Molecular Microbiology (2004) **51**(4), 1003–1014 doi:10.1046/j.1365-2958.2003.03900.x

The *Mycobacterium tuberculosis ino1* gene is essential for growth and virulence

Farahnaz Movahedzadeh,[1] Debbie A. Smith,[2] Richard A. Norman,[3†] Premkumar Dinadayala,[4] Judith Murray-Rust,[3,5] David G. Russell,[6] Sharon L. Kendall,[1] Stuart C. G. Rison,[1] Mark S. B. McAlister,[5†] Gregory J. Bancroft,[2] Neil Q. McDonald,[3,5] Mamadou Daffe,[4] Yossef Av-Gay[7] and Neil G. Stoker[1*]

[1]*Department of Pathology and Infectious Diseases, Royal Veterinary College, London NW1 0TU, UK.*
[2]*Department of Infectious & Tropical Diseases, London School of Hygiene & Tropical Medicine, Keppel Street, London WC1E 7HT, UK.*
[3]*Structural Biology Laboratory, Cancer Research UK London Research Institute, Lincoln's Inn Fields, London WC2A 3PX, UK.*
[4]*Department of Molecular Mechanisms of Mycobacterial Infections, Institut de Pharmacologie et Biologie Structurale, UMR 5089 du CNRS et Université Paul Sabatier, 205, route de Narbonne, 31077 Toulouse cedex 04, France.*
[5]*Bloomsbury Centre for Structural Biology and School of Crystallography, Birkbeck College, Malet Street, London WC1E 7HX, UK.*
[6]*Microbiology and Immunology, College of Veterinary Medicine, Cornell University, Ithaca, NY 14853, USA.*
[7]*Department of Medicine, Division of Infectious Diseases, University of British Columbia, 2733 Heather St., Vancouver, British Columbia, Canada V5Z 3J5.*

Summary

Inositol is utilized by *Mycobacterium tuberculosis* in the production of its major thiol and of essential cell wall lipoglycans. We have constructed a mutant lacking the gene encoding inositol-1-phosphate synthase (*ino1*), which catalyses the first committed step in inositol synthesis. This mutant is only viable in the presence of extremely high levels of inositol. Mutant bacteria cultured in inositol-free medium for four weeks showed a reduction in levels of mycothiol, but phosphatidylinositol mannoside, lipomannan and lipoarabinomannan levels were not altered. The *ino1* mutant was attenuated in resting macrophages and in SCID mice. We used site-directed mutagenesis to alter four putative active site residues; all four alterations resulted in a loss of activity, and we demonstrated that a D310N mutation caused loss of the active site Zn^{2+} ion and a conformational change in the NAD^{+} cofactor.

Accepted 15 October, 2003. *For correspondence. E-mail nstoker@rvc.ac.uk; Tel. (+44) 020 74685272; Fax (+44) 020 74685306. †Present address: AstraZeneca, Alderley Park, Macclesfield, Cheshire SK10 4TG, UK.

Introduction

Tuberculosis is a global public health problem, with one-third of the world's population infected with *Mycobacterium tuberculosis*, and an estimated eight million new cases annually (WHO, 2000). Current control methods of chemotherapy and vaccination are inadequate. Treatment is lengthy, with short-course therapy taking six months, and antibiotic resistance is a serious problem. One approach to addressing this problem is to identify key pathways in *M. tuberculosis* physiology in which enzymes might be novel drug targets. Identification of processes that are essential for *in vivo* survival also provides valuable information about the environments encountered within the host. We have shown previously, for example, that amino acid auxotrophs lacking the *trpD* or *proC* genes are highly attenuated in immunocompromised and immunocompetent mice (Smith *et al*., 2001). Other groups have shown similar results using mutants in amino acid, nucleotide or vitamin biosynthesis (Jackson *et al*., 1999; Chambers *et al*., 2000; Hondalus *et al*., 2000; Sambandamurthy *et al*., 2002). In order to investigate other pathways that might be suitable targets for chemotherapy, we have been studying inositol metabolism in *M. tuberculosis*.

Inositol is an essential polyol in eukaryotes, where phosphatidylinositol (PI) is an important constituent of phospholipid membranes. In addition, PI is cleaved by phospholipase C to generate inositol phosphate and diacyglycerol, both of which are central cell signalling molecules. In prokaryotes, inositol is found in archaebacteria, where derivatives are important in osmotic protection (Santos and da Costa, 2002), and is absent from most eubacteria. The exception is the Actinomycetales group that contains the streptomycetes, corynebacteria and mycobacteria. These all use inositol in their major thiol, mycothiol (Fahey, 2001), which is important in maintaining the redox balance in the cell, in protecting the cell from oxidative stress, and in cysteine storage. In addition, the

corynebacteria and mycobacteria produce cell wall lipoglycans from PI. These include the PI mannosides (PIMs), lipomannan (LM) and lipoarabinomannan (LAM), which are present at high levels. Lipid moieties attached to the PI anchor these molecules in the cell envelope. PI-containing molecules have been shown to be essential for growth of the fast-growing species *Mycobacterium smegmatis*, as mutants lacking PI synthase are not viable (Jackson *et al*., 2000).

Inositol is normally synthesized in two steps from glucose-6-phosphate. The first committed step is the conversion to inositol-1-phosphate (I-1-P) by I-1-P synthase (Ino1; EC 5.5.1.4; *M. tuberculosis* Ino1, tbINO). I-1-P is then dephosphorylated by inositol monophosphate phosphatase to produce inositol. The *M. tuberculosis Rv0046c* gene was identified as *ino1* encoding an inositol phosphate synthase initially on the basis of sequence similarity with yeast INO1. This was confirmed genetically by the demonstration that the mycobacterial gene complemented a *Saccharomyces cerevisiae ino1* mutant (Bachhawat and Mande, 1999). The structures of the *M. tuberculosis* and *Saccharomyces cerevisiae* Ino1 proteins have recently been reported and both contain NAD^+/NADH as a cofactor. While the eukaryotic enzyme requires NH_4^+ (Stein and Geiger, 2002), we showed that the prokaryotic enzyme has a Zn^{2+} ion located in the active site of the holoenzyme, although not in an analogous position to the NH_4^+ ion (Norman *et al*., 2002).

In this study, we have constructed a *M. tuberculosis* mutant lacking *ino1*, and shown that it is attenuated *in vitro* and *in vivo*. We have also carried out a structure-function study to extend our understanding of the mechanism of action of this enzyme.

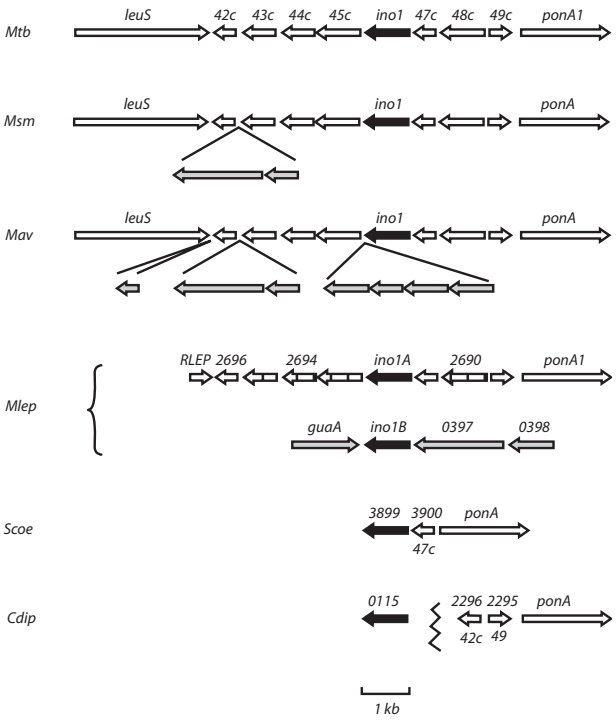


Fig. 1. Comparative genome analysis of actinomycete *ino1* genes. The genetic structure of regions containing the *ino1* gene are shown for *M. tuberculosis* (Mtb), *M. smegmatis* (Msm), *M. avium* (Mav), *M. leprae* (Mlep), *S. coelicolor* (Scoe) and *C. diptheriae* (Cdip). Black arrows, *ino1* gene; white arrows, genes showing homology to those in the *M. tuberculosis leuS-ponA1* region; grey arrows, genes not showing homology to those in the *M. tuberculosis leuS-ponA1* region; white hatched arrows, as for white arrows but pseudogenes. Zigzag line in *Cdip*: indicates the genes on either side are not closely linked. Gene names/numbers above arrows refer to the gene identification for that species. Those below arrows in *Cdip* show the *M. tuberculosis* genes to which the *Cdip* genes are homologous. Numbering of genes refers to the relevant species: thus 42c is short for Rv0042c, 2696 is short for ML2696, 3899 is short for SCO3899, and 2296 is short for CD2296.

