size (competitor 2–4, Fig. 6A). The addition of a 200-fold molar excess of competitor 2, but not competitor 3 or 4 inhibited the formation of a DNA complex with retarded mobility showing that WhiB2Mtb and WhiBTM4 specifically bind to the 5′ end of the WhiB2 promoter (Fig. 6F and G).

Both the conserved cysteines and the C-terminal HTH motif are involved in the WhiBTM4 growth inhibitory effect

To show that growth inhibition induced by WhiBTM4 is not due to a nonspecific effect such as the accumulation of large amounts of protein aggregates, we tried to identify amino-acid substitutions allowing normal growth of *M. smegmatis* cells. Spontaneous resistance towards the growth inhibitory effect in *M. smegmatis* clones expressing WhiBTM4 occurred at a rate of 5×10^{-3}. We created a plasmid library expressing mutated WhiBTM4 proteins and screened for clones with uninhibited growth when expressing proteins harbouring sensitive mutations. Cells transformed with the pSD24WhiBTM4*M*-mutant library and plated on 7H10 containing 0.2% acetamide resulted in an acetamide resistant phenotype at a rate of 1:3. Two dominant phenotypes could be observed on these plates: approximately 50% of the colonies were relatively large

and white resembling wild-type *M. smegmatis* colonies and approximately 50% were somewhat smaller colonies with a red to orange colour (Fig. 7A). Colony-PCR of 100 individual clones amplifying the insert of pSD24 and subsequent sequencing of the products showed that approximately 50% of the white colonies had been selected for re-ligated plasmids without insert. Most of the remaining white colonies had mutations resulting in the insertion of stop codons upstream or within the cluster co-ordinating cysteines (not shown). However, two mutants had single amino acid exchanges of the cysteine at position 39 (Fig. 7B). All plasmids of the red colonies had mutated WhiBTM4 proteins and approximately 50% of these showed single amino acid exchanges resulting in cytotoxicity resistance. These mutations clustered in the C-terminal HTH motif of WhiBTM4 (Fig. 7B). However, we note that six of the 17 single amino acid mutations inducing cytotoxicity resistance affected non-cysteine and non-HTH motif amino acids (Fig. 7B). The role of these amino acids in the function of WhiBTM4 and other WhiB proteins is not clear. Among the orange colonies we found a C45S substitution of the most C-terminal cysteine and C36S/Y substitution. No cytotoxicity resistant clone harbouring mutated C13 (the most N-terminal cysteine) could be recovered from the screen. In WhiB2 this cysteine was shown to be non-essential for the complementation of a WhiB2 mutant which might be an explanation for the lack of this mutation in our cytotoxicity resistance screening (Raghunand and Bishai, 2006b).

Mutated Genes containing the C36S, C39F or C45S amino acid substitution were amplified by PCR and

cloned into pQE80 for expression and UV-visible absorption spectroscopy. Only small amounts of protein carrying the C39F mutation could be purified and there was no UV-visible absorption indicating the presence of a Fe-S cluster (Fig. S3A and B). This mutation seems to destabilize the protein to an extent that prevents sufficient expression indicating the vital role of this amino acid and explaining the white colour of *M. smegmatis* clones expressing this mutated gene. Protein destabilization is a common observation in Fe-S proteins carrying sensitive cysteine mutations (Klinge *et al.*, 2007). The C45S and C36S mutations allowed protein expression and purification comparable to wild-type protein and the eluate of these mutated proteins showed a faint brown colour. UV-visible absorption spectra indicated the presence of a Fe-S cluster with a low A_{423}/A_{280} ratio (Fig. S3A). Compared with the wild-type spectrum, the shoulder at 462 nm was more prominent in the C45S mutant and less prominent in the C36S mutant. Disassembly of the mutated clusters under atmospheric oxygen was approximately twice as fast compared to wild-type protein (not shown).

Mutations of the C-terminal HTH motif affect WhiBTM4 DNA-binding

Since amino acid substitutions of WhiBTM4 allowing normal growth of *M. smegmatis* clustered in the C-terminal part of the protein we were interested in DNA-binding properties of these mutated proteins. Six of these proteins were purified and examined in EMSA-experiments using labelled WhiB2 promoter DNA. Four proteins showed low binding affinity compared to the wild-type protein whereas the V56D substitution abolished DNA binding of WhiBTM4 (Fig. 8). WhiBTM4 with a A49V substitution seemed to be a better DNA-binder than the wild-type protein; however, this protein expressed poorly in *E. coli* and low expression may also be the reason for unaffected growth of *M. smegmatis* in the presence of this altered protein. We note that WhiBTM4 proteins with substitutions of the cysteines at position 36 or 45, which were also tested in this experiment, had no effect on DNA binding (Fig. 8).

WhiBTM4 is dispensable for plaque formation of phage TM4

In the host the WhiBTM4 homologue WhiB2 is an essential protein. To further characterize WhiBTM4 we were interested in the phenotype of a WhiBTM4 mutant. A *whiBTM4*-gene deletion was performed by cloning the *whiBTM4* flanking regions carrying *Mfe*I recognition sites into *Mfe*I cut TM4 genomic DNA (Fig. 9A and B). After transformation the gene deletion was confirmed by
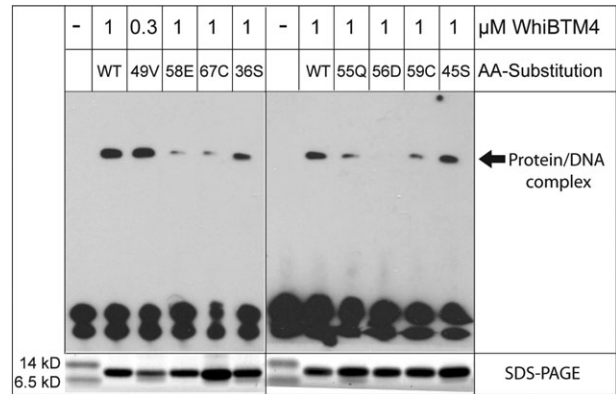

Fig. 8. Electrophoretic mobility shift assay (EMSA) of mutated WhiBTM4. WhiBTM4 proteins with single amino acid substitutions situated in the C-terminal HTH motif were incubated with labelled WhiB2Mtb promoter DNA. The L55Q, V58E and R67C substitutions led to decreased binding affinity compared with the wild-type protein whereas the V56D substitution abolished DNA binding of WhiBTM4. WhiBTM4 with a A49V substitution seemed to be a better DNA-binder than the wild-type protein. This protein expressed poorly in *E. coli* (SDS-PAGE below EMSA film). Substitutions of the cysteines at position 36 and 45 had no effect on DNA binding. The experiment was performed three times to rule out pipetting errors.

plaque-PCR using the forward primer I30 which anneals to gene 48 and the reverse primer I31 (Table S3) which anneals to gene 50. The deletion leads to smaller PCR products compared with PCR performed on wild-type phage (Fig. 9B). Furthermore, primers 49for and 49rev (Table S3) were used to show that *whiBTM4* is missing in the deletion mutant (Fig. 9B). Mutant and wild-type phage showed similar plaque morphology indicating that WhiBTM4 is dispensable for phage growth (Fig. 9C). We also determined the average number of particles per plaque for both, mutant and wild-type phage ($n = 10$ plaques). In the mutant we counted 5.4×10^7 (standard deviation 3.1×10^7) particles per plaque and 7.4×10^7 (standard deviation 3.7×10^7) particles per plaque for the wild-type phage. Using student's *t*-test this difference was not significant ($P = 0.195$, 95% confidence interval) indicating that there is no decrease in phage fitness caused by the *whiBTM4* deletion.

The presence of WhiBTM4 induces superinfection exclusion

Since WhiBTM4 is non-essential for proper plaque formation in top-agar containing *M. smegmatis* cells we were wondering whether the expression of sub-toxic levels of WhiBTM4 allows TM4 an exclusive utilization of its host. Interestingly, we have recently identified another early phage protein of mycobacteriophage L5 that leads to a similar phenotye when expressed in the host bacterium *M. smegmatis* (Rybniker *et al.*, 2008). The alteration of
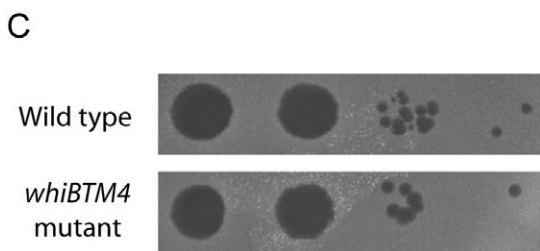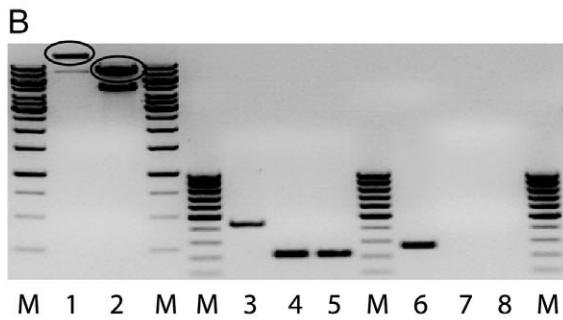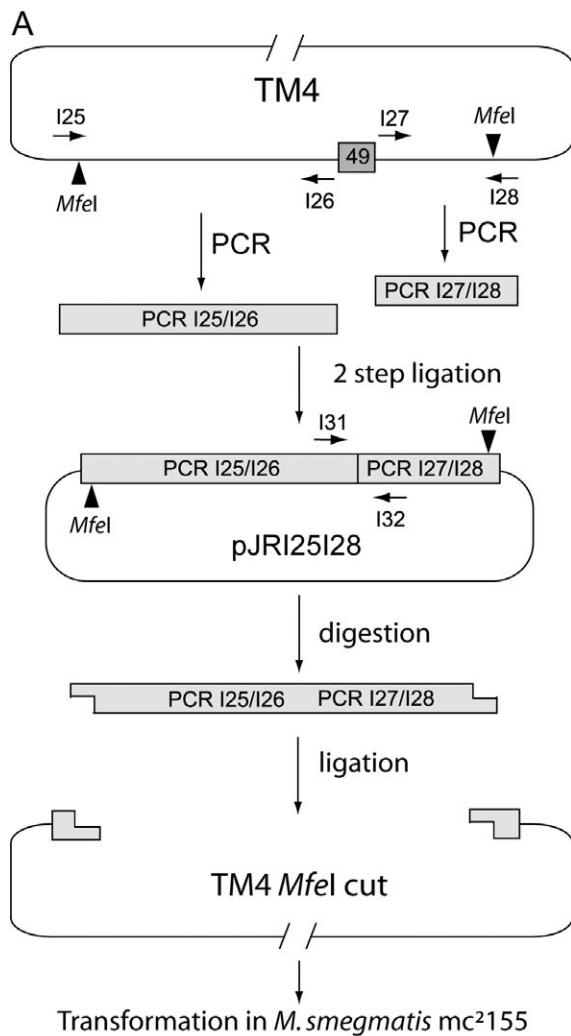
Fig. 9. A. Strategy for creating a *whiBTM4* deletion mutant. PCR products of primers I25/I26 and I28/I29 were sequentially cloned in pQE80 giving pJRI25I28. Phage DNA and pJRI25I28 were MfeI cut, DNA pieces were gel purified, ligated and transformed in *M. smegmatis.*
B. Phage TM4 DNA (lane 1) and pJRI25I28 (lane 2) digested with *Mfe*I, the labelled bands were excised, purified and ligated. The gene deletion was confirmed by PCR using wild-type TM4 (lanes 3, 6), mutant (lanes 4, 7) and pJRI25I28 (lanes 5, 8) as templates. Primers I30/I31 (lanes 3–5) anneal to the flanking genes of *whiBTM4* thus the PCR product of the mutant phage is approximately 260 bp smaller compared with the wild-type PCR product. Primers 49for/49rev (lanes 6–8) amplify *whiBTM4.*
C. Plaque morphology of mutated and wild-type phage.

cell wall components early in the infectious cycle may prevent infection of *M. smegmatis* with additional TM4 particles or other mycobacteriophage species, a phenomenon known as superinfection exclusion. Full expression of WhiBTM4 was lethal for the host. Uninduced cells grew relatively slow due to leakiness of pSD24 (Rybniker *et al.*, 2008) which leads to basal transcription of *whiBTM4*-specific RNA (Table S1B). Spotting phage lysate on topagar containing these cells did not lead to the formation of single clear plaques indicating a TM4 resistant phenotype (Fig. 10A). Thus, a function of the early cell shape altering protein WhiBTM4 may be found in superinfection exclusion. This mechanism seems not to involve altered binding of phage particles. An adsorption assay performed with wild-type *M. smegmatis*, *M. smegmatis* carrying pSD24WhiBTM4 or bacteria expressing mutated WhiBTM4 species gave similar results for the different hosts (Fig. 10B).

Discussion

This work identifies and characterizes the bacteriophage derived Fe-S protein WhiBTM4. To our knowledge this is the first description of a viral protein capable of co-ordinating a Fe-S cluster. Since the sequence of WhiBTM4 is highly identical to the WhiB-protein family of actinomycetes, functional interpretation of this protein also requires substantial knowledge of its bacterial counterparts.

Sequence comparison of WhiBTM4 with the seven WhiB proteins of mycobacteria revealed high sequence identity in 66 amino acids of the C-terminal region of WhiB2 of *M. tuberculosis* and *M. smegmatis*, a protein that is essential for proper septation and cell division (Fig. 1). This C-terminal part represents the active part of WhiB2 which is capable of rescuing a WhiB2Ms deletion mutant (Raghunand and Bishai, 2006b). Our data show that WhiBTM4 expression leads to downregulation of WhiB2 with subsequent filamentation and growth inhibition. The presence of the Fe-S co-ordination cysteines is mandatory for this phenotype. Furthermore, cells
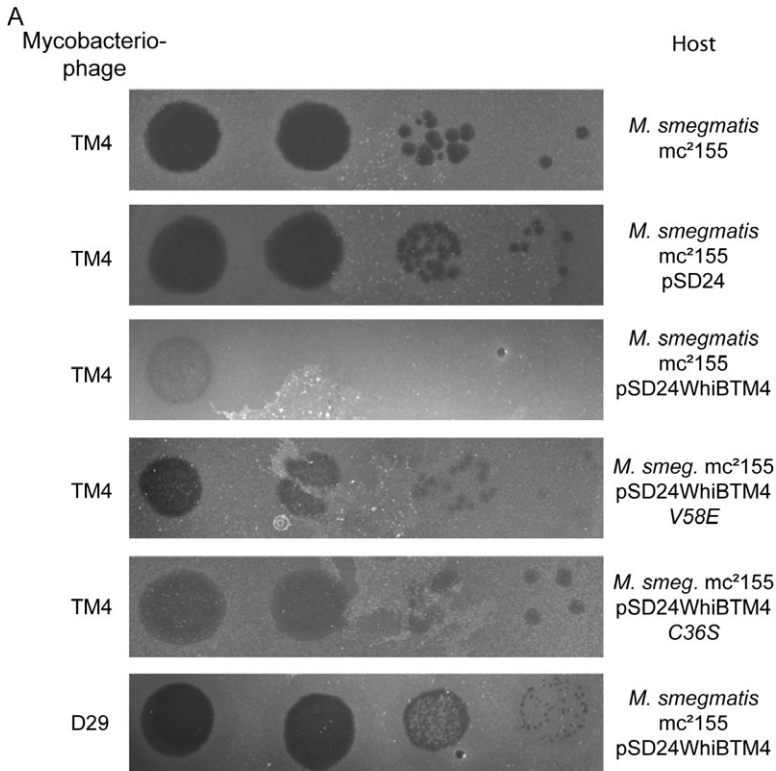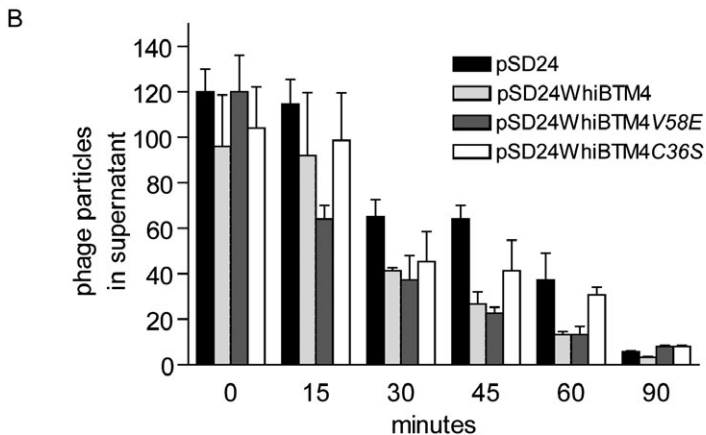
A
Mycobacterio-
phage   Host

TM4   *M. smegmatis*
                                                       mc²155

TM4   *M. smegmatis*
                                                       mc²155
                                                       pSD24

TM4   *M. smegmatis*
                                                       mc²155
                                                       pSD24WhiBTM4

TM4   *M. smeg.* mc²155
                                                       pSD24WhiBTM4
                                                       V58E

TM4   *M. smeg.* mc²155
                                                       pSD24WhiBTM4
                                                       C36S

D29   *M. smegmatis*
                                                       mc²155
                                                       pSD24WhiBTM4

Fig. 10. A. Plaque formation on *M. smegmatis* mc²155 carrying no plasmid, pSD24, pSD24WhiBTM4 or the mutants pSD24WhiBTM4*V58E* and pSD24WhiBTM4*C36S*. In contrast to TM4 grown on the native host, there is no zone of clear plaques visible in cells expressing basal levels of WhiBTM4. Only at the lowest dilution some turbid plaques are visible. Phage grown on bacteria expressing mutated proteins show single turbid plaques at the highest dilution. Mycobacteriophage D29, another lytic mycobacteriophage, forms small clear plaques at the highest dilution.
B. Adsorption assay of phage TM4 incubated with different *M. smegmatis* transformants. Cells in the exponential growth phase were incubated with TM4 at a MOI (multiplicity of infection) of 0.0002. Unbound phage particles were rescued at different time points. There is no significant difference in phage adsorption between the examined transformants. Experiments were performed in triplicates.

B

expressing the TM4-protein showed a loss of acid fast properties indicating a WhiBTM4 mediated modulation of cell wall lipids (Fig. 2).

