

## A novel host factor for integration of mycobacteriophage L5

MARISA L. PEDULLA\*, MONG HONG LEE†, DAWN C. LEVER\*, AND GRAHAM F. HATFULL\*‡

\*Department of Biological Sciences, University of Pittsburgh, Pittsburgh, PA 15260; and †Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, NY 10021

Communicated by Allan M. Campbell, Stanford University, Stanford, CA, October 24, 1996 (received for review July 12, 1996)

**ABSTRACT** Bacterial integration host factors (IHF) play central roles in the cellular processes of recombination, DNA replication, transcription, and bacterial pathogenesis. We describe here a novel mycobacterial IHF (mIHF) of *Mycobacterium smegmatis* and *Mycobacterium tuberculosis* that stimulates integration of mycobacteriophage L5. mIHF is the product of a single gene and is unrelated at the sequence level to other integration host factors. By itself, mIHF does not bind preferentially to *attP* DNA, although it significantly alters the pattern of integrase (Int) binding, promoting the formation of specific integrase–mIHF–*attP* intasome complexes.

The mycobacteria include the important human pathogens *Mycobacterium tuberculosis* and *Mycobacterium leprae*, the causative agents of tuberculosis and leprosy, respectively (1). These pathogens are characterized by unusual cell walls and extremely slow growth rates (1). It is not clear how the cellular processes of DNA replication, cell division, and gene expression affect the rate of mycobacterial growth, or what changes in these events occur during bacterial infection. In other bacterial systems, small heat-stable DNA-binding proteins such as integration host factor (IHF) and HU play central roles in regulating these processes. We describe here the mycobacterial integration host factor (mIHF), a novel DNA-binding protein required for phage L5 integration.

Phage L5, a temperate phage of the mycobacteria, integrates site-specifically into the genomes of *Mycobacterium smegmatis*, *M. tuberculosis*, and bacille Calmette–Guérin (2–6). Integration occurs by integrase (Int)-mediated site-specific recombination between the phage attachment site (*attP*) and the bacterial attachment site (*attB*) (3, 6). Although L5 Int is only distantly related to  $\lambda$  Int (3), both proteins are composed of two domains, each of which recognizes a different type of sequence within *attP*; the smaller N-terminal domains bind to arm-type sites, and the C-terminal domains bind to core-type sites and contain the catalytic residues (5, 33, 34). In both systems, *attB* is small (25–30 bp) relative to *attP* ( $\approx$ 250 bp), and strand exchange involves specific cleavages of *attP* and *attB* DNA within a common core (7–10).

Four well-characterized phage systems (*Escherichia coli* phages  $\lambda$  and P2, *Salmonella typhimurium* phage P22, and the *Haemophilus influenzae* phage HP1) require a factor encoded by their bacterial hosts for integrative recombination (11–14). This IHF is a small, heterodimeric, sequence-specific DNA-binding protein (15) that is involved in numerous disparate cellular processes including bacterial pathogenesis (15–18). In the  $\lambda$  system, IHF binds to three sites within *attP* (19, 20) and stimulates integrative recombination through its ability to introduce specific bends at each site (21, 22); these bends promote the formation of intramolecular protein bridges, in which the two domains of Int are simultaneously bound to core- and arm-type sites (23–25).

Although DNA-bending proteins that bind nonspecifically to DNA (such as HU, HMG1, HMG2, and the histone dimer

H2A-H2B) can substitute for IHF to form intasomes, they do not stimulate integrative recombination of phage  $\lambda$  (26, 27). The failure of these nonspecific DNA-binding proteins to support integration appears to result from their inability to introduce bends of the required magnitude and direction at all three IHF binding sites simultaneously (27, 28). The requirements for  $\lambda$  excisive recombination are less stringent, and the nonspecific DNA-binding proteins can substitute for IHF (26, 27). Similarly, L5 Int-mediated integrative recombination displays a strong requirement for a host factor that is present in extracts of *M. smegmatis* (6) and bacille Calmette–Guérin (data not shown). The need for a mycobacterial extract appears to be quite specific: although the stimulating activity shares with IHF and HU the property of being heat stable, *E. coli* extracts, IHF, or HU do not stimulate L5 integration *in vitro* (6). In this report, we show that mIHF is composed of a single, small, heat-stable polypeptide that binds to DNA without specificity for the *attP* site. mIHF is unrelated to previously described DNA-binding proteins and appears to stimulate recombination by binding cooperatively with L5 integrase to *attP*, forming specific intasomal complexes that Int alone is unable to form. Finally, mIHF is highly conserved between the fast-growing *M. smegmatis* and the slow-growing pathogen, *M. tuberculosis*.

