Genetic and Biochemical Investigation of Bxb1 gp47, An Unusual Recombination Directionality Factor

by

Andrew Savinov

BS in Chemistry and Molecular Biology, University of Pittsburgh, 2011

Submitted to the Graduate Faculty of

University Honors College in partial fulfillment

of the requirements for the degree of

Bachelor of Philosophy

University of Pittsburgh

2011

UNIVERSITY OF PITTSBURGH

University Honors College

This undergraduate thesis was presented

by

Andrew Savinov

It was defended on

April 4, 2011

and approved by

Karen M. Arndt, Professor, University of Pittsburgh Department of Biological Sciences

Nigel D. F. Grindley, Professor, Yale University Department of Molecular Biophysics & Biochemistry

Roger W. Hendrix, Distinguished Professor, University of Pittsburgh Department of Biological Sciences

Thesis Director: Graham Hatfull, Eberly Family Professor, University of Pittsburgh Department of Biological Sciences

Copyright © by Andrew Savinov

2011

Genetic and Biochemical Investigation of Bxb1 gp47,
An Unusual Recombination Directionality Factor

Andrew Savinov

University of Pittsburgh, 2011

The temperate mycobacteriophage Bxb1 infects and forms lysogens of Mycobacterium smegmatis, a fast-growing relative of and model for M. tuberculosis. In Bxb1, as in other bacteriophages, switching between the lysogenic cycle and lytic cycle depends on a site-specific DNA recombinase called an integrase. In Bxb1 the directionality of the DNA recombination process depends on the phage-encoded gp47 protein, which acts as the Recombination Directionality Factor (RDF). A number of lines of evidence suggest that Bxb1 gp47 has some additional biological role as well, however. First, very close homologues are found in 15 other mycobacteriophages, including phage L5, whose system for phage DNA integration / excision does not include the Bxb1 gp47 homologue. Second, Bxb1 gp47 and homologues are found clustered with genes for phage DNA replication.

Here we present our investigation into the putative multi-functionality of Bxb1 gp47. To begin, we performed a bioinformatic analysis which predicts Bxb1 gp47 and its homologues to contain a calcineurin-like phosphoesterase domain; this domain is predicted to confer a metal-dependent phosphatase activity on these proteins. Further, we analyzed the phenotypic repercussions of altering the Bxb1 gp47 gene by site-specific phage mutagenesis, performed using Bacteriophage Recombineering by Electroporation of DNA (BRED). Results from this work were consistent with the hypothesis that Bxb1 gp47 has an essential function in the lytic-cycle replication of Bxb1, and also showed that gp47 RDF activity is separable from the lytic-cycle function inferred for gp47. Finally, a number of variants of Bxb1 gp47 protein were overexpressed and purified for studies of RDF and phosphatase activity. RDF activity assays suggested that Motif I of the calcineurin-like phosphoesterase domain has no role in RDF activity, but that Motif V may be involved in both RDF and phosphatase activities. The phosphatase activity assays we have performed provide support for the hypothesis that Bxb1 gp47 is a manganese-dependent phosphatase enzyme. If this result is confirmed, Bxb1 gp47 will be revealed as a highly novel RDF with a secondary phosphatase activity.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

I would like to extend heartfelt thanks to Graham Hatfull and Pallavi Ghosh for all their support, mentorship, and training, thanks to which I was able to perform this work.

James Pan cloned the Bxb1 gp47 gene into pGEX-2T and overexpressed and purified the gp47-GST fusion protein. Pallavi Ghosh provided the pET28c-gp47 plasmid, the pMOS-attL plasmid, the pAIK5 plasmid, the pAIK6 plasmid, the pInt-H6 plasmid, and the complementation plasmids for *M. smegmatis* (pMS47, pMS54, and pMS(53-55)); Pallavi Ghosh also performed site-directed PCR mutagenesis of the pET28c-gp47 plasmid to produce pET28c-gp47* variants and designed the substrates and primers for the gp47 deletion and gp47-Ser153Ala BRED phage mutagenesis experiments.

1.0 INTRODUCTION

Bacteriophages are the viruses of bacteria. The majority of known phages have double-stranded DNA (dsDNA) genomes, and have a basic virion (viral particle) morphology consisting of a head containing the DNA, a tail, and a baseplate with tail fibers for attachment to the host cell, and can be classified broadly according to tail length (5), though such classification is of dubious use in characterizing phage phylogeny (19). Phages are incredibly numerous, estimated at ~10^{31} particles worldwide (35, 34). Some salient characteristics of the bacteriophages are their high degree of genetic diversity (12) and their potential as a source of novel enzymatic activities and protein folds; for instance, some 80% of known mycobacteriophage protein phamilies—which contain sequences having a certain minimal degree of relatedness (14)—have no predicted function based on comparisons to known proteins (12). From a practical standpoint, phages are of scientific interest for several major reasons. First, many bacteriophage proteins are potentially useful antibacterial toxins, and intact phages are also finding use as systemically or locally administered infective particles targeting pathogenic bacteria in what is called “phage therapy” (5). Second, a number of key proteins and expression systems for biotechnology and molecular biology, including restriction enzymes (11), have emerged out of basic research on phages. This trend has continued for decades, and there are likely many more useful applications to be discovered for phage biology and phage proteins.

One group of phages is the mycobacteriophages, which infect mycobacteria. The genus *Mycobacterium* includes *M. tuberculosis*, causative agent of tuberculosis, and *M. leprae*, which causes leprosy. *Mycobacterium smegmatis* (a fast-growing, non-pathogenic relative of and useful model for *M. tuberculosis*) is a convenient host for the study of mycobacteriophages. A particularly large number of mycobacteriophages have been characterized; 80 mycobacteriophage genomes have now been sequenced and comparatively analyzed (28). Studies of the mycobacteriophages of *M. smegmatis* have revealed a large degree of genetic mosaicism, informing our understanding of bacteriophage evolution (7, 25, 13). Genetic mosaicism, as discussed by Juhala *et. al.*, refers to related genomes each containing a "patchwork" of sequences drawn from a common pool, with a given pair of genomes sharing some closely related sequences and having highly disparate sequences in other genomic regions (15). A number of useful tools have also been developed using mycobacteriophages and mycobacteriophage proteins, including site-specific integration-proficient vectors for mycobacteria (26), a GFP reporter phage for detecting and typing *M. tuberculosis* (27), and a system for the recombineering (DNA recombination-based genetic engineering) of mycobacteria, including *M. tuberculosis* (33).

DNA recombination is a process with numerous important biological roles in life-forms ranging from bacteriophages to higher eukaryotes. To give one example, site-specific DNA recombination (occurring between defined and not necessarily similar DNA sites) is important in the production of the vastly diverse immunoglobulins used in the mammalian adaptive immune system (30). Site-specific DNA recombination is catalyzed by two distinct classes of enzymes, the tyrosine and serine recombinases (32). The current model of serine recombinase action is as follows ((9), and see Fig. **1b**). First, recombinases bind to the specific DNA-recombination sites. Then synapsis needs to occur between the recombinase-bound sites. Next, catalytic serine residues in the recombinases perform nucleophilic strand displacement reactions on the

DNA substrates' phosphodiester backbones, yielding concerted double-stranded cleavage of both substrates. Concerted strand rotation then occurs, followed by re-ligation of two DNA duplexes to yield the recombinant products and finally dissociation of the recombinase enzymes. Site-specific DNA recombination proteins have become powerful tools for the genetic engineering of eukaryotic cells, allowing for genetic manipulation of human cell lines via the Cre-Lox system (see for instance (24)) and of the malarial parasite *Plasmodium falciparum* via a phage recombinase (22, 4).

Phage-encoded DNA recombinases have a central role in the biology of certain bacteriophages, which are classified as temperate. Temperate phages are able to undergo two life-cycles, in both cases utilizing the host cell's metabolism to propagate ((29), and see Fig. **1B**). In both, a phage particle must initially inject its DNA into the host cell. In the lytic cycle, the phage hijacks the host machinery to replicate this DNA, which is then the template for producing phage mRNAs and thence phage proteins. Phage proteins have a variety of functions, with key proteins acting to promote amplification of phage DNA and inhibit normal host processes. Mature phage particles begin to assemble, with phage DNAs and often a few important proteins packaged inside protective protein coats. After many phage copies have been produced, phage lysis proteins break open the cell, and the new phage particles are released. In the lysogenic cycle, on the other hand, phage proteins catalyze integration of the phage DNA into the host chromosome and subsequently production of most phage proteins is suppressed. The virus thus enters a dormant state in which it is replicated each time the host cell copies its DNA and divides. Lysogens form at some frequency during standard infections, but are favored at high phage-to-cell ratios. Transition from the lysogenic cycle to the lytic cycle is mediated by phage systems for monitoring stress conditions on the host. Under such stress conditions, the phage DNA is cut back out of the host DNA and rapid lytic-cycle phage replication begins.

Bxb1 is a lysogenic mycobacteriophage which infects *Mycobacterium smegmatis*. In the Bxb1 lysogeny system, the Bxb1 serine-integrase (a serine recombinase enzyme) is responsible for catalyzing the integration of phage DNA into the host chromosome, as well as its excision out of the host chromosome. Serine integrases have not been studied as extensively as tyrosine integrases, and in particular the regulation of serine integrase site-specificity is one of the less well-understood questions in the DNA recombinase field (9). Bxb1 integrase catalyzes site-specific recombination between phage *attP* and bacterial chromosomal *attB* sites, yielding an integrated phage DNA flanked by recombinant *attL* and *attR* sites. Bxb1 gp47 protein, the Recombination Directionality Factor (RDF) for the Bxb1 system, causes Bxb1 integrase to catalyze excision of phage DNA from the host chromosome by altering the site-specificity of the integrase (8). When gp47 is present, integrase favors recombination of the *attL* and *attR* DNA substrates, resulting in phage DNA excision with reproduction of the *attP* and *attB* sites (Fig. **1A**). There are a number of reasons to use lysogenic bacteriophage recombination systems such as that of Bxb1 as models for site-specific recombinases. First of all, phage integrases and regulators thereof are relatively easy to isolate for study using the power of bacterial genetics. Second, simple genetic screens (20, 8) allow for relatively rapid identification of genes for both integrases and RDFs, and indeed two of the best-characterized recombination systems are from bacteriophages λ and P1 (10). L5 and Bxb1 are two phages whose recombination reactions have been reconstituted *in vitro* and are also considered to be excellent models in the site-specific recombination community. For the purposes of understanding serine integrase regulation, then, the Bxb1 system is an attractive model.

Bxb1 gp47 is an unusual RDF with no functionally characterized homologs. In particular, Bxb1 gp47 is not homologous to any other characterized RDF. However, a BLASTp (protein-protein BLAST) search reveals that 15 other mycobacteriophages, including phage L5, encode closely related uncharacterized proteins (Fig. **2A**). The criterion we used to delineate protein sequences

from the BLASTp search as “closely related” to Bxb1 gp47 was a pairwise comparison E-value < 10^{-80}, which for this set of sequences meant amino acid sequence identity of at least 58% between Bxb1 gp47 and the related sequence. In the case of phage L5 in particular the presence of a Bxb1 gp47 homolog (L5 gp54) is surprising, as L5 encodes a tyrosine-integrase and the known requirements for recombination do not include L5 gp54 (20). This strongly suggests that L5 gp54 has some other role to play in the phage life-cycle. Moreover, Bxb1 gp47 and its homologs are located among DNA replication genes in the phage genomes, which is consistent with a secondary role in replication given that phage genomes tend to be organized into clusters of functionally related genes. And indeed, earlier qPCR work in the lab (Ghosh, P. and G.F. Hatfull, unpublished data) has indicated that L5 gp54 is important for DNA replication in a phage L5 lysogen induced to switch into the lytic life-cycle. Further bioinformatic investigations also led us to predict that Bxb1 gp47 is a phosphoesterase (see Results).