Results

Bioinformatics analysis

Conservation of gene synteny can suggest that there are constraints against rearrangement. For example, genes that are in operons will be less readily separated because they share expression signals. Conservation of synteny can also give clues to function, as it may suggest that the genes surrounding *ino1* have related functions. We therefore carried out comparative genomic analyses of the *ino1* regions of related organisms (Fig. 1). The *M. tuberculosis ino1* (*Rv0046c*) gene lies in a group of eight genes between *leuS* (*Rv0041*), which encodes leucyl-tRNA synthetase, and *ponA1* (*Rv0050*), which encodes a penicillin-binding protein. *Ino1* and six other genes are transcribed in the same direction; there are predicted intergenic gaps of 61 bp upstream and 81 bp downstream of *ino1* itself, so it is not clear from this criterion if it is transcribed as part of an operon. The other genes between *leuS* and *ponA1* do not have specific functions assigned to them, although homologies suggest general functions in some cases. Two (*Rv0042c* and *Rv0043c*) are predicted to encode transcriptional regulators of the MarR and GntR families, respectively. *Rv0044c* encodes a putative oxidoreductase, and *Rv0045c* a hydrolase.

In all the genomes studied except for *Corynebacterium diphtheriae*, *Rv0047c* is maintained upstream of *ino1*. Thus, despite the 61 bp gap between these genes in *M. tuberculosis* (the shortest gap is 44 bp in *Streptomyces coelicolor*), this suggests that they are in a single transcription unit, which may also reflect a functional relationship.

Expression of ino1

In order to determine the level of expression of the *ino1*

gene, we carried out real-time PCR experiments using *M. tuberculosis* H37Rv, and compared the level of expression of *ino1* with that of the gene encoding the major sigma factor, *sigA*. We showed that, in mid-log phase growth, the level of *ino1* is 0.98 that of *sigA* (95% confidence interval 0.8–1.2). The *sigA* gene is a moderately highly expressed gene, normally transcribed at over 10^5 copies per 10 ng total RNA (Manganelli *et al*., 1999).

Construction of an ino1 *mutant*

We carried out targeted mutagenesis of the *M. tuberculosis ino1* gene in order to introduce an unmarked mutation that would have no potential polar effects, using a two-step strategy (Parish and Stoker, 2000). A 3 kb fragment containing *Rv0046c* was cloned into plasmid p2NIL and a 900 bp deletion was generated in the 1101 bp gene (Fig. 2A). Following insertion of a gene cassette carrying *lacZ* and *sacB*, this was introduced into *M. tuberculosis* H37Rv and single-crossovers (SCOs) isolated by selection for blue hyg^R kan^R colonies. One SCO colony was plated onto sucrose to isolate bacteria with a second crossover, which will lead to mutant or wild-type cells depending on the location of the recombination event. In the absence of inositol, only wild-type colonies were isolated. However, in the presence of 77 mM inositol, four out of eight colonies were shown by PCR to be *ino1* mutants, which were confirmed by Southern analysis (Fig. 2B).

One of the mutants (strain Fame5) was studied further. Although it grew on inositol-containing plates, no growth was seen on plates lacking inositol (data not shown). We complemented the mutation by cloning *ino1* (with the two upstream genes as we were unsure of the location of the promoter) into an integrating vector, producing plasmid pFM109. When this was reintroduced into Fame5, inositol was no longer required. This provides strong evidence that the *ino1* gene is essential in the absence of exogenous inositol, but does not formally rule out a contribution by *Rv0047c* or *Rv0048c* which were also on the complementing plasmid.

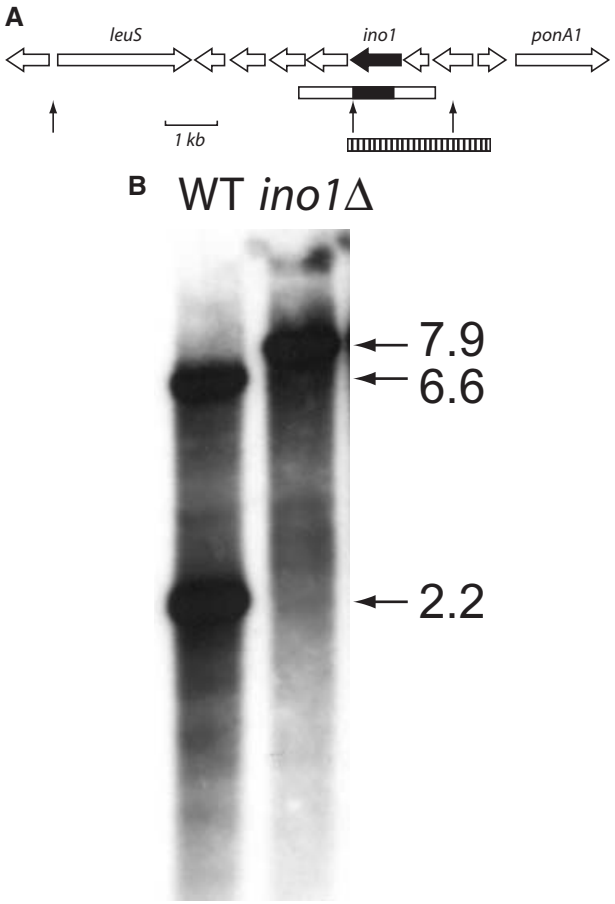


Fig. 2. Southern analysis of *ino1* mutant. DNA was extracted, digested with *Pvu*II, and subjected to Southern analysis. Following probing with the 3 kb *Bam*HI–*Hin*dIII fragment from pFM61 containing the *ino1* gene, two bands of 6.6 and 2.2 kb were seen in DNA from the wild-type strain, and a single band of 7.9 kb in the mutant strain. The 900 bp deletion includes a *Pvu*II site.
A. Diagram of region showing cloned region (open bar) and deleted region (filled bar). Arrows indicate *Pvu*II sites; the middle site is lost in the mutant. The hatched bar shows the fragment used to complement the mutation.
B. Autophotograph of Southern blot. Sizes of bands are indicated in kb.

Fame5 requires a high level of exogenous inositol

We used a high level of inositol in the supplemented medium (77 mM), because we had originally failed to isolate mutants using lower concentrations (data not shown). We therefore cultured Fame5 in media containing different concentrations of inositol. Although growth in 77 mM inositol matched that of wild-type bacteria, it was greatly reduced in 10 mM inositol, and virtually absent at a concentration of 1 mM (Fig. 3). In order to determine whether the bacteria die in the absence of inositol, we cultured mutant bacteria that had been incubated without inositol for four weeks. There was no drop in the CFUs, indicating that they are still viable despite the lack of growth.

Biochemical analyses

In order to look at the effect of removal of inositol from the mutant on mycothiol and inositol-containing lipoglycans, we carried out experiments where the strain was cultured with inositol, then transferred to inositol-free medium. Cells were harvested after two and four weeks, and analysed.

Mycothiol levels fell gradually over the four week period (Fig. 4). In a typical experiment, levels of mycothiol in the mutant grown in the presence of inositol (41.9 pmol/10^9

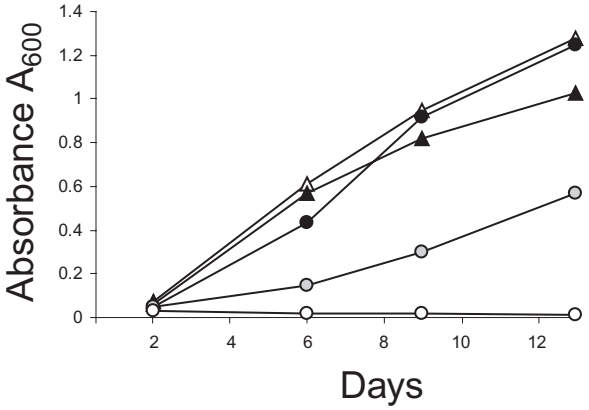


Fig. 3. Growth of the *ino1* mutant construct. Graph showing growth of wild-type and mutant strains in different concentrations of inositol. Filled diamond, WT + 77 mM inositol; empty diamond, WT with no inositol; filled circles, mutant + 77 mM inositol; grey circle, mutant + 10 mM inositol; empty circle, mutant + 1 mM inositol. The results shown are representative of four independent experiments.

cells) exceeded those in wild-type bacteria grown without inositol (31.8 pmol/10^9 cells). By two weeks, this fell to wild-type levels, and after four weeks was still lower (25.5 pmol/10^9 cells).