### MATERIALS AND METHODS

**Plasmid Constructions.** The *M. smegmatis* mIHF gene was isolated from a cosmid library of *M. smegmatis* DNA (kindly provided by Bill Jacobs, Jr., Yeshiva University, New York) using the degenerate oligonucleotides 5'-CCC/gCAGGTC/gACc/gGACGAGCAGCGt/c/g/aGCt/c/g/aGCt/c/g/aGC and 5'-TCGGCIGAGCTc/gAAGGACCGICTc/gAAGCGI-GGIGGIACc/gAACCT (where I is inosine, and positions where base mixtures were used are shown in lowercase letters) that correspond to regions of the N-terminal amino acid sequence of mIHF. DNA fragments from positive cosmid clones were subcloned, and a 1054-bp segment was sequenced using appropriate oligonucleotide primers and single-stranded DNA templates (29); the DNA sequences of both strands were determined. The mIHF overexpression plasmid, pMP21, was generated by PCR amplification of the mIHF gene and insertion into the T7 expression vector, pET21a (Novagen). The predicted protein product is identical to that of the protein isolated from *M. smegmatis*. Sequence analyses were performed using the University of Wisconsin Genetics Computer Group programs including FASTA and BLAST (30–32).

**Southern Hybridization.** Approximately 3  $\mu$ g of *M. smegmatis* chromosomal DNA was digested with 20 units of the appropriate restriction enzyme overnight at 37°C and electrophoresed through a 0.7% agarose gel. DNA was transferred to a GeneScreenPlus membrane (NEN), probed with a 350-bp

Abbreviations: IHF, integration host factor; mIHF, mycobacterial IHF; Int, integrase.

Data deposition: The sequence reported in this paper has been deposited in the GenBank data base (accession no. U75344).

‡To whom reprint requests should be addressed. e-mail: GFH@vms.cis.pitt.edu.

$^{32}\text{P}$ -labeled PCR-generated *mIHF* DNA fragment, washed, and exposed to film.

**Protein Purification.** *mIHF* protein was purified from *M. smegmatis* as follows. Cells from a 35-liter culture of *M. smegmatis* mc<sup>2</sup>155 were pelleted, resuspended in 200 ml of cold TED buffer (20 mM Tris, pH 7.5/10 mM EDTA/1 mM DTT), sonicated, and clarified by centrifugation. The supernatant was extracted in batch by addition of carboxymethyl-Sepharose, which was then collected by centrifugation and extracted with 0.5M NaCl TED. Proteins were precipitated by addition of ammonium sulfate, collected by centrifugation, and resuspended in TED. Following dialysis, the sample was loaded onto an Econo-Pac heparin cartridge (Bio-Rad) connected to an fast protein liquid chromatography system (Pharmacia), and proteins were eluted with a 400-1000 mM NaCl gradient. Active fractions were identified by *in vitro* recombination, pooled, and loaded onto an Econo-Pac S cartridge (Bio-Rad). Proteins were eluted with a 0-1000 mM NaCl gradient, and active fractions were identified using recombination assays. Purification of *mIHF* from *E. coli* strain BL21DE3pLys (Novagen) carrying plasmid pMP21, induced by addition of 0.5 mM isopropyl  $\beta$ -D-thiogalactoside, was accomplished by a similar protocol. Cells from a 14-liter culture were harvested by centrifugation and frozen; thawed pellets were resuspended in 325 ml cold TED and clarified by centrifugation. Following precipitation and removal of nucleic acids with 0.5% polyethyleneimine, proteins were precipitated with ammonium sulfate and *mIHF* purified by chromatography over BioCAD POROS 20 Heparin and carboxymethyl columns (PerSeptive Biosystems, Framingham, MA); active fractions were identified using recombination assays.

**In Vitro Recombination.** Recombination reactions were performed as described (6) and contained  $\approx 0.3$  pmol supercoiled pMH39 DNA (which contains *attP*), either a 3'  $^{32}\text{P}$ -radiolabeled 0.6-kb linear *attB* DNA (obtained by digestion of plasmid pMH57 with *EcoRI* and *HindIII*) or a 3.9-kb unlabeled linear *attB* DNA (obtained by digestion of pMH57 with *HindIII*), 2  $\mu\text{l}$  of an extract containing L5 Int, and 0.1- to 1.0- $\mu\text{l}$  protein fractions. Reactions were incubated for 3 hr at room temperature, stopped by addition of 2  $\mu\text{l}$  of 1% SDS, and electrophoresed on a 0.8% agarose gel. Recombinant products were identified by ethidium bromide staining or autoradiography.

**Electrophoretic Mobility Shift Assay.** Conditions for DNA-binding were similar to those used for recombination (6). DNAs used were as follows: a  $^{32}\text{P}$ -labeled 0.6-kb *attP* DNA fragment derived by digestion of pMH39 with *EcoRI* and *HindIII* (6); a  $^{32}\text{P}$ -labeled 450-bp *Asp-718-HindIII* fragment from pMP3 (containing a segment of mycobacterial DNA that does not include *attP* or *attB*); a  $^{32}\text{P}$ -labeled 364-bp *attP* DNA fragment from plasmid pCP $\Delta$ R11 (34). Unless otherwise noted, all binding reactions contained 1  $\mu\text{g}$  of salmon sperm DNA. Following addition of proteins, reactions were incubated on ice for 15-30 min and electrophoresed through native 5% polyacrylamide gels in 1 $\times$  TBE (100 mM Tris/84 mM borate/1 mM EDTA).

**DNase I Footprinting.** Protein-DNA complexes were formed in binding reactions essentially identical to those used for recombination and electrophoretic mobility shift assays and digested with an appropriate dilution of DNase I at room temperature for 1 min (29). The products were phenol extracted, ethanol precipitated, and analyzed on a 6% sequencing gel in 1 $\times$  TBE buffer.

