Characterization of the mIHF Gene of 
*Mycobacterium smegmatis*

MARISA L. PEDULLA† AND GRAHAM F. HATFULL*

Department of Biological Sciences, University of Pittsburgh, Pittsburgh, Pennsylvania 15260

Received 6 May 1998/Accepted 5 August 1998

Site-specific integration of mycobacteriophage L5 requires the phage-encoded integrase protein and the host-encoded mycobacterial integration host factor (mIHF) protein (13, 14, 17). mIHF is an unusual host factor in that it does not bind specifically to L5 attP DNA (17), but is required for formation of recombinogenic intasomes that contain attP DNA, L5 integrase, and mIHF (17, 21, 22). While its name reflects the requirement of the mIHF protein for L5 integration, mIHF is not closely related at the sequence level to *Escherichia coli* IHF, the HU family of proteins, or any other small DNA-binding proteins (10, 17). IHF is not essential for the viability of *E. coli*, although it is implicated in a variety of cellular processes including gene expression (6, 7), DNA metabolism (2, 8, 9, 12, 16), and pathogenesis (15, 23); it also reaches its highest intracellular level just prior to stationary phase (1, 3, 5) and may be involved in the regulation of genes required for the establishment of stationary phase (11). In this study, we asked whether mIHF is required for the viability of *Mycobacterium smegmatis* and whether its intracellular levels fluctuate with growth of the bacteria.

**Organization of the mIHF locus in M. smegmatis and *Mycobacterium tuberculosis*.** The fast-growing *M. smegmatis* and the slow-growing *Mycobacterium tuberculosis* both contain a single gene encoding mIHF (17). However, the nucleotide sequences beyond the 5′ and 3′ ends of the mIHF genes (with the exception of a small region to the 5′ side) are not closely related (17). Thus, while mIHF may play important roles in the mycobacteria, it is unclear whether the genes occupy similar chromosomal locations. Additional information on the sequence of the mIHF locus of *M. smegmatis* shows that mIHF is located approximately 1 kb downstream of the pyrF gene with no identifiable genes within the intergenic space (Fig. 1A). In *M. tuberculosis* there are two genes in the interval between pyrF and mIHF; one encodes a small protein of unknown function, and the other encodes a protein that, while also of unknown function, bears sequence similarity to a large family of paralogous proteins in *M. tuberculosis*. This family has been designated the PPE family of proteins (4). Members of this family in *M. smegmatis* have yet to be described.

The region of high sequence similarity of the *M. smegmatis* and *M. tuberculosis* mIHF genes extends approximately 125 bp upstream of the mIHF coding regions (Fig. 1B). Part of this region corresponds to the putative ribosome binding sites, and sequences further upstream may be important for promoter activity and regulatory functions. We have identified a putative transcription initiation site by S1 nuclease mapping at position −77 (relative to the start of the mIHF coding sequence), which is within the conserved region (Fig. 1B). Bases at the putative −10 and −35 positions upstream of the transcription initiation site are also well conserved (Fig. 1B).

**Intracellular levels of mIHF vary according to bacterial growth phase.** The abundance of mIHF in *M. smegmatis* and *M. bovis* bacillus Calmette-Guérin (BCG) as a function of the state of bacterial growth was determined as shown in Fig. 2.

---

* Corresponding author. Mailing address: Department of Biological Sciences, University of Pittsburgh, Pittsburgh, PA 15260. Phone: (412) 624-6975. Fax: (412) 624-4870. E-mail: gfh@vms.cis.pitt.edu.

† Present address: Department of Biochemistry and Biophysics, Washington State University, Pullman, WA 99164.
FIG. 2. Growth phase dependency of mIHF. (A) Detection of *M. smegmatis* mIHF by immunoblotting. Following dilution of a saturated culture of *M. smegmatis* into fresh media, samples were removed at the times indicated (in hours), and cells were harvested by centrifugation. Samples were sonicated, normalized for total protein content, and electrophoresed on a sodium dodecyl sulfate–15% polyacrylamide gel. After transfer to polyvinylidene difluoride, the filter was probed with anti-mIHF serum and proteins were visualized by chemiluminescence. The marker lane (M) contains purified mIHF protein. (B) The optical density at 600 nm (OD$_{600}$) of the bacterial culture used for panel A was determined at various times. The mIHF levels shown in panel A were quantitated by using NIH Image and are in arbitrary units. (C) Detection of BCG mIHF by immunoblotting. Samples of *M. bovis* BCG were removed at the indicated times (in days) after dilution of a saturated culture; cells were harvested, normalized for total protein content, and electrophoresed on a sodium dodecyl sulfate–polyacrylamide gel. Following transfer to polyvinylidene difluoride, the filter was probed with anti-mIHF serum and detected by chemiluminescence. (D) The optical density at 600 nm (OD$_{600}$) of the bacterial culture used for panel C was determined at various times. The mIHF levels shown in panel C were quantitated by using NIH Image and are in arbitrary units.

