BMC Microbiology

BioMed Central

Research article

Identification of two *Mycobacterium tuberculosis* H37Rv ORFs involved in resistance to killing by human macrophages

Barbara H Miller[1,2] and Thomas M Shinnick*[1]

Address: [1]Department of Microbiology and Immunology, Emory University School of Medicine and Division of AIDS, STD, and TB Laboratory Research, Centers for Disease Control and Prevention, Atlanta, Georgia, USA and [2]Department of Internal Medicine, University of Michigan School of Medicine, Ann Arbor, MI

E-mail: Barbara H Miller - bhmiller@med.umich.edu; Thomas M Shinnick* - tms1@cdc.gov

*Corresponding author

Published: 17 October 2001

BMC Microbiology 2001, **1**:26

Received: 6 September 2001

Accepted: 17 October 2001

This article is available from: http://www.biomedcentral.com/1471-2180/1/26

Abstract

Background: The ability of *Mycobacterium tuberculosis* to survive and replicate in macrophages is crucial for the mycobacterium's ability to infect the host and cause tuberculosis. To identify *Mycobacterium tuberculosis* genes involved in survival in macrophages, a library of non-pathogenic *Mycobacterium smegmatis* bacteria, each carrying an individual integrated cosmid containing *M. tuberculosis* H37Rv genomic DNA, was passed through THP-1 human macrophages three times.

Results: Two of the clones recovered from this enrichment process, sur2 and sur3, exhibited significantly increased survival relative to wild-type bacteria. In coinfection experiments, the ratio of sur2 colonies to wild-type colonies was 1:1 at 0 hours but increased to 20:1 at 24 hours post phagocytosis. The ratio of sur3 colonies to wild-type colonies was 1:1 at 0 hours and 5:1 at 24 hours. The *M. tuberculosis* ORFs responsible for increased survival were shown to be *Rv0365c* for the sur2 clone and *Rv2235* for the sur3 clone. These ORFs encode proteins with as-of-yet unknown functions.

Conclusions: We identified two *M. tuberculosis* ORFs which may be involved in the ability of tubercle bacilli to survive in macrophages.

Background

Today, an estimated one-third of the world's population is infected with *Mycobacterium tuberculosis*, the causative agent of tuberculosis [1,2], and tuberculosis causes about 2 million deaths annually [1,2]. Many aspects of the interactions between *M. tuberculosis* and its human host remain unclear. Not only is this bacterium able to evade the defenses of the host's immune system, it is also able to persist in the body for years and may reactivate to cause disease decades after the initial infection. A better understanding of the interaction between *M. tuberculosis* and its human host is critical to developing new strategies to control the tuberculosis epidemic.

A key feature of the pathogenicity of *M. tuberculosis* is its ability to evade the antimicrobial processes of the macrophage and replicate intracellularly. Mycobacteria enter macrophages primarily by conventional receptor-mediated phagocytic pathways [3]. Following phagocytosis, phagosomes containing viable tubercle bacilli fail to acidify, apparently because of failure to insert a proton-ATPase pump into the phagosomal membrane [4–6].

This leads to altered vacuolar maturation such that phagosome-lysosome fusion is blocked and the mycobacteria-containing vacuoles end up with an internal pH of about 6 and markers of phagosomes and early and late endosomes [7–11]. The mycobacterial genes responsible for these processes are largely unknown.

One method for identifying bacterial genes involved in pathogenesis is to express these genes in a nonpathogenic host and isolate bacteria with increased virulence. This technique was first used to identify a gene that enables *Yersinia pseudotuberculosis* to invade HEp-2 cells [12]. In these studies, *Escherichia coli* bacteria expressing *Y. pseudotuberculosis* genes were used to infect HEp-2 cells. Only *E. coli* bacteria expressing the *Y. pseudotuberculosis inv* gene were able to invade the animal cells [12]. Using a similar approach, Arruda *et al.* [13] identified an *M tuberculosis* gene responsible for invasion of HeLa cells. We have previously used a similar technique to identify *M. leprae* genes involved in intracellular survival [14]. That is, by expressing *M. leprae* genes in *E. coli,* we were able to isolate recombinant bacteria that exhibited increased resistance to killing by murine bone-marrow derived macrophages [14].

A similar approach was used by Wei et al [15] to isolate 21 *Mycobacterium smegmatis* recombinant clones that displayed a greater than 2-fold enhancement in survival after 48 hours. *M. smegmatis* is a fast-growing, nonpathogenic species of *Mycobacterium* in which *M. tuberculosis* genes can be efficiently expressed [16–18]. An in depth analysis of one recombinant clone revealed that *M. smegmatis* recipients carrying the *M. tuberculosis eis* gene on an extrachromosomal multicopy plasmid displayed 2.4- to 5.3-fold greater survival in U937 macrophages than wild-type *M. smegmatis* bacteria at 24 to 48 hours post-infection [15]. The 42-kDa *eis* gene product has been shown to be associated with the mycobacterial cell surface and is released into extracellular medium, but its precise function is not yet known [19].

In the studies reported here, we used a similar enrichment scheme to identify two additional recombinant *M. smegmatis* clones, sur2 and sur3, that demonstrated enhanced survival during infections of THP-1 human monocyte-derived macrophages. The *M. tuberculosis* ORFs responsible for increased survival were shown to be *Rv0365c* for the sur2 clone and *Rv2235* for the sur3 clone. These ORFs encode proteins with as-of-yet unknown functions.

Results

Cosmids carrying *M. tuberculosis* genes [20] were electroporated into *M. smegmatis* LR222 to create a library of *M. smegmatis* transformants, each of which carries a pYUB178::H37Rv cosmid integrated into its chromosome. The library was generated from about 4000 independent transformants. This represents about 20 *M. tuberculosis* genome-equivalents given that ~225 cosmids contain one genome-equivalent of *M. tuberculosis* [20].

To enrich for clones with increased survival in human macrophages, the library was passed through THP-1 macrophages as shown schematically in Figure 1. THP-1 is a human monocyte-derived macrophage cell line [21] and wild-type *M. smegmatis* bacteria are rapidly killed by THP-1 macrophages [22]. After three rounds of enrichment, the resulting clones were analyzed individually by Southern blot to evaluate the number of different clones present as previously described [20]. Out of the 3000 colonies recovered, genomic DNAs from the bacteria of 90 randomly chosen colonies were digested with *Pst*I and hybridized with the pYUB178 vector. Strains carrying different cosmids should exhibit a different pattern of hybridizing bands because of the presence of different *M. tuberculosis* H37Rv inserts. Two clones, designated sur2 and sur3, were each present three times and were chosen for further study. The remaining 84 clones each displayed unique patterns.

As a first step in the analysis of the sur2 clone, the time course of its survival in THP-1 macrophages was determined by infecting THP-1 macrophages, lysing the infected macrophages at various times after phagocytosis, and enumerating viable intracellular bacteria by plating on solid medium. Both parental and the sur2 bacteria were rapidly killed during the first few hours after phagocytosis (Figure 2). The sur2 bacteria appeared to survive slightly better at the 9 hour time point, but the difference was not statistically significant.

To compare directly the relative ability of the sur2 and wild-type bacteria to survive in macrophages, THP-1 macrophages were infected with a mixture of a genetically marked control strain and the sur2 strain and the survival of each strain was followed independently as previously described [22]. The wild-type strain expresses the *xylE* gene product, catechol 2,3-dioxygenase, such that when its colonies are sprayed with catechol, they turn bright yellow, while wild-type colonies remain white. In essence, each well of the experiment contains an internal standard (the wild-type bacteria) to which to compare the survival of the recombinant bacteria. In experiments in which THP-1 cells were infected at a MOI of 50:1 (results in ~1 phagocytosed *M. smegmatis* bacterium per macrophage) with a mixture containing equal numbers of bacteria of the *xylE*-expressing control strain and a strain carrying the cosmid vector pYUB178, the ratio of recovered white colonies to yellow colonies was 1:1