**Immunoblot of Protein-DNA Complexes.** Rabbit anti-*mIHF* antiserum was prepared by Pocono Rabbit Farms (Canadensis, PA) using *mIHF* purified from *E. coli*. Electrophoretic mobility shift assays were performed as described above, except that  $\approx 0.5$   $\mu\text{g}$  of nonradiolabeled pCP $\Delta$ R11 DNA (34) digested with *EcoRI* (New England Biolabs) and *BamHI* (New England Biolabs) were included. After electrophoresis, the wet gel was

exposed to film for 1.5 hr and the proteins were electroblotted to polyvinylidene difluoride (Bio-Rad) for 2.5 hr at 150 mA. The filter was blocked in PBST (1.5 mM NaH<sub>2</sub>PO<sub>4</sub>/8.1 mM Na<sub>2</sub>HPO<sub>4</sub>/145 mM NaCl/0.05% Tween 20) with 4% milk for 1 hr and incubated with rabbit anti-*mIHF* antiserum diluted 1:1000 in PBST with 4% milk for 2.5 hr. The filter was washed in PBST and incubated in horseradish peroxidase-conjugated goat anti-rabbit IgG (Bio-Rad) diluted 1:5000 in PBST with 4% milk for 1 hr and washed again in PBST. A final wash was performed in PBS (1.5 mM NaH<sub>2</sub>PO<sub>4</sub>/8.1 mM Na<sub>2</sub>HPO<sub>4</sub>/145 mM NaCl), and chemiluminescence was performed with luminol (DuPont) and detected by autoradiography. Autoradiography of the filter for the same period of time in the absence of luminol demonstrated that radioactive DNA did not contribute to the signal.

## RESULTS AND DISCUSSION

It has been shown previously that crude extracts of *M. smegmatis* provide an activity that is required for L5 Int-mediated recombination *in vitro* (6). The protein responsible for this activity was purified from *M. smegmatis* by column chromatography (Fig. 1*a*), using *in vitro* recombination to assay for activity (Fig. 1*b*) (6). Fractions from the final chromatographic separation that strongly stimulate recombination appear to contain a single, small polypeptide approximately 13 kDa in size; N-terminal sequence analysis of this polypeptide produced a unique sequence of 38 amino acids (*N*-ALPQLTD-EQRAAALEKAAAARRARAELKDRLKRGGTNL; see Fig. 2*a*). These observations suggest that the mycobacterial protein that stimulates L5 integrative recombination is com-

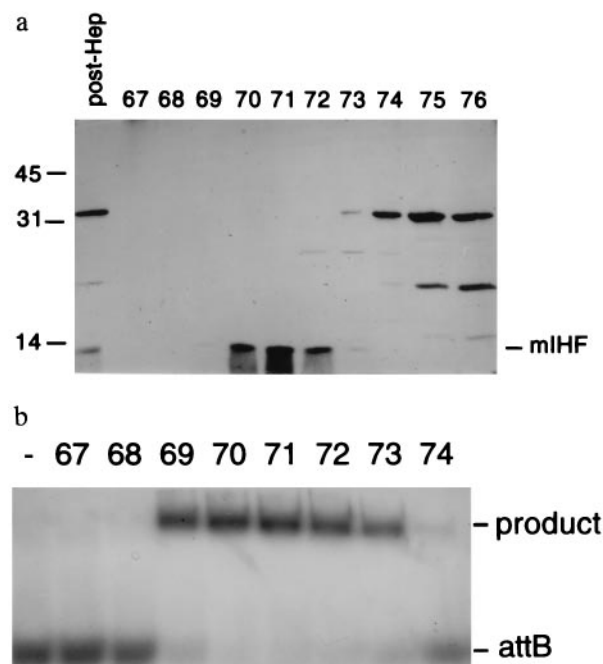


FIG. 1. Purification of *mIHF* from *M. smegmatis*. (a) SDS/PAGE of protein fractions eluted from an Econo-Pac S column. Aliquots of the sample loaded onto an Econo-Pac S column (post-Hep), and fractions 67-76 (as indicated) eluted from the column with a salt gradient are shown. (b) Econo-Pac S fractions 69-73 stimulate recombination. Aliquots of the fractions shown in *a* were added to recombination reactions containing supercoiled *attP* plasmid DNA (pMH39; 3.7 kb), a 0.6-kb linear radiolabeled *attB* DNA and L5 integrase; the products were separated by agarose gel electrophoresis and visualized by autoradiography. The Econo-Pac S fractions used for each reaction and a control reaction without *mIHF* (-) are indicated. Fractions 69-73 stimulate L5 Int-mediated recombination between *attP* and *attB* to produce a 4.3-kb linear product.

posed of a single polypeptide species. We have named this protein mycobacterial Integration Host Factor (mIHF).