FIG. 3. Allelic replacement of mIHF. (A) Strategy for allelic replacement of *M. smegmatis* mIHF. Transformation of *M. smegmatis* mc$^{2}$155 with plasmid pMP32 generates a strain (MP1) containing an integrated copy of the plasmid at the mIHF locus. Strain MP1 is resistant to kanamycin but sensitive to sucrose due to the presence of the sacB gene. Selection of sucrose-resistant colonies can give rise to two alternative products depending on where recombination occurs. If the recombination event is on the same side of aph as the initial integration event, then the original strain (mc$^{2}$155) is regenerated; if it is on the other side, then the mIHF gene is replaced by an interrupted copy. (B) Schematic representation of strains used or generated in allelic replacement experiments. At the top, the mIHF locus of *M. smegmatis* mc$^{2}$155 is represented with the positions of restriction sites for BamHI (B), PstI (P), and SalI (S) indicated; the unlinked attB site is also shown. Also shown are relevant parts of the chromosomes of strains MP1, MP6, MP2, MP3, MP4, and MP5. Strain MP1 is a derivative of mc$^{2}$155 created by insertion of plasmid pMP27 by homologous recombination at the mIHF locus; MP6 is a derivative of MP1 as determined by Southern hybridization and PCR analyses but is sucrose resistant and probably contains an inactivating mutation within the sacB gene (∗). Strain MP2 was derived from MP1 by transformation with an integration-proficient plasmid containing the mIHF gene that integrates site specifically at the attB locus. Strains MP3, MP4, and MP5 are sucrose-resistant derivatives of MP2 that have an inactivating mutation in sacB (MP3) or have undergone recombination at the mIHF locus to leave only the wild-type mIHF gene (MP5) or a replacement by an interrupted mIHF gene (MP4). DNA fragments generated by restriction enzyme digests that hybridize with an mIHF–specific DNA probe are shown as thick horizontal lines with their sizes in kilobases. The positions of primers used in PCR characterization experiments are shown as arrowheads, and the sizes of PCR products are shown in base pairs. Restriction sites originating from plasmid vector sequences are shown with an asterisk. (C) PCR amplification of the mIHF locus in *M. smegmatis* strains. DNAs from various *M. smegmatis* strains were used for PCR amplification with the primers shown in panel B, and the products were separated by agarose gel electrophoresis. DNAs used were from pMP18 (containing the wild-type mIHF gene) (lane 2), pMP32 (containing the aph-interrupted mIHF gene) (lane 3), pMP28 (which contains wild-type mIHF but lacks one of the primer binding sites) (lane 4), mc$^{2}$155 (lane 5), MP1 (lane 6), MP2 (lane 7), MP3 (lane 8), and MP4 (lane 9). Lane 1 contains no DNA. The positions of the 513-bp fragment amplified from the wild-type mIHF gene and the 1,476-bp fragment from the aph-interrupted gene are indicated. Note that when both the wild-type and interrupted mIHF loci are present in the same strain (i.e., MP1, MP2, and MP3) the smaller product is preferentially amplified. Plasmid pMP32 was constructed by insertion of a sacB fragment into pMP27, a pUC119 derivative that contains the mIHF gene interrupted by the aph gene at the EcoNI site.
Following dilution of saturated cultures, samples were removed at various times, and the mIHF levels were determined by Western blotting with anti-mIHF serum (Fig. 2). These data show that the abundance of mIHF is not constant throughout the growth of the bacterial cultures and is most prevalent during late logarithmic growth. Similar patterns were seen for *M. smegmatis* and BCG even though the growth rates for the cultures are very different (Fig. 2). We note that *E. coli* IHF is also most abundant just prior to stationary phase although the magnitude of the effect (5- to 10-fold) (1, 3) is somewhat greater than that observed for mIHF (Fig. 2).