Based on the bioinformatic and genomic information available, we made several hypotheses: first, that Bxb1 gp47 has a secondary role in the phage life-cycle; second, that this role relates to phage DNA replication; third, that Bxb1 gp47 and L5 gp54 have phosphoesterase (phosphatase) activities; and fourth, that the phosphoesterase activity of Bxb1 gp47 is the chemical basis for its putative secondary role in the phage life-cycle. In order to determine the veracity of these hypotheses, we have studied Bxb1 gp47 via (1) genetic engineering of phage Bxb1 and probing the phenotypic repercussions of mutations and (2) biochemical investigation of Bxb1 gp47 protein variants, investigating both RDF and phosphatase activities. We have also been interested in investigating any relationship which may exist between the RDF activity which has been characterized for gp47 and its putative phosphatase activity. It is intriguing that Ser153, despite being located within the sequence region of the predicted phosphoesterase domain (though *not* within any of the motifs) has a key role in promoting excision (*attL* x *attR*) catalysis.

2.0 MATERIALS AND METHODS

Bacterial Strains and Growth Media

M. smegmatis mc^2155 was grown in Middlebrook 7H9 liquid medium and Middlebrook 7H10 solid medium from Difco, supplemented with ADC and kanamycin as needed for selection of strains carrying plasmids for complementation studies. Middlebrook 7H9 medium was further supplemented with 0.05% Tween 80.

E. coli DH5α and *E. coli* BL21(DE3) pLysS (Invitrogen) were grown in LB broth or LB-agarose (Difco) supplemented as needed with antibiotics for selection. 2.0% D-glucose was added to LB broth for expression of H6-MBP-gp47 protein, as described below.

Plasmids

Base Plasmids Used

pLC3 is a modified form of the expression vector pET-21d (Novagen) with a Maltose Binding Protein (MBP)-coding region inserted into the multiple cloning site (MCS). pET-21d and pET-28c were obtained from Novagen. pGEX-2T was obtained from GE LifeSciences. The plasmids referred to here as pMS47, pMS54, and pMS(53-55) contain *M. smegmatis* origins of replication and encode promoterless Bxb1 gp47, L5 gp54, and L5 proteins gp53 – gp55, respectively; these plasmids only have a promoter for their resistance marker (Kan^R). pAIK5 and pAIK6 are previously described derivatives of pJL37 (16); the modified pAIK5 used here is an *M. smegmatis* replicating plasmid which encodes promoterless Bxb1 integrase and Kan^R, and

pAIK6 is a plasmid with no origin of replication which contains a 50-bp *attP* site and encodes Hyg^R.

Cloning and Subcloning

pET28c-gp47 was produced by PCR amplification of gp47 from Bxb1 genomic DNA with added Nde1 and Xho1 sites. The product was digested with Nde1 and Xho1 and ligated into Nde1- and Xho1-digested pET28c, as described for production of pET28a-based pPGgp47 (8).

pLC3-gp47* derivatives (wild-type gp47 and variants) were produced by subcloning from pET28c into pLC3. The Nde1 and Xho1 sites were used for excision of the gp47* coding sequence and ligation into the MCS of digested pLC3 vector.

pGEX-2T-gp47 was produced by PCR amplification of gp47 with flanking HindIII and Nde1 sites and digested with these enzymes, then ligated into digested pGEX-2T vector (GE LifeSciences).

PCR Mutagenesis

pET28c-gp47* point mutants in predicted calcieurin-like phosphoesterase motifs (Motif Knockout or Motif K-O mutants: pET28c-gp47-Motif1K-O, pET28c-gp47-Motif2K-O, pET28c-gp47-Motif3K-O, and pET28c-gp47-Motif5K-O) were produced by standard PCR mutagenesis with one primer containing the desired point mutations (one or two) flanked by complementary bases (2). See Table 4 for specific amino acids mutagenized in the Motif I and Motif V Knockout mutants discussed herein; see also Fig. 2C. Predicted metal-chelating residues were targeted in each motif.

Recombinant Protein Over-Expression and Purification

Bxb1 H6-gp47 and H6-gp47*: as described for pPGgp47 in Ghosh *et. al.* (8), BL21(DE3)pLysS cells were transformed with Bxb1 pET28c-gp47 or one of the various pET28c-gp47* variants. 1 L of LB with added kanamycin and chloramphenicol for selection of pET28c and pLysS was inoculated with BL21(DE3)pLysS:pET28c-gp47 or BL21(DE3)pLysS:pET28c-gp47* cells and grown to OD(600 nm) = 0.6 at 30^{o} C. Cultures were then induced with 0.6 mM IPTG and grown for an additional 4 h at 22^{o} C. After harvesting and freezing at -20^{o} C, cells were thawed and lysed in a 10 mM imidazole purification buffer containing 50 mM Tris-HCl (pH 8.0), 300 mM NaCl, 5% glycerol, and 10 mM imidazole. Sonication was performed in ~10 cycles of 30-s bursts with 60-s breaks, and the crude lysate was then centrifuged at 13,000 rpm for 45 min in a Sorval SLA-600 rotor (centrifuge temperature was maintained at 4^{o} C). The clarified lysate was incubated with Ni-NTA resin (Qiagen) (2 mL column volume) for 1-2 hrs at 4^{o} C and the mixture was then run through disposable affinity chromatography columns (Bio-Rad). The resin was washed with 50 column volumes 10 mM imidazole purification buffer, 10 column volumes 50 mM imidazole purification buffer, and 5 column volumes 75 mM imidazole purification buffer, and the fusion protein of interest was eluted with 5 column volumes 150 mM imidazole purification buffer.

Bxb1 Integrase-H6 was purified by nickel affinity chromatography with an Ni-NTA resin as described previously (16).

Bxb1 gp47-GST was overexpressed from recombinant plasmid pGEX-2T-gp47 and purified according to the recommended procedure for GST fusions from the supplier (GE Healthcare. Affinity Chromatography: Principles and Methods).

H6-MBP-gp47: pLC3-gp47 was transformed into BL21(DE3)pLysS. 1 L of LB + 2.0% D-glucose was inoculated with BL21(DE3)pLysS:pLC3-gp47 and grown to OD(600 nm) = 0.5-0.7 at 30^{o} C. Induction was then performed with 0.5 mM IPTG and cells were grown further for 3 h at 30^{o} C. Cells were harvested and frozen at -20^{o} C. They were then thawed and lysed into a pLC3 buffer containing 50 mM Tris (pH 8.0), 300 mM NaCl, and 5.0% glycerol. Sonication was performed in ~15 cycles of 15-s bursts with 15-s rests. Then centrifugation was performed at 13000 rpm for 45 min (at 4^{o} C) in a Sorvall SLA-600 rotor. Clarified lysate was incubated with ~2 mL amylose resin (Qiagen) for 1-2 h at 4^{o} C and the mixture was loaded onto a disposable column. The resin was washed with 20 column volumes of pLC3 buffer and then elution of the fusion was performed with 10 column volumes of pLC3 buffer + 10 mM maltose.

Bacteriophage Recombineering by Electroporated DNA (BRED)

Mutagenesis of Bxb1 by BRED was performed as described in Marinelli *et. al.* (21). In the case of the point mutant Bxb1 gp47-S153A six generations of lytic propagation were required to obtain a sample of mutant which appeared highly enriched in the desired mutant as per MAMA-PCR and was purely the desired mutant according to whole-genome sequencing. The gp47 deletion mutant was never detected in a total of ~1400 second-generation plaques screened by DADA-PCR.

Whole-Genome Sequencing of Mutant Bacteriophage DNA

Bxb1 gp47-S153A genomic DNA was extracted and then sequenced by Illumina Solexa sequencing technology on a Genome Analyzer IIx platform. Putative single-nucleotide polymorphisms (SNPs) vs. the published genomic sequence of wt phage were confirmed by PCR amplification of 500-bp genomic regions and sequencing though Genewiz.

***In vitro* DNA Recombination Assays**

These assays were performed as described previously (8) with a few modifications. Briefly, a circular DNA substrate containing an *attP* or *attL* site (pAIK6 and pMOS-*attL*) and a 60-bp linear DNA substrate consisting of an *attB* or *attR* site were incubated in a reaction buffer (as described) together with varying concentrations of H6-gp47 and Integrase-H6 variants; reactions were stopped by heat-killing of Integrase at 75^{o} C for 15 min, and products and substrates resolved on a 0.8% agarose gel prior to visualization with ethidium bromide staining and UV exposure.

Phosphatase Activity Assays

We performed fluorimetric assays based on the fluorigenic substrate fluorescein diphosphate (FDP). The assays are based on the commercial Sensolyte FDP Protein Phosphatase Kit from Anaspec (1). 100 µL reactions were performed following the manufacturer's recommendations, with some modifications as noted below. Reaction mixtures containing FDP and a gp47 variant in the provided reaction buffer were incubated for some amount of time in a 96-well plate; Calf Intestinal Phosphatase (CIP) was used as a positive control, and 1 mg/mL BSA was added to all reactions. Divalent metals in the form MCl_2 were added to certain reactions as specified in the Results. Reactions were stopped by the addition of 50 µL of the provided Stop Buffer from Anaspec.

3.0 RESULTS

Analysis of predicted domains: gp47 and homologues may contain a phosphoesterase domain

Using NCBI's CD-Search (Conserved Domain Search) tool, we have found that L5 gp54 is predicted to contain a calcineurin-like phosphoesterase domain (Fig. **2B**, E = 8.04 x 10^{-8}). Calcineurin is a serine/threonine phosphatase dependent on calcium and calmodulin cofactors; it is important in T-cell activation and other cellular processes. Calcineurin's active site contains zinc and iron (17). The calcineurin-like phosphoesterase domain is made up of four main motifs that are often associated with DNA polymerases in other systems; more specifically, calcineurin-like phosphoesterase domains are often found fused to DNA polymerase subunits in bacterial, archaeal, and eukaryotic organisms, as found in an analysis by Aravind and Koonin (3). Aravind and Koonin further hypothesize that the phosphoesterase activity of such proteins might be used to hydrolyze pyrophosphate during DNA replication. The calcineurin-like phosphoesterase domain is made up of 5 conserved motifs—four major ones with at least one universally conserved residue each, and one minor one with no universally conserved residues present. The domain is predicted to confer metal-dependent phosphoesterase activity based on the conserved metal-chelating histidines and glutamates found in the four major motifs. When we compare Bxb1 gp47 to L5 gp54 in a protein-protein BLAST alignment, we find that it contains most of the universally conserved residues of the calcineurin-like phosphoesterase domain motifs, with two conservative substitutions and a single Val—>Phe substitution (Fig. **2C**). The

available bioinformatic evidence suggested to us that Bxb1 gp47 and L5 gp54 may both be phosphatases, though the relevant substrate of these putative phosphatases is not readily apparent.