Degenerate oligonucleotide probes designed from the N-terminal amino acid sequence were used to clone the *mIHF* gene of *M. smegmatis*. Southern hybridization studies show that there is only a single copy of the *mIHF* gene in the *M. smegmatis* genome (Fig. 2a). A segment of  $\approx 1$  kb of *M. smegmatis* DNA was sequenced, within which a single ORF encoding a putative 105-amino acid protein was identified. Residues 2–39 of the predicted amino acid sequence correspond precisely to the empirically determined amino acid sequence, indicating that this ORF is the *mIHF* gene (Fig. 2b).

A search of the sequence data bases identified a closely related gene of *M. tuberculosis*. Alignment of the *M. smegmatis* and *M. tuberculosis* DNA sequences shows that the

*mIHF* genes are very closely related, although the flanking regions are not (Fig. 2c). The mIHF protein sequences are almost identical, with only 6 differences (of a total of 105 residues), 3 of which are at the extreme C-termini (Fig. 2d; by contrast, the DnaA proteins of *M. smegmatis* and *M. tuberculosis* share only 81% identical residues). The high degree of similarity of the *mIHF* genes suggests that the proteins perform important functions in the mycobacteria. However, the mIHF proteins appear to represent a novel class of proteins because no other obvious similarities to known sequences were found, and specific sequence comparisons indicated that the mIHF proteins are not members of either the HU/IHF or H-NS families of proteins. In addition, no obvious similarities to HMG proteins, histones, or other DNA-binding proteins were found (data not shown).