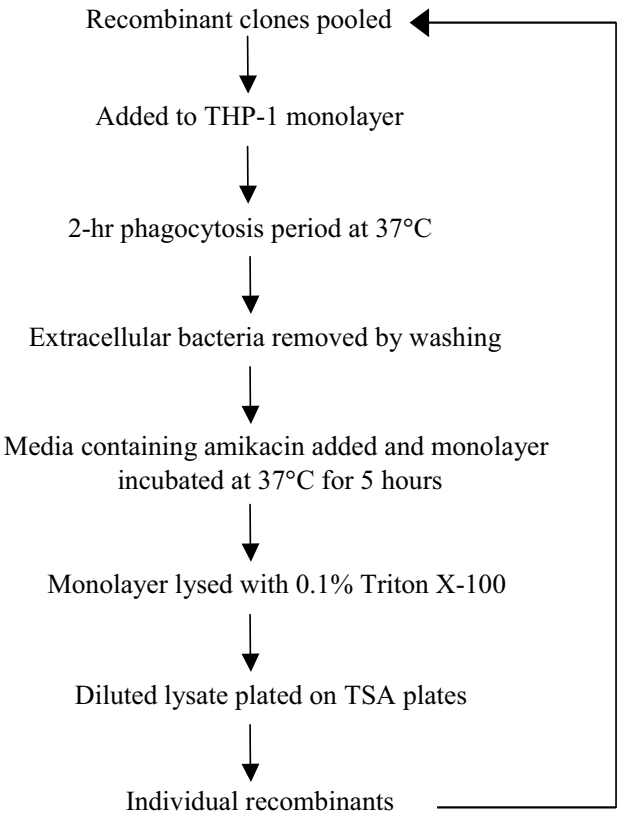

Figure 1
Enrichment procedure.

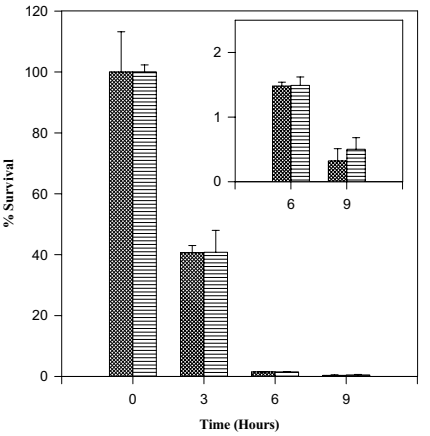

Figure 2
Survival of sur2. THP-1 macrophages were infected with bacteria containing pYUB178 (hatched) or sur2 bacteria (horizontal stripe). Time zero is defined as immediately after the phagocytosis interval. Percent survival at time × was calculated by dividing the number of CFUs recovered at time × by the number of CFU recovered at time zero and multiplying by 100.

at all time points (data not shown). This indicates that the survival of the *xylE*-expressing strain was the same as that of the wild-type and could be used an internal reference by which to measure the survival of other clones.

In coinfection experiments with *xylE*-expressing bacteria, both sur2 and sur3 bacteria exhibited increased survival (Figure 3). Immediately after the 2 hour phagocytosis period (0 hr time point) the ratio of sur2 colonies to *xylE*-expressing colonies was 1:1 and by 12 hours it was about 7:1. By 24 hours, the ratio was approximately 20:1. The differences between the ratios at the zero time point and the subsequent time points were statistically significant ($p < 0.005$) for the 9, 12, and 24 hr time points. The ratio of sur3 colonies to control colonies increased from 1:1 at time 0 to 4:1 at 12 hours and to 5:1 at 24 hours ($p < 0.005$).

Because the recombinant clones contain integrated pYUB178::H37Rv cosmids, the following strategy was used to isolate cosmids corresponding to those in the sur2 and sur3 clones. First, genomic DNA from the sur2 clone was digested with *Pst*I, and genomic DNA from the sur3 clone was digested with *Bam*HI to generate fragments of each integrated cosmid carrying *oriE, aph,* and a portion of the *M. tuberculosis* genomic DNA insert. The presence of *oriE* allows the recombinant to replicate as a plasmid in *E. coli*. The digestion products were treated with T4 DNA ligase and transformed into *E. coli* XL 1-Blue. Plasmid DNA from the resulting kanamycin-resistant colonies were analyzed by restriction site mapping and partial sequencing of the *M. tuberculosis* genomic DNA insert as described in Materials and Methods. For the sur2 recombinant, PCR primers were designed to amplify a 0.8 kb region of the *M. tuberculosis* insert from *M. tuberculosis* genomic DNA for use as a probe in colony blot experiments. Probing colony blots of the *E. coli* (pYUB178::H37Rv) library with the 0.8 kb PCR fragment led to the isolation of a 4 kb plasmid. This plasmid contained an intact *oriE* and *aph* gene, a portion of the integrase gene, and 1.1 kb of *M. tuberculosis* genomic DNA (Figure 4). PCR, DNA sequence, and Southern blot data indicated that the recovered plasmid was the same as the cosmid integrated in the sur2 genome (data not shown). For example, sequencing of PCR amplicons of the junctions between the mycobacterial sequences and vector sequences revealed that the junctions in the recovered plasmid were identical to those in the sur2 genomic DNA.

Comparison of the sequence of the cloned 1.1 kb fragment with the *M. tuberculosis* H37Rv genome sequence [23] revealed that it contains portions of the *Rv0366c* and *Rv0365c* genes (Figure 4). In the sur2 clone, the

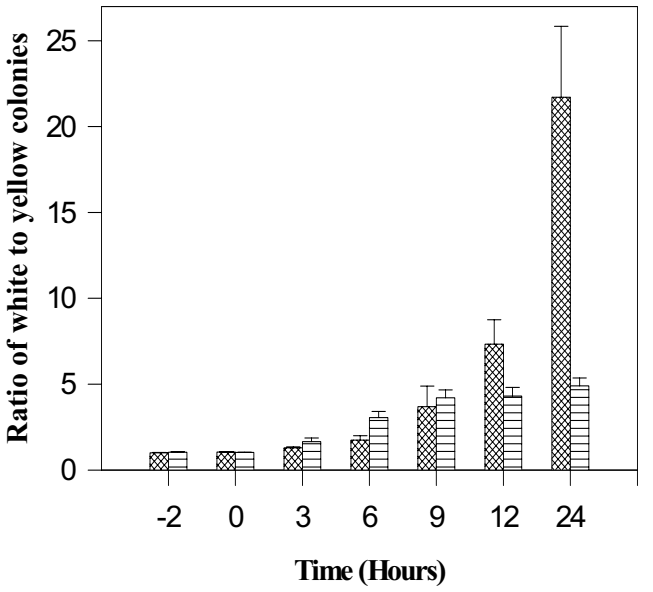

Figure 3
Survival of recombinants relative to wild-type *M. smegmatis*. THP-1 macrophages were infected with an equal mixture of *xylE*-expressing bacteria and sur2 bacteria (hatched) or sur3 bacteria (horizontal stripe). The ratio of the recovered white-to-yellow colonies is shown for -2 hours (initial inoculum), 0 hours (immediately after phagocytosis interval), and at 3, 6, 9, 12, and 24 hours after phagocytosis. The ratios represent the average of the results of three independent experiments. Error bars represent the standard deviation in the ratio of white-to-yellow colonies between experiments.

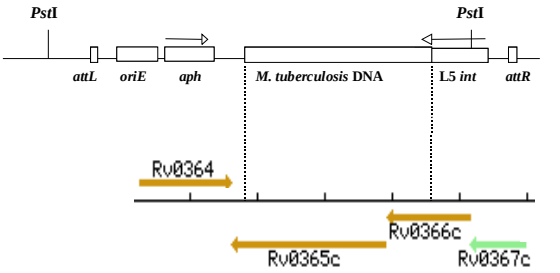

Figure 4
Map of the *M. tuberculosis* H37Rv DNA integrated in the sur2 genome. The integrated 4 kb cosmid is between the two attachment sites, *attL* and *attR*. The 4.4 kb *Pst*I fragment of sur2 is indicated by the two *Pst*I sites. The 1,106 bp *M. tuberculosis* H37Rv genomic DNA insert contains two ORFs designated *Rv0365c* and *Rv0366c* [Reference 23]. The arrows indicate the direction of transcription. Arrows are color coded according to annotations found on the TubercuList server [http://genolist.Pasteur.fr/TubercuList/] : light green (unknown), brown (conserved), yellow (intermediary metabolism), light blue (regulatory), gray (virulence), red (information pathway), dark green (cell wall process), black (lipid metabolism), and dark blue (stable RNA).

amino-terminal 970 bp of the 1128 bp *Rv0365c* ORF is fused to 14 bp from the pYUB178 vector to generate an ORF encoding 328 amino acids (aa), compared to 376 aa encoded by the full-length *Rv0365c* ORF. The insert also contains 104 bp of the *Rv0366c* ORF fused to 220 bp of the L5 integrase ORF. This ORF could encode a 108 aa protein which contains the 73 amino-terminal amino acids of the 344 aa L5 integrase protein fused to the 34 carboxyl-terminal amino acids of the *Rv0366c* ORF.