Deletion study of the Bxb1 gp47 gene: gp47 is implicated in lytic-cycle propagation

If Bxb1 gp47 is solely a recombination directionality factor (RDF), we would expect to see no effect on lytic-cycle propagation from genetic depletion of gp47 in mutant phages. However, if Bxb1 gp47 has a second function used in the lytic lifecycle of Bxb1, a gp47 deletion mutant would be expected to be less viable than the wild-type strain in lytic-cycle propagation. This is especially true if the hypothetical role of this protein in phage DNA replication is indeed present. In order to investigate the proposed role of Bxb1 gp47 in lytic propagation, I performed BRED phage mutagenesis (21) of Bxb1 to study phenotypic repercussions of modifying the gp47 gene.

The first mutation of interest was a deletion of gp47; we chose to make an internal deletion leaving intact 20 bp coding sequence on the 5' end and 42 bp on the 3' end of the gp47 coding strand (see Discussion). In order to study this mutation, I first performed co-electroporation of a 200-bp deletion substrate and wt Bxb1 genomic DNA into electrocompetent *M. smegmatis* mc^2155:pJV53. After a recovery period the culture of treated cells was mixed with a liquid top agar and spread over a bottom agar layer in a standard double agar overlay infectious center assay (18). After a growth period plaques—clearings on the lawn of host cells in the top agar—were recovered. These plaques are referred to here as primary (1^o) or first-generation plaques. A sample taken from one of twenty such primary plaques assayed was found to contain a proportion of gp47 deletion (Δgp47) mutant DNA, as well as wt Bxb1 DNA. Samples were tested for the presence of mutant phage DNA by a modified form of the mismatch amplification mutation assay (MAMA PCR) which has been termed DADA PCR ((21), and see Fig. **3**). Next I made various dilutions of the mutant-positive primary plaque sample, incubated with *M.*

smegmatis mc^{2}155 culture, mixed with a liquid top agar, and plated over a bottom agar layer in an infectious center assay. Hence the first-generation phage population was put through a cycle of lytic-cycle propagation. Deletion mutant Bxb1 could not be detected in ~520 tested secondary (2^o) plaque-forming units (pfu; 14 pfu individual 2^o plaques and ~510 pfu from two lysates) of second-generation phage from wt and *M. smegmatis*(RecA-):pAIK5 host cells (also referred to here as RecA-:pAIK5 cells; see below and Discussion). These data strongly suggested that the deletion mutant phage were lacking in their ability to propagate lytically.

Complementation was then attempted to rescue the Bxb1 gp47 deletion mutant and purify it. Infectious center assays of mutant-positive primary plaque samples were hence performed with *M. smegmatis*(RecA-):pMS47 host cells, which lack the RecA gene (to reduce homologous DNA recombination) and contain a plasmid encoding the gp47 protein. ~870 pfu of second-generation phage recovered after dilution, incubation with host cells, and plating were tested for the presence of deletion mutant DNA (~870 pfu total from individual plaque samples and multi-plaque samples—pools of plaques as well as lysates from an entire plate of plaques—were tested). However, deletion mutant phage could not be detected in any of these second-generation phage samples (see Discussion). The deletion mutation was thus found to be difficult to complement with pMS47. In total, then, ~1400 pfu of second-generation phage derived from primary mutant-positive plaques were tested and found to contain no detectable level of the deletion mutant (Fig. **3**); similar results were found in separate experiments with different initial mutant-positive primary plaques produced via BRED (data not shown).

Site-specific mutagenesis study of the Bxb1 gp47 gene: compromising excision catalysis has no effect on lytic propagation

In prior work we had implicated residue Ser153 of Bxb1 in promoting excisive (*attL* X *attR*) recombination in a random mutagenesis study of Bxb1 gp47 protein (Savinov, A., P. Ghosh,

and G.F. Hatfull unpublished data). Based on our understanding of this residue's role in gp47 and the fact that the mutation S153A greatly reduces excision promotion by gp47 I produced and purified a gp47-Ser153Ala RDF loss-of-function mutant of phage Bxb1. This mutant was made via two point mutations in the gp47 gene, generating the desired mutant protein-coding sequence and also a silent mutation which produced a Bpu10I restriction site useful for double-checking MAMA PCR assays for mutant phage DNA.

Recombineering to produce primary samples of mutant phage and infectious center assays to produce plaques of progeny from a phage sample were performed as described above and in Marinelli *et. al.* (21). Samples were tested for the presence of mutant phage by either selective MAMA-PCR assays or non-selective PCR of part of the gp47 gene followed by Bpu10I restriction digestion (the engineered site thus yields two bands, whereas wt Bxb1 yields a single band from the uncut PCR product). A large number of mutant-positive primary plaques were recovered, and a ~150-pfu second-generation (2^o) lysate sample from one mutant plaque was revealed to contain the designed mutation as well (see Fig. **4A**, **B**). Achieving a pure plaque of mutant phage took 3 more generations of lytic-cycle propagation (Fig. **4C**, **D**, **E**, **F**). Though the initial purification stages (generations 1-4) were carried out on a complementing strain of *M. smegmatis* carrying a plasmid for Bxb1 gp47 (*M. smegmatis*(RecA-):pMS47), the phage attained a homogeneous mutant population (generation 5) only via passage through wt *M. smegmatis*. Homogeneity of the generation 5 mutant plaque was assayed by wt-specific and gp47 Ser153Ala mutant-specific MAMA-PCR, as well as standard PCR of a ~500-bp region about the gp47 Ser/Ala153 coding sequence followed by Bpu10I restriction digestion. Finally, homogeneity of the mutant phage was confirmed by whole-genome sequencing of a purified phage DNA sample derived from this plaque, which yielded in an unambiguous mutant signal for the nucleotides coding for gp47 Ser/Ala153. Putative single-nucleotide polymorphisms in the mutant DNA sample vs. wt Bxb1 genomic DNA were also confirmed by DNA sequencing of

PCR products. The purified gp47 Ser153Ala mutant phage and wt Bxb1 phage were both assayed for lytic-cycle propagation on wt *M. smegmatis* mc^2155 host cells; it was found that there were no observable differences in plaque morphology or growth rate between the point mutant and wt phage (see Fig. **4D**).

Phosphatase and RDF Activity Assays

Given the evidence for Bxb1 gp47 bi-functionality from the BRED experiments, I was interested in determining whether gp47 in fact has the predicted phosphatase activity. In our efforts to purify a stock of gp47 to assay for this activity I used a pET28c-gp47 vector for production of N-terminally 6XHis-tagged gp47 protein (H6-gp47 herein), and we used PCR mutagenesis to produce variants yielding H6-gp47 proteins with universally conserved residues of the calcineurin-like phosphoesterase domain converted to alanines (referred to here as Motif K-O or Motif Knockout mutants; see Table 4), as well as other alanine substitutions based on prior studies (8). I also subcloned from these pET28c vectors into pLC3, a pET21-based vector that adds both an MBP fusion tag and a 6XHis tag to the N-terminus of expressed proteins (with the N→C terminal organization H6-MBP-Protein). In addition, we utilized a pGEX-gp47 plasmid vector to generate a C-terminal gp47-Glutathione S-Transferase fusion.

Using the pET28c-gp47* vectors, I overexpressed and purified H6-gp47, H6-gp47(Motif I K-O), and H6-gp47(Motif V K-O) (Fig. **5**); I also overexpressed and purified Int-H6 (histidine-tagged Bxb1 Integrase) from the pInt-His plasmid developed earlier in the lab (16). In order to investigate the relationship between RDF activity and the putative phosphatase activity I performed RDF function assays. Results from assays of two RDF activities (inhibition of integrative *attP* x *attB* activity and promotion of excisive *attL* x *attR* DNA recombination) are shown in Fig. **6** for the purified wild-type H6-gp47. Active H6-gp47 was successfully produced; the Motif I Knockout mutant was found to be fully active in gp47 RDF function, but the Motif V

Knockout mutant appears to be deficient in promoting *attL* x *attR* recombination. Inhibition of integrative recombination assays of the mutants (data not shown) similarly demonstrated that the Motif I Knockout mutant H6-gp47 does inhibit *attP* x *attB* recombination, while Motif V Knockout mutant H6-gp47 does not.

To investigate Bxb1 gp47 phosphatase activity I have performed fluorimetric assays using the fluorigenic generic substrate fluorescein diphosphate (FDP). The assays are based on the commercial Sensolyte FDP Protein Phosphatase Kit from Anaspec (1). Since the calcineurin-like phosphoesterases are predicted to be metal-dependent, we performed reactions in the absence or presence of various divalent metal ions (Mg++, Ca++, Zn++, Mn++ added as **M**Cl_2). With Calf Intestinal Phosphatase as the positive control, experiments showed significant divalent-metal dependent activity from the H6-gp47 preparations (data not shown), but this result is convoluted by endogenous phosphatases (Table **1**; Figs. **5**, **7**) found to be present in pET28c vector controls for the H6-gp47 purification. However, assays performed with a preparation of gp47-GST show significant phosphatase activity over a pGEX empty vector control prep with 16 mM manganese added (Table **2**). A weaker zinc-stimulated activity is also observed for the GST fusion. Purified H6-MBP-gp47 also has a significant phosphatase activity on the FDP substrate in the presence of divalent manganese (Fig. **8** and Table **3**); more trials are needed to determine whether a weak zinc-stimulated activity is observed with this fusion as well.

4.0 DISCUSSION

Our results support the hypothesis that Bxb1 gp47 has dual functional roles: one in lytic-cycle phage replication, and one as an RDF in the lysogenic-lytic cycle switch. Our results also support the hypothesis that Bxb1 gp47 is a phosphatase enzyme. The BRED experiments showed loss of lytic-cycle viability of a phage Bxb1 mutant with the gp47 gene deleted, but no loss of viability when excision-promoting function of Bxb1 gp47 was removed by the S153A mutation. The RDF assays reveal that Motif I of the calcineurin-like phosphoesterase domain has no role in RDF activity, but that Motif V may be important for RDF activity. The phosphatase activity assays performed show that Bxb1 gp47 has a manganese-dependent phosphatase activity on FDP, with a reasonable ~1/1000th of the activity of Calf Intestinal Phosphatase (CIP) on this same substrate. With regards to the relative activity of gp47 on FDP, it should also be noted that Calf Intestinal Phosphatase is a highly active and promiscuous phosphatase enzyme, with a previously demonstrated μM-range K_m value for a synthetic fluorigenic substrate (6).