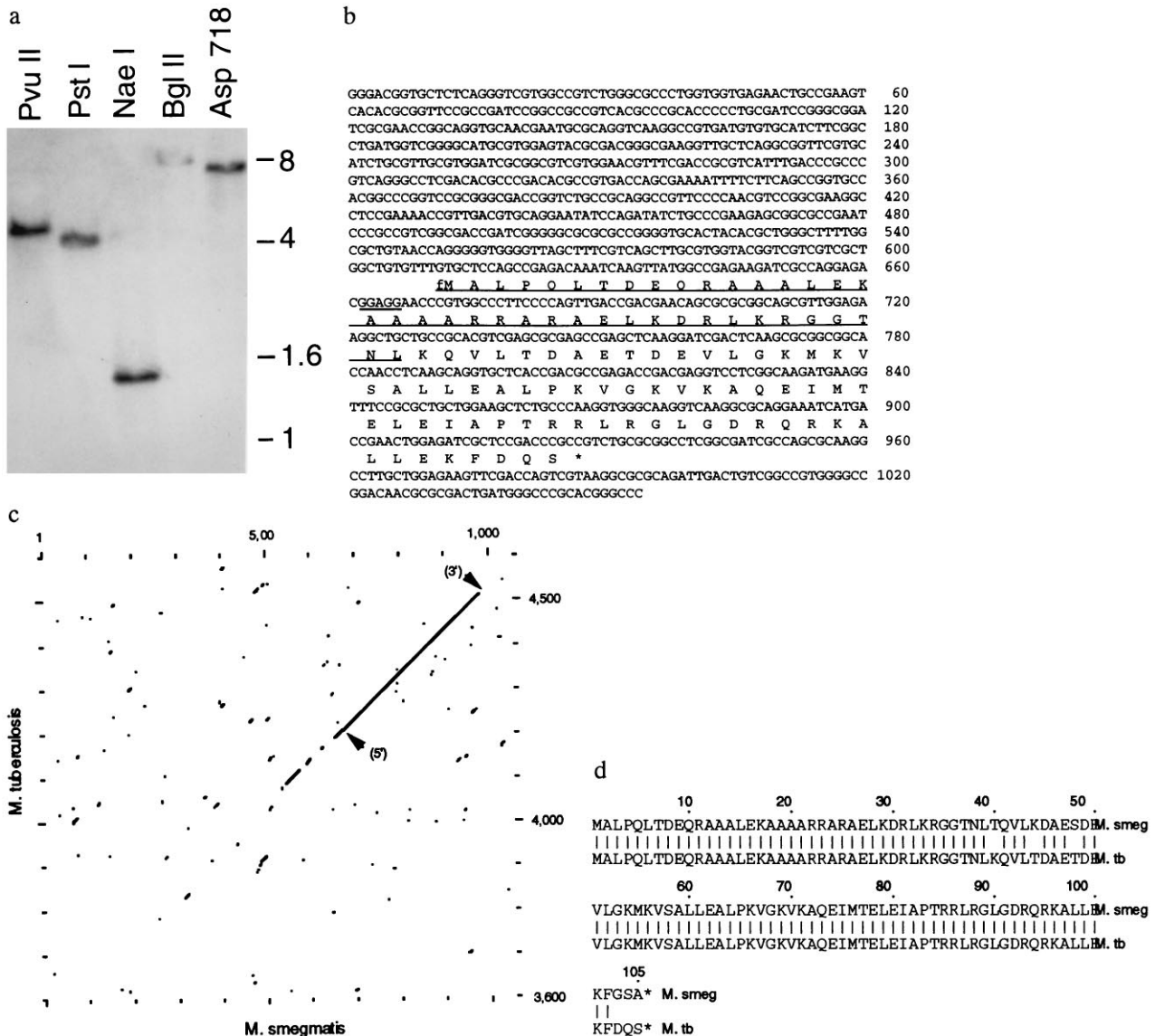


FIG. 2. Characterization of the *mIHF* gene of *M. smegmatis*. (a) *M. smegmatis* *mIHF* is a single-copy gene. DNA isolated from *M. smegmatis* mc<sup>2</sup>155 was digested with the restriction enzymes indicated and hybridized with a radiolabeled mIHF probe. Only one hybridizing band is observed in each lane, indicating that there is only a single *mIHF* gene. (b) Sequence of the *mIHF* gene of *M. smegmatis* and the predicted amino acid sequence of mIHF. The DNA sequence of a 1054-bp segment of the *M. smegmatis* genome (GenBank accession number U75344) and the putative amino acid sequence of the 105-residue mIHF protein are shown. The underlined amino acids (residues 2–39) correspond to those determined by N-terminal amino acid sequencing of the mIHF protein. No other genes were identified in this DNA segment. (c) Sequence alignment of the *mIHF* genes of *M. smegmatis* and *M. tuberculosis*. The DNA sequence shown in b was compared with a segment of *M. tuberculosis* DNA, using a window of 21 nt and a stringency of 14 nt. Arrows indicate the 5' and 3' ends of the mIHF coding regions. The *M. tuberculosis* sequence used was from coordinates 3600 to 4600 of cosmid MTCY21B4 (GenBank accession number Z80108). (d) Sequence alignment of mIHF proteins of *M. smegmatis* and *M. tuberculosis*. The *M. smegmatis* sequence is that shown in b and the *M. tuberculosis* sequence was derived by translation of coordinates 4207–4524 of cosmid MTCY21B4 (GenBank accession number Z80108). Amino acid identities are indicated by a vertical line.

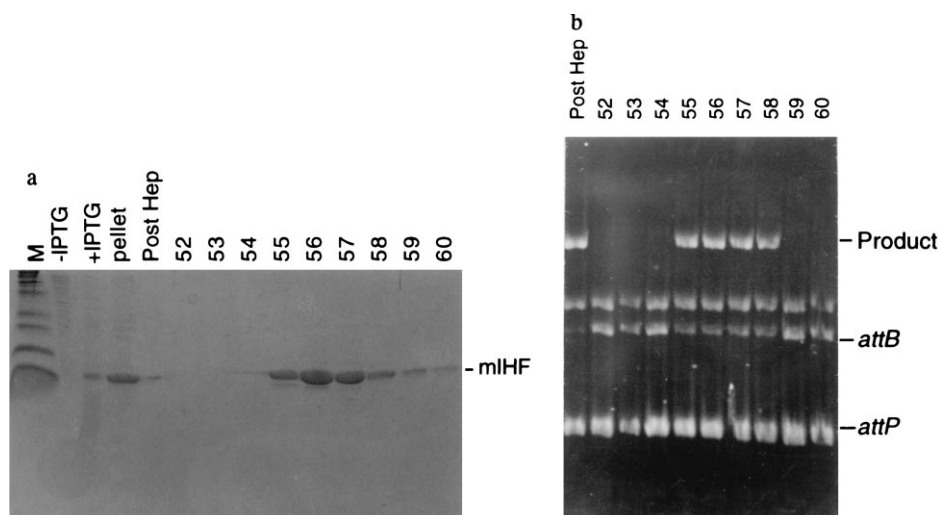


FIG. 3. Expression of mIHF in *E. coli* and purification of active protein. (a) SDS/PAGE showing purification of mIHF from *E. coli*. The *M. smegmatis mIHF* gene was inserted into the expression vector pET21a, and protein expression induced by addition of isopropyl  $\beta$ -D-thiogalactoside (IPTG). Lanes contain markers (M), cells before induction (- IPTG), cells after induction (+ IPTG), ammonium sulfate pellet (pellet), pooled Heparin fractions (Post Hep), and fractions 52-60 obtained by salt elution from a BioCAD carboxymethyl column. (b) mIHF purified from *E. coli* stimulates recombination. Recombination reactions containing an *attP* supercoiled DNA, linear *attB* DNA substrate, L5 Int, and carboxymethyl column fractions (as indicated) were incubated and the products were separated by agarose gel electrophoresis. DNA was identified by ethidium bromide staining and photographed. The recombinant product and the substrates (supercoiled *attP* DNA and linear *attB* DNA) are shown.