We show here, for the first time, that a phage derived WhiB protein and the essential host protein WhiB2 are capable of binding DNA. Both proteins specifically bound the promoter DNA upstream of the *whiB2* gene, strengthening the hypothesis that WhiB proteins are important transcriptional regulators (Fig. 6). Mycobacteriophage TM4 has a broad host range and is lytic in most of the mycobacteria it infects. In contrast to other mycobacteriophages such as L5 and Bxb1, TM4 infection does not lead to a rapid and global shut-off of host protein synthe-

sis immediately after infection (Ford *et al.*, 1998). SDS-PAGE of mycobacterial proteins after infection of the cells with TM4 showed that host gene expression is unchanged throughout the late phase of the cycle which starts approximately 60 min after infection (Ford *et al.*, 1998). This observation is supported by our qRT-PCR experiments where *16S rRNA* and *HSP60*-mRNA-levels were stable during TM4 infection (Fig. 3A). However, *whiB2*-mRNA levels clearly declined. Since *whiB2* is an essential host gene and WhiBTM4 expression is growth inhibitory, the protein functions as a more subtle shut-off protein. Our findings clearly establish WhiBTM4 as a dominant negative regulator of WhiB2.

We also found that host cells expressing WhiBTM4 at a basal, sub-toxic level are resistant to TM4 infection. This implies a function of WhiBTM4 in superinfection exclusion. Though well known in T-even phages of *E. coli*, superinfection exclusion genes have not yet been identified in lytic mycobacteriophages. In the *E. coli* phage T4 the Imm protein induces resistance to most T-even phages but not to phage lambda (Lu and Henning, 1994). The extremely lipophilic Imm protein is associated with the cell membrane and is believed to inhibit the transfer of phage-DNA across the cytoplasmic membrane. It is likely that WhiBTM4 alters the host cell wall in such a way that superinfection by TM4 or other mycobacteriophages is impossible. Since WhiB proteins do not contain membrane spanning domains and are believed to be cytoplasmic proteins, this alteration is most likely achieved by gene regulation rather than direct interaction with the bacterial membrane. Here WhiBTM4 could represent a truncated, non-functional homologue of WhiB2 which acts as a transcriptional antagonist displacing WhiB2 from its promoter region and leading to down regulation of the host protein with subsequent alterations of the cell wall composition. Superinfection exclusion guarantees genetic diversity of phage populations by decreasing the degree of selection for dominant phage species co-infecting hosts or lysogens (Lu and Henning, 1994). Thus in mycobacteriophages, the presence of genes causing superinfection exclusion is consistent with their extraordinarily high genetic diversity.

However, the observed septation deficiency and superinfection exclusion could also be caused by the perturbation of protein levels due to the presence of a fully functional WhiB2 homologue (WhiBTM4). Wild-type bacteria seem to be highly sensitive towards the alteration of WhiB2 levels and the overexpression of WhiB2Ms in *M. smegmatis* leads to filamentation and a decrease in viability (Raghunand and Bishai, 2006a).

Another mycobacterial WhiB protein involved in the regulation of cell wall components is WhiB3. This protein has been shown to co-ordinate an oxygen sensitive [4Fe-4S] cluster with the help of the *M. tuberculosis* cysteine desulfurase IscS (Singh *et al.*, 2007). We provide evidence that WhiBTM4 and WhiB2 co-ordinate a [2Fe-2S] cluster when kept in an oxygen depleted atmosphere (Figs 5, S1, S2B/C). We cannot rule out that WhiBTM4 may build up a [4Fe-4S] cluster *in vivo*, possibly with the help of a host protein such as IscS. Despite several trials under strict anaerobic conditions using different reducing agents such as dithionite and sodium aspartate, we were not able to generate an electroparamagnetic resonance (EPR) signal from either freshly purified or reconstituted WhiBTM4. However, the UV-visible absorbance of WhiBTM4 contained features that are entirely consistent with those of other freshly purified [2Fe-2S] cluster carrying WhiB proteins of the host and those from well-defined [2Fe-2S] cluster co-ordinating proteins such as the ferredoxins (Jakimowicz *et al.*, 2005; Alam *et al.*, 2007; 2009).

A recent publication identified WhiB3 as a regulator of virulence lipid synthesis, most likely by binding to promoter regions upstream of polyketide synthases (Singh *et al.*, 2009). Here, apo-WhiB3 was a much better DNA-binder than the cluster carrying holo-WhiB3 indicating that WhiB3 is a redox-responsive regulatory protein. Additionally, in WhiB3 the conserved cysteines are not only mandatory for Fe-S cluster formation but also for the formation of intramolecular disulfide-bonds after cluster loss. It is the formation of these bonds that enables apo-WhiB3 to form a strong protein-DNA complex (Singh *et al.*, 2009). This mechanism stands in contrast to that of other well described Fe-S proteins such as FNR of *E. coli*. In FNR it rather is the presence of a [4Fe-4S] cluster that activates the expression of genes through binding of the C-terminal HTH motif to specific DNA sequences (Lazazzera *et al.*, 1996; Khoroshilova *et al.*, 1997). Cluster loss through oxygenation leads to the formation of apo-FNR, a biologically inactive molecule without DNA-binding properties (Lazazzera *et al.*, 1996).

It is an intriguing finding that a viral protein expressed early in the infectious cycle is capable of sensing hypoxia, normoxia and the presence of reducing or oxidizing agents with the help of its Fe-S cluster. It is rather unlikely that WhiBTM4 undergoes the cycle of cluster assembly, loss and re-assembly as part of a redox sensing machinery during the brief presence of the protein in the early stage of phage infection. Redox sensing is not considered to be an essential part of the viral infectious cycle. Since it is the apo-forms of WhiB3, WhiB2 and WhiBTM4 that are capable of binding DNA, the Fe-S cluster in WhiBTM4 could rather be a by-product engendered by the presence of four highly conserved cysteines. Here, the cysteine mediated formation of intramolecular disulfide bonds is the prerequisite for DNA-binding and transcriptional regulation. In this scenario the formation of the observed Fe-S cluster is secondary, which may be mirrored in the relative fast cluster loss we observed. Recently published data on the WhiB2 kinetics of cluster degradation through atmospheric oxygen showed a relatively slow cluster loss over a period of 48 h (Alam *et al.*, 2009). This stands in contrast to WhiBTM4 where cluster loss is a matter of minutes. Thus, we hypothesized that WhiB2 downregulation by WhiBTM4 may be achieved through a faster transformation of [2Fe-2S]-WhiBTM4 to apo-WhiBTM4. However, in our hands WhiB2Mtb showed a WhiBTM4-like spectroscopic behaviour under atmospheric oxygen (Fig. S2).

Eight of the 17 mutations leading to normal bacterial growth during WhiBTM4 expression were situated in the

predicted HTH DNA-binding motif represented by the most C-terminal α-helices and an exposed β-turn. We were able to show a decrease in promoter DNA-binding affinity of WhiBTM4 proteins carrying amino acid substitutions in this conserved HTH motif (Fig. 8). These facts strengthen the theory that WhiB proteins transmit their information using this C-terminal structure (Soliveri *et al.*, 2000). In contrast to the highly conserved β-turn, the α-helices are less well conserved among the different WhiB proteins and might have different targets explaining the versatile functions of these otherwise nearly identical proteins.

It was surprising to us that the *whiBTM4* gene is non-essential for proper plaque formation (Fig. 9). It was speculated by others that large portions of the TM4 genome are non-essential due to the relative ease with which foreign genetic elements can be introduced into the TM4 genome (Ford *et al.*, 1998). However, these dispensable genes were expected to be situated downstream of gene 71 and not in the *whiBTM4* (gene 49) region. It is well known that up to 50% of the genes of extensively studied bacteriophages such as T7 are non-essential for plaque formation. These genes, although dispensable for phage growth under laboratory conditions, bear meaningful functions and some of them are known to interact with and inhibit host proteins (Kim and Chung, 1996).

The systematic sequencing of more than 30 mycobacteriophages identified a large amount of interesting genetic data spanning a diverse range of gene families such as modified host-derived genes, clearly bacteriophage derived genes and a substantial number of genes with no match in the database (Pedulla *et al.*, 2003; Hatfull *et al.*, 2006). The analysis of phage genes that are homologous to, but not necessarily identical to host genes is a pioneering way to dissect the function of poorly understood genes of the host. Here we show that the expression and analysis of a phage derived protein provides information not only on the phage itself, but also on essential bacterial proteins. Particularly with regard to the potential drug target WhiB2, we believe that our findings may provide a new tool to target the major pathogen *M. tuberculosis*.

Experimental procedures

Plasmids, primers and bacterial strains

A description of the plasmids, primers and bacterial strains used in this study is given in Tables S2 and S3. A description of the bacteriophage used in this study is given in Table S2.

Bacteriophage methods

Propagation of phage and spot assays were performed as described recently (Rybniker *et al.*, 2006). For spot assays on bacteria carrying pSD24 (Daugelat *et al.*, 2003) or its derivatives cells were grown in 7H9 medium with 50 μg ml^{-1} hygromycin B. Cells were then harvested by centrifugation, washed and plated in top agar without hygromycin B.

Expression of WhibTM4 in M. smegmatis *and microscope techniques*

whibTM4 (ORF49) was amplified by PCR using the primer pair 49for and 49rev (Table S3) and TM4-DNA as a template. The product was cloned into BamHI cut pSD24 using the In-Fusion PCR cloning technique (Clontech). After transformation into the *E. coli* Fusion Blue strain the plasmid pSD24WhibTM4 was isolated, sequenced and transformed into *M. smegmatis* mc²155. Clones were grown on Middlebrook 7H10 agar containing 10% ADC and 50 μg ml^{-1} hygromycin B. For microscopy studies single colonies were picked and grown in 7H9 broth supplemented with 10% ADC, 0.05% Tween 80 and 50 μg ml^{-1} hygromycin B until an OD of 1 at A_{600} was reached, then cells were diluted 1:10 v/v in 7H9 broth containing 0.02% acetamide. Cells were collected at different time points, washed in sterile water and stained.

The TB stain kit (Becton Dickinson) was used for Ziehl-Neelsen acid fast staining. For fluorescence microscopy cells were washed in ultra pure water, the pellets were resuspended in 500 μl SYTO 11 (Molecular Probes) at a final concentration of 5 μM and incubated at 37°C for 30 min. Cells were then washed several times in water until the supernatant was colourless. A total of 20 μl of cells were mounted on microscope slides in ProLong Gold Antifade Reagent (Invitrogen). Images were acquired with an inverted Olympus IX81 microscope (Olympus) equipped with a F-View II Trigger camera and then analysed by analySISD software (Soft Imaging Software, Muenster, Germany).

RT-PCR for the quantification of gene expression during phage infection

Mycobacterium smegmatis mc²155 was grown to late log phase in 7H9 medium containing 10% ADC and 1 mM CaCl$_2$ without Tween 80. The cells were centrifuged and the pellets resuspended in warm 7H9 medium with 1 mM CaCl$_2$ but without ADC. Larger clumps were allowed to settle for 5 min and cells in the supernatant were adjusted to an OD of 0.1 at A_{600} in a total volume of 0.8 litres 7H9 medium with 1 mM CaCl$_2$. (approximately 2×10^{10} cells). Mycobacteriophage TM4 was added at a multiplicity of infection (MOI) of 10 and the cells were shaken slowly at 60 rpm at 37°C. Cells were harvested in 100 ml portions at different time points by centrifugation for 3 min (4000 × g) at 4°C, the pellets were shock frozen and stored at −80°C. Mycobacterial RNA was isolated using the RNeasy MIDI-kit (Qiagen) including the on-column DNAse digestion step. RNA (1 μg) was reverse transcribed with the QuantiTect Reverse Transcription kit (Qiagen). Real-time PCR was performed using the QuantiFast SYBR green PCR kit (Qiagen) on a GeneAmp 5700 sequence detection system (Applied Biosystems). Primers were H27/H28 for the detection of *whiBTM4*, H23/H24 for the mycobacterial *16S rRNA*, H25/H26 for *HSP60* and H31/H32 for the detection of *whiB2* of *M. smegmatis* (primer sequences are listed in

Table S3). Genomic DNA was used to generate a standard for the mycobacterial genes and pSD24WhiBTM4 to standardize the amount of *whiBTM4*.

Purification of WhiBTM4 and WhiB2Mtb

whiBTM4 and *whiB2Mtb* were PCR-amplified using the primer set Q8049for/Q8049rev or Q80WBfor/Q80WBrev (Table S3), respectively, and cloned into the expression plasmid pQE80 (Qiagen) in frame with the hexa-histidine tag. *Escherichia coli* BL21 (DE3) cells were transformed with the respective plasmids and grown in LB-broth containing 100 µg ml^{-1} ampicillin to an OD of 0.6 at A$_{600}$. Expression was induced with 0.5 mM IPTG for 2 h at 30°C without shaking. For purification under native conditions, a cell pellet from a 800 ml culture was resuspended in 20 ml buffer A (200 mM NaCl, 50 mM NaH$_2$PO$_4$, 30 mM imidazol, 10% glycerol, pH 8.0) plus EDTA-free protease inhibitor cocktail (Roche). Cells were disrupted in a FRENCH press (Thermo scientifc) at 1250 psi; the soluble lysate was separated by centrifugation at 17 000 × g for 15 min at 4°C and loaded on a Ni^{2+}-NTA affinity column (HisTrap FF crude, GE-Helthcare). The column was washed with 15 volumes of buffer A and eluted with buffer B (200 mM NaCl, 50 mM NaH$_2$PO$_4$, 300 mM imidazole, 10% glycerol, pH 8.0) on an ÄKTA-FPLC system using the UNICORN workstation software (GE-Healthcare). For purification under denaturing conditions the lysis/wash-buffer was buffer C (8 M urea, 100 mM NaH$_2$PO$_4$, 10 mM Tris-HCL, 30 mM imidazole, pH 8.0) and elution-buffer D (8 M urea, 100 mM NaH$_2$PO$_4$, 10 mM Tris-HCL, 500 mM imidazole, pH 8.0). All buffers were extensively vacuum degassed and purged with nitrogen. Eluted protein was immediately examined by spectroscopy or frozen in liquid nitrogen for storage at −80°C. Protein concentrations were determined with the Compat-Able BCA protein assay and the Pierce 660 nm protein assay (Thermo scientific). Colorimetric tests were performed with apo-WhiBTM4 or apo-WhiBMtb.

Electrophoretic mobility shift assays (EMSA)

Electrophoretic mobility shift assays were performed using the lightshift assay (Pierce) following the recommendations of the manufacturer. In brief, native WhiB2Mtb or WhiBTM4 were desalted using zeba spin columns (Pierce) and incubated with 20 fmol of 5′-biotinylated and annealed WhiB2Mtb promoter DNA (Sigma Aldrich) in binding buffer supplemented with 2% glycerol, 5 mM MgCl$_2$, 0.5% NP-40, 10 mM diamide, 0.5% Tween 20 and 10 ng µl^{-1} poly dI-dC-DNA for 30 min. The reactions were separated using a 4–20% gradient TBE PAGE gel (Invitrogen) and 0.5x TBE buffer as running buffer. DNA was transferred on a nylon membrane and biotinylated DNA was detected by chemoluminescence through exposure to an X-ray film for approximately 2 min.

Reconstitution of the Fe-S cluster

Fully oxygenized protein (0.8 mg ml^{-1}) purified under denaturing conditions was dialysed against buffer E (150 mM NaCl, 50 mM Tris-HCL, 10 mM DTT, 100 µM FeCl$_3$, 100 µM Na$_2$S, pH 7.5) for 12 h at 22°C. The dialysis cassette was rinsed with buffer F (200 mM NaCl, 50 mM Tris-HCL, 10 mM DTT, pH 7.5) and dialysed against buffer F at 4°C for 12 h. Again, all buffers were degassed and nitrogen purged. The dialysis steps were performed under strict anaerobic conditions using anaerobic jars that were depleted from oxygen with an Anoxomat system (Mart Microbiology, Netherlands) which evacuates atmospheric oxygen from the jar and refills with an anaerobic gas mixture (nitrogen 85%, CO$_2$ 10%, hydrogen 5%).

Random mutagenesis of WhibTM4

Low fidelity PCR was performed using the primer pair 49for/49rev (Table S3) and pSD24WhiBTM4 as a template. A 50 µl standard PCR reaction was modified by the addition of 320 µM MgSO$_4$ and 80 µM dGTP (final concentration) and 1 µl of the diversify dNTP-mix (Clontech). Cycle conditions were 94°C for 1 min followed by 25 cycles of 30 s at 94°C and 1 min at 68°C. Using this setting, the mutational bias is approximately 2.7 mutations per 1000 bp. The PCR products were In-Fusion-cloned into BamHI linearized pSD24 and transformed into *E. coli* Fusion-Blue. Approximately 10^4 individual clones were pooled and plasmids isolated (pSD24WhiBTM4*M*). Approximately 0.5 µg of this randomly mutagenized plasmid library was transformed into *M. smegmatis* mc^{2}155 and clones were plated on 7H10 plates with and without 0.2% acetamide. Insertions from mutants growing on 0.2% acetamide were sequenced and the sequences were aligned using ClustalW software. To confirm the acetamide resistant phenotype, all mutated *whiBTM4* genes were re-amplified under stringent PCR conditions and the products were cloned into pSD24. The resulting plasmids were transformed into *M. smegmatis* and the transformants plated on 7H10 plates containing 0.2% acetamide. The inserts of these plasmids were sequenced again as a further proof for a relevant mutation.

Construction of a whiBTM4 *gene deletion mutant*

The TM4 genome carries two MfeI recognition sites at position 28.578 and 36.370 (Fig. 9A). These sites flank the *whiBTM4* gene. Two PCR products were created: one product [primers I25 and I26 (Table S3)] spanning the recognition site at position 28.578 and the gene upstream of *whiBTM4* (gene 48) and a second product [primers I28 and I29 (Table S3)] spanning gene 50 and the downstream recognition site at position 36.370. These products were sequentially cloned into pQE80, the insert was excised and ligated into gel purified MfeI cut TM4 genomic DNA. The ligation reaction was transformed in *M. smegmatis* mc^{2}155. The resulting plaques were plaque purified and the gene deletion was confirmed by PCR using primers 49for/49rev and I30/I31 (Table S3).

Other analytical techniques

For the quantification of iron by atomic absorption spectrometry (AAS) the WhiBTM4 samples were treated with HNO$_3$ and diluted in an appropriate amount of ultrapure water. Samples were analysed on an AAnalyst 800 (Perkin Elmer).

UV-visible absorption spectra were recorded on a Lambda 40 spectrophotometer (Perkin Elmer).