In the BRED mutagenesis experiments with Bxb1 Δgp47 (Fig. **3**), primary mutant plaque samples could be produced but ~520 secondary plaque samples from wt or RecA-:pAIK5 host cells and ~870 secondary plaque samples from complementing strain RecA-:pMS47 cells never showed the presence of mutant phage as assayed by DADA-PCR. Thus the assay of essentiality suggested by Marinelli *et. al.* (21) has been performed, with the results showing that gp47 likely has a role in lytic propagation of Bxb1 phage. It should be noted in addition that one secondary lysate sample from complementing strain host cells did once show a faint mutant

product band in DADA-PCR (data not shown), but this result was not reproduced in a simultaneous trial of the same sample, suggesting that only a minimal amount of mutant phage DNA was present. The RecA-:pAIK5 cells produced a low level of integrase (the pAIK5 construct used here is promoterless, modified from the previously utilized pAIK5 with an *M. bovis* BCG *hsp*60 promoter (16)), which was expected to have no influence on gp47 deletion mutant viability under lytic-cycle dominated phage replication conditions; hence these cells were used as a control on the use of RecA- cells maintaining a Kan^R-marked plasmid, but were otherwise functionally like wt *M. smegmatis* mc^2155. From our results it appears that complementation with RecA-:pMS47 cells was ineffective or at least inefficient, and purification of the deletion mutant was not possible. This result was not entirely unexpected, because the pMS47 plasmid has no promoter on the gp47 gene; pMS47 was used because of previously observed cytotoxicity of gp47 to *M. smegmatis* with higher-expression vectors (Ghosh, P. unpublished data). In addition, the design of the deletion mutation largely ruled out the possibility of effects on neighboring genes due to deletion of gp47. It has long been recognized that overlapping genes occur in dsDNA bacteriophages such as λ (23). In the case of Bxb1 gp47, close inspection of the surrounding genomic region shows that there is open reading frame overlap between gp47 (positive strand 36787 – 36020) and gp48 (upstream of gp47, positive strand 36963 – 36784), with the stop codon of gp48 overlapping with the gp47 coding sequence. Looking downstream of gp47, the coding region is directly adjacent to that for gp46 (positive strand 36015 – 35875). In this case the predicted promoter region (-10 and -35 regions) for gp46 falls into the Bxb1 gp47 coding sequence. Hence, in order to eliminate effects on neighboring genes the gp47 deletion was designed as an internal deletion leaving 20 bp of gp47 intact on the 5' end of the gp47 sense strand (neighboring gp48) and 42 bp of gp47 intact on the 3' end (neighboring gp46). In the future confirmation of the role of Bxb1 gp47 in the lytic life-cycle will be pursued via BRED production of a Bxb1 phage with an internal stop codon in gp47, with stop codon suppression used for complementation. This should allow

complementation and hence purification of the deletion mutant; assaying the purified gp47 nonsense mutant's lytic-cycle viability will allow confirmation of gp47's importance in lytic-cycle replication. I have also designed substrates and primers for BRED phage mutagenesis to produce gp47 calcineurin-like phosphoesterase motif mutant Bxb1 phages; in future work these phage mutagenesis experiments should allow an analysis of whether there is a correlation between lytic-cycle viability and an intact gp47 phosphatase activity.

The BRED mutagenesis results with Bxb1 gp47-S153A (Fig. **4**) demonstrate that an RDF activity deficient mutant of gp47 still allows normal lytic replication, showing that the proposed secondary role of gp47 in the lytic cycle is separable from the RDF function (though this does not eliminate the possibility of some relationship between these functions).

The phosphatase activity assays performed on various gp47 constructs provide evidence that gp47 has a manganese-dependent phosphatase activity (Tables **2**, **3**). If these results are confirmed, I will have shown that Bxb1 gp47 is a novel RDF protein with a secondary phosphatase activity. Such an activity has never been demonstrated for an RDF previously. A few further background controls (vs. a H6-MBP empty vector purification control and an expressor strain background to control for endogenous proteins) with the H6-MBP-gp47 construct are needed to confirm this activity. In a more recent set of experiments I was not able to see the previously observed manganese-stimulated phosphatase activity with a separately prepared, dialyzed H6-MBP-gp47 stock (50% glycerol, 0 mM maltose). However these experiments were inconclusive, as positive controls revealed that the current nearly-depleted stock of FDP has undergone significant degradation or dilution with a strong reduction in the observed activity of CIP control samples (data not shown). Confirmatory experiments will shortly be performed with freshly prepared FDP substrate as well as a new preparation of undialyzed and dialyzed H6-MBP-gp47 and the aforementioned background protein controls. Future

experiments will also investigate dependence of the gp47 phosphatase activity on the calcineurin-like phosphoesterase domain: phosphatase activity assays will be performed on purified H6-MBP-gp47(Motif I K-O), H6-MBP-gp47(Motif II K-O), H6-MBP-gp47(Motif III K-O), and H6-MBP-gp47(Motif V K-O) proteins. The metal-dependent phosphatase activity of Bxb1 gp47 should depend on these motifs, and so a loss of phosphatase activity is expected upon "knockout" mutagenesis of universally conserved residues from the calcineurin-like phosphoesterase domain.

Using a nickel-NTA resin alone to purify H6-gp47 proved ineffective, as metal-dependent phosphatase activities were present in the impurities co-eluting with the gp47 construct (Table **1**). Both H6-gp47 and pET28c background control samples demonstrated activity over buffer background, but the difference in signal between experimental and control samples was not significant. We suspect that these impurities were selected by the nickel in the resin, given the divalent metal requirement for FDP dephosphorylation observed with the background controls (data not shown). However, the H6-gp47 preparations have provided some information on the relationship between the RDF function and the putative phosphatase function of gp47. RDF activity assays with H6-gp47 variants showed that gp47 lacking conserved residues of Motif I is fully active in RDF activity, whereas gp47 lacking conserved residues of Motif V appears to lack RDF functionality (Fig. **6**); these results suggest that some relationship does exist between the two activities. It will be interesting to follow up on this, as the location of Ser153 with respect to the calcineurin-like phosphoesterase domain motifs suggests that the RDF phosphatase functions depend on overlapping domains.

Although a comparison with FDP activity (see above) and a rough estimate of the standard deviation in the noise for these fluorimetric assays (data not shown) indicates that the observed H6-MBP-gp47 and gp47-GST activities should be considered as significant, it should be noted

that some caution is in order with fluorigenic assays for enzymatic activity. In particular, it is important to remember that one characteristic of fluorimetry is a very low detection limit (31), which makes it important to control for even seemingly minor autofluoresence effects due to buffer components in fluorigenic assays. However, great care was taken to have an equivalent reaction buffer between reaction samples and controls, particularly in the Table 3 data where undialyzed protein (in 10 mM maltose elution buffer) was assayed.

Though some further work is needed to confirm the significance of Bxb1 gp47 in lytic-cycle phage replication, our results do support a lytic-cycle role for this protein. Furthermore, the preliminary phosphatase assay data strongly suggest that gp47 has a phosphatase activity on FDP. If this calcineurin-like phosphoesterase domain-based activity is shown to indeed be involved in DNA replication in particular, as hypothesized by Aravind and Koonin (3), gp47 will be shown to be involved in both DNA recombination in the lysogeny-lysis transition and DNA replication in active phage propagation in the lytic cycle. This would make Bxb1 gp47 a highly novel bacteriophage protein. It will be interesting to see how prevalent this type of multi-functionality is amongst the RDFs of mycobacteriophage serine integrases, as well as serine recombinases in general. Additionally, if gp47 has a role in promoting the replication of phage DNA in the host at the expense of host DNA replication, it may have potential as an antimicrobial agent for *M. smegmatis* and perhaps also *M. tuberculosis*.

5.0 TABLES AND FIGURES

Tables 1, 2, 3. Fluorimetric phosphatase assays on fluorescein diphosphate substrate were performed in a 96-well-plate format. See Materials and Methods for details; briefly, a CIP control or gp47 variant sample was incubated for the indicated length of time at 25 ^{o}C with FDP in a reaction buffer from the supplier, supplemented as indicated with divalent manganese, magnesium, calcium, or zinc as **M**Cl2. Reactions were stopped with provided pH-shifting Stop Solution. Detection was performed with a 473 nm laser for excitation and an LBP emission filter. Signal is reported in arbitrary units (a.u.).

Table 1. Phosphatase Activity Assays of H6-gp47 (reaction time = 45 mins).

Sample	Signal – pET Background (a.u.)	[Protein] (μM)	(Signal – pET Background) a.u. / μM protein
gp47 + 16 mM MnCl2	0.23	1.09	0.214
gp47 + 16 mM CaCl2	1.74	1.09	1.60
gp47 + 16 mM MgCl2	0.86	1.09	0.794
gp47 + 16 mM ZnCl2	1.25	1.09	1.14

Table 2. Phosphatase Activity Assays of gp47-GST (reaction time = 45 mins).

Sample	Signal – pGEX Background (a.u.)	[Protein] (μM)	(Signal – pGEX Background) a.u. / μM protein
gp47-GST	0.89	2.92	0.303
gp47-GST + 16 mM CaCl2	2.00	2.92	0.686
gp47-GST + 16 mM MgCl2	-0.37	2.92	-0.125
gp47-GST + 16 mM MnCl2	66.10	2.92	22.7
gp47-GST + 16 mM ZnCl2	11.52	2.92	3.95
CIP Control	417.05	0.00414	101000

Table 3. Phosphatase Activity Assays of H6-MBP-gp47 (reaction time = 30 mins).

Sample	Signal – Buffer Background (a.u.)	[Protein] (μM)	(Signal – Buffer Background) a.u. / μM protein
H6-MBP-gp47	-0.94	0.3026	-3.12
H6-MBP-gp47 + 8 mM ZnCl2	0.88	0.3026	2.90
H6-MBP-gp47 + 8 mM MgCl2	-0.98	0.3026	-3.25
H6-MBP-gp47 + 8 mM CaCl2	0.103	0.3026	0.340
H6-MBP-gp47 + 8 mM MnCl2	22.20	0.3026	73.4
Calf Intestinal Phosphatase (CIP)	427.83	0.0041	103000
Buffer Control	0.00	0.0000	N/A
Buffer Control+ 8 mM MnCl2	-3.60	0.0000	N/A

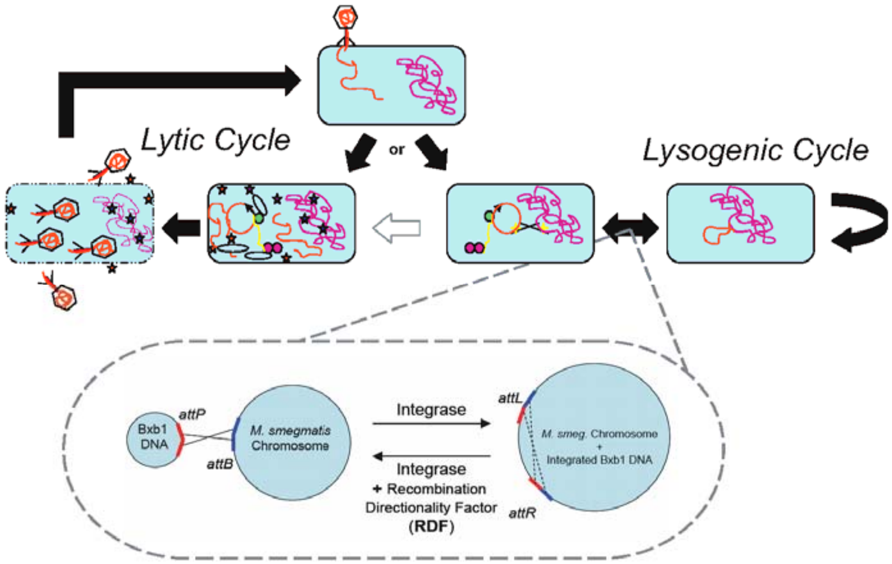
Table 4. Table of gp47 protein variants used in this work.