Analyses of the cell wall showed no difference in levels of PIM_2, PIM_6, LM or LAM, even after four weeks growth in the absence of inositol (data not shown).

Survival of M. tuberculosis ino1 *in the host*

Mycobacterium tuberculosis is able to survive in resting macrophages (Russell *et al*., 1997). We infected bone marrow-derived murine macrophages with the *ino1* mutant and measured survival by assaying colony-forming units (CFUs). Although CFUs in macrophages infected with wild-type bacteria remained stable, those in mutant-infected macrophages fell sharply, and were virtually cleared from the cell cultured by seven days (Fig. 5A).

These data were supported by results in mice. We used the SCID mouse model, which is useful for looking at the innate immune response to infection. Although they lack an acquired immune response, they exhibit initial stages of inflammatory cell recruitment and T cell-independent IFNγ-mediated macrophage activation (Hansch *et al*., 1996; Smith *et al*., 1997). We and others have previously shown that the SCID mouse model is effective at predicting phenotypes in immunocompetent mice (Smith *et al*., 2001; Gordhan *et al*., 2002; McAdam *et al*., 2002; Pym *et al*., 2002; Sambandamurthy *et al*., 2002; Parish *et al*., 2003). We infected mice i.v. with 10^6 CFU *M. tuberculosis*. Whereas all mice infected with wild-type bacteria died between 38 and 44 days (median survival time, MST: 38), all mice infected with mutant bacteria were alive and healthy at 56 days, at which time the experiment was terminated (Fig. 5B). All mice infected with the complemented strain died between days 45–47 with an MST of 45 days. Thus the mutant was significantly attenuated compared to both other strains ($P > 0.001$).

Structure-function studies

We have previously reported the structure of wild-type tbINO in the presence of the NAD cofactor and an active site zinc ion (Norman *et al*., 2002), the positions of which revealed the location of the active site. However we were unable to obtain insights into the catalytic mechanism or catalytic residues because of our inability to co-crystallise the protein with substrate or a substrate analogue. In order to identify key catalytic residues, we selected four residues within the active site, and carried out site-directed mutagenesis. The rationale was firstly to determine if they did indeed play an important role in activity and secondly to identify a mutant which could trap the substrate in the active site, allowing a detailed study of the mechanism of action.

From the 21 residues that are completely conserved in Ino1 homologues (Fig. 6A) we chose to mutate D197, K284, D310 and K346, which all lie close to the predicted active site. We constructed mutants in which each residue had been converted to alanine (D197A, K284A, D310A and K346A), and in addition we made more subtle mutants of the aspartate residues, converting them to asparagine (D197N and D310N). The mutations were introduced into the complementing plasmid, and were

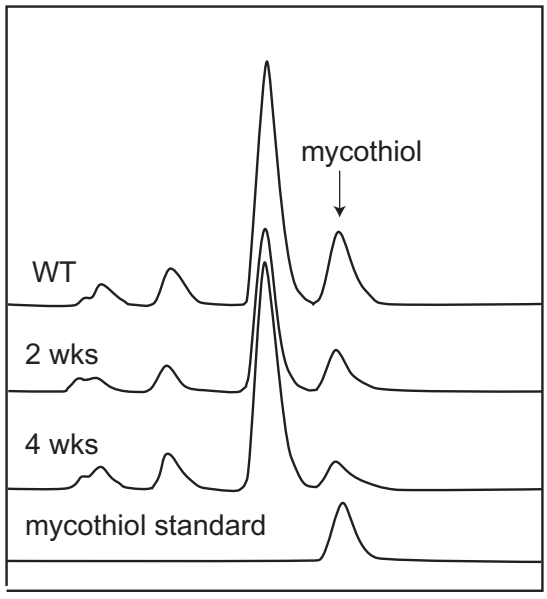


Fig. 4. Mycothiol levels in FAME5 following inositol removal. Mycothiol levels were measured in cultures of wild-type *M. tuberculosis*, and strain FAME5 following two and four weeks growth after removal of inositol.

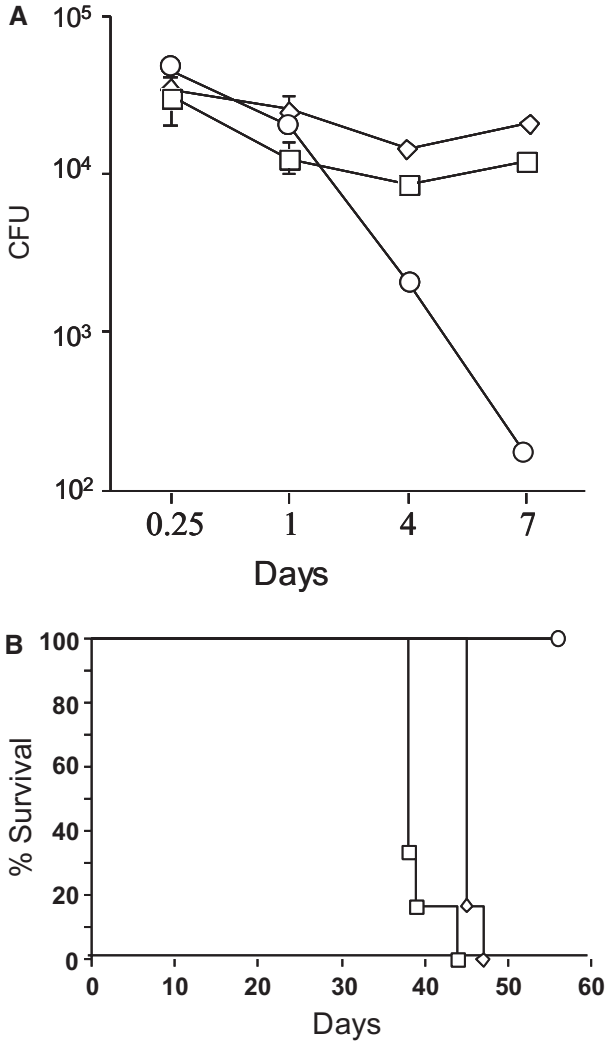


Fig. 5. Residual virulence of *M. tuberculosis ino1*.
A. Survival graph of *M. tuberculosis* following infection of macrophages. Murine bone marrow-derived macrophages were infected with wild-type H37Rv (squares), Δ*ino1* bacteria (circles), and complemented bacteria (diamonds). Comparable bacterial survival profiles were observed in two independent experiments.
B. Six SCID mice per group were infected intravenously with 1×10^6 of either wild-type, mutant (ino1Δ) or complemented (ino1Δ::pFM109) *M. tuberculosis*. Mice were monitored for survival over 2 months. Square, wild-type; circle, mutant; diamond, complemented strain. These data were representative of two separate experiments.

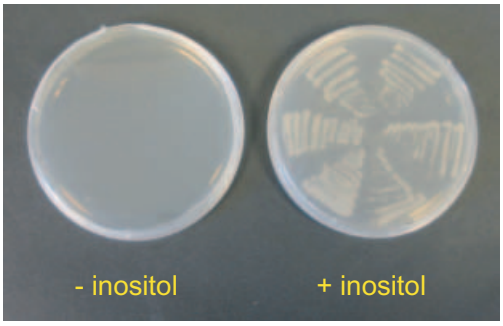
introduced into Fame5, selecting on plates containing inositol. The strains were then cultured in inositol-free media. Growth was negative for all mutants (Fig. 6B), demonstrating that each mutation had caused reduced enzymatic function. This also formally demonstrates that the complementation we obtained earlier was due to the presence on the plasmid of *ino1* rather than of *Rv0047c* or *Rv0048c*.