The PMF was performed as described recently (El Mourabit *et al.*, 2004). In brief, proteins were pre-treated and trypsin digested at 37°C over night. The digest was stopped by the addition of 5–20 μl 1% trifluoraceticacid (TFA) in water and peptides were extracted for 30 min at 37°C. Positive ion spectra were acquired on a Reflex IV MALDI-TOF mass spectrometer (Bruker Daltonics, Bremen, Germany) in the reflectron mode. A peptide calibration standard (Bruker Daltonics, Bremen, Germany) was used for external calibration of the mass range from m/z 1046 to m/z 3147. The FlexAnalysis postanalysis software was used for optional internal recalibration on trypsin autolysis peaks and the generation of peaklists. Biotools 3.0 (Bruker Daltonics, Bremen, Germany) was used for interpretation of mass spectra with regard to the expected sequence of recombinant 6xHIS-tagged WhiBTM4.

Gelfiltration was performed using a superdex 75 column (GE Healthcare) operated at a flow rate of 0.05 ml min^{-1} with 100 mM NaCl, 50 mM Tris pH 8.0. The column was calibrated using a low molecular weight protein standard (GE Healthcare).

Acknowledgements

References

Alam, M.S., Garg, S.K., and Agrawal, P. (2007) Molecular function of WhiB4/Rv3681c of *Mycobacterium tuberculosis* H37Rv: a [4Fe-4S] cluster co-ordinating protein disulphide reductase. *Mol Microbiol* **63:** 1414–1431.

Alam, M.S., Garg, S.K., and Agrawal, P. (2009) Studies on structural and functional divergence among seven WhiB proteins of *Mycobacterium tuberculosis* H37Rv. *FEBS J* **276:** 76–93.

Beinert, H., Holm, R.H., and Munck, E. (1997) Iron-sulfur clusters: nature's modular, multipurpose structures. *Science* **277:** 653–659.

Cole, S.T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., *et al.* (1998) Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* **393:** 537–544.

Crack, J.C., Hengst, C.D., Jakimowicz, P., Subramanian, S., Johnson, M.K., Buttner, M.J., *et al.* (2009) Characterization of [4Fe-4S]-containing and cluster-free forms of Streptomyces WhiD. *Biochemistry* **48:** 12252–12264.

Daugelat, S., Kowall, J., Mattow, J., Bumann, D., Winter, R., Hurwitz, R., and Kaufmann, S.H. (2003) The RD1 proteins of *Mycobacterium tuberculosis*: expression in *Mycobacterium smegmatis* and biochemical characterization. *Microbes Infect* **5:** 1082–1095.

El Mourabit, H., Muller, S., Tunggal, L., Paulsson, M., and Aumailley, M. (2004) Analysis of the adaptor function of the LIM domain-containing protein FHL2 using an affinity chromatography approach. *J Cell Biochem* **92:** 612–625.

Ford, M.E., Stenstrom, C., Hendrix, R.W., and Hatfull, G.F. (1998) Mycobacteriophage TM4: genome structure and gene expression. *Tuber Lung Dis* **79:** 63–73.

Gomez, J.E., and Bishai, W.R. (2000) whmD is an essential mycobacterial gene required for proper septation and cell division. *Proc Natl Acad Sci USA* **97:** 8554–8559.

Hatfull, G.F., Pedulla, M.L., Jacobs-Sera, D., Cichon, P.M., Foley, A., Ford, M.E., *et al.* (2006) Exploring the mycobacteriophage metaproteome: phage genomics as an educational platform. *PLoS Genet* **2:** e92.

Jakimowicz, P., Cheesman, M.R., Bishai, W.R., Chater, K.F., Thomson, A.J., and Buttner, M.J. (2005) Evidence that the Streptomyces developmental protein WhiD, a member of the WhiB family, binds a [4Fe-4S] cluster. *J Biol Chem* **280:** 8309–8315.

Khoroshilova, N., Popescu, C., Munck, E., Beinert, H., and Kiley, P.J. (1997) Iron-sulfur cluster disassembly in the FNR protein of *Escherichia coli* by O2: [4Fe-4S] to [2Fe-2S] conversion with loss of biological activity. *Proc Natl Acad Sci USA* **94:** 6087–6092.

Kim, S.H., and Chung, Y.B. (1996) Isolation of a mutant bacteriophage T7 deleted in nonessential genetic elements, gene 19.5 and m. *Virology* **216:** 20–25.

Klinge, S., Hirst, J., Maman, J.D., Krude, T., and Pellegrini, L. (2007) An iron-sulfur domain of the eukaryotic primase is essential for RNA primer synthesis. *Nat Struct Mol Biol* **14:** 875–877.

Lazazzera, B.A., Beinert, H., Khoroshilova, N., Kennedy, M.C., and Kiley, P.J. (1996) DNA binding and dimerization of the Fe-S-containing FNR protein from *Escherichia coli* are regulated by oxygen. *J Biol Chem* **271:** 2762–2768.

Lu, M.J., and Henning, U. (1994) Superinfection exclusion by T-even-type coliphages. *Trends Microbiol* **2:** 137–139.

Messerschmidt, A. (2001) *Handbook of Metalloproteins*. Chichester; New York: Wiley.

Morris, P., Marinelli, L.J., Jacobs-Sera, D., Hendrix, R.W., and Hatfull, G.F. (2008) Genomic characterization of mycobacteriophage Giles: evidence for phage acquisition of host DNA by illegitimate recombination. *J Bacteriol* **190:** 2172–2182.

Morris, R.P., Nguyen, L., Gatfield, J., Visconti, K., Nguyen, K., Schnappinger, D., *et al.* (2005) Ancestral antibiotic resistance in *Mycobacterium tuberculosis*. *Proc Natl Acad Sci USA* **102:** 12200–12205.

Pedulla, M.L., Ford, M.E., Houtz, J.M., Karthikeyan, T., Wadsworth, C., Lewis, J.A., *et al.* (2003) Origins of highly mosaic mycobacteriophage genomes. *Cell* **113:** 171–182.

Raghunand, T.R., and Bishai, W.R. (2006a) *Mycobacterium smegmatis* whmD and its homologue *Mycobacterium tuberculosis* whiB2 are functionally equivalent. *Microbiology* **152:** 2735–2747.

Raghunand, T.R., and Bishai, W.R. (2006b) Mapping essen-

tial domains of *Mycobacterium smegmatis* WhmD: insights into WhiB structure and function. *J Bacteriol* **188:** 6966–6976.

Rouhier, N., Unno, H., Bandyopadhyay, S., Masip, L., Kim, S.K., Hirasawa, M., *et al.* (2007) Functional, structural, and spectroscopic characterization of a glutathione-ligated [2Fe-2S] cluster in poplar glutaredoxin C1. *Proc Natl Acad Sci USA* **104:** 7379–7384.

Rybniker, J., Kramme, S., and Small, P.L. (2006) Host range of 14 mycobacteriophages in *Mycobacterium ulcerans* and seven other mycobacteria including *Mycobacterium tuberculosis* – application for identification and susceptibility testing. *J Med Microbiol* **55:** 37–42.

Rybniker, J., Plum, G., Robinson, N., Small, P.L., and Hartmann, P. (2008) Identification of three cytotoxic early proteins of mycobacteriophage L5 leading to growth inhibition in *Mycobacterium smegmatis*. *Microbiology* **154:** 2304–2314.

Singh, A., Guidry, L., Narasimhulu, K.V., Mai, D., Trombley, J., Redding, K.E., *et al.* (2007) *Mycobacterium tuberculosis* WhiB3 responds to O2 and nitric oxide via its [4Fe-4S] cluster and is essential for nutrient starvation survival. *Proc Natl Acad Sci USA* **104:** 11562–11567.

Singh, A., Crossman, D.K., Mai, D., Guidry, L., Voskuil, M.I., Renfrow, M.B., and Steyn, A.J. (2009) *Mycobacterium tuberculosis* WhiB3 maintains redox homeostasis by regulating virulence lipid anabolism to modulate macrophage response. *PLoS Pathog* **5:** e1000545.

Soliveri, J.A., Gomez, J., Bishai, W.R., and Chater, K.F. (2000) Multiple paralogous genes related to the *Streptomyces coelicolor* developmental regulatory gene whiB are present in Streptomyces and other actinomycetes. *Microbiology* **146:** 333–343.

Steyn, A.J., Collins, D.M., Hondalus, M.K., Jacobs, W.R., Jr, Kawakami, R.P., and Bloom, B.R. (2002) *Mycobacterium tuberculosis* WhiB3 interacts with RpoV to affect host survival but is dispensable for *in vivo* growth. *Proc Natl Acad Sci USA* **99:** 3147–3152.

Tscherne, D.M., Evans, M.J., von Hahn, T., Jones, C.T., Stamataki, Z., McKeating, J.A., *et al.* (2007) Superinfection exclusion in cells infected with hepatitis C virus. *J Virol* **81:** 3693–3703.

Van Dessel, W., Van Mellaert, L., Liesegang, H., Raasch, C., De Keersmaeker, S., Geukens, N., *et al.* (2005) Complete genomic nucleotide sequence and analysis of the temperate bacteriophage VWB. *Virology* **331:** 325–337.

Supporting information

Additional supporting information may be found in the online version of this article.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Identification of three cytotoxic early proteins of mycobacteriophage L5 leading to growth inhibition in *Mycobacterium smegmatis*

Jan Rybniker,[1] Georg Plum,[2] Nirmal Robinson,[2] Pamela L. Small[3] and Pia Hartmann[1,4]

Correspondence
Jan Rybniker
jan.rybniker@uk-koeln.de

[1]1st Department of Internal Medicine, University of Cologne, 50924 Cologne, Germany

[2]Institute for Medical Microbiology, Immunology and Hygiene, University of Cologne, 50924 Cologne, Germany

[3]Department of Microbiology, University of Tennessee, Knoxville, TN 37996, USA

[4]Department of Internal Medicine 1, Division of Infectious Diseases, University of Regensburg, 93042 Regensburg, Germany

Mycobacteriophage L5 is a temperate phage with a broad host range among the fast- and slow-growing mycobacteria such as *Mycobacterium smegmatis*, *Mycobacterium tuberculosis*, *Mycobacterium avium* and *Mycobacterium ulcerans*. L5 switches off host protein synthesis during the early stage of lytic growth, as was previously shown by protein expression profiling. Also, lethal genetic elements have been identified in L5 based on the fact that transformants could not be obtained with these genes. Using an inducible mycobacterial shuttle vector, we have identified three ORFs within an early operon of mycobacteriophage L5 which encode gene products (gp) toxic to the host *M. smegmatis* when expressed. These ORFs, coding for gp77, gp78 and gp79, presumably function as shut-off genes during early stages of phage replication. There is evidence that cell division is affected by one of the proteins (gp79). The transcription of the cytotoxic polypeptides is directed by a promoter situated in ORF83 and transcription control is achieved through the phage repressor gp71, which is shown by co-expression of this protein. The findings presented here should provide useful tools for the molecular genetics of mycobacteria. Further analysis of these and other mycobacteriophage-derived toxic polypeptides, together with the identification of their cellular targets, might provide a tool for the rapid identification of promising drug targets in emerging and re-emerging mycobacterial pathogens.

INTRODUCTION

To date more than 250 mycobacteriophages have been isolated and the genomic sequences of more than 30 phages are available in the NCBI database. Taken together, these relatively large genomes code for more than 3000 putative proteins, many of which are without a match in the database (Hatfull *et al.*, 2006; Pedulla *et al.*, 2003). The functions of the vast majority of these mycobacteriophage ORFs remain unknown and only a few proteins have been expressed and examined in detail. Even more striking is the genomic diversity of mycobacteriophages, and exchange of genes among the phages themselves and among their hosts has led to a vivid and probably highly influential evolutionary relationship between virus and bacterial host.

Extensive comparison of the mycobacteriophage genomes and amino acid sequences and their organization in

so-called 'phamilies' (putative proteins with amino acid identity of 25 % or more) showed that almost half of the phage-derived sequences match non-phage and predominantly bacterial genes (Hatfull *et al.*, 2006). Furthermore, some of these non-phage-associated genes are homologues of suspected mycobacterial virulence genes and genes that are involved in bacterial cell-wall synthesis (Pedulla *et al.*, 2003). On the host's side, the sequencing of mycobacterial genomes led to the identification of many prophages, which presumably influence the phenotype and virulence of some mycobacteria (Stinear *et al.*, 2007).

Mycobacteriophage L5, a temperate phage isolated from *Mycobacterium smegmatis*, was the first mycobacteriophage to be sequenced (Hatfull & Sarkis, 1993). L5 forms stable lysogens in *M. smegmatis* and has a broad host range among the pathogenic mycobacteria (Rybniker *et al.*, 2006). Although L5 is a temperate phage, the following evidence suggests that the phage switches off host protein

Abbreviations: gp, gene product; PBP, penicillin-binding protein.

synthesis during early lytic growth. First, a quantitative decrease of host proteins is seen minutes after induction of a thermo-inducible L5 lysogen (Hatfull & Sarkis, 1993), and second, it has been reported that a genetic element spanning ORF83 to ORF77 does not transform in *M. smegmatis* (Donnelly-Wu *et al.*, 1993). These ORFs are situated on the right arm of the L5 genome, which encodes primarily early regulatory proteins.

The potential presence of host shut-off proteins in the temperate phage L5 implies an extremely tight silencing of the genes responsible by the L5 repressor. The repressor encoded by ORF71 is responsible for the temperate phenotype and the superinfection immunity of lysogens. Transcriptional silencing is achieved by binding of the repressor protein within or downstream of the four early promoter sequences identified so far (Brown *et al.*, 1997; Donnelly-Wu *et al.*, 1993).

The host shut-off phenomenon is well known for many lytic bacteriophages as well as for viruses of the eukaryotic cell (Fenwick & Clark, 1982; Sharma *et al.*, 2004; Svenson & Karlstrom, 1976). It has recently been shown that the systematic exploration of staphylococcal bacteriophage derived shut-off proteins and their targets in *Staphylococcus aureus* can lead to the identification of new antibiotic chemical compounds (Liu *et al.*, 2004).

For the identification of phage-derived toxic proteins, the existence of tightly regulated inducible expression vectors is a prerequisite. The recent development of inducible mycobacterial shuttle plasmids carrying the acetamidase promoter of *M. smegmatis* enabled us to screen for potentially toxic ORFs within the regulatory region of mycobacteriophage L5 (Daugelat *et al.*, 2003).

Candidate genes were amplified from the L5 genetic region that does not transform in *M. smegmatis* and were inserted in a set of modified acetamidase-inducible expression plasmids. Here we present data showing that the L5 gene products (gp) 77, 78 and 79 are toxic to its natural host. Furthermore, we provide new insight into the regulation of these proteins by the description of a new promoter element and transcription control through the L5 repressor gp71.

METHODS

Bacterial strains, bacteriophages and plamids. Details of the bacterial strains, phage and plasmids used in this study are given in Table 1. Since the pSD24 acetamidase promoter shows intrinsic activity in *Escherichia coli*, amplification of some plasmid constructs carrying toxic proteins was not possible. To address this problem, the two *lac* operator elements of plasmid pQE80 (Qiagen) were amplified using the primers lacOfor (5′-TTTTATCCATGGATCGTTGTGA-GCGGATAA-3′) and lacOrev (5′-TCTGCAGCTGGATCCGT-GATGGTGATGGTG-3′) and cloned between the acetamidase promoter and multiple cloning site of pSD24, giving pJR5 (Table 1, Fig. 1). This allowed the propagation of otherwise toxic plasmids in *E. coli* strains constitutively expressing the *lac* repressor such as Fusion-Blue (Clontech). Insertion of this DNA fragment had no influence on the function of the acetamidase promoter in *M. smegmatis*.

Growth conditions. *M. smegmatis* was grown using Middlebrook 7H10 agar or Middlebrook 7H9 broth (Difco) supplemented with 10 % OADC (oleic acid, albumin fraction V, dextrose, catalase),

Table 1. Strains, bacteriophages and plasmids used in the study

See Fig. 2 and Fig. 5 for additional plasmids generated in this study.

	Description	Reference or source
Escherichia coli		
Fusion blue	endA1 hsdR17 (r_{K12}^- m_{K12}^+) supE44 thi-1 recA1 gyrA96 relA1 lac F′[proA⁺B⁺ lacI�q ZΔM15∷Tn10(Tet^R)]	Clontech
DH5α	F⁻, ϕ80dlacZΔM15 Δ(lacZYA-argF)U169 deoR recA1 endA1 hsdR17(r_k^- m_k^+) phoA supE44 λ⁻ thi-1 gyrA96 relA1	Lab collection
M. smegmatis mc²155	High-efficiency transformation mutant	Snapper *et al.* (1990)
Bacteriophage		
Mycobacteriophage L5	Temperate phage of *M. smegmatis*	
Plasmids		
pQE80L	Expression vector carrying the *cis-laqI*�q repressor gene	Qiagen
pSD24	Acetamide-inducible mycobacterial expression vector	Daugelat *et al.* (2003)
pJR5	pSD24 carrying two *lac* operators between acetamidase promoter and multiple cloning site	This study
pJR6	pJR5 with L5 repressor gp71 including its promoter elements P1 and P2 in opposite direction to the acetamidase promoter	This study
pMC1871	*E. coli* promoter-probe vector	Shapira *et al.* (1983)
pJR7	pSD24 carrying the *lacZ* gene of pMC1871 (Fig. 1)	This study
pJR6	pJR7 with L5 repressor gp71	This study
pJR10	pJR7 without acetamidase promoter (Fig. 1)	This study

0.05 % Tween 80 and 0.5 %, v/v, glycerol. *E. coli* strains were grown using LB agar or broth. Cells transformed with the derivatives of pSD24 were grown on medium containing 50 µg hygromycin B ml^{-1} (Roth) for *M. smegmatis* or 100 µg hygromycin B ml^{-1} for *E. coli*. Ampicillin (Roth) was used at a final concentration of 100 µg ml^{-1} and tetracycline (Roth) at a concentration of 20 µg ml^{-1}.

DNA manipulation. DNA manipulation, electrophoresis and PCRs were carried out using standard techniques (Sambrook & Russell, 2001). All PCRs for subsequent cloning were performed using Phusion high fidelity polymerase (Finnzymes) according to the instructions provided by the manufacturer.

DNA cloning. All cloning steps were done by homologous recombination using the In-Fusion PCR cloning kit (Clontech). All primers had 15 bp homology with each end of the linearized vector at their 5′-end followed by 15 bp specific sequence homologues to the L5-ORF amplified as recommended by the In-Fusion protocol. Primers for the amplification and cloning of ORF77 in pJR5 are given as examples: forward primer 5′-TCACCATCACGGATCC*catgATC-GACGGCAA*-3′, reverse primer 5′-TCTGCAGCTGGATCC*tca-GCCCGCCAGCGC*-3′; sequence with homology to the linearized vector is underlined whereas sequence with homology to ORF77 is shown in italics (*atg*, start codon; *tca*, stop codon). With regard to pSD24, pJR5, pJR6, pJR7, pJR8 and pJR10 the single *Bam*HI restriction site was used for all insertional cloning steps with subsequent expression, whereas pQE80 was linearized with *Bam*HI and *Hin*dIII. Sequence identity was confirmed for all clones by sequencing the first 250–500 bp downstream of the restriction enzyme site that was used for the cloning procedure.