To determine if the *Rv0365c* gene was responsible for the increased survival of sur2 bacteria, the full-length *M. tuberculosis Rv0365c* ORF as well as a truncated ORF similar to that present in the insert in the sur2 clone were subcloned into the pHIP vector downstream of the *M. tuberculosis hsp65* promoter. The truncated ORF contained the 969 bp of *Rv0365c* present in the sur2 insert followed by a stop codon but did not contain the 5 amino acids corresponding to the vector contribution to the ORF. In coinfection experiments, the ratio of *M. smegmatis* bacteria expressing the full-length *Rv0365c* ORF to *xylE*-expressing wild-type bacteria was 1:1 at time zero and increased to approximately 10:1 at 12 hours and to 11:1 at 24 hours (Figure 5). The clone expressing the truncated *Rv0365c* ORF exhibited slightly less, but not statistically significantly different, increased survival (1:1 at time zero, 8.6:1 at 24 hrs).

To investigate the difference in survival at 24 hours of the full-length *Rv0365c* ORF expressing bacteria and sur2 bacteria, a strain expressing *Rv0365c* and *xylE* was constructed and used in THP-1 coinfections with the original sur2 clone. In this coinfection, the ratio of white (sur2) to yellow (*xylE* and *Rv0365c* expressing) colonies remained 1:1 through 12 hours and then increased to ~3:1 at 24 hrs, consistent with the above-described observations.

A cosmid corresponding to the one in the sur3 clone was isolated from the *E. coli* (pYUB178::H37Rv) library by probing colony blots with the 4.5 kb *Bam*HI/*Eco*Rl fragment of the *M. tuberculosis* genomic DNA insert in the plasmid recovered from the sur3 clone. PCR, DNA sequence, and Southern blot data indicated that the recovered cosmid was the same as the cosmid integrated in the sur3 genome (data not shown). The ~10.7 kb cosmid in the sur3 chromosome contains a 5.76 kb fragment of *M. tuberculosis* H37Rv genomic DNA (Figure 6) [23]. This region encodes eight potential ORFs designated *Rv2233-Rv2240c* as well as the tRNA for valine.

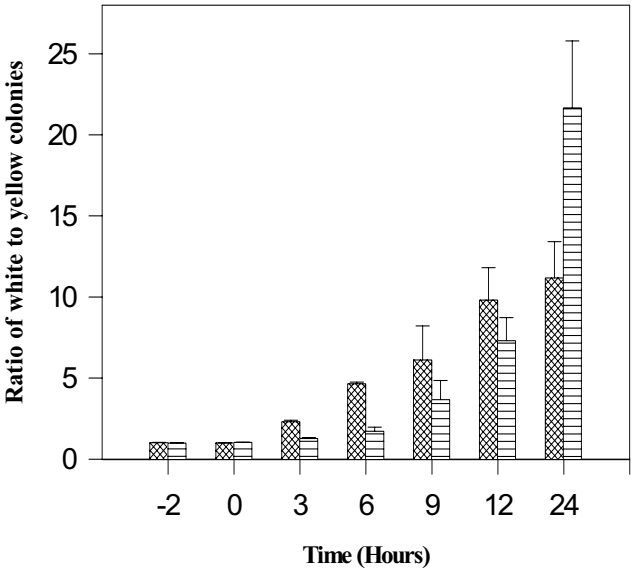

Figure 5
Survival of bacteria expressing ORF *Rv0365c*. THP-1 macrophages were infected with an equal mixture of *xylE*-expressing bacteria and bacteria expressing the full length *Rv0365c* ORF under the control of the *hsp65* promoter (hatched) or sur2 bacteria (horizontal stripes). The ratios of white-to-yellow colonies represent the average of at least three independent experiments. Error bars represent the standard deviation in the ratio of white-to-yellow colonies between experiments.

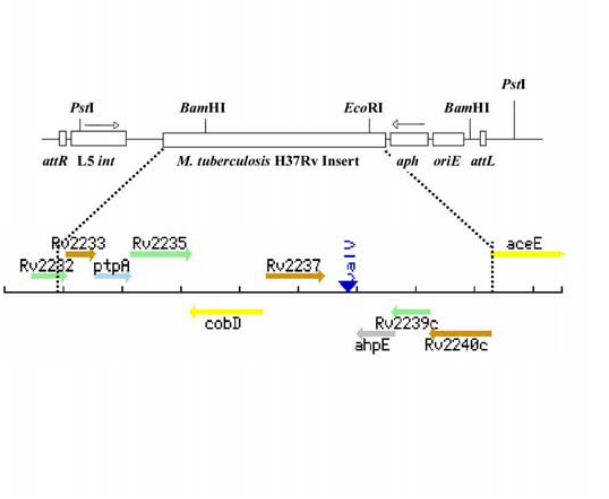

Figure 6
Map of the *M. tuberculosis* H37Rv DNA integrated in the sur3 genome. The integrated 10.7 kb cosmid is between the two attachment sites, *attL* and *attR*. The ~7.4 kb *Bam*HI and ~4.5 kb *Bam*HI/*Eco*RI fragments are indicated by the *Bam*HI and *Eco*RI sites. The 5.76 kb fragment of *M. tuberculosis* H37Rv genomic DNA contains the 34 bp of DNA upstream of *Rv2233*, *Rv2233-Rv2240c*, and the tRNA for valine [Reference 23]. The arrows indicate the direction of transcription. Arrows are color coded according to annotations found on the TubercuList server [http://genolist-Pasteur.fr/TubercuList/] : light green (unknown), brown (conserved), yellow (intermediary metabolism), light blue (regulatory), gray (virulence), red (information pathway). dark green (cell wall process), black (lipid metabolism), and dark blue (stable RNA).

To determine which ORF(s) was responsible for the increased survival exhibited by sur3 bacteria, each ORF was subcloned individually into the expression vector pHIP. The two potential operons, *Rv2233-Rv2235* and *Rv2238c-Rv2240c,* were also subcloned into pHIP. The recombinant bacteria were examined for survival in the macrophage in coinfection experiments at 0, 6, and 12 hours post phagocytosis (Figure 7). Each of the recombinants tested exhibited a 1:1 ratio of white-to-yellow colonies at 0 hours. Two of the recombinants demonstrated an increase in the ratio of white-to-yellow colonies over time, while the ratio remained at 1:1 for the other eight recombinants. One of the two recombinants contained the potential operon of *Rv2233-Rv2235,* and the other contained *Rv2235* only. The ratio of white-to-yellow colonies for bacteria expressing ORFs *Rv2233-Rv2235* was ~2:1 at 6 hours and ~6:1 at 12 hours. For bacteria expressing *Rv2235,* the ratio of white-to-yellow colonies was ~2:1 at 6 hours and >3:1 at 12 hours. The survival of recombinant bacteria expressing the *Rv2235* ORF relative to wild-type was more directly compared to the relative survival of sur3 by doing the coinfections in parallel. The ratios of white-to-yellow for both recombinant bacteria and sur3 bacteria were 1:1 at 0 hours, ~5:1 at 12 hours, and >5:1 at 24 hours. The differences in the relative survival of the three strains were not statistically significant.

A PCR fragment containing the *M. tuberculosis* ORFs *Rv2233*, *Rv2234*, and *Rv2235* without the upstream promoter region, was cloned into the pBPhin vector, which does not contain a promoter to express the inserted DNA. In coinfections with the *xylE*-expressing bacteria, the recombinant bacteria containing *Rv2233*, *Rv2234*, and *Rv2235* did not exhibit the same increase in macrophage survival as the sur3 clone (data not shown), suggesting that the *Rv2235* ORF is expressed using signals upstream of the *Rv2233* ORF.