Bxb1 gp47 Variant	Features
H6-gp47	N-terminal 6XHistidine tag
H6-gp47-(Motif I K-O)	D10A mutation in predicted domain Motif I, N-terminal 6XHistidine tag
H6-gp47-(Motif V K-O)	H175A, H177A mutations in predicted domain Motif V, N-terminal 6XHistidine tag
H6-MBP-gp47	N-terminal 6XHistidine-tagged Maltose Binding Protein tag
gp47-GST	N-terminal Glutathione S-Transferase tag

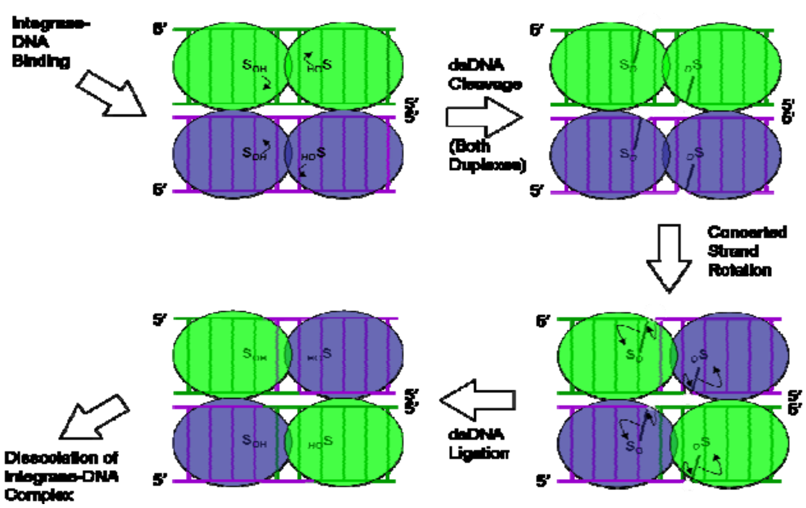
Figure 1.

(A)Temperate phage life-cycles and the lytic/lysogenic cycle transition mechanism in Bxb1.
(B) The mechanism of serine recombinase-catalyzed DNA recombination.
(C) Sequences of the *attP*, *attB*, *attL*, and *attR* DNA sites of Bxb1 and *M. smegmatis*, and their component half-sites. **Red** lines indicate the central dinucleotide of each *att* site.

A



B



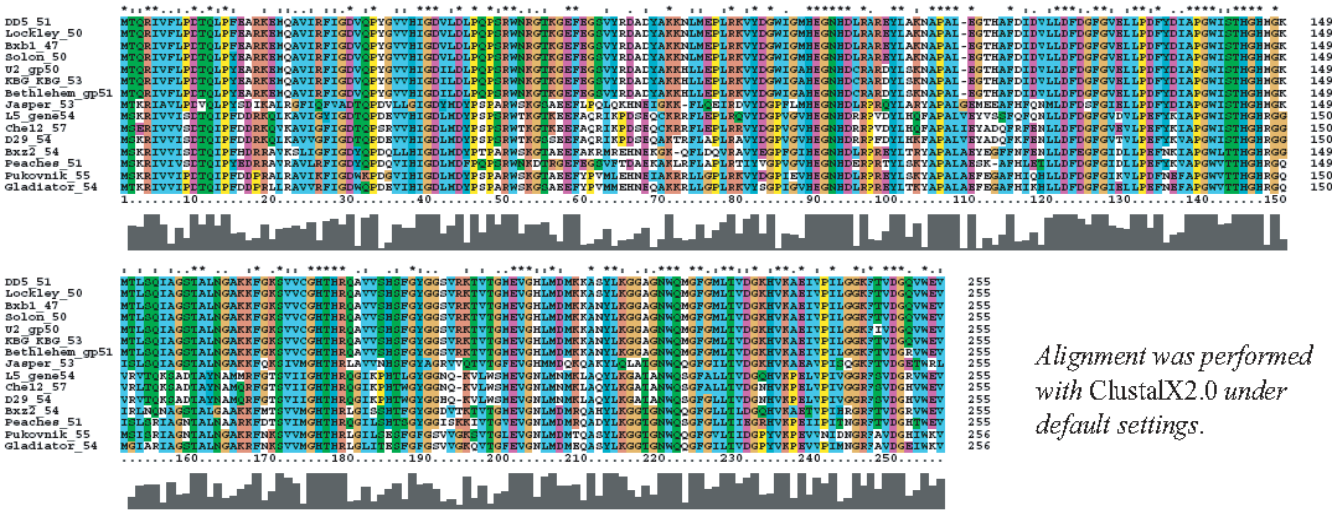
C

Figure 2.

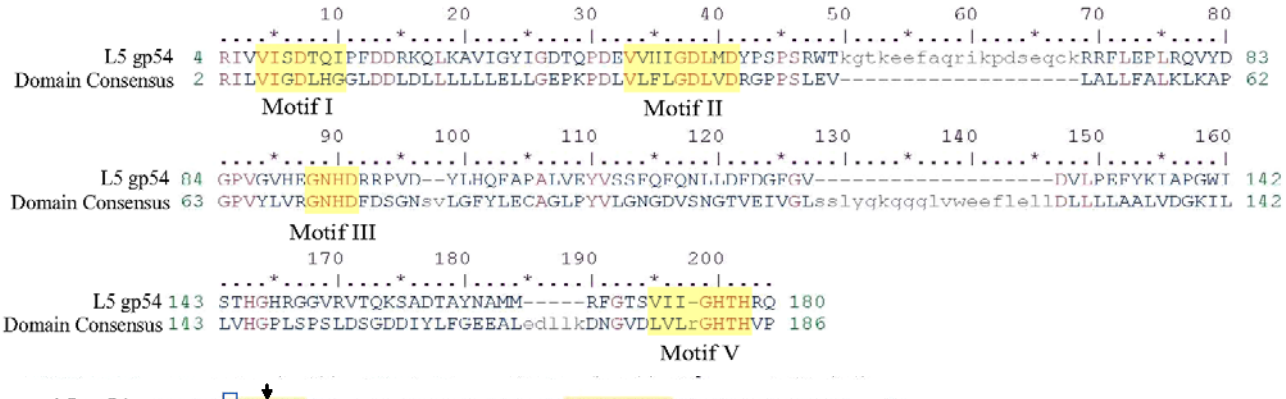
Sequence alignments providing bioinformatic evidence of a secondary function for Bxb1 gp47.

(A) Bxb1 gp47 vs. selected homologs.
(B) L5 gp54 vs. the calcineurin-like phosphoesterase domain consensus.
(C) L5 gp54 vs. Bxb1 gp47; the **blue** box indicates the Bxb1 gp47 Phe residue replacing the conserved Val of the domain consensus sequence, and the **magenta** box locates Bxb1 gp47 residue Ser153. The **black** arrows indicate residues mutated to alanines in the Motif I and Motif V K-O variants of gp47 (see also Table 4).

A



B



C

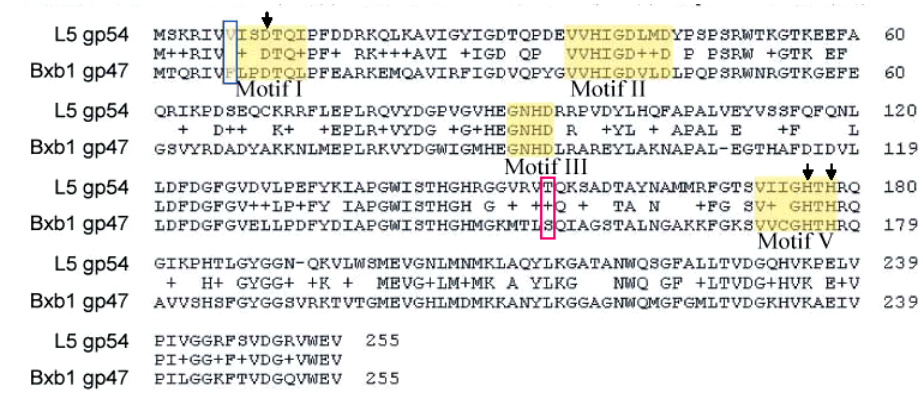


Figure 3.

BRED mutagenesis results for Bxb1 gp47 deletion mutant. Phage samples were screened by DADA-PCR; PCR products were separated by agarose gel electrophoresis and EtBr-stained.
(A) Primary plaque picks. The red box indicates the primary mutant plaque picked after recombineering which was positive for the mutation, as assayed by DADA-PCR; the red arrows indicate the position of the expected PCR product size for a sample containing deletion mutant phage DNA.
(B) A representative sample of results from second-generation plaque analysis is shown; the "mutant + 1°" control is a mutant-positive primary plaque sample, as is the "Δgp47, earlier recombineering" control in (A).