To examine these mutant proteins further, we determined the structures of D310A and D310N at 2.1 Å and 2.3 Å respectively (Table 1). Each mutant has an overall fold identical to that of the wild-type holoenzyme (pdb code 1 g0) with rms deviations in Cα positions of 0.19 Å for 325 atoms (D310A) and 0.15 Å for 328 atoms (D310N). We had anticipated that these conservative mutations would not perturb the active site but would abolish cata-

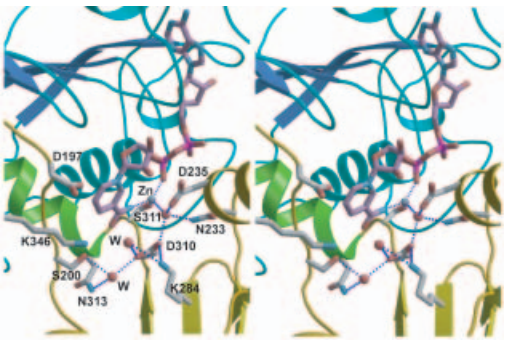
A

```
  1 - MSEHQSLPAP EASTEVRVAI VGVGNCASSL VQGVEYYYNA DDTSTVPGLM
 51 - HVRFGPYHVR DVKFVAAFDV DAKKVGFDLS DAIFASENNT IKIADVAPTN
101 - VIVQRGPTLD GIGKYYADTI ELSDAEPVDV VQALKEAKVD VLVSYLPVGS
151 - EEADKFYAQC AIDAGVAFVN ALPVFIASDP VWAKKFTDAR VPIVGDDIKS
201 - QVGATITHRV LAKLFEDRGV QLDRTMQLNV GGNMDFLNML ERERLESKKI
251 - SKTQAVTSNL KREFKTKDVH IGPSDHVGWL DDRKWAYVRL EGRAFGDVPL
301 - NLEYKLEVWD SPNSAGVIID AVRAAKIAKD RGIGGPVIPA SAYLMKSPPE
351 - QLPDDIARAQ LEEFIIG
```

B



C



D

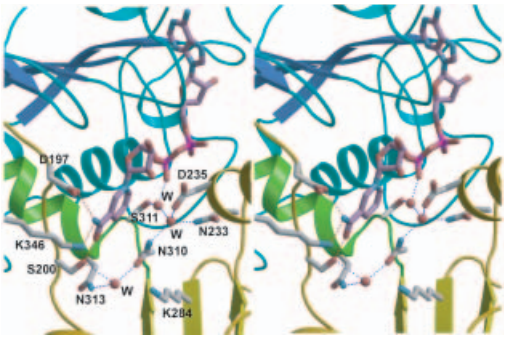


Fig. 6. Site-directed mutagenesis of tbINO.
A. Sequence of tbINO. Conserved amino acids are highlighted. Mutated amino acids are shown in bold-type.
B. Effect of mutations on growth. *M. tuberculosis* FAME5 derivatives carrying integrating plasmids pFM130–135 (strains FAME14–19) were plated in the presence and absence of 77 mM inositol.
C and D. Effect of altering D310. Stereo images of active sites are shown of WT tbINO (C) and of tbINO carrying a D310N mutation (D). Pictures were made with Bobscript (Esnouf, 1999) and Raster3D (Merritt and Murphy, 1994).

Table 1. Structural data collection and refinement.

	D130A	D130N
Data collection		
Resolution (Å)	2.3	2.1
Spacegroup	$P42_12$	$P42_12$
Z	1	1
Cell dimensions (Å)		
a, b	116.8	116.3
c	65.46	64.28
Wavelength (Å)	0.934	0.934
Observed reflections	62 117	164,238
Unique reflections	15 533	26,347
Highest resolution shell (Å)	2.43–2.30	2.22–2.10
Redundancy[a]	3.6 (3.7)	6.2 (6.0)
I/σ(i) [a]	7.8 (2.8)	7.9 (2.7)
Completeness (%)[a]	89.4 (89.4)	98.1 (90.5)
Rmerge (%)[a]	7.4 (23.0)	7.1 (28.0)
Refinement		
Rwork (%)	21.1	22.5
Rfree (%)[b]	25.9	24.5
Rms bond lengths (Å)	0.006	0.007
Rms bond angles (°)	1.5	1.5
Number of non-H atoms	2530	2572
Number of water molecules	44	111

a. Values in parentheses are those for the highest resolution shell.
b. R_{free} was calculated against 5% of the complete dataset excluded from refinement.

lytic activity by removing an essential sidechain. However, both mutants have two major changes in their active sites. First, the zinc ion is no longer observed and is instead replaced by a water molecule. Second, the nicotinamide part of the NAD cofactor, which coordinates the active site zinc ion in the wild-type Ino1 (Fig. 6C), rotates by approximately 180 degrees around the C1–N1 bond, allowing it to hydrogen bond to the sidechains of D197 and K346 (Fig. 6D). The sidechain of K284 becomes disordered in the D310A mutant. These mutants were therefore predicted to be catalytically inactive but are not suitable for pursuing the structure of enzyme–substrate complexes due to the conformational rearrangement induced by the mutation.

Discussion

Analyses of gene synteny can be useful in identifying genes that are maintained together. We have shown that the gene upstream of *ino1*, *Rv0047c*, is retained in the same position in all mycobacteria and in *S. coelicolor*, which suggests that they are co-transcribed. There is also the possibility that *Rv0047c* has a related function. *Rv0048c (*100 bp upstream of *Rv0047c*) and the genes downstream of *ino1* are not maintained close to *ino1* in the same way, and we suggest that these are less likely to be in an operon with *ino1*. The fact that surrounding genes are pseudogenes in *M. leprae* (see below) also implies that they are not essential. It was interesting that *ino1* was reported to be non-essential using a saturation transposon mutagenesis strategy where inserts were localized using microarrays, and essentiality inferred statistically (Sassetti *et al*., 2003). Our results demonstrate that this gene is essential under the conditions they used, and suggests that some essential genes may be missed using their method.

Remarkably, *M. leprae* (Cole *et al*., 2001) has two apparently functional *ino1* genes (Fig. 1). One (*ML2692*) lies in the same genetic context as in the mycobacterial species described above, and we have called this *ino1A*. The second (*ML0399*) lies in a different part of the genome, but the sequence of the entire genes (extending 18 bp upstream and 58 bp downstream) are 100% identical, suggesting that a gene duplication has occurred relatively recently. We have called this *ino1B*. The genes surrounding *ino1A* appear to be orthologues of those in *M. tuberculosis*, with homologues from *Rv0042c* to *ponA1* evident. Of these, at least four are pseudogenes, containing multiple stop codons. Synteny stops downstream of the *Rv0042c* homologue where there is an *RLEP* element. The *M. leprae leuS* is found at *ML0032*, near the chromosomal origin in a context similar to that in *M. tuberculosis*, except that an RLEP sequence lies downstream of it. This suggests that at some point an RLEP element has inserted between *leuS* and the *Rv0042c* orthologue, and recombination between this copy of RLEP and another one elsewhere on the chromosome caused the *ino1A* region to be displaced. This phenomenon was proposed by Cole *et al*. (2001) as a major factor in genome rearrangement in *M. leprae*.

The *ino1* gene was shown to be expressed at a moderately high level, equalling the level of *sigA*. This is supported by the fact that it has been identified by in proteomic analyses of *M. tuberculosis* (Rosenkrands *et al*., 2000). Recently reported microarray data show that the gene appears to differentially expressed, with lowered levels in macrophages (resting, activated; wild-type, NOS2), but increased *in vitro* following exposure to H_2O_2 or palmitic acid (Schnappinger *et al*., 2003; supplementary data). These data are somewhat surprising, as it might be expected that H_2O_2 and palmitic acid would both mimic intracellular conditions, and would therefore result in a similar phenotype.

The *M. tuberculosis ino1* mutant required an extremely high level of inositol. A concentration of 77 mM led to normal growth, but growth in 10 mM was greatly reduced, whereas no growth was seen with 1 mM. This is in stark contrast to the equivalent *S. cerevisiae* mutant, which required only 27 μM inositol supplement (Henry *et al*., 1977), and to *M. smegmatis* which requires 0.1 mM supplementation (H. Billman-Jacobe, pers. comm.). The difference with the fast-growing mycobacterium is particularly intriguing; one possibility is that uptake is much less efficient. tbINO is a potential drug target if

inhibitors can be identified that do not affect the eukaryotic homologue; however, the cause of the requirement for such a high level of exogenous inositol would need to be established if a competitive inhibitor were to be sought. The bioinformatics analysis shows that the region surrounding *ino1* in *M. tuberculosis* and *M. smegmatis* is identical except that the fast grower has two extra genes downstream. It is possible that these two predicted ABC transporters play a role in inositol uptake.