Southern blots of genomic DNA from *M. tuberculosis*, *M. smegmatis*, *Mycobacterium leprae*, and *Mycobacterium avium* were probed with *Rv0365c* and *Rv2235*. *Rv0365c* hybridized to bands in *M. tuberculosis*, *M. avium*, and *M. smegmatis* under high stringency conditions (data not shown). *Rv2235* hybridized to a band in *M. tu-*

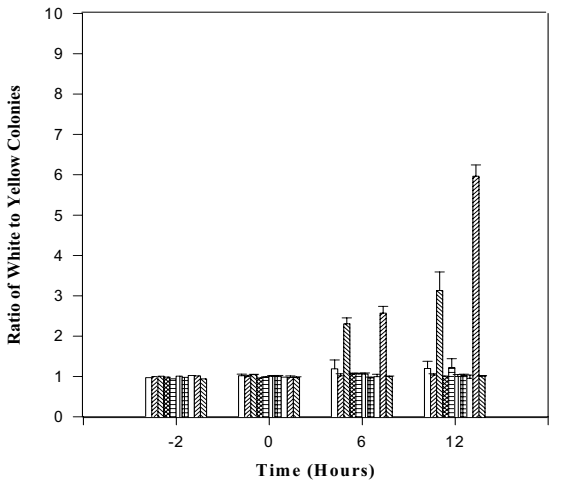

Figure 7
Survival of *M. smegmatis* LR222 bacteria expressing ORFs in the sur3 insert. THP-1 macrophages were infected with an equal mixture of *xylE*-expressing bacteria and bacteria expressing one or more of the sur3 ORFs. The ratio of the recovered white-to-yellow colonies is shown for -2 hours (initial inoculum), 0 hours (immediately after phagocytosis interval), and at 6 and 12 hours after phagocytosis. THP-1 macrophages were infected with an equal mixture of *xylE*-expressing bacteria and bacteria expressing *Rv2233* (first open column), *Rv2234* (first /// lines), *Rv2235* (first \\\ lines), *Rv2236c* (hatched), *Rv2237* (horizontal stripes), *Rv2238c* (vertical stripes), *Rv2239c* (box pattern), *Rv2240c* (second open column), *Rv2233-Rv2235* (second /// lines), or *Rv2238c-Rv2240c* (second \\\ lines). Error bars represent the standard deviation in the ratio of white-to-yellow colonies between experiments.

berculosis, and under low stringency conditions recognized a band in *M. leprae* (data not shown).

Database searches revealed that *Rv0365c* encodes an ~41-kDa protein which displays significant homology only to a truncated *Corynebacterium glutamicum* hypothetical protein of unknown function (42% identity in a 296 aa overlap) located upstream of a gene encoding a fructose-bisphosphate aldolase. *Rv2235* encodes a conserved hypothetical membrane protein of about 30-kDa and shares a motif with the SURF-1 family of proteins. The other two ORFs in the putative *Rv2233-Rv2235* operon, Rv2233 and Rv2234, share homology with several proteins in data base searches. *Rv2233* may encode a putative phosphatase, and *Rv2234* may encode a low molecular weight protein tyrosine phosphatase [23].

Discussion

There are several potential limitations of isolating *M. tuberculosis* genes involved in intracellular survival using an enrichment procedure. One limitation is that this protocol is biased towards the recovery of clones with the greatest increase in survival relative to wild-type [14]. Thus, the recovered clones are not a random collection of genes involved in intracellular survival and hence the number of genes involved for survival cannot be calculated. This approach is also biased towards identifying genes expressing proteins that directly interfere with the antimicrobial processes of the macrophage. Some types of genes involved in resistance to killing, such as genes that are part of a multi-enzyme pathway, may not be isolated. Also, genes necessary for intracellular survival as opposed to resistance to killing are not likely to be isolated.

Two general classes of clones might be recovered following enrichment of the *M. smegmatis* recombinant library for clones with increased intracellular survival. One type might be clones that carry *M. tuberculosis* genes that confer enhanced resistance to the antimicrobial processes of the macrophage. Another type might be clones whose *M. tuberculosis* gene products are involved in attachment or invasion or increase phagocytosis. The two genes that were isolated in the studies reported here confer enhanced resistance rather than increased uptake. That is, the ratios of bacteria expressing either *Rv0365c* or *Rv2235* to wild-type bacteria were 1:1 in both the initial mixture and inside the macrophages at the end of the phagocytosis period. Differences in survival compared to wild-type did not become apparent until about 9 hours post-phagocytosis for bacteria expressing *Rv0365c* and about 6 hours post-phagocytosis for those expressing *Rv2235*.

In this study, the enrichment process resulted in the isolation of two strains containing small, integrated plasmids rather than the expected 30–50 kb cosmids [20]. Small plasmids corresponding to the integrated ones were present in the original *E. coli* (pYUB178:H37Rv) cosmid library. PCR, sequencing, and Southern blot analysis demonstrated that the plasmid integrated in the sur2 clone was approximately 4 kb containing 1.1 kb of *M. tuberculosis* DNA and the plasmid integrated in the sur3 clone was approximately 10.7 kb containing about 5.7 kb of *M. tuberculosis* DNA. Cosmid libraries frequently contain clones without DNA inserts [24], so it is not too surprising that clones with small DNA fragments were present in the *E. coli* library after infection with λ phage. The small plasmids may have a growth or replication advantage causing them to be over-represented in the library following the various amplification steps.

The ~2 kb deletion of vector DNA in the sur2 clone results in a fusion of the ORF of the L5 integrase with that of the *M. tuberculosis* insert such that the fused ORF encodes a protein that contains only the amino-terminal 73 aa of the 344 aa L5 integrase. The observation that the sur2 clone contains an integrated plasmid corresponding to the plasmid isolated from the cosmid library suggests that either a) the hybrid protein retains integrase activity, b) the plasmid inserted into the attachment site by homologous recombination between *attP* and *attB,* or c) a functional integrase was provided in *trans* by a second cosmid transiently present in the original transformant or by a gene present in the *M. smegmatis* LR222 genome. However, repeated attempts to electroporate the 4 kb plasmid into *M. smegmatis* LR222 did not generate any stable kanamycin-resistant transformants (data not shown). These results suggests that the hybrid protein does not retain integrase activity, that integration by homologous recombination into *attP* is unlikely, and that provision of integrase activity by an *M. smegmatis* gene chromosomal does not occur. The simplest explanation is that integrase was provided *in trans* by a second cosmid transiently present in the original transformant, and indeed, *M. smegmatis* bacteria carrying an integrated copy of the 4-kb plasmid can be readily isolated following electroporation of a mixture of the 4-kb plasmid and a plasmid that expresses integrase (unpublished results).

Data base searches did not reveal any homologies that could be used to predict functions for the gene products of ORFs *Rv0365c* or *Rv2235*. *Rv0365c* encodes a hypothetical protein of 376 amino acids which displays significant homology only to a *Corynebacterium glutamicum* hypothetical protein of unknown function (42% identity in a 296 aa overlap) [23]. *Rv2235* encodes a hypothetical protein of 271 aa with three putative transmembrane domains and which displays significant homology only to hypothetical protein MLCB1243.32c in *M. leprae,* to which it is 74% identical [23,25]. This *M. leprae* homologue was evident in Southern blot experiments done under low stringency conditions. Rv2235 also contains the SURF-1 signature sequence [26] and modest homology to members of the SURF-1 family such as the SURF-1 protein of *Caulobacter crescentus* (25% identity; 40% similarity) [27]. SURF-1 proteins are ~33-kDa, integral membrane proteins whose precise function is not known. In eukaryotic cells, SURF-1 proteins are involved in the assembly and maintenance of mitochondrial respiratory chain complexes including cytochrome oxidase [26,28]. The homology to SURF-1 proteins raises the possibility that Rv2235 could play a role in resisting the antimicrobial activities of macrophages by helping to maintain the stability or function of an important cellular process, akin to the stabilizing role of chaperonins during a heat shock.