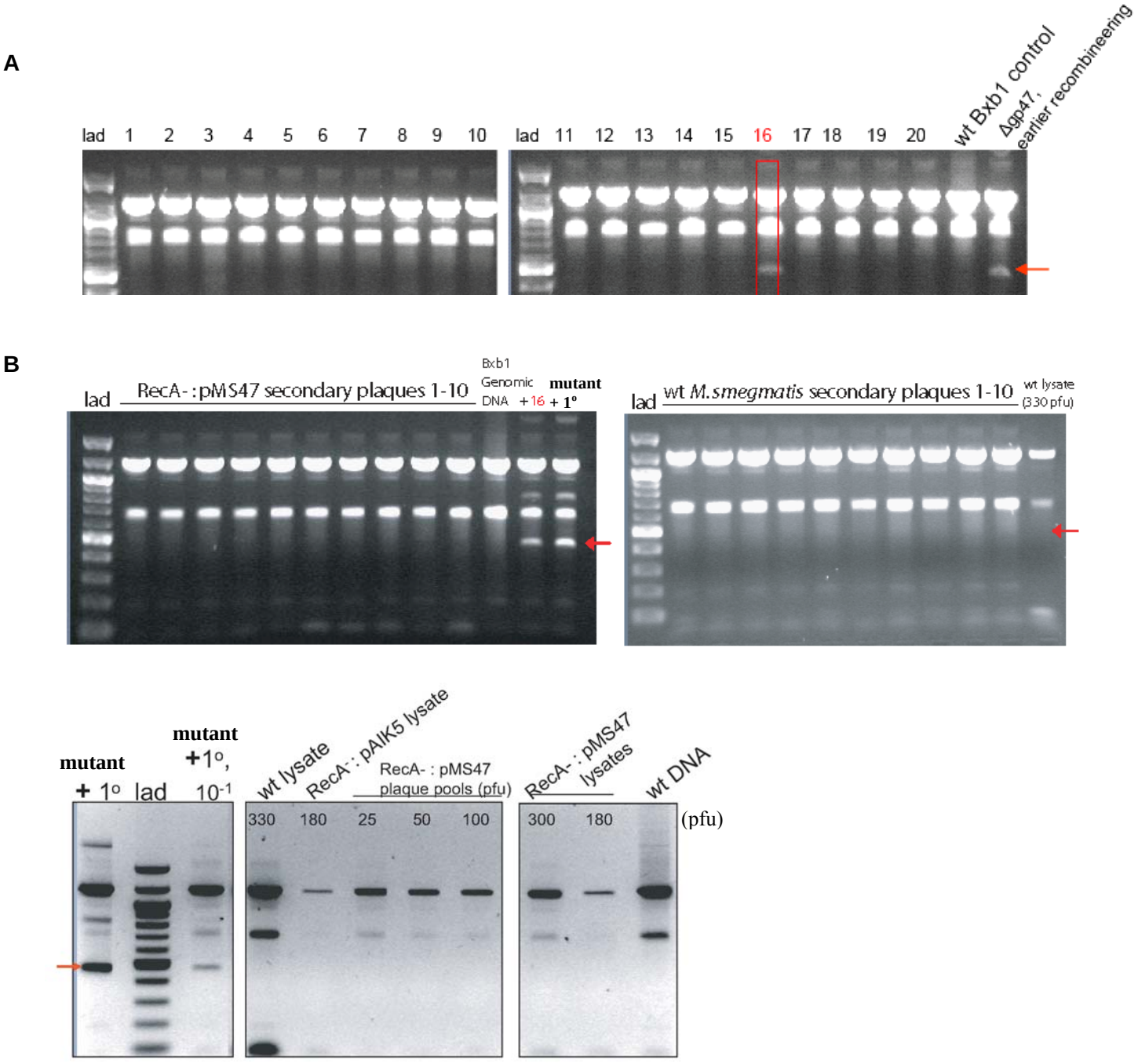
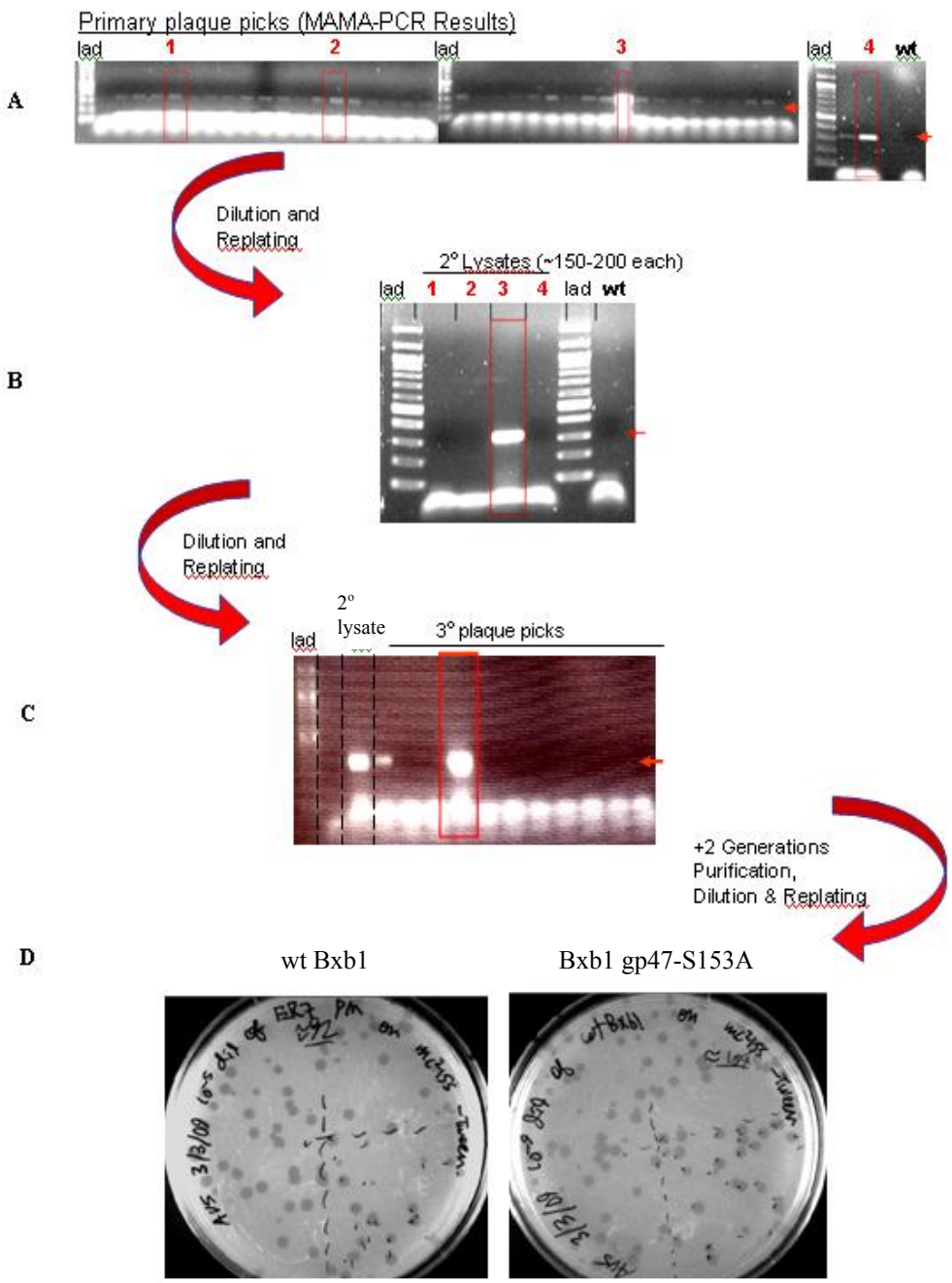


Figure 4.

BRED mutagenesis results for Bxb1 gp47-S153A mutant. Samples were screened by PCR assays, and the products run on agarose gels and EtBr-stained. Red boxes indicate phage samples diluted and replated to produce next-generation samples.

(A-C) Primary picks 1, 2, 3, and 4 were made off of wt *M. smegmatis*, RecA-:pMS47 cells, RecA-:pMS54 cells, and RecA-:pMS53-55 cells respectively and replated onto these same cell lawns to produce secondary lysates of 150-200 pfu each. The RecA-:pMS54 primary plaque ultimately yielded pure mutant phage after 5 generations of dilution & replating; the first three generations of the purification are shown here. Mutant purity was tested by whole-genome sequencing and local sequencing analysis for possible SNPs. Red arrows indicate expected position of mutant-specific MAMA-PCR product band.
(D) Pure mutant phage vs. wt phage compared in a lytic propagation assay; 2 representative plates are shown.
(E) 4th-, 5th-, and 6th-generation plaque sample analysis by mutant- and wt-specific MAMA-PCR. Expected positions of mutant/ wt MAMA-PCR products are indicated by red arrows.
(F) 4th- and 5th-generation plaque sample analysis via restriction digestion by Bpu10I.



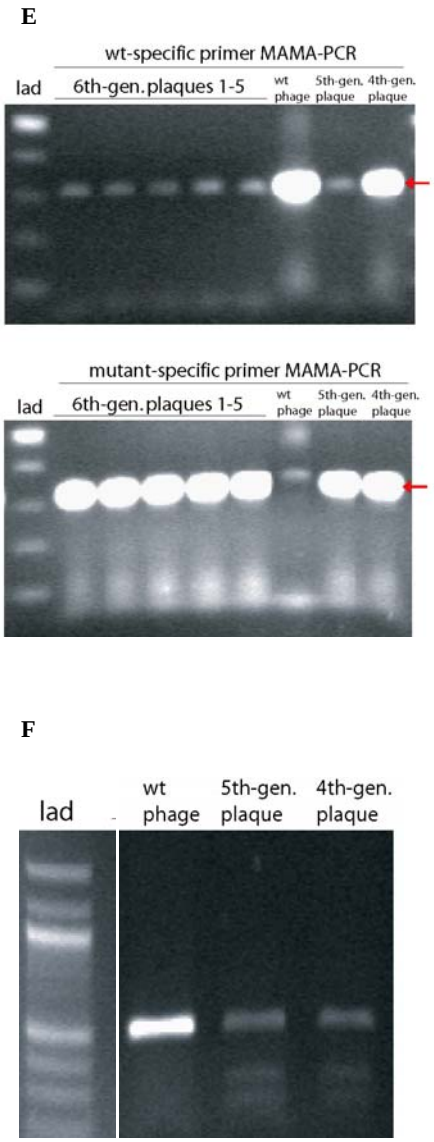


Figure 5.

Purifications of Int-H6 and H6-gp47 protein variants. Yellow arrows indicate the expected position of the recombinant protein band in the SDS-PAGE analysis. F-T is used as an abbreviation for “flow-through” fraction of the column purification.