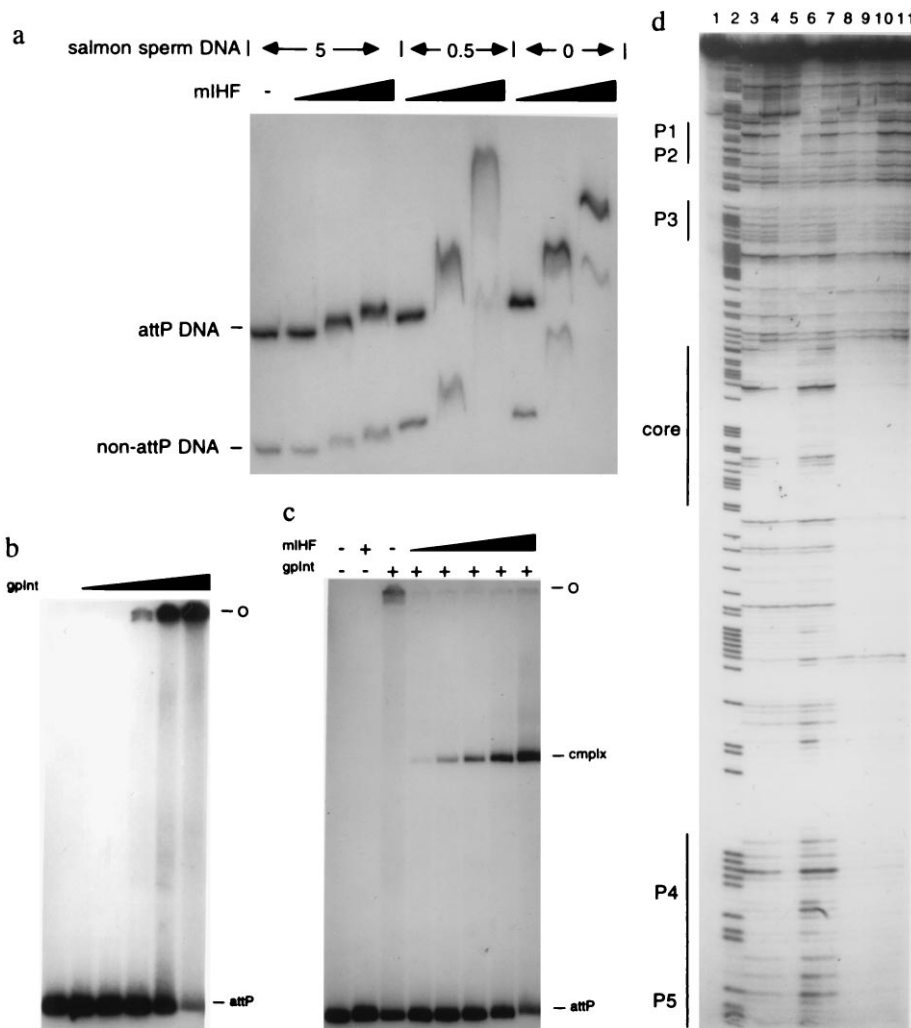


FIG. 4. DNA-binding properties of mIHF. (a) Native gel electrophoresis of mIHF-DNA complexes. Three sets of binding reactions were performed in which increasing concentrations of mIHF ( $\approx 0.75$ , 1.5, and 3  $\mu$ M) were incubated with a mixture of radiolabeled DNA fragments. The sets of reactions differed by the amount of salmon sperm DNA included (5  $\mu$ g, 0.5  $\mu$ g, or none, as indicated). Unbound DNAs containing *attP* or without *attP* are shown; protein-DNA complexes migrate slower than the unbound DNA. (b) Native gel electrophoresis of L5 Int-*attP* DNA complexes. Lanes show reactions containing radiolabeled *attP* DNA with either no Int (-) or with  $\approx 2.4$ , 7.3, 24, 73, or 240 nM Int. The position of free *attP* DNA and the origin of electrophoresis (O) are shown. (c) Native gel electrophoresis of L5 Int-mIHF-*attP* DNA complexes. DNA-binding reactions contained radiolabeled *attP* DNA, and either mIHF alone ( $\approx 870$  nM), Int alone ( $\approx 24$  nM), or with both mIHF and Int, as indicated. When both proteins were present, the concentration of Int was constant at 24 nM; mIHF was present at 8.7, 26, 87, 260, or 870 nM. The positions of free *attP* DNA, the origin of electrophoresis (O) and a protein-DNA complex (cmplx) are shown. (d) DNase I footprinting of L5 Int-mIHF-*attP* DNA complexes. Reactions in each lane are as follows: lane 1, 240 nM Int, no DNase I; lane 2, A+G marker; lane 3, DNase I only; lane 4, 73 nM Int; lane 5, 240 nM Int; lane 6, 1  $\mu$ M mIHF; lane 7, 300 nM mIHF; lane 8, 240 nM Int and 1  $\mu$ M mIHF; lane 9, 240 nM Int and 300 nM mIHF; lane 10, 73 nM Int and 1  $\mu$ M mIHF; lane 11, 73 nM Int and 300 nM mIHF. The positions of the arm-type sites P1-P5 and core-type sites for integrase binding are indicated.