In addition to the genes described in this report, other studies with *M. smegmatis* recombinants and/or mutants have implicated 11 other *M. tuberculosis* genes in intracellular survival. Genes identified using enrichment or screening protocols include Rv2962c and Rv2958c (probable glucuronsyl transferases), Rv2220 (glutamine synthetase A1), Rv3913-Rv3914 (thioredoxin, thioredoxin reductase) and Rv2416c (*eis,* unknown function) [14,15,22,29]. By screening insertional mutants of *M. smegmatis,* Lagier et al [30] isolated 8 mutants with impaired ability to survive in human peripheral blood monocyte-derived macrophages and identified the *M. tuberculosis* gene corresponding the mutated *M. smegmatis* gene for five of them. The genes included: Rv3052c (probable *nrdI*) which is postulated to be involved in deoxynucleotide production under stressed conditions; Rv0101 which is a nonribosomal peptide synthetase that displays strong homology with a *Pseudomonas* nonribosomal peptide synthetase required for the synthesis of the pyoveridine, a siderophore involved in iron uptake; Rv3420c which displays homology with the S18 ribosomal protein acetyltransferase which behaves as a heat shock protein in *Chlamydia trachomatis;* and Rv0497 and Rv3604c which are hypothetical conserved membrane proteins of unknown function.

It should be noted that these *M. tuberculosis* genes generally confer only a limited enhancement of the survival of *M. smegmatis* bacteria in the human macrophages. That is, usually only a few per cent of the recipients are viable 24 to 48 hours post-infection. This reinforces the concept that the intracellular survival of a pathogenic mycobacteria is a complex multifactoral process. The precise role(s) of any of the identified genes in the intracellular survival of mycobacteria is not yet known, although several of the identified genes have features of stress response genes. This is not unexpected given the relatively small enhancements of the survival of the *M. smegmatis* recombinants and the numerous environmental stresses encountered in the macrophage. Additional studies, such as the construction and characterization of targeted knock-out mutants, will be needed to determine the roles of the proteins encoded by these genes in the survival of *M. tuberculosis* in human macrophages.

Conclusions

Using an enrichment and screening procedure, two *M. tuberculosis* genes, Rv0365c and Rv2235, were identified that could confer an enhanced ability to survive in human macrophages to normally susceptible *M. smegmatis* recipients. The functions of these two proteins are not known. This study brings the number of *M. tuberculosis* genetic loci that have been implicated in enhancing the intracellular survival of *M. smegmatis* cells to 13. The

precise role(s) of any of the identified genes in the intracellular survival of mycobacteria remain to be elucidated.

Materials and methods

Bacterial strains, plasmids, and growth conditions

The bacterial strains and plasmids used in this study are listed in Table 1. The E. coli strain XL 1-Blue was obtained from Stratagene (La Jolla, Calif). The M. smegmatis strain LR222 was obtained from Dr. Jack Crawford, Tuberculosis and Mycobacteriology Branch, Centers for Disease Control and Prevention (CDC), Atlanta, GA. The pYUB178 plasmid and the λ phage library of pYUB178::H37Rv cosmids were generously provided by Dr. William Jacobs, Albert Einstein University, New York, NY [20]. The cosmid library contains 30–50 kb fragments of M. tuberculosis genomic DNA generated by partial Sau3A digestion cloned into BclI-digested pYUB178. In the λ phage library, ~225 cosmids represent one genome-equivalent of M. tuberculosis [20]. An E. coli (pYUB178::H37Rv) library was created by infecting E. coli XL1-Blue with the λ phage library. Bacteria from ~4000 colonies were recovered and pooled, and cosmid DNA was isolated. The pooled pYUB178::H37Rv cosmid DNAs were electroporated into M. smegmatis LR222, and kanamycin-resistant colonies were isolated. Bacteria from ~4,000 colonies (representing ~20 genome equivalents) were recovered and pooled to generate the M. smegmatis (pYUB178::H37Rv) library. Because the cosmids integrate into the mycobacteriophage L5 attachment site in the M. smegmatis genome, a single copy of the M. tuberculosis DNA is maintained in the M. smegmatis transformants [20].

The *E. coli* (pYUB178::H37Rv) library was grown in Luria broth (LB) (Difco Laboratories, Detroit, Mich.) containing 50 μg kanamycin/mL (Sigma Chemical Company, St. Louis, Mo.). The *M. smegmatis* (pYUB178::H37Rv) library was grown in Middlebrook 7H9 media (Difco) containing 10 μg kanamycin/mL and 0.05% (v/v) Tween 80 (Sigma) or on tryptic soy agar (TSA) (Difco) containing 10 μg kanamycin/mL. *E. coli* bacteria containing pHIP-based plasmids were grown in LB containing 200 μg hygromycin/mL (Boehringer Mannheim Biochemicals, Indianapolis, Ind.). *M. smegmatis* bacteria containing pHIP-based plasmids were grown on TSA containing 50 μg hygromycin/mL or in Middlebrook 7H9 media containing 50 μg hygromycin/mL and 0.05% (v/v) Tween 80.

Table 1: Bacterial strains and plasmids used in this study.

Strain	Relevant Characteristic/Use	Source/Reference
E. coli XL 1 -Blue	Laboratory strain	Stratagene
M. smegmatis LR222	Laboratory strain	[40]
M. smegmatis LR222 (pYUB178)	pYUB178	[20]
M. smegmatis LR222 (pHIP1)	*xylE*	This study
M. smegmatis LR222 (pHIP2)	*Rv0365c*	This study
M. smegmatis LR222 (pHIP3)	truncated *Rv0365c*	This study
M. smegmatis LR222 (pHIP4)	*xylE, Rv0365c*	This study
M. smegmatis LR222 (pHIP5)	*Rv2233*	This study
M. smegmatis LR222 (pHIP6)	*Rv2234*	This study
M. smegmatis LR222 (pHIP7)	*Rv2235*	This study
M. smegmatis LR222 (pHIP8)	*Rv2236c*	This study
M. smegmatis LR222 (pHIP9)	*Rv2237*	This study
M. smegmatis LR222 (pHIP10)	*Rv2238c*	This study
M. smegmatis LR222 (pHIP11)	*Rv2239c*	This study
M. smegmatis LR222 (pHIP12)	*Rv2240c*	This study
M. smegmatis LR222 (pHIP13)	*Rv2233, Rv2234, Rv2235*	This study
M. smegmatis LR222 (pHIP14)	*Rv2238c, Rv2239c, Rv2240c*	This study
M. smegmatis LR222 (pBPhin1)	*Rv0365c*	This study
M. smegmatis LR222 (pBPhin2)	*Rv2233, Rv2234, Rv2235*	This study
Plasmids pTKmx	*xylE*	[32]
pYUB178	integrating cosmid	[20]
pBPhin	integrating plasmid	[31]
pHIP	integrating, *hsp65* promoter	This study

An amplicon containing the *hsp65* promoter was generated by PCR from the *M. tuberculosis* H37Rv genome. This PCR fragment was cloned into the *Bam*HI site of pBPhin [31] to generate pHIP. An amplicon carrying the complete *xylE* ORF without the *xylE* promoter was generated by PCR of pTKmx [32], cleaved with *Bam*HI, and cloned into the *Bam*HI site of pHIP downstream from the *hsp65* promoter to generate pHIP1.

DNA manipulations

All enzyme reactions were performed as recommended by the manufacturers (GIBCO Bethesda Research Laboratories, Inc., Gaithersburg, Md., New England Biolabs, Beverly, Mass.). Cosmid DNA from *E. coli* was prepared using the Qiagen Plasmid Maxi Kit (Qiagen, Valencia, Calif.) according to manufacturer's instructions for low-copy number plasmids. Wizard Plus Minipreps Kits (Promega, Madison, Wis.) were used to isolate plasmid DNA from *E. coli* strains. Mycobacterial genomic DNA was isolated as previously described [33].

PCR

The primers used in this study and their sequences are listed in Table 2. Primers were synthesized on a 381A DNA synthesizer (Applied Biosystems, Foster City, Calif.) at the Biotechnology Core Facility, National Center for Infectious Diseases, CDC. Amplifications were done using either a Perkin-Elmer Amp PCR System 2400 or Model 480 DNA Thermal Cycler (Perkin-Elmer Cetus, Norwalk, Conn). Each 100 µl PCR contained 3–5 µl of template DNA, 5 µl of dimethyl sulfoxide (DMSO), and 90–92 µl of a reaction mixture (200 µM (each) deoxynucleotide triphosphates, 1.0 µM (each) primer, 1.25 U of AmpliTaq DNA Polymerase (Perkin-Elmer Cetus), 10 mM Tris hydrochloride pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$, and 0.01% (w/v) gelatin). Each sample was amplified for 30 cycles of denaturation at 94°C for 1.5 minutes, annealing at 60°C for 1.75 minutes, and extension at 72°C for 2.5 minutes.