We removed inositol from the medium of *M. tuberculosis ino1* and measured the levels of inositol-containing molecules in the bacteria over a four week period. We saw no clear reduction in cell wall lipoglycans, which may reflect how stable these molecules are. This agrees with the observation that LAM is stable in the tissues of leprosy patients (Hunter *et al*., 1986). Mycothiol levels did fall, but there were still high level after two weeks, which accords with previous reports of its high stability (Newton *et al*., 1995). The bacteria during this period were still viable.

We investigated the residual virulence of the *ino1* mutant, and showed severe attenuation in macrophages, and in a SCID mouse model. This is not surprising given the level of inositol supplement needed, as this is unlikely ever to be encountered in the host. It has been shown that inositol levels in normal mouse liver are approximately 100 μM (Berry *et al*., 2003), far lower than the levels required by this mutant. The rapid killing in the macrophages may result in part from a reduction in mycothiol levels, rendering the bacteria more susceptible to oxidative stress.

Our structural studies demonstrated the key importance of certain sidechains in the active site. The D310A and D310N tbINO mutant proteins are biologically inactive, and their active sites differ from those of the wild-type enzyme most notably by loss of the zinc ion and a conformational change in the NAD cofactor. In the wild-type enzyme, D310 is involved in a network of hydrogen bonds, so that the aspartate sidechain makes contact with the sidechain of K248, and water-mediated contacts to the zinc ion, the sidechain of D235 and the sidechain of N313. The alanine sidechain of D310A is incapable of making any of these interactions, so that some rearrangement of local sidechains is not surprising. In D310N the sidechain can and does form many of the hydrogen bonds equivalent to those formed by the aspartyl sidechain of the native enzyme, although one water molecule is absent (see below). However, it does not interact with K284, which in the wild-type structure interacts with D310 both directly and via a water molecule. In the D310N mutant the K284 sidechain is in an extended, ordered conformation, directed towards a region of electron density in the active site which we have tentatively interpreted as a glycerol molecule. Glycerol was used as cryoprotectant, and is capable of competing with glucose analogues at the active site of an enzyme (Tsitsanou *et al*., 1999). However, its presence seems insufficient to trigger ordering of the entire protein similar to that seen in the inhibitor complex of the yeast enzyme (Stein and Geiger, 2002). Changes in the exact orientation of K284 may also permit the observed change in sidechain torsion angles, most notably chi1, for F236, and the absence of an active site water molecule analogous the one of those interacting with OD2 of D310. We conclude that replacement of the negatively charged sidechain at D310 with the isosteric neutral asparagine disturbs a delicate charge balance in the active site, and is thus responsible for the loss of the zinc ion, which in turn removes stabilising bonds to the NAD, allowing it to rotate into the new conformation.

In this paper we have shown that the *M. tuberculosis ino1* gene is necessary for growth *in vivo* and *in vitro*. Its central role in the synthesis of diverse molecules that are important to the cell makes it an attractive target for the development of novel antibiotics, and the structural experiments we describe are a step towards understanding the mechanism of action of the enzyme. However, any antibiotic would have to penetrate the cell wall, and it will therefore be necessary to understand the reason for the requirement for such a high concentration of exogenous inositol.

Experimental procedures

Bacterial strains, plasmids and media

Bacterial strains and plasmids used are shown in Table 2. *Mycobacterium tuberculosis* was cultured on Middlebrook 7H10 agar. For liquid cultures, Middlebrook 7H9 broth (Difco) plus 10% OADC supplement (Becton Dickinson) and 0.05% Tween 80 was used. Hygromycin was added to 100 μg ml^{-1}, kanamycin to 20 μg ml^{-1}, gentamicin to 10 μg ml^{-1}, and Xgal to 50 μg ml^{-1}, where appropriate. For supplementation with inositol, a 14% stock (w/v) (77 mM) of myo-inositol (Sigma) was prepared and filter-sterilised. *Escherichia coli* DH5α was used for all plasmid constructions.

Bioinformatics

Genomes and gene predictions. For *M. tuberculosis* H37Rv (Cole *et al*., 1998), genome and predicted protein encoding genes were downloaded from the EBI (http://www.ebi.ac.uk/genomes/index.html). The genome and predicted protein encoding genes for *S. coelicolor* (Bentley *et al*., 2002) and *C. diphtheriae* (Ana Cerdeño-Tarraga, pers. comm.) were obtained from the Sanger Centre (http://www.sanger.ac.uk/Projects/Microbes/) on 7 February 2003. Sequence data for *M. smegmatis* and *M. avium* were obtained from TIGR (http://www.tigr.org/) on 4 December 2002. The shotgun sequencing of *M. smegmatis* is complete, but closure of gaps still under way. Sixteen unjoined contigs were therefore downloaded. Preliminary BLAST searches

Table 2. Strains, plasmids and primers used in this study.

Strains	Characteristics	Source
M. tuberculosis H37Rv	wild-type laboratory strain	ATCC 25618
FAME5	*M. tuberculosis ino1Δ*	This study
FAME 6	FAME5::pFM109	This study
FAME 8	*M. tuberculosis* H37Rv::pFM76(SCO)	This study
FAME14	FAME5::pFM130	This study
FAME15	FAME5::pFM131	This study
FAME16	FAME5::pFM132	This study
FAME17	FAME5::pFM133	This study
FAME18	FAME5::pFM134	This study
FAME19	FAME5::pFM135	This study
E. coli DH5		Invitrogen
E. coli BL21 (DE3)		Studier *et al*. (1990)
Plasmids		
pBluescript II SK +		Stratagene
pET15b		Novagen
pUC-Gm-int	pUC-based plasmid carrying gentamycin resistance gene and phage L5 *int* gene	Lee *et al*. (1991)
p2NIL	manipulation vector	Parish and Stoker (2000)
pGOAL19	delivery cassette vector	Parish and Stoker (2000)
pFM55	*ino1* in pET15b	This study
pFM61	*ino1* and flanking area in p2NIL	This study
pFM62	pFM61 with deletion in *ino1*	This study
pFM76	pFM62 with *Pac*I cassette of pGoal19	This study
pFM108	*ino1* with 2 kb upstream in pBluescript SK +	This study
pFM109	pFM108 containing *Hin*dIII cassette of pUC-Gm-int	This study
pFM117	Mutation of D197-A on pFM55	This study
pFM118	Mutation of D197-N on pFM55	This study
pFM119	Mutation of K284-A on pFM55	This study
pFM120	Mutation of D310-A on pFM55	This study
pFM121	Mutation of D310-N on pFM55	This study
pFM122	Mutation of K346-A on pFM55	This study
pFM130	PFM124 containing *Hin*dIII cassette of pUC-Gm-int	This study
pFm131	pFM125 containing *Hin*dIII cassette of pUC-Gm-int	This study
pFM132	pFM126 containing *Hin*dIII cassette of pUC-Gm-int	This study
pFM133	pFM127 containing *Hin*dIII cassette of pUC-Gm-int	This study
pFM134	pFM128 containing *Hin*dIII cassette of pUC-Gm-int	This study
pFM135	pFM129 containing *Hin*dIII cassette of pUC-Gm-int	This study
Primers		
tb_lmsy1 (ATACCTCGTTTTCGCAGGTG)		
tb_lmsy2 (GGAAAGAAGTGGTCCCAGGT)		
tb_lms3 (CGCGGATCCGCCAGCAGGATCAGTTCTTC)		
tb_lms4 (CGCGGATCCTACCAGGCAATTGTGCGAGT)		
D197A-for (GGTACCCATCGTCGGTGATGCCATCAAGAGCCAGGTCGG)		
D197A-rev (CCGACCTGGCTCTTGATGGCATCACCGACGATGGGTACC)		
D197N-for (GGTACCCATCGTCGGTGATAACATCAAGAGCCAGGTCGG)		
K284A-for (GGATGGCTCGACGATCGCGCATGGGCCTATGTCCG)		
K284A-rev (CGGACATAGGCCCATGCGCGATCGTCGAGCCATCC)		
D310A-for (CAAGCTCGAGGTGTGGGCCTCGCCAAACTCGGCCGG)		
D310A-rev (CCGGCCGAGTTTGGCGAGGCCCACACCTCGAGCTTG)		
D310N-for (CAAGCTCGAGGTGTGGAACTCGCCAAACTCGGCCGG)		
D310N-rev (CCGGCCGAGTTTGGCGAGAGTTCCACACCTCGAGCTTG)		
K346A-for (CGTCGGCCTACCTGATGGCGAGCCCGCCCGAGCAGC)		
K346A-rev (GCTGCTCGGGCGGGCTCGCCATCAGGTAGGCCGACG)		

(Altschul *et al*., 1997) indicated that *M. smegmatis* homologues of *M. tuberculosis ino1* and surrounding genes were all located within a single *M. smegmatis* contig, Ms3312. The *M. avium* sequence was downloaded as a single complete sequence with gaps closed. Predicted protein-encoding genes were not available for *M. smegmatis* and *M. avium*, so we used the GeneMark.hmm gene prediction algorithm based on an *M. tuberculosis* model (Lukashin and Borodovsky, 1998) (http://opal.biology.gatech.edu/GeneMark/gmhmm2_prok.cgi). The algorithm identified 5244 putative genes in *M. avium*, and 1667 putative genes in contig 3312 of *M. smegmatis*. For all these organisms, known and predicted genes were translated into proteins, and stored in a database suitable for BLAST searching (Altschul *et al*., 1997).