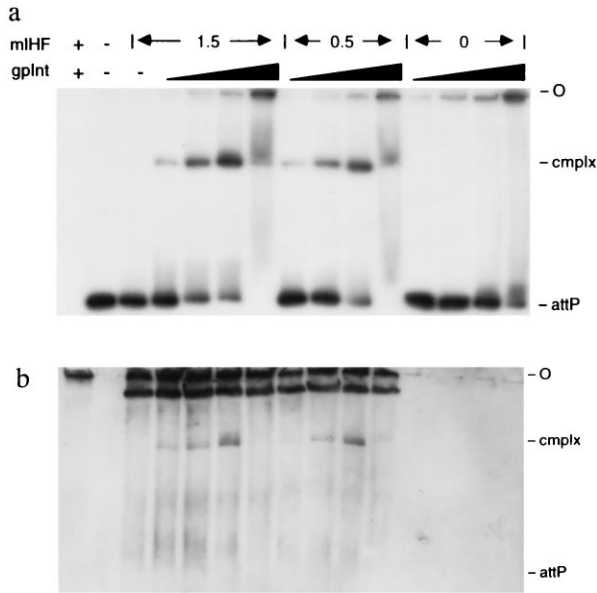


FIG. 5. Presence of mIHF in Int-mIHF-*attP* DNA complexes. Complexes were formed by addition of Int and mIHF to *attP* DNA, and the complexes were analyzed by native gel electrophoresis. All reactions contained 1  $\mu$ g salmon sperm DNA and with the exception of the leftmost lane, contained  $\approx 0.5$   $\mu$ g *EcoRI*-*Bam*HI digested pCPAR11 DNA of which the 364-bp *attP* fragment was radiolabeled. The three sets of reactions shown contain either 1.5  $\mu$ M, 0.5  $\mu$ M or no mIHF as indicated, and increasing amounts of Int (14, 48, 144, and 480 nM). The leftmost lane (no DNA control) contained 0.5  $\mu$ M mIHF and 144 nM Int. An autoradiograph of the wet gel immediately after electrophoresis is shown in *a*, and an immunoblot of the same gel with anti-mIHF serum is shown in *b*. The origin of electrophoresis (O) and the positions of *attP* DNA and Int-mIHF-DNA complex (cmplx) are shown. Other bands in the immunoblot probably correspond to mIHF nonspecifically bound to DNA. The reactions in the absence of mIHF (last four lanes) show that the antiserum does not cross-react with Int.

The *M. smegmatis* mIHF gene was inserted into an *E. coli* expression vector, and the mIHF protein was overexpressed and purified (Fig. 3*a*). The mIHF protein produced in *E. coli* stimulates recombination (Fig. 3*b*) and has a similar specific activity to the protein isolated from *M. smegmatis*;  $\approx 0.2$  pmol of mIHF from either source fully stimulates recombination of  $\approx 0.05$  pmol *attP* plasmid DNA (data not shown). These observations confirm that the 13-kDa protein isolated from *M. smegmatis* (Fig. 1) is the stimulating factor and show that the product of only a single gene is required for this activity (in contrast to heterodimeric *E. coli* IHF and HU) (15).

Purified mIHF does not appear to bind specifically to *attP* DNA *in vitro* (Fig. 4*a*). At high concentrations mIHF binds to DNA, but mIHF-DNA complexes are formed equivalently with both *attP* and a non-*attP* DNA fragment when little or no salmon sperm carrier DNA is present (Fig. 4*a*). In other experiments, mIHF exhibited no preference for binding to *attP* DNA over more than a dozen DNA fragments produced by digestion of plasmid pMH94 DNA (3) with *Bst*YI and *Sty*I even over a range of salt conditions (data not shown). In contrast, L5 Int does bind specifically to *attP* DNA but does not produce a single complex with well-defined electrophoretic mobility; instead, Int-*attP* complexes remain at the origin of electrophoresis or produce a smear (Fig. 4*b*). The nature of these complexes is not known, but we speculate that they are complicated networks formed by simultaneous binding of the two domains of Int to different *attP* DNA molecules. When both Int and mIHF are present with *attP* DNA, however, a complex with well-defined mobility is observed (Fig. 4*c*). Moreover, this complex is formed at concentrations of mIHF that are orders of magnitude lower than those needed to

produce the nonspecific shift seen in Fig. 4*a*. Thus, mIHF acts to stimulate the formation of a specific intasome complex.

DNase I footprinting experiments show that the presence of both Int and mIHF results in protection of parts of *attP* DNA not seen with either protein alone (Fig. 4*d*). In particular, most of the 70-bp region between the core and the P4 arm-type site is protected, and there may be some additional protection between the core and the P3 arm-type site (Fig. 4*d*). These protections could result from direct contact of mIHF with the DNA, or from an altered pattern of Int-binding in an mIHF-dependent fashion, but clearly do not result from independent binding of mIHF. The absence of any obvious protection by mIHF alone is consistent with its apparent lack of preference for *attP* DNA in the gel-shift assays but we cannot rule out the possibility that mIHF binds to *attP* DNA in regions where there is a paucity of DNase cleavage sites and that these interactions are unstable in the gel conditions used in Fig. 4*a*.

To address the question of whether mIHF is a constituent of the Int-*attP* complexes seen in Fig. 4*c* and *d*, we used anti-mIHF serum to immunoblot complexes separated by native gel electrophoresis (Fig. 5). These results demonstrate that mIHF is indeed an integral component of this complex (Fig. 5). We favor the parsimonious explanation that mIHF is in direct contact with the DNA, and as such provides the protection from DNase I seen in Fig. 4*d*. However, we cannot rule out the possibility that mIHF is only in contact with Int, such that the mIHF-dependent DNase I protection is a consequence of altered Int binding. Thus, mIHF could act by binding and stabilizing transient DNA bends, or through direct contact with Int (these are not mutually exclusive possibilities). We note, however, that mIHF is able to promote the formation of lambda *attL* intasome complexes, which require stabilization of lambda Int-mediated protein bridges (A. Segall, M.L.P., H. Nash, and G.F.H., unpublished observations). Nevertheless, integrative recombination of both L5 and  $\lambda$  is strongly dependent on the cognate integration host factor, which cannot be replaced by other DNA-binding proteins (6, 27).