Table 2: PCR and sequencing primers used in this study. Underlined bases are restriction enzyme sites used in cloning. Bold bases represent either the start codon or stop codon of the gene being cloned.

Primer Sequence (5' to 3')	Location/ Function
GGATAGATCTAGTTGCTGCAGCGT	5' end of *HSP65* promoter
GAAGTGGATCCTCCGATCGGGGATG	3' end of *HSP65* promoter
GACGGATCC**ATG**ACGTCATGAAC	5' end of *xylE*
GACGGATCCAAGCTTGCATGCC	3' end of *xylE*
CGGTGTCGCGGGCGGCGGCGTCG	5' end of sur2 probe
GATCGCATTATGAATCTGGCAAACC	3' end of sur2 probe
CAGGGATCCATCGCATT**ATG**AATCTGGCAAACCG	5' end of *Rv0365c*
CAGGGATCCGTG**TTA**CCCGAGTGAGCTGACC	3' end of *Rv0365c*
CAGGGATCC**TTA**ATCCCAGAACGCCCCGAACACC	sur2 end of truncated *Rv0365c*
GCGGCCGCAGATCTAGTTGCTGCAGCGTGAC	5' end of *HSP65* promoter
GCGGCCGCGTGTTACCCGAGTGAGCTGACC	3' end of *Rv0365c*
GGATCCGACAGGCTAGGGCAGGATCGC	5' end of *Rv0365c* promoter
GACGGATCCGTTGGGCG**ATG**AAACAGCTTG	5' end of *Rv2233*
GACGGATCCGGA**TCA**GACACCTAGCGCCTC	3' end of *Rv2233*
GACAGATCTGTTGGGCG**ATG**AAACAGCTTG	5' end of *Rv2233*
GACAGATCTGGCGCTAG**GTG**TCTGATCCGC	5' end of *Rv2234*
GACAGATCTGCA**TCA**ACTCGGTCCGTTCCG	3' end of *Rv2234*
GACGGATCCCCGAGTTG**ATG**CCCCCGCC	5' end of *Rv2235*
GACGGATCCGGT**TTA**CCGCCGGCGGCC	3' end of *Rv2235*
GACAGATCTGGT**TTA**CCGCCGGCGGCC	3' end of *Rv2235*
GACGGATCCCTGCCGGA**GTG**TTTGCATCG	5' end of *Rv2236c*
GACGGATCCCCG**CTA**CGGCCGCCGGC	3' end of *Rv2236c*
GACGGATCCCCAGGTCG**ATG**CAAACACTCC	5' end of *Rv2237*
GACGGATCCGCT**CTA**TCAGACGATTCGGCG	3' end of *Rv2237*
GAGGGATCCAAGCGCTG**ATG**CTGAACGTCG	5' end of *Rv2238c*
GAGGGATCCCCCAAAACC**TTA**GGCCGTAAG	3' end of *Rv2238c*
GACGGATCCCGACGATA**ATG**CCCATCGCG	5' end of *Rv2239c*
GACGGATCCGCA**TCA**GCGCTTGCCAGCC	3' end of *Rv2239c*
GACGGATCCGCCAGATC**GTG**GCGGGCG	5' end of *Rv2240c*
GACGGATCCGCC**TCA**GAAGGCGGCCACG	3' end of *Rv2240c*

THP-1 human macrophages

The THP-1 cell line was obtained through the Biological Products Branch of the National Center for Infectious Diseases, CDC. THP-1 cells were grown in RPMI 1640 media (GIBCO BRL) containing 10% fetal calf serum (FCS) (GIBCO BRL) at 37°C in 5% CO_2 [21]. To differentiate the THP-1 cells into macrophage-like cells, the cells were treated with 10 μM phorbol myristate acetate (PMA) (Sigma) as follows [21]. All PMA manipulations were done under low light conditions. THP-1 cells were harvested by centrifugation for 10 minutes at 228 × *g*, and the pellet was resuspended in RPMI 1640/10% FCS/10 μM PMA to give a cell density of approximately 1×10^6 THP-1 cells/mL. Three milliliters of the suspension was added to each well of a six-well tissue culture plate (Costar, Corning, NY). The plates were incubated for 48 hours at 37°C in 5% CO_2. The medium was removed from each well, the attached cells were washed once with RPMI 1640/10% FCS, and 3 ml of fresh RPMI 1640/10% FCS was added. The cultures were incubated at 37°C in 5% CO_2 for an additional 48 hours. Immediately prior to infection, cells in each well were washed once with fresh RPMI 1640/10% FCS.

Enrichment protocol

About 10^6 bacteria of the *M. smegmatis* (pYUB178::H37Rv) library were inoculated into Middlebrook 7H9 media containing 10 μg kanamycin/mL and then were grown to midlog phase (OD_{600} ~ 0.3). The bacteria were harvested by centrifugation for 1 minute at 16,000 × *g* and washed twice with RPMI 1640/10% FCS. The bacteria were suspended in RPMI 1640/10% FCS at 5×10^8 bacteria/mL. The enrichment process was as follows (Figure 1): differentiated THP-1 macrophages were infected by adding 3 mL of the bacterial suspension to each well of a six-well plate. The multiplicity of infection (MOI) was ~500 bacteria per THP-1 cell. The culture was left at 37°C in 5% CO_2 for 2 hours, which resulted in the phagocytosis of about 10 bacteria per macrophage. After the phagocytosis period, each well was washed twice with RPMI 1640/10% FCS to remove free bacteria. To kill any remaining extracellular bacteria, 3 mL of fresh RPMI 1640/10% FCS/200 μg amikacin/mL (Sigma) was added to each well. The infected THP-1 cultures were incubated at 37°C for an additional 5 hours. Each well was then washed twice with 3 mL of RPMI 1640/10% FCS, and then 1 mL of 0.1% (v/v) Triton X-100 (Sigma) was added to each well to lyse the macrophages. The wells were scraped with a rubber policeman, and the lysates were removed and diluted for plating on TSA containing 10 μg kanamycin/mL. After 3 days of incubation at 37°C, bacteria from the resulting colonies were harvested, suspended in Middlebrook 7H9 media containing 10 μg kanamycin/mL and 0.05% (v/v) Tween 80, and pooled for reinfection of THP-1 macrophages. The enrichment process was done for a total of three cycles. After the third round of enrichment, individual clones were isolated and analyzed.

Southern blots

*Pst*I-digested genomic DNA was electrophoresed through a 1.0% agarose gel, denatured, neutralized, and transferred by capillary blotting to a Hybond -N^+ membrane (Amersham, Arlington Heights, IL). The blots were hybridized to pYUB178 plasmid DNA labeled using the ECL Direct Nucleic Acid Labeling and Detection System (Amersham). All hybridization and washing steps were done at 42°C under either stringent (0.1 M NaCl) or non-stringent conditions (0.5 M NaCl) according to kit instructions.

Coinfection assay to measure survival

Coinfection assays were done as previously described [22]. Briefly, separate cultures of recombinant and of *xylE*-expressing bacteria were grown to midlog phase (OD_{600} = ~0.3). The bacteria from each culture were harvested by centrifugation for 1 minute at 16,000 × *g*, washed twice with RPMI 1640/10% FCS, and resuspended in RPMI 1640/10% FCS at a concentration of 1.5×10^8 bacteria/mL. Equal volumes of the two bacterial suspensions were mixed to produce a suspension containing a 1:1 ratio of recombinant-to-control bacteria. A portion of the combined mixture was plated onto TSA plates to determine the number of colony forming units (CFUs) of each strain in the initial inoculum (the -2 hour time point in figures). The bacterial suspension was diluted with RPMI 1640/10% FCS to give approximately 5×10^7 bacteria/mL, and 3 mL was added to each well containing 1 × 10^6 THP-1 macrophages (MOI of 50 bacteria/macrophage). The cultures were incubated for 2 hours at 37°C in 5% CO_2 to allow phagocytosis to occur, and then each well was washed twice with RPMI 1640/10% FCS to remove unphagocytosed bacteria. Typically, this results in one phagocytosed bacterium per macrophage. To kill extracellular bacteria, 3 mL of RPMI 1640/10% FCS containing 200 μg amikacin/mL was added to each well. Cultures were incubated at 37°C in 5% CO_2. At various times, the medium was removed from each of three wells, and 1 mL of 0.1% (v/v) Triton X-100 in H_20 was added to each well to lyse the macrophages. Each lysate was diluted as necessary, and portions were plated on TSA plates. The cultures which were assayed immediately after the addition of the media with amikacin serve to measure of the number of phagocytosed viable bacteria; the time at which these cultures were assayed was considered time zero (t_0).