BLAST searches

The gene products of the ten *M. tuberculosis* genes illustrated in Fig. 1 (*leuS, Rv0042c, Rv0043c, Rv0044c, Rv0045c, ino1, Rv0047c, Rv0048c, Rv0049* and *ponA1*) were all searched against the protein databases for the other

organisms illustrated in Fig. 1. Searches were performed using BLASTP (Altschul *et al*., 1997) with default settings. Further *ad hoc* BLASTP searches were performed, such as those done to confirm that the two extra genes identified in *M. smegmatis* are likely orthologues of those in *M. avium* (see Fig. 1 and *Results*).

Generation of ACT comparison files

To help with the process of identifying the cross-species homologies for *M. tuberculosis ino1* and surrounding genes, the Artemis Comparison Tool (ACT, http://www.sanger.ac.uk/Software/ACT/) was used. Artemis Comparison Tool requires at least two genomes, and a file describing homology between these genomes, usually generated using the BLASTN (complete nucleotide genome versus complete nucleotide genome) or TBLASTX (target and query complete nucleotide genomes both translated in all six reading frames and compared). The ACT files were generated for all pairwise comparisons between *M. tuberculosis* and the other organisms shown in Fig. 1 using BLASTN, and for a subset of these, using TBLASTX. For comparisons with *M. leprae*, we used comparison files generated at the Sanger Centre using TBLASTX (http://www.sanger.ac.uk/Projects/M_leprae/comparison.shtml)

Real-time PCR

RNA was prepared from a 7-day rolling culture as described previously (Betts *et al*., 2002). cDNA synthesis was carried out using Superscript II (Invitrogen) according to the manufacturer's protocol. Real-time quantitative PCR reactions were set up using the DyNAmo SYBR Green qPCR kit (MJ Research) with a DNA Engine Opticon® 2 System (GRI). Reactions were set up on ice containing 1 × DNA Master SYBR Green I mix, 1 μl of cDNA product and 0.4 μM of each primer in a final volume of 20 μl. These were heated to 95°C for 10 min before cycling for 35 cycles of 95°C for 30 s, 60°C (*ino*) or 62°C (*sigA*) for 20 s, and 72°C for 20 s. Fluorescence was measured at the end of each cycle after heating to 80°C to ensure the denaturation of primer-dimers. Triplicate reactions for each gene were carried out on cDNA from two separate RNA preparations. Confidence intervals were calculated using the residual error of a factorial ANOVA analysis for log transformed mRNA levels with factors gene and preparation.

Cloning, mutagenesis and complementation of M. tuberculosis ino1

Mutagenesis was carried out essentially as described previously (Parish and Stoker, 2000). The coding sequence of *M. tuberculosis* H37Rv *ino1* (*Rv0046c*) (with flanking DNA, 758 bp upstream and 1144 bp downstream of *ino1*) was amplified by PCR using *Pfu Turbo* DNA polymerase enzyme (Stratagene). The primers used, each at 300 nM, were tb_Imsy1 and tb_Imsy2 (Table 2), and cycling conditions were: an initial 3 min at 95, then 35 cycles of 1 min at 95°C, 1 min at 57°C and 3 min at 72°C, and finally a last extension step of 72°C for 10 min to complete primer extension. The PCR product was phosphorylated using polynucleotide kinase and cloned into the *Pml*I site of plasmid p2NIL, producing pFM61. A 900 bp deletion was generated in the 1101 bp gene by digestion with *Pvu*II and *Pml*I followed by religation. Following insertion of a gene cassette carrying *lacZ* and *sacB* from pGOAL19 into the vector's *Pac*I site, the DNA was UV-irradiated (100 Jm^{-2}) and introduced into *M. tuberculosis* H37Rv by electroporation. Cells carrying single-crossovers (SCOs) were isolated by selection for blue hyg^R kan^R colonies on Middlebrook 7H10 agar containing OADC supplement (Difco). One SCO colony was plated onto agar containing sucrose (2%) to isolate bacteria with a second crossover, which will lead to mutant or wild-type cells depending on the location of the recombination event. In the absence of inositol, only wild-type colonies were isolated. However, in the presence of 77 mM inositol, four out of eight colonies were shown by PCR to be *ino1* mutants, which were confirmed by Southern analysis. One of these (FAME5) was analysed further.

In order to complement the *ino1* mutant, *Rv0046c* was cloned together with *Rv0047c* and *Rv0048c*, to ensure its promoter was present. Polymerase chain reaction was carried out as above except that 2% DMSO was included, the primers used were tb_Ims3 and tb_Ims4, and the temperature cycle used was: an initial 3 min at 95°C, then 40 cycles of 1 min at 95°C, 1 min at 54°C and 3 min 20 s at 72°C, and finally a last extension step of 72°C for 10 min to complete primer extension. The PCR product was phosphorylated and cloned into *Sma*I site of pBluescript to produce pFM108. Then *Hin*dIII cassette of pUC-Gm-int which has the integration part and gentamicin gene cloned in *Hin*dIII site of pFM108 to produce pFM109.

Growth of FAME5

Strains were grown in 7H9 media + OADC (and 77 mM inositol if needed) to an A_{600} of 0.2–0.8, then subcultured into 1 : 100 in 100 ml of the appropriate medium, and incubated in rolling culture. The ODs were read at the times shown.

Extraction and analysis of PIMs

Cells (0.2 g) were delipidated with chloroform/methanol (1/1, v/v) for 48 h at room temperature with continuous stirring. Lipids were separated from the delipidated cells by centrifugation (3000 r.p.m., 15 min, 2600 *g*) and analysed by TLC on silica gel-coated plates developed with chloroform/methanol/water (60 : 35 : 8, v/v/v). The various PIMs were identified by their mobilities on TLC and their positive reactivity with both a sugar-specific reagent (0.2% anthrone in concentrated H_2SO_4 followed by heating) and the Dittmer-Lester reagent that specifically detects phosphorous-containing lipids (Dittmer and Lester, 1964).

Production and analysis of LAM and LM

Delipidated cells were washed and disrupted using a cell disrupter (2 kbars, Constant System Ltd; one shot model). The resulting material was extracted with 40 ml of ethanol/water (1/1, v/v) for 8 h at 65°C; the bacterial residues were discarded and the supernatant was dried. Six ml hot phenol/

water (1/1, v/v) was added and the mixture was heated for 1 h at 70°C under continuous stirring followed by a two-phase partition. The phenol phase was discarded and the upper phase extensively washed and dried. The extract was solubilized in water and Triton X114 (2% wt/v) was added to the cooled suspension. The mixture was stirred for 10 min and then heated at 50°C until two phases formed. The detergent phase was recovered, diluted by adding 1 ml water and washed three times with $CHCl_3$. The resulting aqueous phase was dried to evaporate the chloroform and resuspended in water (0.2 ml). This portion was analysed by SDS-PAGE with a 5% stacking gel and a 15% running gel. Samples were denatured in the presence of 2% SDS in 50 mM Tris-HCl (pH 6.8). After electrophoresis, gels were treated with periodate/ethanol/acetic acid (0.7/40/5, w/v/v), and silver-stained. Authentic samples of mycobacterial LAM and LM from *Mycobacterium bovis* BCG were used as standard.

Sugar compositional analysis

The sugar constituents of the various materials were determined either after acid hydrolysis with 2 M CF_3COOH at 110°C for 1 h; the mixture of hydrolysed products was dried, treated with trimethylsilyl reagents (Sweeley *et al*., 1963) to derivatize monosaccharides and analysed by gas chromatography (GC) for their sugars.