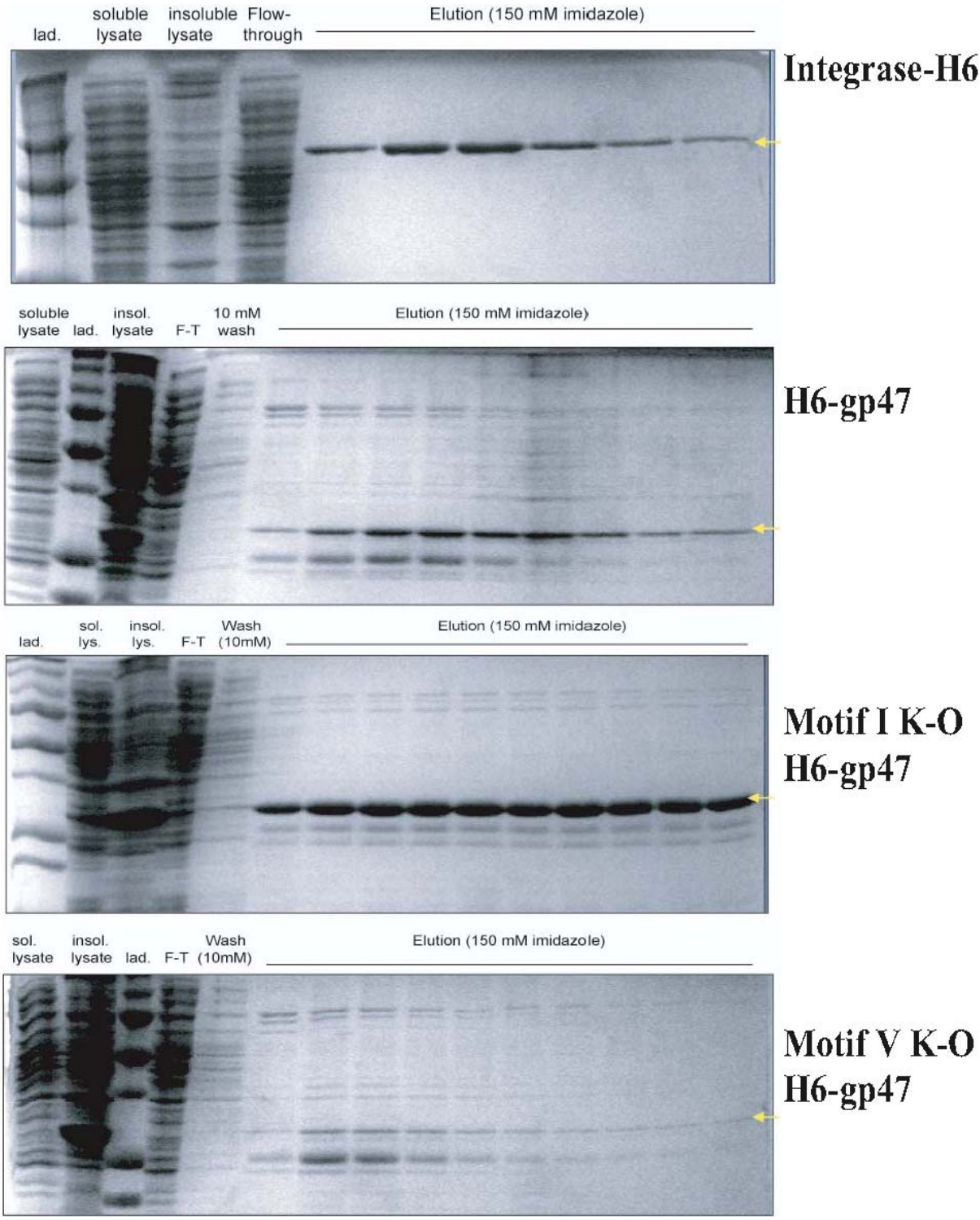
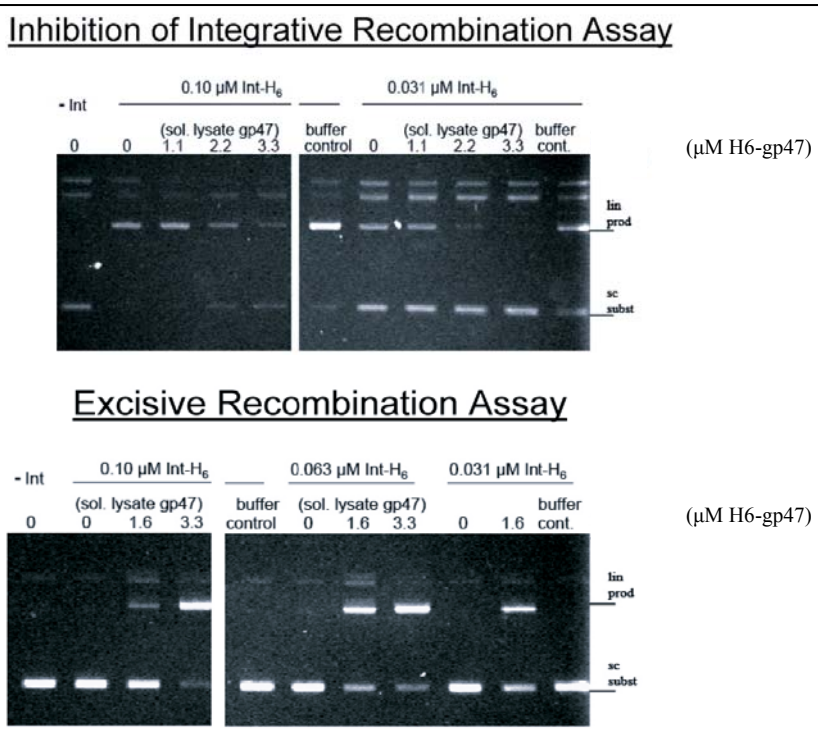


Figure 6.

Integrative DNA recombination reactions (*attP* X *attB*) and excisive DNA recombination reactions (*attL* X *attR*) were performed as described in Materials and Methods; briefly, a plasmid DNA substrate containing *attL* or *attP* and a 50-bp *attR* or *attB* dsDNA oligonucleotide were incubated with the indicated concentrations of integrase-H6 and H6-gp47 for 1 hr (integrative reactions) or 3 hrs (excisive reactions) at 25 °C, then stopped by heat-killing integrase (75 °C, 15 mins) and run out on 0.8% agarose gels. Gels were run without EtBr, and then EtBr-stained for analysis. A linear product band of known size was the indicator of recombination reaction progress.

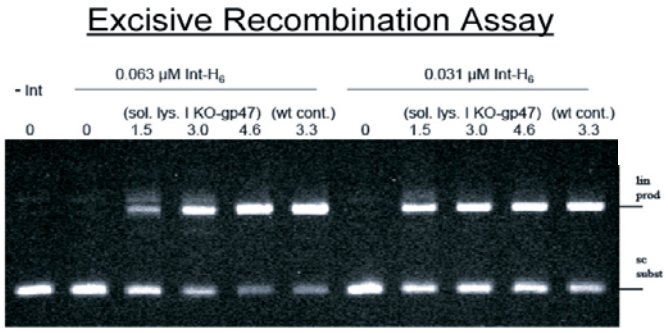
A

wt



B

H6-gp47



C

H6-gp47

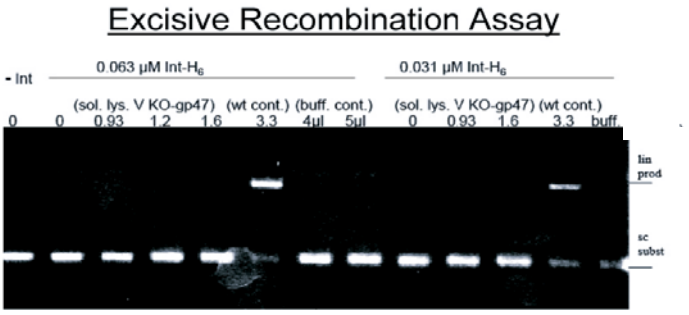
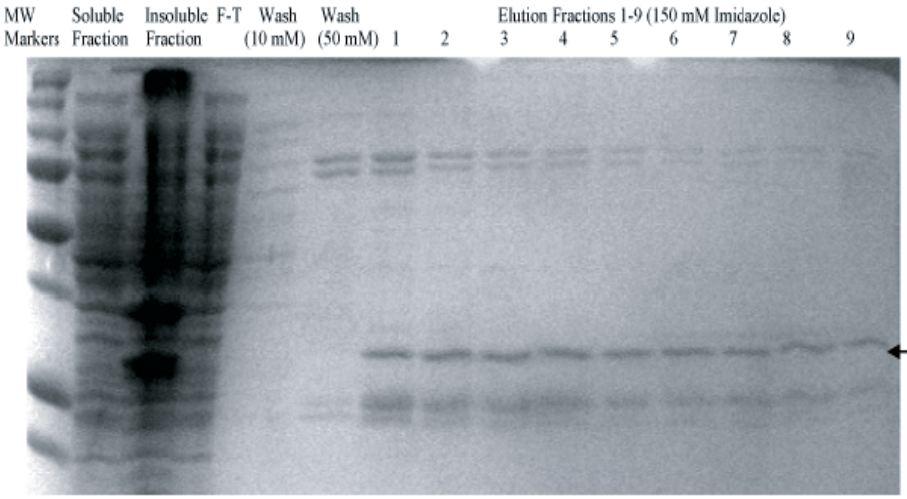


Figure 7.

Purification of H6-gp47 protein and the pET28c control purification used to determine presence of contaminant phosphatase activity in the H6-gp47 preparation (see Table 1).

A

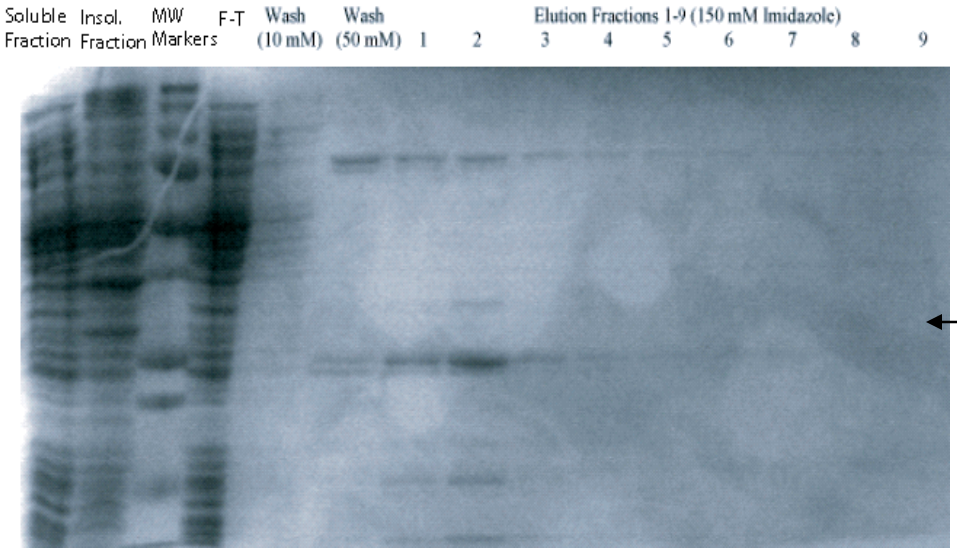
gp46-H6



H6-gp47

B

pET28c

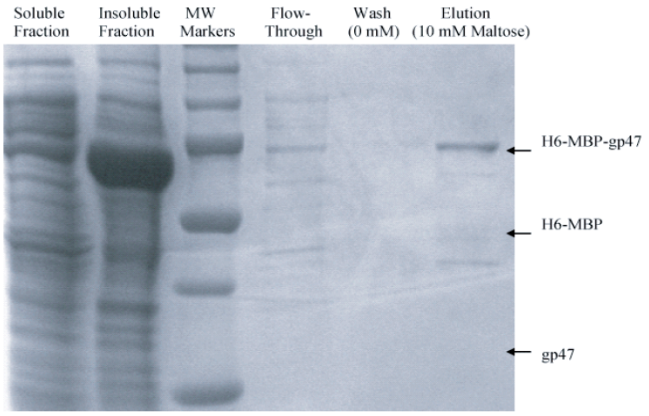


H6-gp47

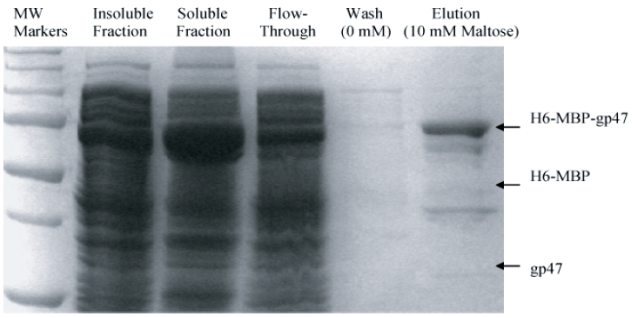
Figure 8.

Purifications of H6-MBP-gp47, pLC3 control, and BL21(DE3):pLysS expressor strain cells control used to determine phosphatase activity of H6-MBP-gp47 (see Table 3).