Electroporation of mycobacteria. Electrocompetent *M. smegmatis* cells were prepared by growing the bacteria in 7H9 broth (containing OADC and Tween 80) to an OD$_{600}$ of 1.0. The cells were kept on ice for 90 min and harvested by centrifugation at 4000 ***g*** for 15 min. The cells were washed three times in ice-cold 10 % glycerol and resuspended in a 1/100 volume of 10 % glycerol. Settings for electroporation were 2.5 kV, 25 µF and a resistance of 200 Ω.

Identification of toxic phage ORFs. Fig. 2 illustrates the plasmids that were created for the identification of toxic phage ORFs. *M. smegmatis* clones were subcultured on 7H10 agar containing 0.1 % acetamide (Sigma) and *E. coli* DH5α clones on LB agar supplemented with 0.2 mM IPTG (Roth). Clones showing growth inhibition on selective media were further examined using serial dilutions of acetamide for *M. smegmatis* and IPTG for *E. coli* followed by OD$_{600}$ readings after 4 h for *E. coli* and after 12 h for *M. smegmatis*. Also, standard survival curves were created for the clones of interest (see Fig. 4). For survival curves and OD$_{600}$ readings, subculturing of *M. smegmatis* clones with derivatives of pSD24 carrying toxic ORFs was kept to a minimum since pSD24 plasmid propagation is not fully stable over several subculturing steps.

Fluorescence microscopy. This was performed as described previously (Robinson *et al.*, 2007). In brief, individual colonies were scraped off 7H10 agar plates and washed in 1× PBS. Then 180 µl 0.1 M NaHCO$_3$ and 20 µl FITC (100 mg ml^{-1}; Sigma) were added to approximately 25 mg bacteria (wet wt). After incubation for 30 min at 37 °C the stained cells were washed three times with PBS and mounted on a glass slide. Microscopy was done on an Olympus IX81 inverted fluorescence microscope.

Insertion of the repressor gene ORF71 in pJR5 and co-expression studies. A sequence spanning bp 44331 to bp 45045 of the L5 genome was amplified by PCR using the forward primer 5′-

GCGCGGCCGCGGTACCTTAGGGCTCCGAGA-3′ and the reverse primer 5′-CACTTCTTTATCTAGATAACGCCAGTCGAT-3′. This amplified sequence covers ORF71, which codes for the L5 repressor, and the two promoters P1 and P2, which are situated upstream of ORF71. The PCR product was cloned into pJR5 linearized with *Xba*I and *Acc*65I, giving pJR6 (Fig. 1).

A PCR product spanning ORF83 to ORF77 was cloned in pJR5 and pJR6 respectively (Fig. 2). Plasmid DNA was amplified in *E. coli* Fusion-Blue and 250 ng of each construct was electroporated into 80 µl competent *M. smegmatis* cells. Transformants were plated onto 7H10 plates containing 50 µg hygromycin B ml^{-1}.

Generation of mycobacterial promoter-probe vectors and promoter identification. The *lacZ* gene of pMC1871 was amplified by PCR using the primers pSDlacFO (5′-TTTTATCCATGGAT-CCGTCGTTTTACAACG-3′) and pSDlacRE (5′-TCTGCAGCT-GGATCTCCCCTGCCCGCTTAT-3′). Primer pSDlacRE deletes the second *Bam*HI site, leaving only one functional *Bam*HI site upstream of the *lacZ* gene. The PCR product was purified and cloned into *Bam*HI-linearized pSD24, giving pJR7 (Fig. 1). The acetamidase promoter of pJR7 was deleted by digestion with *Bam*HI and *Xba*I and subsequent religation of the vector ends using an oligo-linker made by annealing the oligos LinkXbaBamFO (5′-CTAGCCCGAGGTA-TCTAGATATCTTTG-3′) and LinkXbaBamRE (5′-GATCCAAAGA-TATCTAGATACCTCGGG-3′), resulting in the plasmid pJR10 (Fig. 1). To generate pJR8, which is pJR7 constitutively expressing the L5 phage repressor gp71, a similar procedure was chosen as described for pJR5 and pJR6. These three plasmids, pJR7, pJR8 and pJR10, gave the basis for several constructs used for the identification of promoter activity in ORF83, which was amplified by PCR together with 12 bp of the 5′-end of ORF82, leading to a fusion protein of the gene product of ORF82 and the β-galactosidase protein (see Fig. 5a). For pJR7∷8382 and pJR8∷8382 the genetic element spanning ORF83, ORF82, the intergenic region of ORF82/ORF81 and the start codon of ORF81 was amplified and cloned in-frame with the *lacZ* gene. The 21 bp sequence between the stop codon of ORF82 and the start codon of ORF81 contains a repressor-binding site (Brown *et al.*, 1997).

Quantitative β-galactosidase assays. *M. smegmatis* clones were grown to mid-exponential phase in 5 ml 7H9 broth, the OD$_{600}$ was measured and the cells were harvested and resuspended in 500 µl cold PBS. As controls all constructs were grown in parallel in the presence of 0.1 % acetamide for 6 h. Glass beads (0.1 mm) were added and the cell suspension was disrupted by using a bead mill (MM200, Retsch) for a total of 15 min at maximum speed with 5 min on ice after every 5 min of disruption. After centrifugation for 5 min at 4 °C and 10 000 ***g***, β-galactosidase activity was measured by adding 30 µl supernatant to 210 µl Z-buffer (0.06 M Na$_2$HPO$_4$.7H$_2$O, 0.04 M NaH$_2$PO$_4$.H$_2$O, 0.01 M KCl, 0.001 M MgSO$_4$, 0.05 M β-mercaptoethanol, pH 7.0). The reaction was started by the addition of 60 µl ONPG (Sigma) at 4 mg ml^{-1} in 0.1 M sodium phosphate (pH 7.5) and subsequently incubated at 37 °C. After development of a visible colour change to yellow the reaction was stopped by the addition of 500 µl 1 M Na$_2$CO$_3$ and A_{420} was measured. β-Galactosidase activity was calculated in Miller units using the following equation: $1000 \times A_{420}/(t \times v \times OD_{600})$ with t being the time from the addition of ONPG to the addition of Na$_2$CO$_3$ and v the culture volume (5 ml) (Zhang & Bremer, 1995).

Primer extension. Five picomoles of primer Seclaq (5′-TTGGGTAACGCCAGGGTTTTC-3′) was end-labelled with [γ-^{32}P]ATP using 20 units T4-polynucleotide kinase (Fermentas). Labelled primer was passed through a Sephadex-G50 NICK-column (GE-Healthcare) to remove unincorporated nucleotides. Total RNA was isolated from *M. smegmatis* clones carrying either pJR10∷83 or
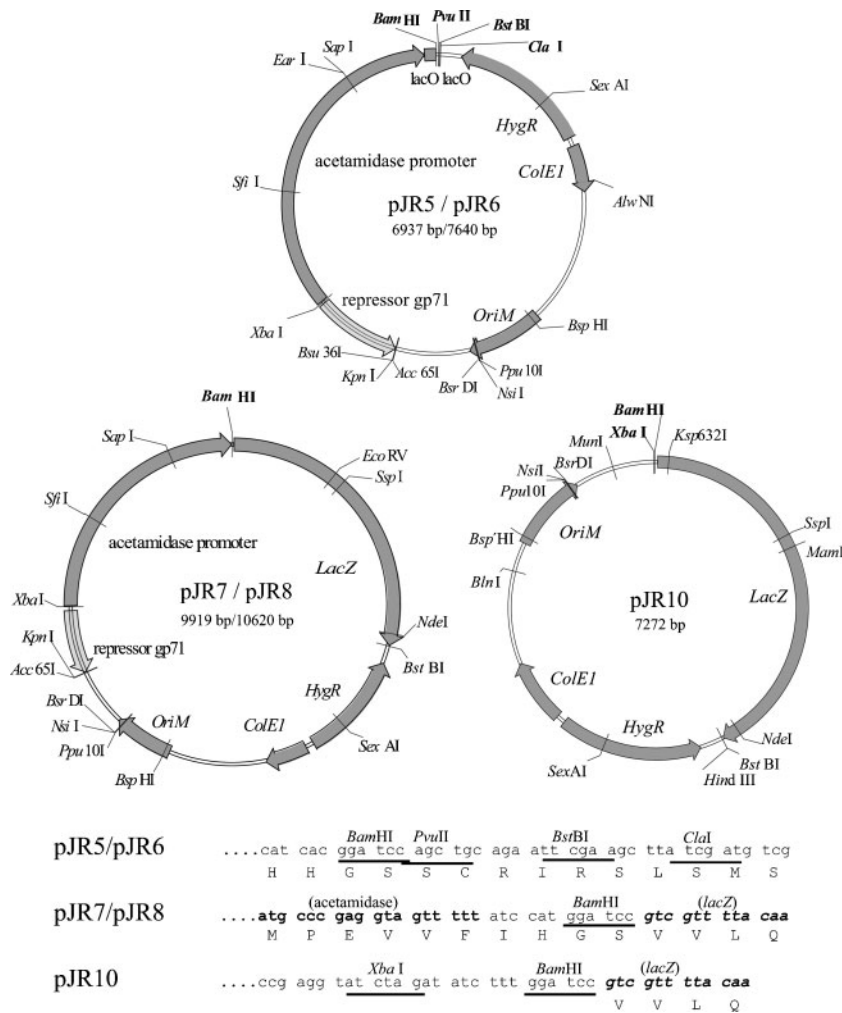
Fig. 1. Map of the modified acetamidase expression plasmid pJR5 and the new promoter-probe plasmids pJR7/pJR8 and pJR10. Only pJR6 and pJR8 carry the L5 repressor gene gp71 used for co-expression studies. The lower part shows the multiple cloning sites.

pJR10 as a control. Five micrograms of total RNA was incubated with approximately 50 fmol labelled primer for 5 min at 65 °C in a total volume of 10 μl. Two microlitres of a 10 mM dNTP mix, 1 μl 0.1 M DTT, 4 μl 5× cDNA buffer, 2 μl water and 15 units Thermoscript reverse transcriptase (Invitrogen) were added to each sample and the tubes were incubated at 50 °C for 45 min. After heating the samples to 85 °C for 5 min, phenol/chloroform extraction and ethanol precipitation were performed and the pellets were resuspended in 5 μl water together with 5 μl stop solution (T7-sequencing kit). A sequencing ladder was generated using the T7-sequencing kit (USB) following the instructions of the manufacturer. In brief, 2 μg plasmid DNA derived from pJR10 :: 83 were denatured with 2 M NaOH, ethanol precipitated and resuspended in 10 μl water. Primer Seqlac (10 pmol in 2 μl) and 2 μl annealing buffer were added and incubated for 5 min at 65 °C followed by 10 min at 37 °C and 5 min at room temperature. Then 3 μl labelling mix, 1 μl [α-^{32}P]dCTP and 2 μl diluted T7-DNA polymerase were added. After 5 min incubation at room temperature, 4.5 μl aliquots were added to 2.5 μl of the A, C, G and T termination mixes, respectively. The reaction was stopped after 5 min by the addition of 5 μl stop solution, and loaded next to the primer-extended samples on a denaturing sequencing gel (6 % acrylamide : bis-acrylamide 19 : 1, 7 M urea, 0.9 × TBE). The bands were visualized on storage phosphor screens using a Typhoon imager (GE-Healthcare).

RESULTS

Identification of toxic phage ORFs in M. smegmatis

As shown previously, the L5 sequence spanning bp 48750 to bp 46034, which covers ORF83 to ORF77, does not produce transformants when introduced into *M. smegmatis* mc²155 (Donnelly-Wu *et al.*, 1993). We could confirm this finding using the construct pJR5 :: 8377 (Fig. 2), which could not be transformed into *M. smegmatis* mc²155. We then deleted ORF83 from this construct, leading to the plasmid pJR5 :: 8277, which was transformable, giving relatively small *M. smegmatis* colonies on 7H10 agar after 5 days incubation at 37 °C. When transformants were subcultured on 7H10 agar containing 0.1 % acetamide there was no visible growth, suggesting that the phenotype of clones carrying pJR5 :: 8277 is due to one or more inhibitory ORFs within the phage sequence spanning ORF82 to ORF77. One could also assume that in pJR5 :: 8277 these toxic ORFs are decoupled from a promoter sequence upstream of ORF82 (Fig. 2). To define
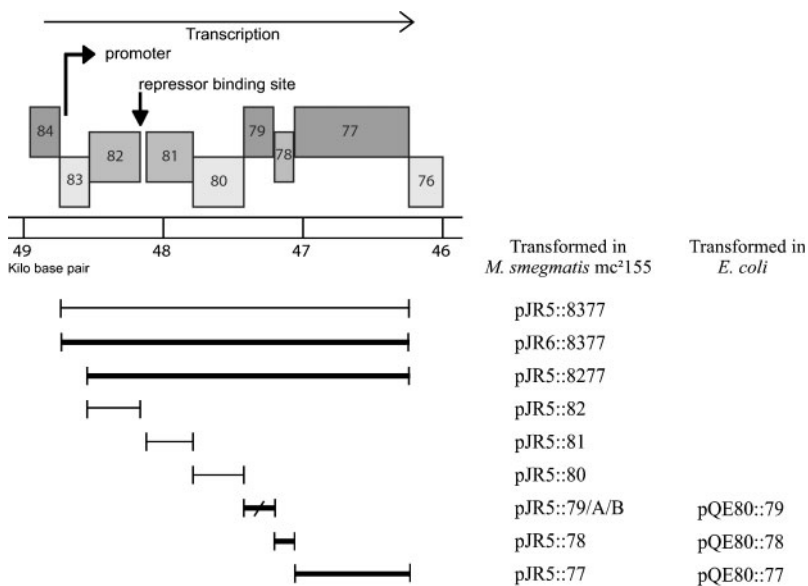
Fig. 2. Recombinant plasmids created in this study. The upper part illustrates L5 ORFs of the right arm, which predominantly carries early genes. Bold lines indicate inserts with growth-inhibitory properties when clones are grown in the presence of acetamide. pJR5::8377 does not produce transformants in *M. smegmatis* whereas pJR5::8277 does, leading to the assumption that ORF83 contains a putative promoter sequence. pJR6::8377 allows transformation since this construct constitutively expresses the phage repressor gp71. pJR5::79/A expresses the N-terminal putative signal peptide of ORF79 and pJR5::79/B the remaining C-terminal part. In contrast to pJR5::79, these constructs lead to wild-type colony morphology when induced.

the exact ORF responsible for toxicity, six plasmids were designed, individually carrying ORF82, 81, 80, 79, 78 or 77 under control of the acetamidase promoter (Fig. 2). Using these plasmids we were able to show that gp77, gp78 and gp79 possess growth-inhibitory properties, with gp77 being the most potent (Fig. 3a, b).

Uninduced *M. smegmatis* clones containing pJR5::77 clearly showed smaller colonies on 7H10 agar compared to the other plasmid-carrying clones (Fig. 3a). This is most likely due to the expression of minute amounts of gp77 from an incompletely repressed acetamidase promoter. This incomplete silencing of the acetamidase promoter in pSD24 was also shown in the ONPG assay as described further below.

Acetamide titration and survival curves of clones carrying toxic phage ORFs

To identify the minimum amount of acetamide necessary for the growth-inhibitory effect, individual clones were grown to an OD_{600} of 0.1 and acetamide titration was performed. Here, acetamide concentrations as low as 0.01 % were sufficient to inhibit the growth of clones containing pJR5::77, pJR5::78 and pJR5::79. The growth inhibition was directly proportional to the concentration of acetamide as determined by OD_{600} measurements after 12 h in a shaking incubator at 37 °C (Fig. 4a). The *M. smegmatis* clone carrying pJR7, which expresses LacZ under control of the acetamidase promoter, was used as a control.

To further study the effects of L5 toxic gene expression in *M. smegmatis* (bacteriostatic vs bacteriocidal) survival curves of induced (0.05 % acetamide) and uninduced clones grown in 7H9 broth were created by plating serial dilutions every 6 h for a period of 18 h. Expression of gp77 led to a bacteriostatic effect on *M. smegmatis*, with growth arrest for all time points examined and recovery of the

bacteria after plating on 7H10 agar without acetamide for enumerating c.f.u. (Fig. 4c). Expression of gp78 and gp79 allowed growth of *M. smegmatis*, although at a significantly lower rate compared to the uninduced clones, gp78 being the least potent inhibitor.

Expression of gp79 alters cell morphology of *M. smegmatis*

Growing induced *M. smegmatis* clones carrying pJR5::79 to stationary phase in liquid medium led to the development of non-homogeneous turbidity and the formation of small clumps despite the fact that Tween 80 was added to the medium (Fig. 3b). When grown on solid medium containing 0.1 % acetamide these clones had a rough surface morphology compared to that of wild-type *M. smegmatis*. Fluorescence microscopy of induced pJR5::79 clones showed filamentation, with most of the bacteria being more than three times longer than uninduced clones and *M. smegmatis* expressing LacZ, indicating impaired cell division (Fig. 3c). Interestingly, some pJR5::79-bearing cells showed a branching phenotype (Fig. 3c).