We have described here a novel DNA-binding protein that is utilized by phage L5 as an integration host factor. The role of mIHF in L5 integrative recombination appears to be similar to that of *E. coli* IHF in  $\lambda$  integration: to promote the formation of specific Int-*attP* DNA higher-order structures that recombine with *attB* DNA. However, mIHF differs significantly from the IHF proteins of *E. coli* and other bacteria in that it stimulates recombination and intasome formation without sequence-specific recognition of *attP* DNA. Further studies will elucidate the role of mIHF in the regulation of the central physiological processes of the pathogenic mycobacteria.

We thank J. Hempel for determination of amino acid sequences and composition of mIHF. We also thank A. Segall, H. Nash, L. Barsom, G. Sarkis, and C. Peña for helpful comments. This work was supported by grants from the National Institutes of Health.

- Bloom, B., ed. (1994) *Tuberculosis* (Am. Soc. for Microbiol., Washington, DC).
- Snapper, S., Lugosi, L., Jekkel, A., Melton, R., Keiser, T. Bloom, B. R. & Jacobs W. R., Jr. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 6987-6991.
- Lee, M. H., Pascopella, L., Jacobs, W. R. Jr., & Hatfull, G. F. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 3111-3115.
- Hatfull, G. F. & Sarkis, G. (1993) *Mol. Microbiol.* **7**, 395-405.
- Hatfull, G. F. (1994) *ASM News* **60**, 255-260.
- Lee, M. H. & Hatfull, G. F. (1993) *J. Bacteriol.* **175**, 6838-6841.
- Mizuuchi, M. & Mizuuchi, K. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 3220-3224.
- Mizuuchi, M. & Mizuuchi, K. (1985) *Nucleic Acids Res.* **13**, 1193-1208.
- Mizuuchi, K., Weisberg, R., Enquist, L., Mizuuchi, M., Buraczynska, M., Foeller, C., Hsu, P. L., Ross, W. & Landy, A. (1981) *Cold Spring Harbor Symp. Quant. Biol.* **45**, 429-437.

10. Peña, C. E. A., Stoner, J. & Hatfull, G. F. (1996) *J. Bacteriol.* **178**, 5533–5536.
11. Nash, H. A. & Robertson, C. A. (1981) *J. Biol. Chem.* **256**, 9246–9253.
12. Hakimi Astumian, J. H., Waldman, A. S. & Scoocca, J. J. (1988) *J. Bacteriol.* **171**, 1747–1750.
13. Yu, A. & Haggard-Ljungquist, E. (1993) *J. Bacteriol.* **175**, 1239–1249.
14. Smith-Mungo, L., Chan, I. T. & Landy, A. (1994) *J. Biol. Chem.* **269**, 20798–20805.
15. Nash, H. (1990) *Trends Biochem. Sci.* **15**, 222–227.
16. Oberto, J., Drlica, K. & Rouviere-Janiv, J. (1994) *Biochimie* **76**, 901–908.
17. Friedman, D. I. (1988) *Cell* **55**, 545–554.
18. Mahan, M., Slauch, J. & Mekalanos, J. (1993) *Science* **259**, 686–688.
19. Craig, N. L. & Nash, H. A. (1984) *Cell* **39**, 707–716.
20. Gardner, J. F. & Nash, H. A. (1986) *J. Mol. Biol.* **191**, 181–189.
21. Robertson, C. A. & Nash, H. A. (1988) *J. Biol. Chem.* **263**, 3554–3557.
22. Goodman, S. D. & Nash, H. A. (1989) *Nature (London)* **341**, 251–254.
23. Moitoso de Vargas, L., Kim, S. & Landy, A. (1989) *Science* **244**, 1457–1461.
24. Kim, S., Moitoso de Vargas, L., Nunes-Düby, S. & Landy, A. (1990) *Cell* **63**, 773–781.
25. Kim, S. & Landy, A. (1992) *Science* **256**, 198–203.
26. Goodman, S. D., Nicholson, S. C. & Nash, H. A. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 11910–11914.
27. Segall, A. M. & Nash, H. A. (1993) *EMBO J.* **12**, 4567–4576.
28. Nunes-Düby, S. E., Smith-Mungo, L. I. & Landy, A. (1995) *J. Mol. Biol.* **253**, 228–242.
29. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY).
30. Devereux, J., Haeberli, P. & Smithies, O. (1984) *Nucleic Acids Res.* **12**, 387–395.
31. Pearson, W. R. & Lipman, D. J. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 2444–2448.
32. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990) *J. Mol. Biol.* **215**, 403–410.
33. Moitoso de Vargas, L., Pargellis, C. A., Hasan, N., Bushman, E. & Landy, A. (1988) *Cell* **54**, 923–929.
34. Peña, C. E. A., Lee, M. H., Pedulla, M. L. & Hatfull, G. F. (1997) *J. Mol. Biol.*, in press.