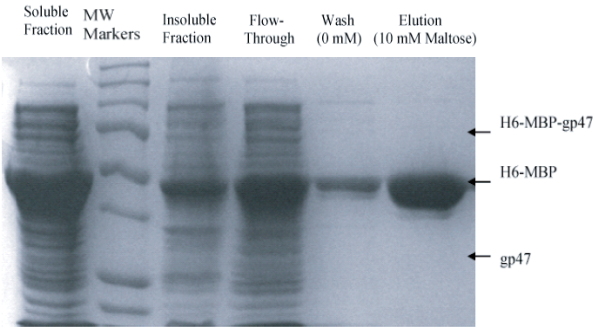
A
H6-MBP-47
(small-scale test)



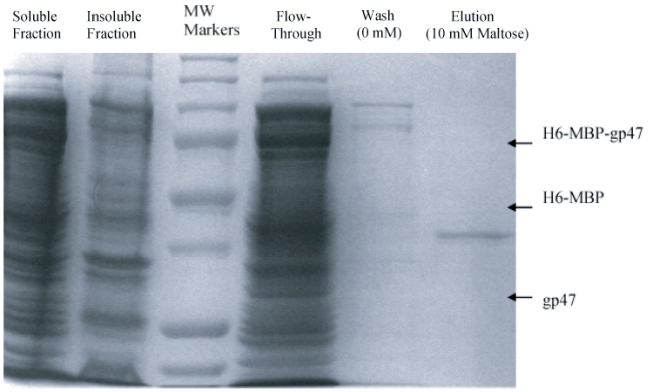
B
H6-MBP-47
(1-L purification)



C
pLC3 Vector Control
(1-L purification)



D
Expressor Strain Control
(1-L purification)



LITERATURE CITED

1. 2010. SensoLyte® FDP Protein Phosphatase Assay Kit *Fluorimetric*. [Online.]
2. **Allemandou, F., J. Nussberger, H. R. Brunner, and N. Brakch.** 2003. Rapid Site-Directed Mutagenesis Using Two-PCR-Generated DNA Fragments Reproducing the Plasmid Template. J Biomed Biotechnol **2003:**202-207.
3. **Aravind, L., and E. V. Koonin.** 1998. Phosphoesterase domains associated with DNA polymerases of diverse origins. Nucleic Acids Res **26:**3746-52.
4. **Crabb, B. S., and P. R. Gilson.** 2007. A new system for rapid plasmid integration in Plasmodium parasites. Trends Microbiol **15:**3-6.
5. **Deresinski, S.** 2009. Bacteriophage therapy: exploiting smaller fleas. Clin Infect Dis **48:**1096-101.
6. **Fernley, H. N., and P. G. Walker.** 1965. Kinetic behaviour of calf-intestinal alkaline phosphatase with 4-methylumbelliferyl phosphate. Biochemical Journal **97:**95-103.
7. **Ford, M. E., G. J. Sarkis, A. E. Belanger, R. W. Hendrix, and G. F. Hatfull.** 1998. Genome structure of mycobacteriophage D29: implications for phage evolution. J Mol Biol **279:**143-64.
8. **Ghosh, P., L. R. Wasil, and G. F. Hatfull.** 2006. Control of phage Bxb1 excision by a novel recombination directionality factor. PLoS Biol **4:**e186.
9. **Grindley, N. D., K. L. Whiteson, and P. A. Rice.** 2006. Mechanisms of site-specific recombination. Annu Rev Biochem **75:**567-605.
10. **Groth, A. C., and M. P. Calos.** 2004. Phage integrases: biology and applications. J Mol Biol **335:**667-78.
11. **Hartwell, L. H., L. Hood, M.L. Goldberg, A.E. Reynolds, L.M. Silver, and R.C. Veres.** 2008. Genetics: From Genes to Genomes. McGraw Hill Higher Ed., New York, NY.
12. **Hatfull, G. F.** 2010. Mycobacteriophages: genes and genomes. Annu Rev Microbiol **64:**331-56.
13. **Hatfull, G. F., S. G. Cresawn, and R. W. Hendrix.** 2008. Comparative genomics of the mycobacteriophages: insights into bacteriophage evolution. Res Microbiol **159:**332-9.
14. **Hatfull, G. F., M. L. Pedulla, D. Jacobs-Sera, P. M. Cichon, A. Foley, M. E. Ford, R. M. Gonda, J. M. Houtz, A. J. Hryckowian, V. A. Kelchner, S. Namburi, K. V. Pajcini, M. G. Popovich, D. T. Schleicher, B. Z. Simanek, A. L. Smith, G. M. Zdanowicz, V. Kumar, C. L. Peebles, W. R. Jacobs, Jr., J. G. Lawrence, and R. W. Hendrix.** 2006. Exploring the mycobacteriophage metaproteome: phage genomics as an educational platform. PLoS Genet **2:**e92.

15. **Juhala, R. J., M. E. Ford, R. L. Duda, A. Youlton, G. F. Hatfull, and R. W. Hendrix.** 2000. Genomic sequences of bacteriophages HK97 and HK022: Pervasive genetic mosaicism in the lambdoid bacteriophages. Journal of Molecular Biology **299:**27-51.
16. **Kim, A. I., P. Ghosh, M. A. Aaron, L. A. Bibb, S. Jain, and G. F. Hatfull.** 2003. Mycobacteriophage Bxb1 integrates into the Mycobacterium smegmatis groEL1 gene. Mol Microbiol **50:**463-73.
17. **Kissinger, C. R., H. E. Parge, D. R. Knighton, C. T. Lewis, L. A. Pelletier, A. Tempczyk, V. J. Kalish, K. D. Tucker, R. E. Showalter, E. W. Moomaw, and et al.** 1995. Crystal structures of human calcineurin and the human FKBP12-FK506-calcineurin complex. Nature **378:**641-4.
18. **Kropinski, A. M., A. Mazzocco, T. E. Waddell, E. Lingohr, and R. P. Johnson.** 2009. Enumeration of bacteriophages by double agar overlay plaque assay. Methods Mol Biol **501:**69-76.
19. **Lawrence, J. G., G. F. Hatfull, and R. W. Hendrix.** 2002. Imbroglios of viral taxonomy: genetic exchange and failings of phenetic approaches. J Bacteriol **184:**4891-905.
20. **Lewis, J. A., and G. F. Hatfull.** 2000. Identification and characterization of mycobacteriophage L5 excisionase. Mol Microbiol **35:**350-60.
21. **Marinelli, L. J., M. Piuri, Z. Swigonova, A. Balachandran, L. M. Oldfield, J. C. van Kessel, and G. F. Hatfull.** 2008. BRED: a simple and powerful tool for constructing mutant and recombinant bacteriophage genomes. PLoS One **3:**e3957.
22. **Nkrumah, L. J., R. A. Muhle, P. A. Moura, P. Ghosh, G. F. Hatfull, W. R. Jacobs, Jr., and D. A. Fidock.** 2006. Efficient site-specific integration in Plasmodium falciparum chromosomes mediated by mycobacteriophage Bxb1 integrase. Nat Methods **3:**615-21.
23. **Normark, S., S. Bergstrom, T. Edlund, T. Grundstrom, B. Jaurin, F. P. Lindberg, and O. Olsson.** 1983. Overlapping Genes. Annual Review of Genetics **17:**499-525.
24. **Paillard, F.** 1999. Reversible cell immortalization with the Cre-lox system. Hum Gene Ther **10:**1597-8.
25. **Pedulla, M. L., M. E. Ford, J. M. Houtz, T. Karthikeyan, C. Wadsworth, J. A. Lewis, D. Jacobs-Sera, J. Falbo, J. Gross, N. R. Pannunzio, W. Brucker, V. Kumar, J. Kandasamy, L. Keenan, S. Bardarov, J. Kriakov, J. G. Lawrence, W. R. Jacobs, Jr., R. W. Hendrix, and G. F. Hatfull.** 2003. Origins of highly mosaic mycobacteriophage genomes. Cell **113:**171-82.
26. **Pham, T. T., D. Jacobs-Sera, M. L. Pedulla, R. W. Hendrix, and G. F. Hatfull.** 2007. Comparative genomic analysis of mycobacteriophage Tweety: evolutionary insights and construction of compatible site-specific integration vectors for mycobacteria. Microbiology **153:**2711-23.
27. **Piuri, M., W. R. Jacobs, Jr., and G. F. Hatfull.** 2009. Fluoromycobacteriophages for rapid, specific, and sensitive antibiotic susceptibility testing of Mycobacterium tuberculosis. PLoS One **4:**e4870.
28. **Pope, W. H., D. Jacobs-Sera, D. A. Russell, C. L. Peebles, Z. Al-Atrache, T. A. Alcoser, L. M. Alexander, M. B. Alfano, S. T. Alford, N. E. Amy, M. D. Anderson, A. G. Anderson, A. A. Ang, M. Ares, A. J. Barber, L. P. Barker, J. M. Barrett, W. D. Barshop, C. M. Bauerle, I. M. Bayles, K. L. Belfield, A. A. Best, A. Borjon, C. A. Bowman, C. A. Boyer, K. W. Bradley, V. A. Bradley, L. N. Broadway, K. Budwal, K. N. Busby, I. W. Campbell, A. M. Campbell, A. Carey, S. M. Caruso, R. D. Chew,**

C. L. Cockburn, L. B. Cohen, J. M. Corajod, S. G. Cresawn, K. R. Davis, L. Deng, D. R. Denver, B. R. Dixon, S. Ekram, S. C. Elgin, A. E. Engelsen, B. E. English, M. L. Erb, C. Estrada, L. Z. Filliger, A. M. Findley, L. Forbes, M. H. Forsyth, T. M. Fox, M. J. Fritz, R. Garcia, Z. D. George, A. E. Georges, C. R. Gissendanner, S. Goff, R. Goldstein, K. C. Gordon, R. D. Green, S. L. Guerra, K. R. Guiney-Olsen, B. G. Guiza, L. Haghighat, G. V. Hagopian, C. J. Harmon, J. S. Harmson, G. A. Hartzog, S. E. Harvey, S. He, K. J. He, K. E. Healy, E. R. Higinbotham, E. N. Hildebrandt, J. H. Ho, G. M. Hogan, V. G. Hohenstein, N. A. Holz, V. J. Huang, E. L. Hufford, P. M. Hynes, A. S. Jackson, E. C. Jansen, J. Jarvik, P. G. Jasinto, T. C. Jordan, T. Kasza, M. A. Katelyn, J. S. Kelsey, L. A. Kerrigan, D. Khaw, J. Kim, J. Z. Knutter, C. C. Ko, G. V. Larkin, J. R. Laroche, A. Latif, et al. 2011. Expanding the diversity of mycobacteriophages: insights into genome architecture and evolution. PLoS One **6:**e16329.

29. **Ptashne, M.** 1992. A genetic switch : phage [lambda] and higher organisms, 2d ed. Cell Press : Blackwell Scientific Publications, Cambridge, Mass.
30. **Roth, D. B., and N. L. Craig.** 1998. VDJ recombination: a transposase goes to work. Cell **94:**411-4.
31. **Skoog, D. A., F. J. Holler, and S. R. Crouch.** 2007. Principles of instrumental analysis, 6th ed. Thomson Brooks/Cole, Belmont, CA.
32. **Smith, M. C., and H. M. Thorpe.** 2002. Diversity in the serine recombinases. Mol Microbiol **44:**299-307.
33. **van Kessel, J. C., and G. F. Hatfull.** 2007. Recombineering in Mycobacterium tuberculosis. Nat Methods **4:**147-52.
34. **Weinbauer, M. G.** 2004. Ecology of prokaryotic viruses. FEMS Microbiol Rev **28:**127-81.
35. **Whitman, W. B., D. C. Coleman, and W. J. Wiebe.** 1998. Prokaryotes: the unseen majority. Proc Natl Acad Sci U S A **95:**6578-83.