After a 3-day incubation at 37°C, the TSA plates from each time point were stored overnight at 4°C. The following day, the plates were sprayed with 0.5 M catechol (Sig-

ma) in 50 mM potassium phosphate (pH 7.5) to distinguish the *xylE*-expressing colonies (yellow) from the recombinant colonies (white). Storing the plates overnight at 4°C results in a stronger yellow color. To determine the percent survival of a particular clone at time point X, the number of CFUs at time × was divided by the number of CFUs at to and multiplied by 100.

Recovery of the cloned M. tuberculosis genomic DNA fragment

One μg of sur2 genomic DNA was digested with the restriction enzyme *Pst*I then ethanol precipitated. The precipitated DNA was resuspended in 20 μL T4 ligase buffer (GIBCO BRL) and then ligated for 2 hours at room temperature with 0.5 units of T4 ligase (GIBCO BRL). The ligation mixture was electroporated into electrocompetent *E. coli*, and kanamycin-resistant transformants were isolated. After sequencing the *M. tuberculosis* H37Rv insert in the recovered plasmid, PCR primers were designed to amplify a product containing a portion of the cloned *M. tuberculosis* H37Rv genomic DNA. The PCR product was used as a probe for colony blots to find cosmids in the *E. coli* (pYUB178::H37Rv) library carrying at least a portion of the cosmid integrated in the sur2 clone.

A plasmid carrying a portion of the *M. tuberculosis* H37Rv insert in the sur3 clone was isolated in a similar manner from a *Bam*HI digest of sur3 genomic DNA. The *Bam*HI/*Eco*RI fragment of the *M. tuberculosis* H37Rv insert of the recovered plasmid was used as a probe in colony blots to find cosmids in the *E. coli* (pYUB178::H37Rv) library.

Colony blots

Portions of the *E. coli* (pYUB178::H37Rv) cosmid library were plated on LB agar containing 50 μg kanamycin/mL, and colony blots were performed with the ECL Direct Nucleic Acid Labeling & Detection System (Amersham) according to manufacturer's instructions. Colonies hybridizing with the probe of interest were removed as plugs and incubated in LB containing 50 μg kanamycin/mL for approximately 30 minutes at 37°C. Dilutions of this culture were plated on LB agar containing 50 μg kanamycin/mL to give well-separated colonies. Colony blots were performed and positive clones were selected for further study.

Subcloning ORFs into pHIP and pBPhin

Unless otherwise stated, all *M. tuberculosis* open-reading frames (ORFs) were generated by PCR from *M. tuberculosis* H37Rv genomic DNA as full-length ORFs without their natural promoters. The PCR primers (Table 2) were designed to contain restriction enzyme sites for cloning of the amplicon into the pHIP vector downstream of the *hsp65* promoter as well as the eight base-pairs (bp) upstream of the start codon of the gene being cloned. Because the *hsp65* promoter in pHIP contains a ribosome binding site (rbs), the spacing between the rbs and the start codon of the cloned gene is about the same in these constructs as the spacing between the rbs and the start codon of the *hsp65* gene in wild-type *M. tuberculosis*.

A truncated form of the *Rv0365c* gene was generated by PCR to contain 969 bp of the 970 bp of the truncated ORF present in the sur2 clone followed immediately by a stop codon. This fragment was *Bam*HI-digested and ligated to *Bam*HI-digested pHIP to form pHIP3. An amplicon containing the *hsp65* promoter and full-length *Rv0365c* gene was generated by PCR amplification of the pHIP2 plasmid. This amplicon was *Not*I-digested and cloned into the *Not*I site of the pHIP1 plasmid containing the *xylE* gene to generate pHIP4. Plasmid pBPhin1 was constructed by ligating a *Bam*HI-digested PCR fragment containing the *Rv0365c* gene plus 24 bp located upstream of it to *Bam*HI-digested pBPhin. Plasmid pHIP13 was created by cloning a *Bgl*II-digested PCR fragment containing *Rv2233*,*Rv2234*, and *Rv2235* into the *Bam*HI site downstream of the *hsp65* promoter of pHIP. Plasmid pHIP14 was created by cloning a *Bam*HI-digested PCR fragment containing *Rv2238c, Rv2239c,* and *Rv2240c* into the *Bam*HI site downstream of the *hsp65* promoter of pHIP such that *Rv2240c* was proximal to the *hsp65* promoter. Plasmid pBPhin2 was created by ligating the *Bgl*II-digested *Rv2233-Rv2235* fragment into the *Bam*HI site of pBPhin.

Electroporation

All electroporations were conducted using a Bio-Rad Pulse Controller (Bio-Rad, Hercules, Calif). Preparation and electroporation of competent *E. coli* XL 1-Blue cells were done according to Bio-Rad instructions. Competent *M. smegmatis* LR222 cells were prepared and electroporated as described by Jacobs *et al.* [34].

DNA sequencing

All sequencing reactions were prepared with the Applied Biosystems, Inc. (ABI) PRISM Dye Terminator Cycle Sequencing Ready Reaction Kit (Applied Biosystems) according to the manufacturer. All sequencing was conducted using an ABI 373 DNA Sequencing System (Applied Biosystems).

Statistical analysis

Results were analyzed by the two-sample T test.

DNA and protein homology analyses

DNA and protein database searches were performed using the BLAST services (blastn, blastp, and psi-blast) at the National Center for Biotechnology Information (NC-

BI) [35,36] and the genomes site at The Institute for Genomic Research [http://www.tigr.org] . Protein domain/motif searches were performed using PSI-BLAST [37], and CD-Search (RPS-BLAST) [36] at NCBI and the SMART utility at the European Molecular Biology Laboratories [38,39].

Acknowledgements

We thank Dr. William Jacobs for generously providing the pYUB178::H37Rv cosmid library and Dr Jack Crawford for providing strain LR222.