Activity of phage ORFs in *E. coli*

Since some toxic proteins of bacteriophages infecting Gram-positive bacteria show a growth-inhibitory effect on *E. coli*, we were interested in the phenotype of *E. coli* DH5α clones expressing L5 proteins encoded by ORF79 to ORF77. In contrast to pSD24 or pJR5, the expression vector pQE80 is tightly regulated, allowing normal growth of uninduced clones. Growth of clones carrying pQE80::77, pQE80::78 and pQE80::79 in the presence of IPTG was clearly inhibited (Fig. 4b, d). With respect to gp79, the effect of expression in *E. coli* was different from that observed in *M. smegmatis*. Expression of gp79 is clearly
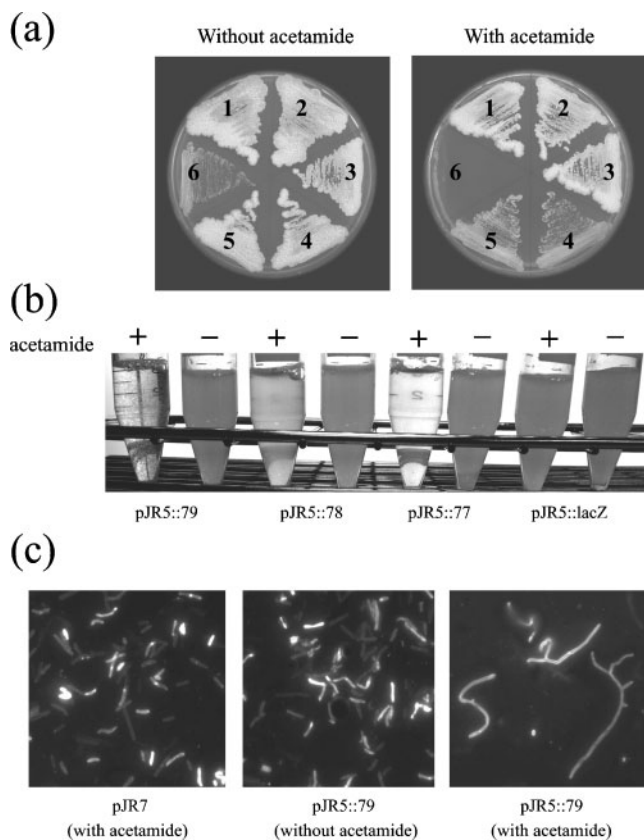
Fig. 3. (a) Recombinant plasmids carrying ORF82 to ORF77 (1–6) under control of the acetamidase promoter cloned into *M. smegmatis* mc^{2}155 and subcultured with or without the presence of inducing agent. (b) *M. smegmatis* clones grown to stationary phase in Middlebrook 7H9 medium plus Tween 80 with or without acetamide (0.05 %). Expression of gp77, gp78 and 79 is growth inhibitory; the presence of gp79 leads to clumping. (c) Fluorescence microscopy of FITC-stained pJR5 :: 79 clones grown with or without acetamide (0.05 %); a clone expressing LacZ was used as a control. Expression of gp79 leads to elongation of the individual bacteria and a branching phenotype (photographs taken at ×1000 magnification, reproduced here at ×650).

bactericidal to *E. coli*, with almost a 2 log loss of viable cells after 3 h growth in the presence of 0.1 mM IPTG (Fig. 4d). The expression of gp77 in *E. coli* led to a bacteriostatic effect whereas the expression of gp78 resulted in a significantly lower growth rate (Fig. 4b, d).

Co-expression of the L5 repressor allows transformation of ORF83 to ORF77

In contrast to pJR5 :: 8377, which does not transform competent *M. smegmatis* cells, electroporation of pJR6 :: 8377, which constitutively expresses the L5 repressor gp71, resulted in approximately 2.4×10^4 transformants per μg plasmid DNA. Interestingly, the colonies obtained

varied widely in colony size and morphology. Subcultures of these clones on media containing acetamide led to growth inhibition, showing that the extremely strong acetamidase promoter may overcome transcriptional silencing of the L5 repressor (data not shown).

ORF83 contains a putative promoter that is functional in *M. smegmatis* but not in *E. coli*

Transformability of pJR5 :: 8277 but not of pJR5 :: 8377 indicates the presence of a positive regulatory element within ORF83, which possibly induces the transcription of the three cytotoxic genes. To confirm this hypothesis two new mycobacterial promoter-probe vectors were created using pSD24 as a backbone. The insertion of *lacZ* from pMC1871 downstream of the acetamidase promoter led to pJR7. This construct allows the insertion of candidate promoter sequences between the acetamidase promoter and the *lacZ* gene using the single *Bam*HI site (Fig. 1). pJR7 grown without acetamide shows a relatively low but measurable β-galactosidase activity whereas addition of acetamide to the medium results in an extremely high activity (Fig. 5a). This is consistent with recently published data indicating that the acetamidase promoter in pSD24 is about three times stronger than the mycobacterial hsp60 promoter (Daugelat *et al.*, 2003). Induced pJR7 was used as a positive control in subsequent experiments for the evaluation of the promoter activity of L5 genes. The presence of a functional promoter inserted between the acetamidase promoter and *lacZ* gene results in an increased β-galactosidase activity compared to uninduced clones carrying pJR7.

To abolish any background activity of the acetamidase promoter the complete element was excised, resulting in pJR10 (Fig. 1). This construct had no detectable β-galactosidase activity, like wild-type *M. smegmatis* mc^{2}155.

The insertion of ORF83 together with a small part of the 5′-end of ORF82 into the linearized promoter-probe vectors (leading to a fusion protein of gp82 and LacZ) gave pJR7 :: 83 and pJR10 :: 83 (Fig. 5a). *M. smegmatis* transformants harbouring these plasmids showed an increase of β-galactosidase activity compared to clones containing pJR10 and uninduced pJR7 (Fig. 5a). These constructs mimic the situation seen in pJR5 :: 8377, with the toxic ORFs being exchanged for the *lacZ* gene. Compared to fully induced pJR7, promoter activity was about three times lower in (uninduced) pJR7 :: 83 and about four times lower in pJR10 :: 83, suggesting that the promoter within ORF83 is of moderate strength (Fig. 5a).

Uninduced pJR7 had measurable activity, indicating incomplete suppression of the acetamidase promoter in pSD24 and related constructs (Fig. 5a). This explains the slower growth of uninduced clones containing pJR5 :: 77.

Interestingly, *E. coli* DH5α containing pJR10 :: 83 (which has shuttle-vector properties) or pMC1871 :: 83, an *E. coli* promoter-probe vector carrying ORF83, did not show any
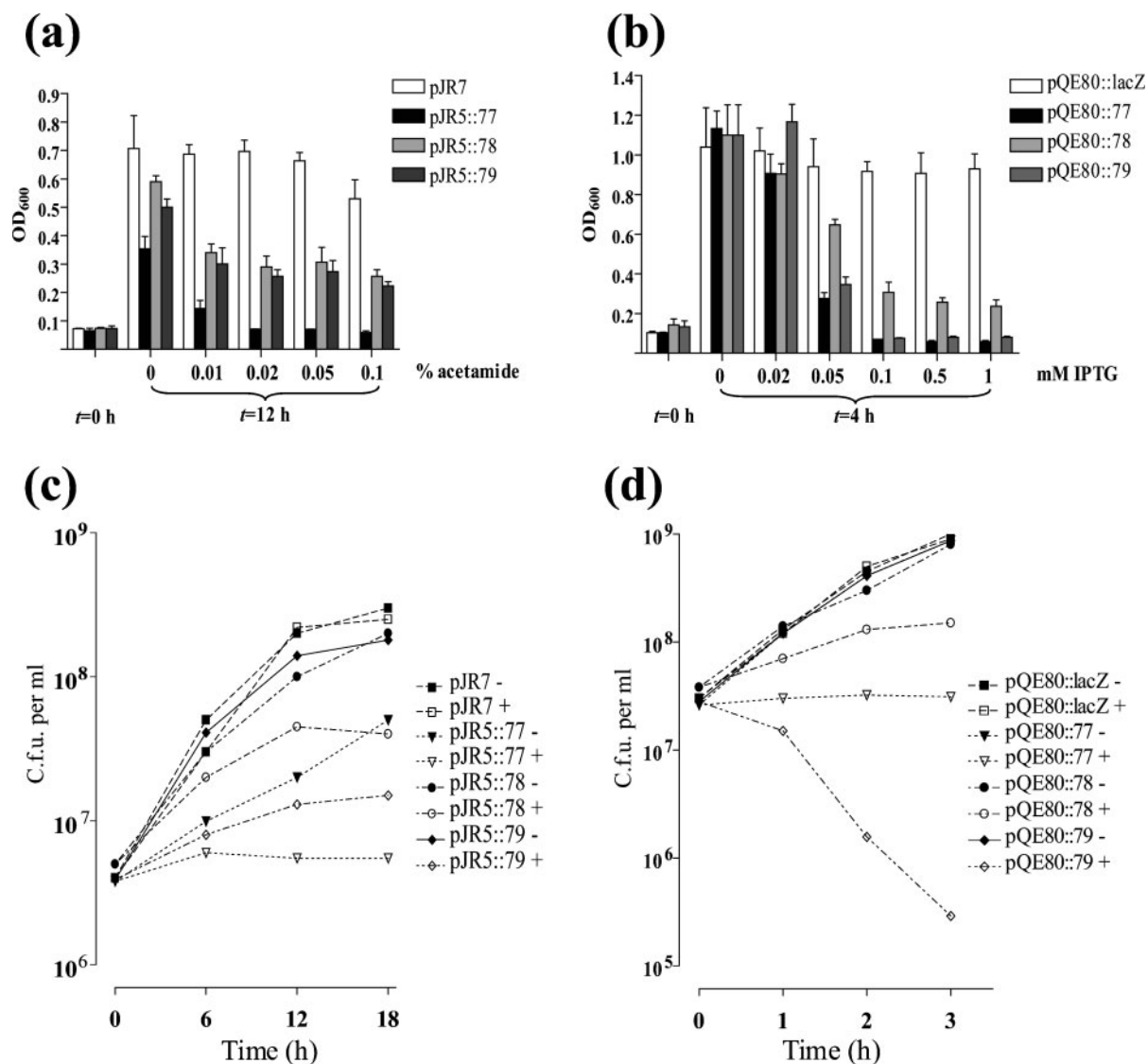
Fig. 4. Different clones expressing toxic phage proteins or LacZ as a control. (a) *M. smegmatis* clones were grown to early exponential phase and different concentrations of acetamide were added followed by measurement of the OD_{600} after 12 h. (b) *E. coli* DH5α clones were grown for 4 h in the presence of different molarities of IPTG followed by measurement of the OD_{600}. Relatively small amounts of inducing agents are needed for growth inhibition, which shows dose dependence. (c) Survival curves of *M. smegmatis* clones grown in 7H9 medium with 0.05 % acetamide (+) or without acetamide (−). (d) Survival curves of *E. coli* DH5α clones grown in the presence of 0.1 mM IPTG (+) or without IPTG (−).

β-galactosidase activity in the ONPG assay (data not shown). This indicates that promoter activity within ORF83 is restricted to mycobacteria.

Detection of the transcriptional start site within ORF83

Primer extension analysis was performed on *M. smegmatis* RNA containing either pJR10 :: 83 or pJR10 as a negative control. As shown in Fig. 5(b), a transcriptional start point was detected mapping to position 37 of ORF83. A putative core promoter sequence with −35 (GTGACC) and −10 (TCAAAT) hexamers is situated upstream of this position. The −35 site is homologous with the −35 site of promoter P4 of mycobacteriophage L1, a closely related homo-immune phage of L5 (Chattopadhyay *et al.*, 2003). The experiment was performed with three individual sets of total RNA, which showed identical results.
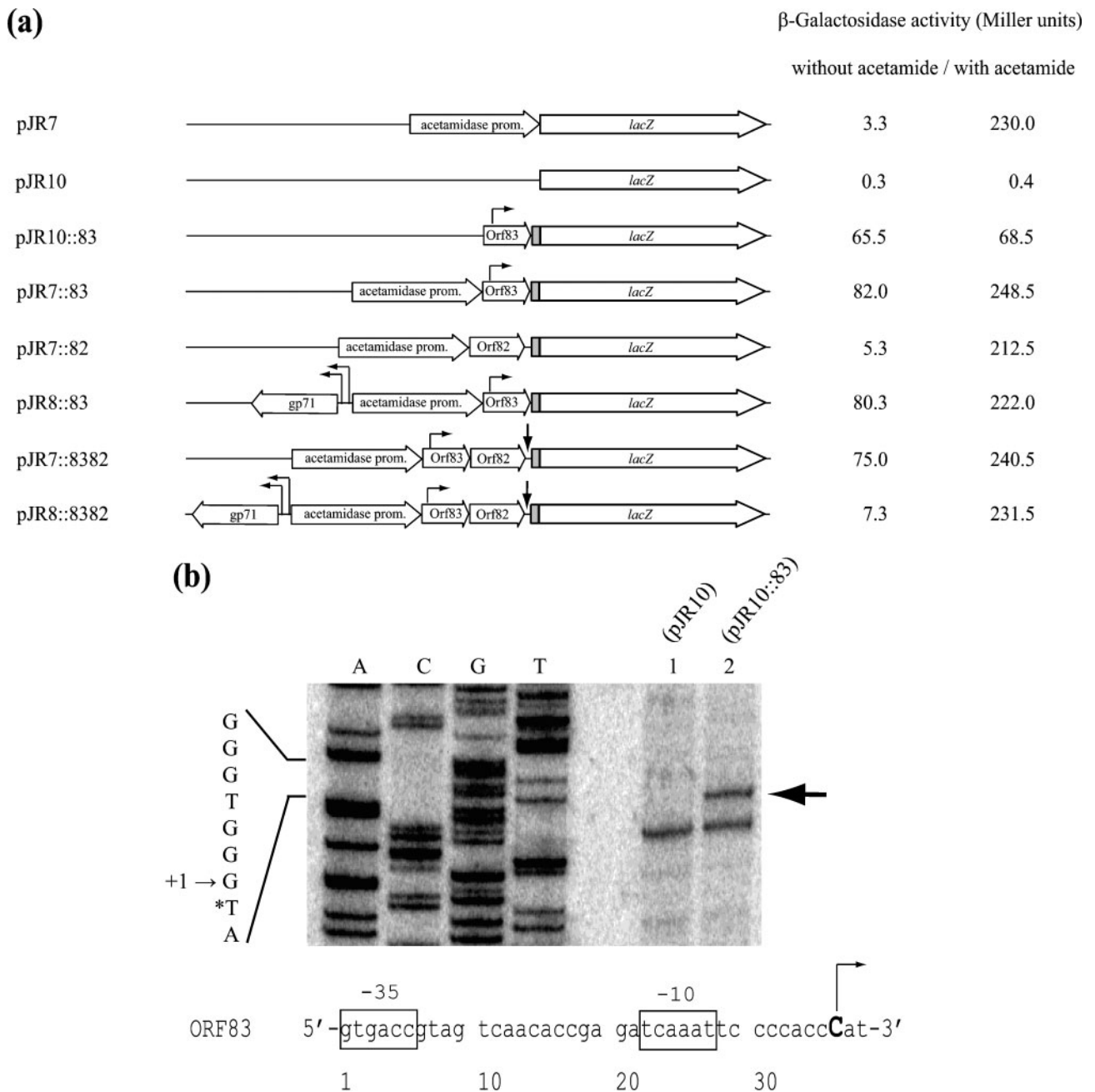
Fig. 5. (a) β-Galactosidase assay of *M. smegmatis* with different plasmid constructs carrying the *lacZ* gene. The top two lines show results for cells carrying the new promoter-probe vectors pJR7 and pJR10 grown with and without acetamide (0.1 %). Induced pJR7 functions as a positive control. Uninduced pJR7 shows some activity whereas deletion of the promoter element (pJR10) leads to no detectable β-galactosidase activity, indicating leakiness of the acetamidase promoter in pSD24. Constructs carrying ORF83 upstream of the *lacZ* gene show promoter activity within this ORF. Only the presence of both the repressor gp71 and the repressor binding site (↓) between ORF82 and ORF81 (pJR8 : : 8382) leads to a 10-fold reduction of promoter activity, which can be overcome by the strong acetamidase promoter. (b) Primer extension analysis. Lanes A, C, G, T, sequencing ladder generated with pJR10 : : 83 DNA. Lanes 1 and 2, primer extension products obtained with RNA isolated from *M. smegmatis* clones harbouring pJR10 (promoterless control) and pJR10 : : 83, respectively. A single band was detected matching position 38 (*) of the sequencing ladder and position 37 of ORF83 (one base upstream due to use of [α-^{32}P]dCTP for the sequencing ladder and the end-labelled primer for the primer extension experiment). The two lower bands observed in lanes 1 and 2 are presumably non-specific cDNA products. The lower part of the figure shows the DNA sequence of ORF83, with the transcriptional start site at position 37 and putative hexameric −35 and −10 sites indicated by boxes.

Transcriptional regulation of the putative promoter is achieved through repressor binding between ORF82 and ORF81

Lysogeny of mycobacteriophage L5 is achieved through binding of the repressor gp71 to asymmetric DNA sites with the consensus sequence GGTGG(C/A)TGTCAAG, which are situated either within some promoter sites or downstream of promoters (Brown *et al.*, 1997). The binding sites within promoters represent true operators that negatively regulate transcription initiation whereas the downstream ones act as stoperators inhibiting transcription elongation, leading to a downregulation of gene expression (Brown *et al.*, 1997). One such stoperator is situated downstream of ORF83 between the stop codon of ORF82 and the start codon of ORF81 (Fig. 2). To show that gp71 and its binding site have a regulatory influence on the promoter element in ORF83, the repressor gene 71 was cloned into the promoter-probe plasmid pJR7, leading to pJR8 (Fig. 1). Insertion of ORF83 upstream of the *lacZ* gene of pJR8 led to similar *β*-galactosidase activity as seen for pJR7::83 in the ONPG assay whereas integration of ORF83 together with the intergenic region of 82 and 81 (pJR8::8382) led to a 10-fold reduction of *β*-galactosidase activity (Fig. 5a). Addition of acetamide to the medium abolished gp71 activity. A single stoperator site downstream of the extremely strong acetamidase promoter seems to have only a slight negative regulatory effect on promoter activity and thus may not be detectable in this assay.

DISCUSSION

In this study we have shown that the temperate mycobacteriophage L5 produces toxic proteins leading to growth inhibition of the natural host *M. smegmatis*. The ORFs encoding the toxic gene products are located in the right arm of the phage genome, which primarily codes for regulatory proteins such as DNA helicase, primase and the repressor (gp71). The finding that a temperate phage expresses early toxic proteins is remarkable and difficult to interpret, especially since the closest relative of L5, mycobacteriophage D29, a lytic phage with high sequence similarity and a shared host range, lacks these early genes and the repressor gene. The fact that the promoter region upstream of the toxic proteins is only of moderate strength does not necessarily mean weak transcription and expression of downstream ORFs since there is at least one additional promoter (P_left) upstream of ORF83 (Nesbit *et al.*, 1995). Thus tight regulation of these early gene products is a prerequisite to achieve successful L5 lysogeny. Our co-expression results are in favour of the explanation that transcriptional silencing of the downstream genes during lysogeny is exerted by the phage repressor gp71. These results emphasize the major role of gp71 for globally silencing prophage gene expression, a crucial mechanism for a temperate virus. However, co-expression of gp71 only led to suppression of promoter activity to values just above

the background activity of the leaky acetamidase promoter but not to complete silencing of expression. Also, *M. smegmatis* clones carrying pJR6::8377, a plasmid containing the repressor gene, the putative promoter sequence in ORF83, the repressor-binding site and the downstream toxic ORFs, showed variable colony size and morphology. This may reflect unstable expression or a variable half-life of the repressor protein, which could be one explanation for the high rate of L5 infections which lead to lytic growth (80 %) rather than lysogeny (Sarkis *et al.*, 1995). However, it is not clear how the promoter within ORF83 is regulated in true L5 lysogeny. In contrast to P_left, a promoter upstream of ORF88 that has a gp71-binding site reaching into its -35 sequence controlling the initiation of transcription, there are no such sites in ORF83 (Brown *et al.*, 1997). There may be other, so far undescribed regulatory genes in L5, not present in our plasmid construct, which control transcription during lysogeny.