**mIHF is essential for *M. smegmatis* viability.** To determine whether mIHF is essential for mycobacterial viability, allelic replacement experiments were performed. The strategy used was similar to that described previously (18–20) in which the *Bacillus subtilis* sacB gene is used as a counterselectable marker; *M. smegmatis* is normally resistant to sucrose but becomes sucrose sensitive when the *sacB* locus (e.g., MP4) (Fig. 3B and C). Selection of sucrose-resistant derivatives of this strain can either regenerate the initial strain or—by recombination on the side of the *aph* gene opposite to that which gave rise to the integrant—generate a strain having only an interrupted copy of the gene (Fig. 3A). However, the replacement of *mIHF* by an interrupted gene will result in viable cells only if mIHF is not required for growth of the bacteria; if the gene is essential, then the replacement will produce viable cells only if a second copy of *mIHF* is present elsewhere on the chromosome.

A plasmid (pMP32) which cannot replicate in mycobacteria and contains the *M. smegmatis* mIHF gene with the *aph* kanamycin-resistance gene inserted within the coding region was constructed (Fig. 3A). This plasmid was introduced into *M. smegmatis* mc²155 by electroporation (24), and kanamycin-resistant transformants were recovered. One of these (MP1) (Fig. 3B) was characterized further and was shown to be sucrose sensitive and to contain a single copy of the plasmid integrated at the *mIHF* locus (Fig. 3C and data not shown). When MP1 was cultured and plated onto solid media containing sucrose (but without kanamycin), sucrose-resistant colonies were generated at a frequency of approximately 10⁻³ (Table 1). When these were tested for the *aph* phenotype, 56% were found to be kanamycin resistant as well. However, when examined by PCR amplification of the *mIHF* locus (Fig. 3C), all of the more than 100 individual sucrose-resistant kanamycin-resistant colonies tested retained the wild-type *mIHF* locus (data not shown). Southern hybridization of a subset of these colonies indicated that there were no additional recombination events in this region and that these colonies most likely arose from either point mutations within *sacB* or suppressor mutations elsewhere in the chromosome (e.g., MP6) (Fig. 3B).

While the frequency of these events is somewhat higher than expected (Table 1), the recombinant strains containing the *sacB* gene grow noticeably slower than the parent strain, even in the absence of sucrose, providing a selective advantage for *sacB* mutants.

The inability to isolate sucrose-resistant recombinants that have lost the wild-type *mIHF* gene suggests that either *mIHF* is an essential gene or the recombination events that give rise to the *mIHF* replacement are very infrequent events. To address this issue, we constructed a strain (MP2) that contains an additional copy of the *mIHF* gene integrated at the phage L5 *attB* attachment site (Fig. 3B). This was accomplished by introduction of an integration-proficient plasmid containing the wild-type *mIHF* gene (pMP28) into strain MP1 and selection of transformants resistant to kanamycin and hygromycin (Fig. 3B). MP2 was cultured and plated on solid media to select sucrose-resistant colonies as described above; 58% of these were shown to be kanamycin resistant (Table 1). The *mIHF* loci of 12 of these sucrose-resistant kanamycin-resistant colonies were tested by PCR, and 50% of these were shown to have lost the normal wild-type *mIHF* locus (e.g., MP4) (Fig. 3B and C). The remaining 50% presumably have mutations within the *sacB* gene (e.g., MP3) (Fig. 3B). These experiments show that the recombination events that give rise to a replacement of *mIHF* can indeed occur but produce viable cells only if an additional wild-type copy of *mIHF* is present. We conclude that *mIHF* is an essential gene in *M. smegmatis*.

These experiments show that mIHF plays an important role in the mycobacteria. It is clearly essential for the growth of *M. smegmatis*, and it seems likely that it is also essential in slow-growing mycobacteria. Assuming that *M. smegmatis* also contains HU-like and HupB DNA-binding proteins similar to those identified in *M. tuberculosis* (4), these do not appear to compensate for the loss of mIHF. It thus seems likely that mIHF performs specialized functions in the mycobacteria. Since it is most abundant prior to entry into the stationary phase, one of these functions may be to regulate the expression of stationary-phase-specific genes in a manner similar to that of *E. coli* IHF.

This work was supported by NIH GM49647.

We thank Erica Shepard for assistance with DNA sequencing, G. Sarkis for integrating plasmid pG567, L. Pascopella and W. R. Jacobs, Jr., for the pYUB415::mc2155 cosmid library, and M. Pavelka and W. R. Jacobs, Jr., for the *sacB*-containing plasmid pYUB657. We also thank Carol Peña for helpful comments on the manuscript.