References

1. Dye C, Scheele S, Dolin P, Pathania V, Raviglione MC: **Consensus statement. Global burden of tuberculosis: estimated incidence, prevalence, and mortality by country. WHO Global Surveillance and Monitoring Project.** *JAMA* 1999, **282**:677-686
2. Murray CJ, Salomon JA: **Modeling the impact of global tuberculosis control strategies.** *Proc Natl Acad Sci USA* 1998, **95**:13881-13886
3. Schlesinger LS: **Entry of Mycobacterium tuberculosis into Mononuclear Phagocytes.** *In: Tuberculosis Edited by Shinnick TM, vol. 215. pp. 71–96. Berlin: Springer-Verlag;* 199671-96
4. Crowle AJ, Dahl R, Ross E, May MH: **Evidence that vesicles containing living, virulent Mycobacterium tuberculosis or Mycobacterium avium in cultured human macrophages are not acidic.** *Infect Immun* 1991, **59**:1823-1831
5. Sturgill-Koszycki S, Schlesinger PH, Chakraborty P, Haddix PL, Collins HL, Fok AK, Allen RD, Gluck SL, Heuser J, Russell DG: **Lack of acidification in Mycobacterium phagosomes produced by exclusion of the vesicular proton-ATPase.** *Science* 1994, **263**:678-681
6. Oh YK, Straubinger RM: **Intracellular fate of Mycobacterium avium: use of dual-label spectrofluorometry to investigate the influence of bacterial viability and opsonization on phagosomal pH and phagosome-lysosome interaction.** *Infect Immun* 1996, **64**:319-325
7. Armstrong JA, D'Arcy Hart P: **Response of cultured macrophages to Mycobacterium tuberculosis, with observations on fusion of lysosomes with phagosomes.** *J Exp Med* 1971, **134**:713-740
8. Sturgill-Koszycki S, Schaible UE, Russell DG: **Mycobacterium-containing phagosomes are accessible to early endosomes and reflect a transitional state in normal phagosome biogenesis.** *Embo J* 1996, **15**:6960-8
9. Clemens DL, Horwitz MA: **Characterization of the Mycobacterium tuberculosis phagosome and evidence that phagosomal maturation is inhibited.** *J Exp Med* 1995, **181**:257-270
10. Clemens DL, Lee BY, Horwitz MA: **Deviant expression of Rab5 on phagosomes containing the intracellular pathogens Mycobacterium tuberculosis and Legionella pneumophila is associated with altered phagosomal fate.** *Infect Immun* 2000, **68**:2671-2684
11. Deretic V, Via LE, Fratti RA, Deretic D: **Mycobacterial phagosome maturation, rab proteins, and intracellular trafficking.** *Electrophoresis* 1997, **18**:2542-2547
12. Isberg RR, Falkow S: **A single genetic locus encoded by Yersinia pseudotuberculosis permits invasion of cultured animal cells by Escherichia coli K-12.** *Nature* 1985, **317**:262-264
13. Arruda S, Bomfim G, Knights R, Huima-Byron T, Riley LW: **Cloning of an M. tuberculosis DNA fragment associated with entry and survival inside cells.** *Science* 1993, **261**:1454-1457
14. Mundayoor S, Shinnick TM: **Identification of genes involved in the resistance of mycobacteria to killing by macrophages.** *Annals NY Acad Sci* 1994, **730**:26-36
15. Wei J, Dahl JL, Moulder JW, Roberts EA, O'Gaora P, Young DB, Friedman RL: **Identification of a Mycobacterium tuberculosis gene that enhances mycobacterial survival in macrophages.** *J Bacteriol* 2000, **182**:377-384
16. Zhang Y, Heym B, Allen B, Young D, Cole S: **The catalase-peroxidase gene and isoniazid resistance of Mycobacterium tuberculosis.** *Nature* 1992, **358**:591-593
17. Garbe T, Harris D, Vordermeier M, Lathigra R, Ivanyi J, Young D: **Expression of the Mycobacterium tuberculosis 19-kilodalton antigen in Mycobacterium smegmatis. immunological analysis and evidence of glycosylation.** *Infect Immun* 1993, **61**:260-267
18. Harth G, Horwitz MA: **Expression and efficient export of enzymatically active Mycobacterium tuberculosis glutamine synthetase in Mycobacterium smegmatis and evidence that the information for export is contained within the protein.** *J Biol Chem* 1997, **272**:22728-22735
19. Dahl JL, Wei J, Moulder JW, Laal S, Friedman RL: **Subcellular localization of the intracellular survival-enhancing Eis protein of Mycobacterium tuberculosis.** *Infect Immun* 2001, **69**:4295-4302
20. Pascopella L, Collins FM, Martin JM, Lee MH, Hatfull GF, Stover CK, Bloom BR, Jacobs WR Jr: **Use of in vivo complementation in Mycobacterium tuberculosis to identify a genomic fragment associated with virulence.** *Infect Immun* 1994, **62**:1313-1319
21. Asseffa A, Dickson LA, Mohla S, Bremner TA: **Phorbol myristate acetate-differentiated THP-1 cells display increased levels of MHC class I and class II mRNA and interferon-gamma-inducible tumoricidal activity.** *Oncology Res* 1993, **5**:11-18
22. Miller BH, Shinnick TM: **Evaluation of Mycobacterium tuberculosis genes involved in resistance to killing by human macrophages.** *Infect Immun* 2000, **68**:387-390
23. Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, Gordon SV, Eiglmeier K, Gas S, Barry CE 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, Barrell BG, *et al*: **Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence.** *Nature* 1998, **393**:537-544
24. Sambrook J, Fritsch EF, Maniatis T: *Molecular Cloning: A Laboratory Manual. Cold Spring Harbor: Cold Spring Harbor Laboratory Press;* 1989
25. Cole ST, Eiglmeier K, Parkhill J, James KD, Thomson NR, Wheeler PR, Honore N, Garnier T, Churcher C, Harris D, Mungall K, Basham D, Brown D, Chillingworth T, Connor R, Davies RM, Devlin K, Duthoy S, Feltwell T, Fraser A, Hamlin N, Holroyd S, Hornsby T, Jagels K, Lacroix C, Maclean J, Moule S, Murphy L, Oliver K, Quail MA, Rajandream MA, Rutherford KM, Rutter S, Seeger K, Simon S, Simmonds M, Skelton J, Squares R, Squares S, Stevens K, Taylor K, Whitehead S, Woodward JR, Barrell BG: **Massive gene decay in the leprosy bacillus.** *Nature* 2001, **409**:1007-1011
26. Zhu Z, Yao J, Johns T, Fu K, De Bie I, Macmillan C, Cuthbert AP, Newbold RF, Wang J, Chevrette M, Brown GK, Brown RM, Shoubridge EA: **SURFI, encoding a factor involved in the biogenesis of cytochrome c oxidase, is mutated in Leigh syndrome.** *Nat Genet* 1998, **20**:337-343
27. Nierman WC, Feldblyum TV, Laub MT, Paulsen IT, Nelson KE, Eisen J, Heidelberg JF, Alley MR, Ohta N, Maddock JR, Potocka I, Nelson WC, Newton A, Stephens C, Phadke ND, Ely B, DeBoy RT, Dodson RJ, Durkin AS, Gwinn ML, Haft DH, Kolonay JF, Smit J, Craven MB, Khouri H, Shetty J, Berry K, Utterback T, Tran K, Wolf A, Vamathevan J, Ermolaeva M, White 0, Salzberg SL, Venter JC, Shapiro L, Fraser CM: **Complete genome sequence of Caulobacter crescentus.** *Proc Natl Acad Sci USA* 2001, **98**:4136-4141
28. Nijtmans LG, Artal Sanz M, Bucko M, Farhoud MH, Feenstra M, Hakkaart GA, Zeviani M, Grivell LA: **Shy1p occurs in a high molecular weight complex and is required for efficient assembly of cytochrome c oxidase in yeast.** *FEBS Lett* 2001, **498**:46-51
29. Wieles B, Ottenhoff TH, Steenwijk TM, Franken KL, de Vries RR, Langermans JA: **Increased intracellular survival of Mycobacterium smegmatiscontaining the Mycobacterium leprae thioredoxin-thioredoxin reductase gene.** *Infect Immun* 1997, **65**:2537-2541
30. Lagier B, Pelicic V, Lecossier D, Prod'hom G, Rauzier J, Guilhot C, Gicquel B, Hance AJ: **Identification of genetic loci implicated in the survival of Mycobacterium smegmatis in human mononuclear phagocytes.** *Mol Microbiol* 1998, **29**:465-475
31. Plikaytis BB, Crawford JT, Shinnick TM: **IS1549 from Mycobacterium smegmatis forms long direct repeats upon insertion.** *J Bacteriol* 1998, **180**:1037-1043
32. Kenney TJ, Churchward G: **Genetic analysis of the Mycobacterium smegmatis rpsL promoter.** *J Bacteriol* 1996, **178**:3564-3571
33. Wilson K: **Preparation of genomic DNA from bacteria.** *In: Current Protocols in Molecular Biology Edited by Ausubel F, Brent R, Kinston R, Moore D, Seidman I, Smith J, Struhl K, vol. 1. pp. 2.4.1–2.4.2. New York: Wiley Interscience;* 19902.4.1-2.4.2

34. Jacobs WR Jr, Kalpana GV, Cirillo JD, Pascopella L, Snapper SB, Udani RA, Jones W, Barletta RG, Bloom BR: **Genetic systems for mycobacteria.** *Methods Enzymol* 1991, **204**:537-555
35. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: **Basic local alignment search tool.** *J Mol Biol* 1990, **215**:403-410
36. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ: **Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.** *Nucleic Acids Res* 1997, **25**:3389-3402
37. Schaffer AA, Aravind L, Madden TL, Shavirin S, Spouge JL, Wolf YI, Koonin EV, Altschul SF: **Improving the accuracy of PSI-BLAST protein database searches with composition-based statistics and other refinements.** *Nucleic Acids Res* 2001, **29**:2994-3005
38. Schultz J, Milpetz F, Bork P, Ponting CP: **SMART, a simple modular architecture research tool: identification of signaling domains.** *Proc Natl Acad Sci USA* 1998, **95**:5857-5864
39. Schultz J, Copley RR, Doerks T, Ponting CP, Bork P: **SMART: a web-based tool for the study of genetically mobile domains.** *Nucleic Acids Res* 2000, **28**:231-234
40. Miller LP, Crawford JT, Shinnick TM: **The rpo B Gene of Mycobacterium tuberculosis.** *Antimicrob Agents Chemoth* 1994, **38**:805-811