Although the amino acid sequences of gp77, gp78 and gp79 do not show any sequence homologies with gene products of phages infecting *E. coli*, we showed that these L5 proteins are inhibitory to the growth of this Gram-negative bacterium. This is not unprecedented since other phage-derived shut-off genes such as e3 of phage SPO1, a bacteriophage infecting the Gram-positive *Bacillus subtilis*, also lead to growth inhibition when expressed in *E. coli* (Wei & Stewart, 1993).

The distinct interaction of the three proteins with the host remains unclear. The predicted 28.2 kDa protein encoded by ORF77, which possesses the strongest activity in *M. smegmatis*, does not contain an excess of positively charged amino acids and we were unable to identify any DNA-binding motifs such as helix–turn–helix structures enabling DNA-binding activity. Membrane association of gp77 is unlikely since there is a lack of hydrophobic regions, signal peptides or predicted transmembrane helices. Using p-BLAST to compare the amino acid sequence of gp77 with the available database revealed sequence homology not only to other mycobacteriophages but also to hypothetical proteins of actinomycetes such as *Frankia* spp., *Streptomyces* spp. and *Mycobacterium* spp. (Table 2). This indicates that L5 or an L5-related phage may have acquired gp77 from a former host by illegitimate recombination. The toxic protein may have been retained, conferring a selective advantage upon the host.

The N-terminal segment of gp79 (amino acids 1–41) shares sequence similarity with the signal peptide of the D-alanyl-D-alanine carboxypeptidase of *Bacillus licheniformis*. This enzyme removes C-terminal D-alanyl residues from sugar-peptide cell-wall precursors and is also a penicillin-binding protein (PBP). It is synthesized with a hydrophobic signal peptide to target it to the cell wall. In most bacterial species, including mycobacteria, PBPs are essential for the biogenesis of the peptidoglycan layer of the cell wall. Knockout of PBP-A in *M. smegmatis* results in hindered growth and defective cell septation and division (Dasgupta

Table 2. Molecular mass of gp77, gp78 and gp79, and p-BLAST results

	L5 gp77	L5 gp78	L5 gp79
Molecular mass	28.2 kDa	5.5 kDa	7.8 kDa
p-BLAST mycobacteriophage proteins (% identities)	gp86 Mycophage Che12 (70%) gp79 Mycophage Bxz2 (67%)	gp95 Mycophage Llij (75%) gp99 Mycophage PMC (78%) gp107 Mycophage Che8 (75%)	gp84 Mycophage Bxb1 (76%) gp235 Mycophage Omega (74%) gp86 Mycophage Bethlehem (66%) gp88 Mycophage Che12 (58%)
p-BLAST non-phage proteins (% identities)	Hypothetical protein, *Frankia* sp. CcI3 (29%) Hypothetical protein *Streptomyces kanamyceticus* (25%) Hypothetical protein *Mycobacterium* sp. JLS (24%)		Carboxypeptidase, amino acids 1–42, *Bacillus licheniformis* (41%)

et al., 2006). PBPs are highly conserved across slow- and fast-growing species and interact with several other cell-wall-associated proteins such as FtsZ and FtsW. These proteins form a structure called Z-ring, a cytoskeletal scaffold mediating cell division (Datta *et al.*, 2006). Alteration of Z-ring proteins also leads to growth inhibition and impaired cell division in both *E. coli* and *M. smegmatis* (Belhumeur & Drapeau, 1984; Dziadek *et al.*, 2003). Furthermore, simultaneous deletion of several low-molecular-mass PBPs of *E. coli* leads to elongation and branching of individual mutants (Nelson & Young, 2000). This study presents evidence that gp79 interferes with the cell membrane or cell-wall synthesis of *M. smegmatis*. First, colonies of clones carrying pJR5::79 grown for several days on 7H10 agar containing 0.1% acetamide have a rough surface compared to the smooth phenotype of wild-type *M. smegmatis* and clones expressing other L5 proteins. Second, the same clones grown in 7H9 broth containing acetamide form relatively large clumps even in the presence of Tween 80, again in contrast to wild-type *M. smegmatis* (Fig. 3b). Third, microscopy of induced clones shows elongated rods compared to the uninduced control, and in some bacterial cells branching can be observed, indicating that gp79 interferes with cell septation (Fig. 3c).

Expression of only the N-terminal hydrophobic part (pJR5::79A, Fig. 2) or the C-terminal part (pJR5::79B, Fig. 2) of gp79 eliminates the toxic effect (data not shown). This indicates that the complete polypeptide is necessary for proper function despite the fact that there is a predicted cleavage site between the alanine at position 34 and the aspartic acid at position 35.

In addition to the fact that PBP knockout mutants in several bacterial species show a similar phenotype, the homology of the hydrophobic N-terminal part of gp79 to a PBP signal peptide may indicate an interaction of gp79 with proteins or metabolites involved in the peptidoglycan synthesis of *M. smegmatis*.

The synthesis of the outer cell envelope lipids seems not to be affected by gp79 since their analysis by TLC revealed no difference between induced and uninduced *M. smegmatis* clones (data not shown).

The reason for L5 interfering with the cell wall of its host in the early stage of phage propagation is not clear. The genes responsible for lysis of the cell wall after completion of the lytic cycle are situated in the far left arm of the phage genome.

The smallest and least active of the toxic proteins described is gp78. This polypeptide, with a predicted molecular mass of 5.5 kDa, has detectable sequence homology only to other mycobacteriophage-derived proteins (Table 2), leaving little option for the interpretation of its function.

The data presented here address the intriguing question of the mechanism by which mycobacteriophages 'take over' the host machinery during early lytic growth to replicate to higher levels or to evade host defence. With gp77 and gp79, neighbouring L5 proteins have been described that interfere with the host biogenesis, most likely through two completely different mechanisms.

A potential application of the inducible expression of toxic phage proteins could be as counter-selectable markers for the genetic manipulation of mycobacteria. Furthermore, we note that gp71 expression and binding features combined on a single plasmid provide powerful and convenient applications for the controlled expression of foreign genes in mycobacteria.

In view of the large number of mycobacteriophages in the ecosystem and their extraordinarily high diversity, the approach described in this paper may also allow the further identification of proteins targeting important metabolic pathways of mycobacteria. Target identification is a major step in the drug-discovery process (Terstappen & Reggiani, 2001). With regard to the pathogenic mycobacteria such as *Mycobacterium tuberculosis*, determination of the interacting partner proteins on the host's side may lead to the identification and exploitation of as yet unknown key proteins of mycobacterial metabolism and biogenesis.

ACKNOWLEDGEMENTS

REFERENCES

Belhumeur, P. & Drapeau, G. R. (1984). Regulation of cell division in *Escherichia coli*: properties of new *ftsZ* mutants. *Mol Gen Genet* **197**, 254–260.

Brown, K. L., Sarkis, G. J., Wadsworth, C. & Hatfull, G. F. (1997). Transcriptional silencing by the mycobacteriophage L5 repressor. *EMBO J* **16**, 5914–5921.

Chattopadhyay, C., Sau, S. & Mandal, N. C. (2003). Cloning and characterization of the promoters of temperate mycobacteriophage L1. *J Biochem Mol Biol* **36**, 586–592.

Dasgupta, A., Datta, P., Kundu, M. & Basu, J. (2006). The serine/threonine kinase PknB of *Mycobacterium tuberculosis* phosphorylates PBPA, a penicillin-binding protein required for cell division. *Microbiology* **152**, 493–504.

Datta, P., Dasgupta, A., Singh, A. K., Mukherjee, P., Kundu, M. & Basu, J. (2006). Interaction between FtsW and penicillin-binding protein 3 (PBP3) directs PBP3 to mid-cell, controls cell septation and mediates the formation of a trimeric complex involving FtsZ, FtsW and PBP3 in mycobacteria. *Mol Microbiol* **62**, 1655–1673.

Daugelat, S., Kowall, J., Mattow, J., Bumann, D., Winter, R., Hurwitz, R. & Kaufmann, S. H. (2003). The RD1 proteins of *Mycobacterium tuberculosis*: expression in *Mycobacterium smegmatis* and biochemical characterization. *Microbes Infect* **5**, 1082–1095.

Donnelly-Wu, M. K., Jacobs, W. R., Jr & Hatfull, G. F. (1993). Superinfection immunity of mycobacteriophage L5: applications for genetic transformation of mycobacteria. *Mol Microbiol* **7**, 407–417.

Dziadek, J., Rutherford, S. A., Madiraju, M. V., Atkinson, M. A. & Rajagopalan, M. (2003). Conditional expression of *Mycobacterium smegmatis ftsZ*, an essential cell division gene. *Microbiology* **149**, 1593–1603.

Fenwick, M. L. & Clark, J. (1982). Early and delayed shut-off of host protein synthesis in cells infected with herpes simplex virus. *J Gen Virol* **61**, 121–125.

Hatfull, G. F. & Sarkis, G. J. (1993). DNA sequence, structure and gene expression of mycobacteriophage L5: a phage system for mycobacterial genetics. *Mol Microbiol* **7**, 395–405.

Hatfull, G. F., Pedulla, M. L., Jacobs-Sera, D., Cichon, P. M., Foley, A., Ford, M. E., Gonda, R. M., Houtz, J. M., Hryckowian, A. J. & other authors (2006). Exploring the mycobacteriophage metaproteome: phage genomics as an educational platform. *PLoS Genet* **2**, e92.

Liu, J., Dehbi, M., Moeck, G., Arhin, F., Bauda, P., Bergeron, D., Callejo, M., Ferretti, V., Ha, N. & other authors (2004). Antimicrobial drug discovery through bacteriophage genomics. *Nat Biotechnol* **22**, 185–191.

Nelson, D. E. & Young, K. D. (2000). Penicillin binding protein 5 affects cell diameter, contour, and morphology of *Escherichia coli*. *J Bacteriol* **182**, 1714–1721.

Nesbit, C. E., Levin, M. E., Donnelly-Wu, M. K. & Hatfull, G. F. (1995). Transcriptional regulation of repressor synthesis in mycobacteriophage L5. *Mol Microbiol* **17**, 1045–1056.

Pedulla, M. L., Ford, M. E., Houtz, J. M., Karthikeyan, T., Wadsworth, C., Lewis, J. A., Jacobs-Sera, D., Falbo, J., Gross, J. & other authors (2003). Origins of highly mosaic mycobacteriophage genomes. *Cell* **113**, 171–182.

Robinson, N., Wolke, M., Ernestus, K. & Plum, G. (2007). A mycobacterial gene involved in synthesis of an outer cell envelope lipid is a key factor in prevention of phagosome maturation. *Infect Immun* **75**, 581–591.

Rybniker, J., Kramme, S. & Small, P. L. (2006). Host range of 14 mycobacteriophages in *Mycobacterium ulcerans* and seven other mycobacteria including *Mycobacterium tuberculosis* – application for identification and susceptibility testing. *J Med Microbiol* **55**, 37–42.

Sambrook, J. & Russell, D. W. (2001). *Molecular Cloning: a Laboratory Manual*, 3rd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Sarkis, G. J., Jacobs, W. R., Jr & Hatfull, G. F. (1995). L5 luciferase reporter mycobacteriophages: a sensitive tool for the detection and assay of live mycobacteria. *Mol Microbiol* **15**, 1055–1067.

Shapira, S. K., Chou, J., Richaud, F. V. & Casadaban, M. J. (1983). New versatile plasmid vectors for expression of hybrid proteins coded by a cloned gene fused to *lacZ* gene sequences encoding an enzymatically active carboxy-terminal portion of beta-galactosidase. *Gene* **25**, 71–82.

Sharma, R., Raychaudhuri, S. & Dasgupta, A. (2004). Nuclear entry of poliovirus protease-polymerase precursor 3CD: implications for host cell transcription shut-off. *Virology* **320**, 195–205.

Snapper, S. B., Melton, R. E., Mustafa, S., Kieser, T. & Jacobs, W. R., Jr (1990). Isolation and characterization of efficient plasmid transformation mutants of *Mycobacterium smegmatis*. *Mol Microbiol* **4**, 1911–1919.

Stinear, T. P., Seemann, T., Pidot, S., Frigui, W., Reysset, G., Garnier, T., Meurice, G., Simon, D., Bouchier, C. & other authors (2007). Reductive evolution and niche adaptation inferred from the genome of *Mycobacterium ulcerans*, the causative agent of Buruli ulcer. *Genome Res* **17**, 192–200.

Svenson, S. B. & Karlstrom, O. H. (1976). Bacteriophage T4-induced shut-off of host-specific translation. *J Virol* **17**, 326–334.

Terstappen, G. C. & Reggiani, A. (2001). *In silico* research in drug discovery. *Trends Pharmacol Sci* **22**, 23–26.

Wei, P. & Stewart, C. R. (1993). A cytotoxic early gene of *Bacillus subtilis* bacteriophage SPO1. *J Bacteriol* **175**, 7887–7900.

Zhang, X. & Bremer, H. (1995). Control of the *Escherichia coli rrnB* P1 promoter strength by ppGpp. *J Biol Chem* **270**, 11181–11189.

Edited by: P. R. Herron

Research Paper

The cytotoxic early protein 77 of mycobacteriophage L5 interacts with MSMEG_3532, an L-serine dehydratase of *Mycobacterium smegmatis*

Jan Rybniker[1], Karin Krumbach[3], Edeltraud van Gumpel[1], Georg Plum[2], Lothar Eggeling[3] and Pia Hartmann[1]

[1] 1st Department of Internal Medicine, University of Cologne, Cologne, Germany
[2] Institute for Medical Microbiology, Immunology and Hygiene, University of Cologne, Cologne, Germany
[3] Institut für Biotechnologie 1, Forschungszentrum Jülich GmbH, Jülich, Germany

Mycobacteriophage L5 is a temperate phage infecting a broad range of mycobacterial species. Upon induction of lytic growth, L5 rapidly switches off host protein synthesis. We have recently identified the mycobacteriophage L5 early protein gp77 as a host shut-off protein that acts growth inhibitory in the mycobacterial host when expressed through the corresponding phage promoter. Here we present data showing that this purified phage protein of unknown function specifically binds to protein MSMEG_3532 when incubated with cell lysates of *Mycobacterium smegmatis*. This interaction was confirmed by pull-down assays using purified MSMEG_3532 as bait which co-purified with gp77. The amino acid sequence of MSMEG_3532 is nearly identical to that of threonine dehydratases, serine dehydratases and an L-threo-3-hydroxyaspartate dehydratase. An enzymatic assay identified this host protein as a pyridoxal-5′-phosphate-dependent L-serine dehydratase (SdhA) which converts L-serine to pyruvate. This is the first biochemical characterization of a SdhA derived from mycobacteria. Though the addition of purified gp77 to the established *in vitro* assay had no influence on SdhA activity at a saturating L-serine concentration, the specific interaction of phage protein and dehydratase *in vivo* may well have a role in altering the amino acid pool or the products of amino acid metabolism in favour of phage maturation.

Introduction

The systemic isolation and characterization of mycobacteriophages has led to the development of many genetic tools for the dissection and understanding of their mycobacterial hosts [1, 2]. The complete sequencing and annotation of more than 50 mycobacteriophages now allows the further investigation and exploitation of these viruses [3–6]. So far, the functions of the vast majority of proteins encoded by the numerous annotated open reading frames (ORFs) identified re-main unknown since only a few proteins have been expressed and examined in detail. Recently, we have focussed on a set of early proteins of mycobacteriophage L5 (gp77, 78 and 79) which are cytotoxic for the host *Mycobacterium smegmatis* upon expression from an inducible mycobacterial shuttle vector and from a newly identified L5 promoter. It is most likely that these proteins function as shut-off proteins during the early stages of phage propagation [7]. Mycobacteriophage L5 is a temperate phage with a broad host range including pathogens of clinical relevance such as *Mycobacterium tuberculosis, Mycobacterium avium* and *Mycobacterium ulcerans* [8]. In our attempt to identify the functions of L5 toxic proteins we focussed on gp77. Intracellular minute amounts of this protein act bacteriostatically on the natural host *M. smegmatis*. The amino

Correspondence: Jan Rybniker, 1st Department of Internal Medicine, Division of Infectious Diseases, University Hospital of Cologne, 50924 Cologne, Germany
E-mail: jan.rybniker@uk-koeln.de
Phone: ++49 221 478 88835
Fax: ++49 221 478 3470

acid sequence of gp77 shows homology with hypothetical proteins of actinomycetes suggesting that this protein was acquired by the phage from a former host. Sequence analysis of gp77 identified homology to the conserved domain DUF2786. This domain which is approximately 40 amino acids in length is found in a family of bacterial proteins with unknown function [9].

To further characterize gp77, we have tried to identify interacting host proteins by incubating affinity purified gp77 with soluble proteins of *M. smegmatis*. Here we show that gp77 specifically interacts with the host protein MSMEG_3532 *in vitro*. We identified MSMEG_3532 as a pyridoxal-5'-phosphate-dependent L-serine dehydratase (SdhA) by an enzymatic assay. We also analyzed the Michaelis-Menten kinetics of this enzyme and show that the protein forms a homodimer in its active state. To our knowledge this is the first identification and characterization of a serine-dehydratase derived from the genus *Mycobacterium*. In *Escherichia coli* L-serine dehydratase deficiency has recently been shown to impair growth through the inhibition of cell division [10]. Though the addition of gp77 to the enzymatic assay did not affect the activity of the amino acid dehydratase *in vitro*, the specific interaction of phage protein and dehydratase may play a role in providing amino acids or products of the amino acid metabolism required for phage maturation.

Materials and methods

Bacterial strains, bacteriophages and plamids
Details of the bacterial strains, phage and plasmids used in this study are given in Table 1.

DNA manipulation and cloning
DNA manipulation, electrophoresis and PCRs were carried out using standard techniques [11]. All PCRs for subsequent cloning were performed using phusion high fidelity polymerase (Finnzymes, Espoo, Finland) according to the instructions provided by the manufacturer.