Gas chromatography and mass spectrometry

Gas chromatography was performed using a Hewlett Packard HP4890A equipped with a fused silica capillary column (25 m length × 0.22 mm i.d) containing WCOT OV-1 (0.3 mm film thickness, Spiral). A temperature gradient of 100–290°C at 5°C min^{-1}, followed by a 10 min isotherm plateau at 290°C, was used.

Mycothiol assay

Labelling of cell extracts with monobromobimane (mBBr) to determine thiol content was performed with modifications to previously published protocols (Newton *et al*., 1996; Anderberg *et al*., 1998). Cell pellets from 3 ml culture were resuspended in 0.5 ml of warm 50% acetonitrile-water containing 2 mM mBBr (Cal Biochem), and 20 mM Hepes-HCl, pH 8.0. The suspension was incubated for 15 min in a 60°C water bath and then cooled on ice. 2–5 μl of 5 M HCl or 5 M trifluoracetic acid was added to produce a final acidic pH.

The control samples were extracted with 0.5 ml of warm 50% acetonitrile-water containing 5 mM NEM and 20 mM Hepes-HCl, pH 8.0 The suspension was incubated for 15 min in a 60°C water bath and then cooled on ice. Two millimolar mBBR was added to the solution followed by a second incubation for 15 min at 60°C. The control sample was cooled but not acidified. Cell debris were pelleted in each sample by centrifugation (5 min 14 000 *g*).

High performance liquid chromatography analysis of thiols was carried out by injecting 25 μl of 1 : 4 dilution of samples in 10 mM HCl on to a Beckman Ultrasphere IP 5_ (250 mm × 4.6 mm) column using 0.25% glacial acetic acid pH 3.6 (buffer A) and 95% methanol (buffer B). The gradient was: 0 min, 10% B; 15min, 18% B; 30min, 27% B; 32min, 100% B; 34min,10% B; and 60min, 10% B (reinjection). The flow rate was 1 ml min^{-1}, and the fluorescence detection was accomplished on a Varian Fluorichrom model 430020 with a 370-nm excitation filter and a 418–700 nm emission filter. Data collection and analysis was performed on Dynamax Mac Integrator (Rainin Instruments).

Macrophage infections

Monolayers of bone marrow-derived macrophages were established in the wells of 24-well plates (10^5 cells per well). Bacteria were grown with supplement where necessary, and washed and frozen before infection. The macrophages were infected with bacteria at a multiplicity of 2.5 : 1 in triplicate for 2 h. The monolayers were washed and incubated for the time period indicated. Samples were harvested by rinsing the well and lysing the cells in 1 ml 0.5% Tween 20 + 0.05% SDS in H_20. The cell lysate was subjected to reciprocal dilution in Middlebrook medium and plated on quadrant plates with MB7H11 agar (supplemented with 70 mM inositol). The colonies were scored and displayed as a mean for the triplicate wells with standard deviation indicated by vertical lines.

In vivo *studies*

The methodology for infecting SCID mice has been described in detail previously (Smith *et al*., 2001). Briefly, mice were infected with 1×10^6 viable mycobacteria in 200 μl of pyrogen-free saline via a lateral tail vein. At the time of inoculation or *in vitro* experimentation the mycobacteria were routinely plated to verify input colony forming units (CFUs). Where appropriate, infected mice were killed by cervical dislocation in accordance with humane endpoint protocols under the Animals Scientific Procedures Act, 1986 (UK). Median survival times were calculated for each group and statistical analysis performed using the Log Rank tests of survival.

Site-directed mutagenesis

Site-directed mutagenesis was carried out using the non-PCR-based Quickchange kit (Stratagene) to introduce six point mutations into the pFM108, producing pFM124-129. Following sequence-confirmation, the *Hin*dIII cassette of pUC-Gm-int were cloned into the above constructs, producing pFM130-135. These constructs were transformed into FAME5 in the presence of 77 mM inositol.

The same primers were used to introduce the six mutations into the expression construct pFM55, which produces a his-tagged tbINO. After an initial 30 s at 95°C, 12 cycles (95°C 30 s, 55°C 1 min, 68°C 13 min) were carried out. This was digested with *Dpn*I, transformed into *E. coli* and plasmids (pFM117-122) from resultant colonies sequenced to confirm the presence of the mutation.

Structural analyses

Mutant and wild-type tbINO were prepared as described previously (Norman *et al*., 2002). In brief, both mutant and wild-type proteins were prepared in *E. coli* BL21(DE3) cells using a pET-15b expression vector. Purification involved a Ni-NTA affinity column followed by size exclusion chromatography.

Finally a Q-sepharose ion exchange column gave pure mutant proteins that crystallised under similar conditions to the wild-type protein. Mutant datasets were collected at the ESRF synchrotron on stations 14.1 and 14.4 using an ADSC CCD detector. Data were processed with DENZO-SCALEPACK package (Otwinowski and Minor, 1997). The structures were solved by molecular replacement using the program MOLREP as part of the CCP4 package (1994), and refined with CNS (Brunger *et al*., 1998).

The D310A mutant was refined to 2.3 Å with an R factor 21.1 of and an R_{free} of 25.9. The D310N mutant was refined to 2.1 Å with an R factor 22.6 of and an R_{free} of 25.0. Atomic coordinates will be deposited at proof stage.

Acknowledgements

We would like to thank Jane Turner, Heidi Alderton and the staff of the Biological Services Facility at the London School of Hygiene and Tropical Medicine. Lorenz Wernisch provided statistical advice. This work was supported by Wellcome Trust grants 058810 (F.M.), 062508 (S.C.G.R.) and 068969, the European Union Vth Framework Program consortia QLK2-CT-1999–01093 (F.M.) and QLRT-1999–00099 (J.M.-R.), and BBSRC grant 48/P18545 (S.L.K.). D.S. was supported by funding from GlaxoSmithKline. We gratefully acknowledge the help of staff and use of facilities at ESRF, Grenoble, France. Y. A.-G. was supported by the TB Veterans Association. Preliminary sequence data was obtained from The Institute for Genomic Research through the website at http://www.tigr.org. Preliminary *C. diphtheriae* sequence data, gene predictions and comparison files were obtained from Ana Cerdeño-Tarraga and Julian Parkhill at the Sanger Centre.

References

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25:** 3389–3402.

Anderberg, S.J., Newton, G.L., and Fahey, R.C. (1998) Mycothiol biosynthesis and metabolism. Cellular levels of potential intermediates in the biosynthesis and degradation of mycothiol in *Mycobacterium smegmatis*. *J Biol Chem* **273:** 30391–30397.

Bachhawat, N., and Mande, S.C. (1999) Identification of the INO1 gene of *Mycobacterium tuberculosis* H37Rv reveals a novel class of inositol-1-phosphate synthase enzyme. *J Mol Biol* **291:** 531–536.

Bentley, S.D., Chater, K.F., Cerdeno-Tarraga, A.M., Challis, G.L., Thomson, N.R., James, K.D., *et al.* (2002) Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3 (2). *Nature* **417:** 141–147.

Berry, G.T., Wu, S., Buccafusca, R., Ren, J., Gonzales, L.W., Ballard, P.L., *et al.* (2003) Loss of murine Na+/myo-inositol cotransporter leads to brain myo-inositol depletion and central apnea. *J Biol Chem* **278:** 18297–18302.

Betts, J.C., Lukey, P.T., Robb, L.C., McAdam, R.A., and Duncan, K. (2002) Evaluation of a nutrient starvation model of *Mycobacterium tuberculosis* persistence by gene and protein expression profiling. *Mol Microbiol* **43:** 717–731.

Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., *et al.* (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* **54:** 905–921.

CCP4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr* **D50:** 760–763.

Chambers, M.A., Williams, A., Gavier-Widen, D., Whelan, A., Hall, G., Marsh, P.D., *et al.* (2000) Identification of a *Mycobacterium bovis* BCG auxotrophic mutant that protects guinea pigs against M. bovis and hematogenous spread of *Mycobacterium tuberculosis* without sensitization to tuberculin. *Infect Immun* **68:** 7094–7099.

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.

Cole, S.T., Eiglmeier, K., Parkhill, J., James, K.D., Thomson, N.R., Wheeler, P.R., *et al.* (2001) Massive gene decay in the leprosy bacillus. *Nature* **409:** 1007–1011.