All cloning steps were done by homologous recombination using the In-Fusion PCR cloning kit (Clontech, Mountain View, CA). Cloning of pQE80::77 and pQE80::79 was performed as described recently [7]. Primers for amplifying gene 77 for subsequent cloning in pET29b were 5'-GGAGATATACATATGATCGACGGCAAGACG-3' and 5'-GTGCGGCCGCAAGCTTCAGCCCGCCAGCGC-3'. Primers amplifying MSMEG_3532 were 5'-TCACCATCACGGATCCGACCTGGTGACCCT-3' and 5'-TCAGCTAATTAAGCTTCTAACCGCCCGCGA-3' for pQE80::sdhA and 5'-GGAGATATACATATGGACCTGGTGACCCTC-3' and 5'-GTGCGGCCGCAAGCTTTCTAACCGCCCGCG-3' for pET29b::sdhA. Genomic DNA of *M. smegmatis* mc²155 was used as a template for PCR. pET29b was linearized with *Nde*I and *Hind*III, primers were designed in such a way that there was no fusion with the C-terminal His-tag.

Pull-down assays
For protein pull downs the Ni-NTA magnetic agarose bead kit (Qiagen) was used according to the instructions provided by the manufacturer. In brief $6 \times$His tagged gp77 was expressed in *E. coli* BL21(DE3). The strain was grown in LB to an OD_{600} of 0.6, IPTG was added to a final concentration of 0.5 mM and after incubation for 2 h at 37 °C 50 ml cells were harvested by centrifugation. The pellet was washed once with cold PBS and resuspended in 1 ml cold lysis buffer (50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidazole, 0.05% Tween 20, pH 8.0, $1 \times$EDTA free protease inhibitor cocktail (Roche)). 250 µl of 0.1 mm glass beads were added and the cell-suspension was disrupted by using a bead mill (MM200, Retsch) for a total of 5 min at maximum speed. After centrifugation for 10 min at 4 °C and 15,000 g the supernatant was brought to a fresh tube and 50 µl magnatic

Table 1. Strains, bacteriophage and Plasmids used in the study.

Strains, bacteriophage, plasmids	Description	Reference or source
Escherichia coli		
Fusion blue	*end*A1, *hsd*R17 (r_{K12-}, m_{K12+}), *sup*E44, *thi*-1, *rec*A1, *gyr*A96, *rel*A1, lac F'[*proA*⁺B⁺, *lacI*qZΔM15::Tn10(tet^R)]	Clontech
BL21(DE3)	F– *ompT hsd*SB(rB–, mB–) *gal dcm* (DE3)	Lab collection
M. smegmatis mc²155	High efficiency transformation mutant	[28]
Bacteriophage		
Mycobacteriophage L5	Temperate phage of *M. smegmatis*	Lab collection
Plasmids		
pQE80L	Expression vector carrying the *cis-laqIq* repressor gene and an N-terminal His-tag sequence	Qiagen
pET29b	Expression vector carrying an N-terminal S-tag/thrombin configuration plus an optional C-terminal His-tag sequence	Novagen

beads were added. After incubation for 30 min at 4 °C on a rotary shaker the beads were washed three times with wash buffer (lysis buffer + imidazole, 20 mM final concentration) and 500 µl cell lysate of *M. smegmatis* or *E. coli* BL21(DE3) expressing SdhA from pET29b were added. After incubation for 30 min beads were washed again and eluted with 50 µl elution buffer (lysis buffer + imidazole, 250 mM final concentration). Eluates were examined by SDS-PAGE gel electrophoresis and subsequent staining using the colloidal blue staining kit (Invitrogen).

M. smegmatis cell lysates were produced by growing cells to late log phase in Middlebrook 7H9 medium enriched with 10% ADC and 0.05% Tween 80. 50 ml of these cells were harvested by centrifugation. The pellet was washed once in PBS and resuspended in 1 ml cold lysis buffer. Cells were again disrupted with 0.1 mm glass beads using the bead mill for 10 min at maximum speeds. All steps were performed on ice, the grinding jars were kept at −20 °C before use. After disruption the cell walls, debris and unsoluble matter were pelleted by centrifugation for 10 min at 20,000 g and the supernatants containing soluble proteins were stored at −20 °C until use.

Peptide mass fingerprint

Peptide mass fingerprint (PMF) was performed as described recently [12]. In brief, proteins were pre-treated and trypsin digested at 37 °C over night. The digest was stopped by the addition of 5–20 µl 1% trifluoracetic acid (TFA) in water and peptides were extracted for 30 min at 37 °C. Positive ion spectra were acquired on a Reflex IV MALDI-TOF mass spectrometer (Bruker Daltonics, Bremen, Germany) in the reflectron mode. A peptide calibration standard (Bruker Daltonics, Bremen, Germany) was used for external calibration of the mass range from m/z 1046 to m/z 3147. The FlexAnalysis postanalysis software was used for optional internal recalibration on trypsin autolysis peaks and the generation of peaklists. Biotools 3·0 (Bruker Daltonics, Bremen, Germany) was used for interpretation of mass spectra.

Serine dehydratase isolation

E. coli BL21 (DE3) containing plasmid pQE80L::sdhA was grown in 50 ml LB at 37 °C and at an OD$_{600}$ of 0.5 IPTG was added to a final concentration of 0.5 mM. After two hours cells were harvested, washed once with 0.9% NaCl and aliquots frozen at −20 °C in 1.5 ml Eppendorf reaction tubes. For protein isolation the thawed cells were disrupted by sonication and centrifuged for 10 min at 13,000 rpm. The resulting clear supernatant

was applied to a 1.5 ml Ni-NTA column and protein eluted according to the manufacturer (Quiagen, Hilden, Germany). In first trial experiments for activity determination protein in the elution buffer containing 250 mM imidazole was used. However, this was found to result in rapid inactivation of the enzyme. Therefore, in subsequent experiments the enzyme isolated via Ni-NTA chromatography was immediately transferred into 0.1 mM potassium phosphate buffer pH 8.2, 0.1 mM pyridoxal-5′-phosphate, 10% glycerol using PD10 columns.

Gelfiltration

The column used was a HiLoad 16/60 Superdex 200 prep grade (GE Healthcare) operated at a flow rate of 1 ml min^{-1} with 0.1 M KPP pH 8.2, 0.1 mM pyridoxal-5′-phosphate, 10% glycerol. The column was calibrated with aprotinin (6.6 kDa), cytochrome C (12.4 kDa), carboanhydrase (29 kDa), and albumin (66 kDa).

Dehydratase enzyme assay

The assay was performed in 0.1 mM potassium phosphate buffer pH 8.2 containing 0.1 mM pyridoxal-5′-phosphate and 3–60 µg protein in a final volume of 1 ml. For the activity screening L-threonine, D-threonine, L-serine, together with D-serine were used in a concentration of 40 mM each. At defined time points 100 µl-samples were removed and transferred to tubes containing 10 µl 36% trichloroacetic acid to stop the reaction. Subsequently, 110 µl derivatization reagent (4 mg 1.2-diamino-4.5-dimethoxybenzol, 3.5 ml mercaptoethanol, 4.1 ml 37% HCl ad H$_2$O to a total volume of 50 ml [13]) was added and the samples were incubated for 2 h at 105 °C. Finally the products were analysed by HPLC (Agilent series 1100) using a gradient of 20 mM KH$_2$PO$_4$ at pH 3.2 with increasing methanol concentrations (fluorometric detection settings were: excitation 243 nm, emission 436 nm).

Results

gp77 interacts with MSMEG_3532 a host protein of *M. smegmatis*

To identify host proteins that may interact with gp77 of mycobacteriophage L5 we purified 6×His-tagged gp77 on Ni-NTA magnetic beads. Loaded beads were incubated with soluble protein lysates of *M. smegmatis*, washed extensively and eluted in buffer containing a high concentration of imidazole. Eluates were examined by SDS-PAGE gel electrophoresis and bands present only in the gp77 (bait) plus *M. smegmatis* lysate
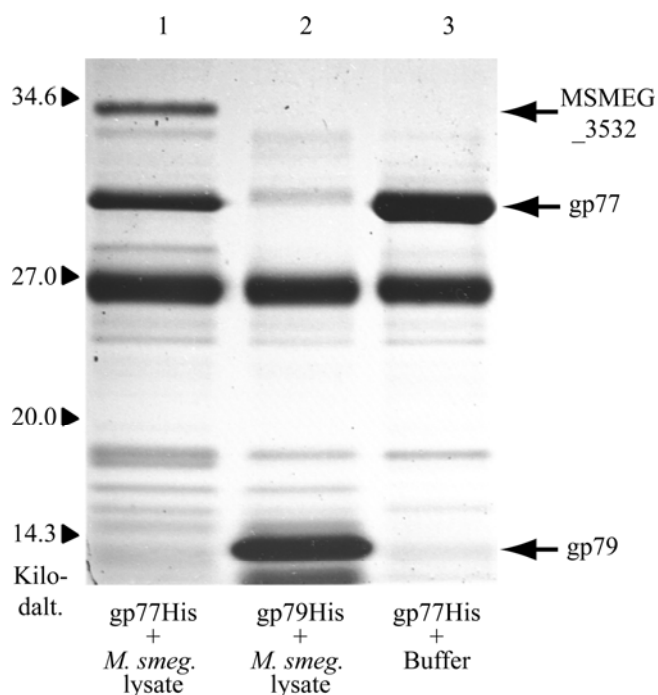
Figure 1. SDS-PAGE of eluates from pull down assay of purified gp77 incubated with *M. smegmatis* lysates. gp77 (lane 1) and gp79 (control; lane 2) of mycobacteriophage L5 were bound to Ni-NTA magnetic beads and incubated with *M. smegmatis* lysates for 30 min at 4 °C. Beads were washed and bound proteins eluted. Lane 3: gp77 bound to beads and incubated with lysis buffer instead of *M. smegmatis* lysate (second control). MSMEG_3532 co-purifies with gp77. The additional single band at approximately 18 kD in lane 1 is an unspecific binding product derived from *E. coli* (as determined by PMF).

(prey) lane were excised. In three independent experiments a specific mycobacterial protein co-purified with gp77. This protein was not present in the controls (Fig. 1). PMF analysis of the excised band identified the protein as MSMEG_3532. To verify the interaction MSMEG_3532 was expressed in *E. coli* BL21(DE3) and *E. coli* lysates were incubated with gp77 coupled on magnetic beads. Again MSMEG_3532 co-purified with the phage protein (Fig. 2). As a final proof the experiment was also performed with 6×His-tagged MSMEG_3532 bound to Ni-NTA magnetic beads as bait and untagged gp77 as prey. Here gp77 co-purified with MSMEG_3532 (Fig. 2).

Identification and characterisation of MSMEG_3532

The protein encoded by MSMEG_3532 exhibits high identities to pyridoxal-5′-phosphate-dependent threonine dehydratases, serine dehydratases and an L-threo-3-hydroxyaspartate dehydratase. These lyases are specific for their substrate as well as to the enantiomer they attack. In order to identify the function of MSMEG_3532, His-tagged protein was isolated and an enzymatic assay performed using a mixture of D-serine, L-serine, D-threonine and L-threonine as described in Material and Methods. After 20 min reaction time, 2-oxoacids were derivatized with 1.2-diamino-4.5-di-methoxybenzol and products analysed by HPLC together with pyruvate and 2-oxobutyrate as standard [13]. Only pyruvate was detectable, whereas no 2-oxo-
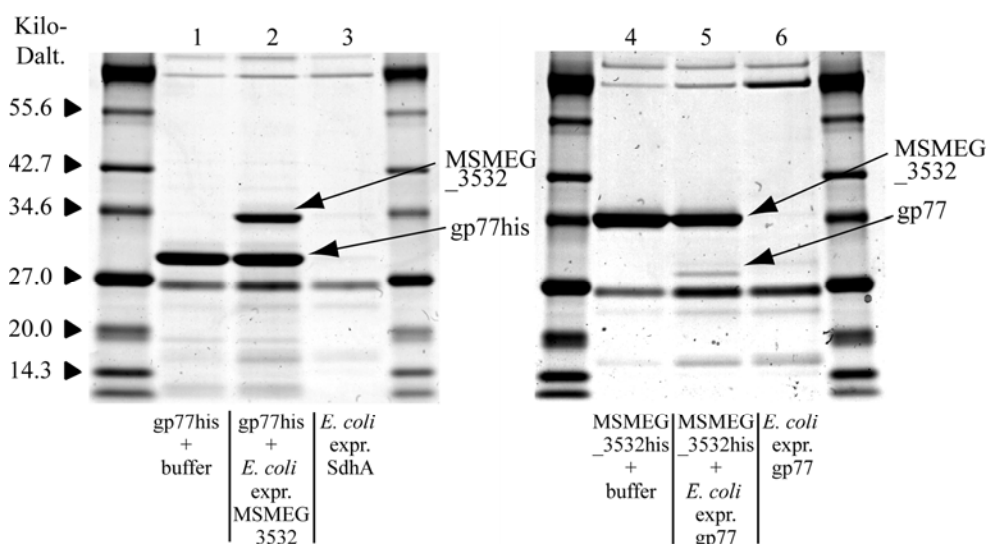

Figure 2. SDS-PAGE of pull down assays for confirmation. All proteins were expressed in *E. coli* BL21. gp77his was bound to magnetic beads (lanes 1 and 2) and incubated with lysate of *E. coli* expressing MSMEG_3532 from pET29b (lane 2), in lane 3 this lysate was incubated with beads only (control). MSMEG_3532 expressed in *E. coli* binds to and co-purifies with gp77. MSMEG_3532-his was bound to magnetic beads (lanes 4 and 5) and incubated with lysate of *E. coli* expressing gp77 from pET29b (lane 5), in lane 6 this lysate was incubated with beads only (control). gp77 binds to and co-purifies with MSMEG_3532.
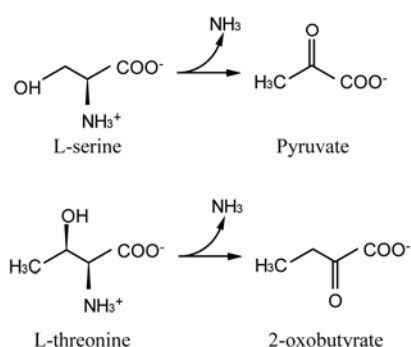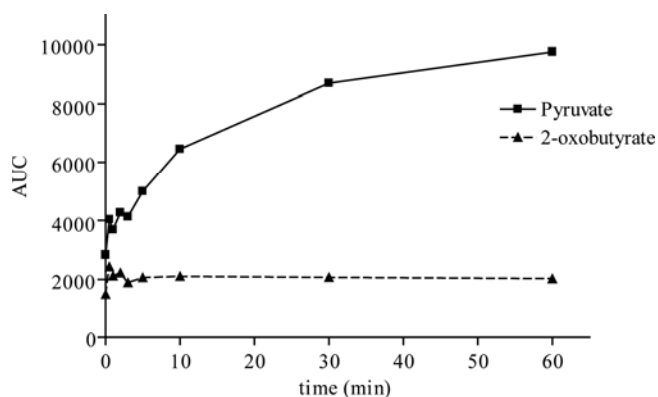
Figure 3. Enzyme assay performed using a mixture of D-serine, L-serine, D-threonine and L-threonine added to recombinant SdhA of *M. smegmatis.* In this assay pyruvate but no 2-oxobutyrate could be detected making SdhA a serine-dehydratase. AUC: area under the curve. Amounts of pyruvate and 2-oxobutyrate are shown as area under the concentration curve as determined by the HPLC software.
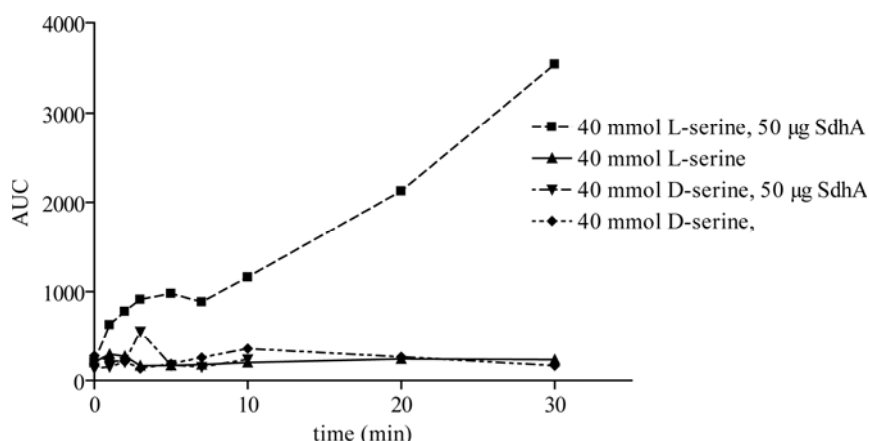
butyrate was formed (Fig. 3). Subsequent assays analyzing the time-dependent conversion of L-serine but not D-serine to pyruvate clearly identified MSMEG_3532 as an L-serine dehydratase (SdhA) despite its annotation as a threonine dehydratase (Fig. 4). For a further characterization of the enzyme steady-state kinetics were determined with 56 μg protein and various L-serine concentrations. Assuming Michaelis-Menten kinetics the enzyme showed a K_m of 2.58 ± 1.22 μM and a maximal reaction velocity, v_{max}, of 11.7 ± 0.21 min^{-1} mg^{-1} (Fig. 5). In addition, we determined the molecular weight of the native enzyme. Using gel-chromatography and calibration with the appropriate standards, a mass of 57 kDa was determined. Since the mass of the His-tagged protein used is 35.68 kDa, we conclude that the protein in its active state is a homodimer, as also known for rat liver serine dehydratase [14].