Dittmer, J.C.F., and Lester, R.L. (1964) A simple specific spray for the detection of phospholipids on thin layer chromatography. *J Lipid Res* **5:** 126–127.

Esnouf, R.M. (1999) Further additions to Molscript, Version 1.4. including reading and contouring of electron-density maps. *Acta Crystallogr D Biol Crystallogr* **55:** 938–940.

Fahey, R.C. (2001) Novel thiols of prokaryotes. *Annu Rev Microbiol* **55:** 333–356.

Gordhan, B.G., Smith, D.A., Alderton, H., McAdam, R.A., Bancroft, G.J., and Mizrahi, V. (2002) Construction and phenotypic characterization of an auxotrophic mutant of *Mycobacterium tuberculosis* defective in L-arginine biosynthesis. *Infect Immun* **70:** 3080–3084.

Hansch, H.C., Smith, D.A., Mielke, M.E., Hahn, H., Bancroft, G.J., and Ehlers, S. (1996) Mechanisms of granuloma formation in murine *Mycobacterium avium* infection: the contribution of CD4+ T cells. *Int Immunol* **8:** 1299–1310.

Henry, S.A., Atkinson, K.D., Kolat, A.I., and Culbertson, M.R. (1977) Growth and metabolism of inositol-starved *Saccharomyces cerevisiae*. *J Bacteriol* **130:** 472–484.

Hondalus, M.K., Bardarov, S., Russell, R., Chan, J., Jacobs, W.R. Jr and Bloom, B.R. (2000) Attenuation of and protection induced by a leucine auxotroph of *Mycobacterium tuberculosis*. *Infect Immun* **68:** 2888–2898.

Hunter, S.W., Gaylord, H., and Brennan, P.J. (1986) Structure and antigenicity of the phosphorylated lipopolysaccharide antigens from the leprosy and tubercle bacilli. *J Biol Chem* **261:** 12345–12351.

Jackson, M., Phalen, S.W., Lagranderie, M., Ensergueix, D., Chavarot, P., Marchal, G., *et al.* (1999) Persistence and protective efficacy of a *Mycobacterium tuberculosis* auxotroph vaccine. *Infect Immun* **67:** 2867–2873.

Jackson, M., Crick, D.C., and Brennan, P.J. (2000) Phosphatidylinositol is an essential phospholipid of mycobacteria. *J Biol Chem* **275:** 30092–30099.

Lee, M.H., Pascopella, L., Jacobs, W.R. Jr and Hatfull, G.F. (1991) Site-specific integration of mycobacteriophage L5: integration- proficient vectors for *Mycobacterium smegmatis*, *Mycobacterium tuberculosis*, and bacille

Calmette-Guerin. *Proc Natl Acad Sci USA* **88:** 3111–3115.

Lukashin, A.V., and Borodovsky, M. (1998) GeneMark.hmm: new solutions for gene finding. *Nucleic Acids Res* **26:** 1107–1115.

Manganelli, R., Dubnau, E., Tyagi, S., Kramer, F.R., and Smith, I. (1999) Differential expression of 10 sigma factor genes in *Mycobacterium tuberculosis*. *Mol Microbiol* **31:** 715–724.

McAdam, R.A., Quan, S., Smith, D.A., Bardarov, S., Betts, J.C., Cook, F.C., *et al.* (2002) Characterization of a *Mycobacterium tuberculosis* H37Rv transposon library reveals insertions in 351 ORFs and mutants with altered virulence. *Microbiology* **148:** 2975–2986.

Merritt, E.A., and Murphy, M.E.P. (1994) Raster3d, version 2.0: A program for photorealistic molecular graphics. *Acta Cryst D Biol Crystallogr* **50:** 869–873.

Newton, G.L., Bewley, C.A., Dwyer, T.J., Horn, R., Aharonowitz, Y., Cohen, G., *et al.* (1995) The structure of U17 isolated from *Streptomyces clavuligerus and* its properties as an antioxidant thiol. *Eur J Biochem* **230:** 821–825.

Newton, G.L., Arnold, K., Price, M.S., Sherrill, C., Delcardayre, S.B., Aharonowitz, Y., *et al.* (1996) Distribution of thiols in microorganisms: mycothiol is a major thiol in most actinomycetes. *J Bacteriol* **178:** 1990–1995.

Norman, R.A., McAlister, M.S., Murray-Rust, J., Movahedzadeh, F., Stoker, N.G., and McDonald, N.Q. (2002) Crystal structure of inositol 1-phosphate synthase from *Mycobacterium tuberculosis*, a key enzyme in phosphatidylinositol synthesis. *Structure (Camb)* **10:** 393–402.

Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. In *Macromolecular Crystallography, Part A*, Vol. 276. Carter, C.W. Jr and Sweet, R.M., (eds). New York: Academic Press, pp. 307–326.

Parish, T., and Stoker, N.G. (2000) Use of a flexible cassette method to generate a double unmarked *Mycobacterium tuberculosis tlyA plcABC* mutant by gene replacement. *Microbiology* **146:** 1969–1975.

Parish, T., Smith, D.A., Kendall, S.L., Casali, N., Bancroft, G.J., and Stoker, N.G. (2003) Deletion of two-component regulatory systems increases virulence of *Mycobacterium tuberculosis*. *Infect Immun* **71:** 1134–1140.

Pym, A.S., Brodin, P., Brosch, R., Huerre, M., and Cole, S.T. (2002) Loss of RD1 contributed to the attenuation of the live tuberculosis vaccines *Mycobacterium bovis* BCG and *Mycobacterium microti*. *Mol Microbiol* **46:** 709–717.

Rosenkrands, I., King, A., Weldingh, K., Moniatte, M., Moertz, E., and Andersen, P. (2000) Towards the proteome of *Mycobacterium tuberculosis*. *Electrophoresis* **21:** 3740–3756.

Russell, D.G., Sturgill-Koszycki, S., Vanheyningen, T., Collins, H., and Schaible, U.E. (1997) Why intracellular parasitism need not be a degrading experience for *Mycobacterium*. *Philos Trans R Soc Lond B Biol Sci* **352:** 1303–1310.

Sambandamurthy, V.K., Wang, X., Chen, B., Russell, R.G., Derrick, S., Collins, F.M., *et al.* (2002) A pantothenate auxotroph of *Mycobacterium tuberculosis* is highly attenuated and protects mice against tuberculosis. *Nat Med* **8:** 1171–1174.

Santos, H., and da Costa, M.S. (2002) Compatible solutes of organisms that live in hot saline environments. *Environ Microbiol* **4:** 501–509.

Sassetti, C.M., Boyd, D.H., and Rubin, E.J. (2003) Genes required for mycobacterial growth defined by high density mutagenesis. *Mol Microbiol* **48:** 77–84.

Schnappinger, D., Ehrt, S., Voskuil, M.I., Liu, Y., Mangan, J.A., Monahan, I.M., *et al.* (2003) Transcriptional adaptation of *Mycobacterium tuberculosis* within macrophages: Insights into the phagosomal environment. *J Exp Med* **198:** 693–704.

Smith, D., Hansch, H., Bancroft, G., and Ehlers, S. (1997) T-cell-independent granuloma formation in response to *Mycobacterium avium*: role of tumour necrosis factor-alpha and interferon-gamma. *Immunology* **92:** 413–421.

Smith, D.A., Parish, T., Stoker, N.G., and Bancroft, G.J. (2001) Characterization of auxotrophic mutants of *Mycobacterium tuberculosis* and their potential as vaccine candidates. *Infect Immun* **69:** 1142–1150.

Stein, A.J., and Geiger, J.H. (2002) The crystal structure and mechanism of 1-L-myo-inositol-1-phosphate synthase. *J Biol Chem* **277:** 9484–9491.

Studier, F.W., Rosenberg, A.H., Dunn, J.J., and Dubendorff, J.W. (1990) Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol* **185:** 60–89.

Sweeley, C.C., Bentley, R., Makita, M., and Wells, W.W. (1963) Gas-liquid chromatography of trimethylsilyl derivatives of sugars and related substances. *J Am Chem Soc* **85:** 2497–2507.

Tsitsanou, K.E., Oikonomakos, N.G., Zographos, S.E., Skamnaki, V.T., Gregoriou, M., Watson, K.A., *et al.* (1999) Effects of commonly used cryoprotectants on glycogen phosphorylase activity and structure. *Protein Sci* **8:** 741–749.

WHO (2000) *Tuberculosis Fact Sheet no. 104.* Geneva: World Health Organization.