Addition of gp77 to SdhA has no effect on enzymatic activity *in vitro*

In an attempt to assay whether the interaction of gp77 with SdhA has an influence on L-serine dehydratase activity, both proteins were incubated together in 0.1 M potassium phosphate buffer pH 8.2, containing 0.1 mM pyridoxal-5′-phosphate, 10% glycerol. Three assays were set up at a constant concentration of 15.6 μg ml^{-1} gp77, 17.8 μg SdhA protein was added, corresponding to a molar ratio of 1:1. The same reaction was performed with a SdhA/gp77 molar ratio of 1:2 or 1:0.2. The reaction mixture was incubated for 30 min at RT, and sub-

Figure 4. Pyruvate formation with the recombinant serine dehydratase, SdhA, of *M. smegmatis.* Samples from assays containing L-serine plus protein (■), and a control without protein (▲), as well as samples containing D-serine plus protein (▼) and D-serine without protein (♦) were taken at different time intervals and the pyruvate formed was quantified. Only the addition of L-serine to SdhA led to a measurable increase of pyruvate making SdhA an L-serine dehydratase. AUC: area under the curve. Amounts of pyruvate are shown as area under the concentration curve as determined by the HPLC software.
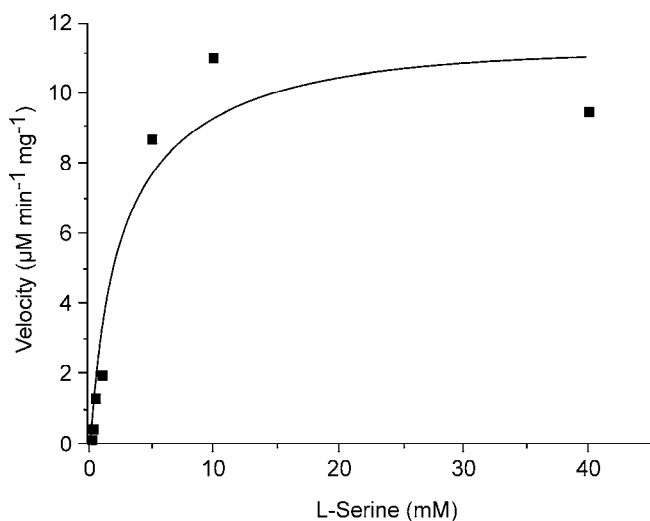
Figure 5. Michaelis Menten kinetics of SdhA of *M. smegmatis.* The enzyme showed a K_m of 2.58 ± 1.22 µM and a maximal reaction velocity, v_{max}, of 11.7 µmol ± 0.21 min^{-1} mg^{-1}.

sequently L-serine dehydratase activities at saturating L-serine concentrations were determined. As a result no difference in enzyme activity was seen between the three assays compared to the SdhA-only control (Fig. 6).

Discussion

In this study we identified a protein of *M. smegmatis* which interacts with the early cytotoxic protein gp77 of mycobacteriophage L5. This host protein (MSMEG_3532) is a pyridoxal-5′-phosphate-dependent L-serine dehydratase, with the gene accordingly termed *sdhA*. This catabolic enzyme converts L-serine to pyruvate through

dehydration and deamination steps (Fig. 3). Consumption of L-serine provides pyruvate as a substrate for the tricarboxylic acid cycle and the acetate switch in bacteria [15]. In *E. coli* the L-serine dehydratase is induced under stressful conditions like exposure to DNA damaging agents such as UV-light and mitomycine, or an increase in growth temperature [16]. The deletion of the three *E. coli* genes involved in the L-serine to pyruvate conversion leads to a specific growth defect. It has been assumed that the accumulation of serine starves the cell for S-adenosylmethionine, an important methyl donor in bacterial replication [10]. In the annotated genome of *M. smegmatis* a peptide transporter is located adjacent to the *sdhA* gene. This supports the hypothesis that SdhA might enable the organism to utilize L-serine as a component released from extracellular peptides. Uptake and hydrolysis of the latter would provide convenient access to L-serine as carbon and energy source. Similar metabolic mechanisms have been observed for selected strains of *E. coli* [17] as well as for *Corynebacterium glutamicum* in its stationary phase, ultimately resulting in an improved growth yield [18]. In *M. smegmatis sdhA* is erroneously annotated as threonine dehydratase, *ilvA*. The serine dehydratase exhibits about 33–36% identities over its entire length to a second *potential* threonine dehydratase that can be found in many mycobacterial species. In *M. smegmatis* this *ilvA* gene is encoded by MSMEG_3183. Furthermore, this locus is syntenic in *Mycobacterium* species including *M. leprae* with its largely decayed genome [19], indicative of its essential function in producing of 2-oxobutyrate for L-isoleucine synthesis.

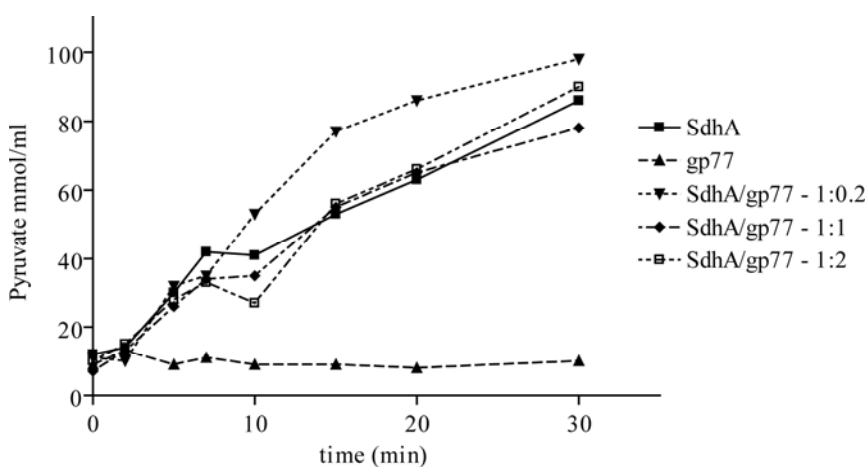In addition to these pyridoxal-5′-phosphate-dependent serine and threonine dehydratases, mycobacteria

Figure 6. Effect of gp77 on SdhA *in vitro*. Different amounts of gp77 were added to SdhA and pyruvate was quantified after the addition of L-serine. The addition of the phage protein had no statistical significant effect on the function of SdhA.

have a predicted serine dehydratase (annotated *sdaA*) catalysing the identical chemical dehydration reaction but using a different mechanism involving an iron-sulfur cluster [20]. Neither a serine dehydratase nor a threonine dehydratase has yet been characterized in mycobacteria, albeit it is known that *Mycobacterium bovis* BCG is unable to catabolize L-serine, which is suggested to be due to an inadequate expression of a corresponding serine dehydratase gene [21].

Little is known about the amino acid requirements of mycobacteriophages during lytic growth and the present knowledge is mostly based on investigations that were carried out in the 1950s. For bacteriophages infecting gram negative bacteria, it is well known that the quality of the host's amino acid pool has major influence on phage replication and some data indicate that bacteriophages actively modify this pool according to their needs. In *E. coli*, phage development is inhibited in the absence of amino acids such as leucine, valine, isoleucine and threonine [22]. Addition of pyruvate to the growth medium of lysogenic *E. coli* K12 led to increased survival of bacteria after induction of the phage. Furthermore *E. coli*-cells grown on L-serine rich media accumulate pyruvate after infection with phage T7 indicating that the host's pyruvate metabolism is altered by bacteriophages [23]. There are also data showing that the activity of catabolic enzymes such as the serine dehydratase is decreased upon bacteriophage infection of *E. coli* [24]. Furthermore, the infection of *E. coli* with the single stranded bacteriophage ΦX174 leads to the repression of catabolite sensitive enzymes like the D-serine dehydratase and it was shown that this repression is not due to a down-regulation of enzyme specific RNA [25].

It is an intriguing finding that a mycobacteriophage protein such as gp77 specifically interacts with the catabolic enzyme SdhA. Currently we can not determine how gp77 modifies SdhA during lytic phage replication since the addition of the phage protein had no influence on the conversion of serine to pyruvate in our *in vitro* assay. The reason for this may be found in general limitations of simplified *in vitro* assays where important co-factors or accessory proteins for the alteration of enzymatic processes by secondary binding proteins may be missing. Also binding of gp77 may inhibit the dimerization of SdhA *in vivo* which could alter enzymatic activity – a mechanism that can not be detected *in vitro*. Additionally the role of SdhA in interacting with gp77 could be to provide L-serine or amino groups for further metabolic processes involving gp77 without influencing enzymatic activity of SdhA.

Copies of gp77 are present in some other mycobacteriophages. In addition, the amino acid sequence is homologous to mostly hypothetical proteins of actinomycetes and the short conserved domain DUF2786 of unknown function which can be found in both, gram negative and gram positive bacteria [9]. A careful sequence analysis showed partial identity to cytosolic catabolic enzymes such as a putative amino acid decarboxylase of *Streptomyces ambofaciens* or enzymes involved in peptidoglycan catabolism such as the N-acetylmuramyl-L-alanine amidase of *Bacillus cereus*. In prokaryotes the interaction of enzymes involved in the processing of amino acids with other cytosolic proteins is a well known phenomenon [26, 27]. Presumably the phage gene 77 was acquired from a former host of L5 by homologous recombination and modified in such a way that binding to SdhA leads to a favourable situation for the phage, for example, by increasing the amino acid pool for the generation of new phage particles or modifying energy recruiting cell cycles in favour of the phage production machinery.

Acknowledgements

Conflict of Interest Statement

The authors declare that there is no financial or commercial conflict of interest associated with this work.

References

[1] van Kessel, J.C., Marinelli, L.J., Hatfull, G.F., 2008. Recombineering mycobacteria and their phages. Nat. Rev. Microbiol., **6**, 851–857.

[2] Bardarov, S., Kriakov, J., Carriere, C., Yu, S. *et al.*, 1997. Conditionally replicating mycobacteriophages: a system for transposon delivery to *Mycobacterium tuberculosis*. Proc. Natl. Acad. Sci. USA, **94**, 10961–10966.

[3] Hatfull, G.F., Pedulla, M.L., Jacobs-Sera, D., Cichon, P.M. *et al.*, 2006. Exploring the mycobacteriophage metaproteome: phage genomics as an educational platform. PLoS Genet., **2**, e92.

[4] Morris, P., Marinelli, L.J., Jacobs-Sera, D., Hendrix, R.W., Hatfull, G.F., 2008. Genomic characterization of mycobacteriophage Giles: evidence for phage acquisition of host DNA by illegitimate recombination. J. Bacteriol., **190**, 2172–2182.

[5] Pedulla, M.L., Ford, M.E., Houtz, J.M., Karthikeyan, T. *et al.*, 2003. Origins of highly mosaic mycobacteriophage genomes. Cell, **113**, 171–182.

[6] Pham, T.T., Jacobs-Sera, D., Pedulla, M.L., Hendrix, R.W., Hatfull, G.F., 2007. Comparative genomic analysis of my-cobacteriophage Tweety: evolutionary insights and cons-truction of compatible site-specific integration vectors for mycobacteria. Microbiology, **153**, 2711–2723.

[7] Rybniker, J., Plum, G., Robinson, N., Small, P.L., Hart-mann, P., 2008. Identification of three cytotoxic early proteins of mycobacteriophage L5 leading to growth inhi-bition in *Mycobacterium smegmatis*. Microbiology, **154**, 2304–2314.

[8] Rybniker, J., Kramme, S., Small, P.L., 2006. Host range of 14 mycobacteriophages in *Mycobacterium ulcerans* and se-ven other mycobacteria including *Mycobacterium tuberculo-sis*-application for identification and susceptibility testing. J. Med. Microbiol., **55**, 37–42.

[9] Marchler-Bauer, A., Anderson, J.B., Chitsaz, F., Derby-shire, M.K. *et al.*, 2009. CDD: specific functional annota-tion with the Conserved Domain Database. Nucleic Acids Res., **37**, D205–D210.

[10] Zhang, X., Newman, E., 2008. Deficiency in l-serine dea-minase results in abnormal growth and cell division of *Escherichia coli* K-12. Mol. Microbiol., **69**, 870–881.

[11] Sambrook, J., Russell, D.W., 2001. Molecular Cloning: a Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

[12] El Mourabit, H., Muller, S., Tunggal, L., Paulsson, M., Aumailley, M., 2004. Analysis of the adaptor function of the LIM domain-containing protein FHL2 using an affinity chromatography approach. J. Cell. Biochem., **92**, 612–625.

[13] Hara, S.T., Iwata, T., Yamaguchi, M., Nakamura, M., 1985. Fluorometric determination of alpha-keto acids with 4,5-dimethoxy-1,2-diaminobenzene and its application to high-performance liquid chromatography. Anal. Chim. Acta, **172**, 167–173.

[14] Ogawa, H., Gomi, T., Takusagawa, F., Masuda, T. *et al.*, 2002. Evidence for a dimeric structure of rat liver serine dehydratase. Int. J. Biochem. Cell Biol., **34**, 533–543.

[15] Wolfe, A.J., 2005. The acetate switch. Microbiol. Mol. Biol. Rev., **69**, 12–50.

[16] Su, H.S., Lang, B.F., Newman, E.B., 1989. L-serine degrada-tion in *Escherichia coli* K-12: cloning and sequencing of the sdaA gene. J. Bacteriol., **171**, 5095–5102.

[17] Zinser, E.R., Kolter, R., 1999. Mutations enhancing amino acid catabolism confer a growth advantage in stationary phase. J. Bacteriol., **181**, 5800–5807.

[18] Netzer, R., Peters-Wendisch, P., Eggeling, L., Sahm, H., 2004. Cometabolism of a nongrowth substrate: L-serine utilization by *Corynebacterium glutamicum*. Appl. Environ. Microbiol., **70**, 7148–7155.

[19] Cole, S.T., Eiglmeier, K., Parkhill, J., James, K.D. *et al.*, 2001. Massive gene decay in the leprosy bacillus. Nature, **409**, 1007–1011.

[20] Hofmeister, A.E., Textor, S., Buckel, W., 1997. Cloning and expression of the two genes coding for L-serine dehy-dratase from *Peptostreptococcus asaccharolyticus*: relation-ship of the iron-sulfur protein to both L-serine dehydrata-ses from *Escherichia coli*. J. Bacteriol., **179**, 4937–4941.

[21] Wheeler, P.R., Brosch, R., Coldham, N.G., Inwald, J.K. *et al.*, 2008. Functional analysis of a clonal deletion in an epidemic strain of *Mycobacterium bovis* reveals a role in li-pid metabolism. Microbiology, **154**, 3731–3742.

[22] Gots, J.S., Hunt, G.R. Jr., 1953. Amino acid requirements for the maturation of bacteriophage in lysogenic *Escheri-chia coli*. J. Bacteriol., **66**, 353–361.

[23] Borek, E., Rockenbach, J., Ryan, A., 1956. The effect of lipoic acid on recovery in induced lysogenic organisms. Proc. Natl. Acad. Sci. USA, **42**, 708–710.

[24] Pardee, A.B., Kunkee, R.E., 1952. Enzyme activity and bacteriophage infection. II. Activities before and after vi-rus infection. J. Biol. Chem., **199**, 9–24.

[25] Ghosh, A., Pal, S.K., Poddar, R.K., 1985. Modulation of gene expression in *Escherichia coli* infected with single-stranded bacteriophage phi X174. Mol. Gen. Genet., **198**, 304–308.

[26] Arifuzzaman, M., Maeda, M., Itoh, A., Nishikata, K. *et al.*, 2006. Large-scale identification of protein-protein interac-tion of *Escherichia coli* K-12. Genome Res., **16**, 686–691.

[27] Parrish, J.R., Yu, J., Liu, G., Hines, J.A. *et al.*, 2007. A pro-teome-wide protein interaction map for *Campylobacter je-juni*. Genome Biol., **8**, R130.

[28] Snapper, S.B., Melton, R.E., Mustafa, S., Kieser, T., Jacobs, W.R. Jr., 1990. Isolation and characterization of efficient plasmid transformation mutants of *Mycobacterium smeg-matis*. Mol. Microbiol., **4**, 1911–1919.

12.   Erklärung

Ich versichere, dass ich die von mir vorgelegte Dissertation selbständig angefertigt, die benutzten Quellen und Hilfsmittel vollständig angegeben und die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen -, die anderen Werken im Wortlaut oder dem Sinn nach entnommen sind, in jedem Einzelfall als Entlehnung kenntlich gemacht habe; dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie – abgesehen von unten angegebenen Teilpublikationen – noch nicht veröffentlicht worden ist sowie, dass ich eine solche Veröffentlichung vor Abschluss des Promotionsverfahrens nicht vornehmen werde. Die Bestimmungen dieser Promotionsordnung sind mir bekannt. Die von mir vorgelegte Dissertation ist von PD Dr. Pia Hartmann betreut worden.

Curriculum Vitae

Personal data:

Name:	Jan Rybniker
Degrees:	MD (Dr. med.)
Address:	Rheinbacher Str. 26
	50937 Köln

Phone:	+49 179 971 4043
Email:	jan.rybniker@uk-koeln.de
Date of birth:	May 23rd 1975
Place of birth:	Karsruhe / Germany
Nationality:	german
Marital status:	single

Professional career:

Starting February 2006 — Medical Doctor/post doctoral researcher at the 1st Dept. of Internal Medicine, University of Cologne

August 2004 through January 2006 — MD and post doc position at the Bernhard-Nocht-Institute for Tropcial Medicine, Hamburg, Germany

MD thesis:
2000 until 2004

"Transposon mutagenesis of *Mycobacterium ulcerans* and *Mycobacterium marinum* using temperature- sensitive mycobacteriophages "
summa cum laude

- At the Institute for medical Microbiology, Immunology and Hygiene (Univ. of Cologne)
- At the Prince Leopold Institute of Tropical Medicine, Antwerp, Institute scientifique de la Santé, Brussels

Study at University:
starting1996

University of Cologne, School of Medicine
- final state exam march 2004
- final grade 2 (upper second class)

Practical year:
May 2003 - April 2004

- Paediatrics at the University of Cologne
- Internal medicine at the Univ. of Cologne
- Surgery at Hôtel Dieu, Service de Chirurgie générale, Paris / France

Electives:

August – September 1998	Agogo Hospital, dept. of surgery, Ghana
February – April 1999	Clinic for child and adolescent psychiatry, Bonn
March 2000	Training in molecular biology methods at the Institute of Pathology, University of Cologne
July – August 2000	Clinic for Neurosurgery, University of Cologne
November 2001	Aid in establishing a small microbiological lab in Benin / West-Africa (IMPETUS-project)

Educational:

1981-1985	Elementary school in Karlsruhe / Germany
1985-1994	Helmholtzgymnasium in Karlsruhe, Abitur 1994

Funding, Fellowships:
- Researcher position, six year period starting in 2007. Funded by the German federal ministry of education and research (BMBF)
- Doctorate studentship Maria Pesch foundation (2000 through 2002)
- Foundation Köln Fortune - fellowship for scientists abroad (2002 through 2003)

Research abroad: May 2002 until June 2003

Visiting scientist at the University of Tennessee, Dept. of Microbiology, USA (Prof. Pamela Small)

Projects: Potential of Bacteriophage-therapy for *Mycobacterium ulcerans* Infection, host range of several mycobacteriophages, function and structure of mycobacteriophage D29

Awards
- first winner "young investigator award", 116th Congress Of The German Society For Internal Medicine (DGIM). Wiesbaden 2010
- second winner "young investigator award", 114th Congress Of The German Society For Internal Medicine (DGIM). Wiesbaden 2008
- first place poster award, congress for infectious diseases and tropical medicine, Innsbruck 2008