AD _____

# Cloning Sequencing and Structural Manipulation of the Enterotoxin D and E Genes from Staphylococcus aureus

Annual Summary Report

John J. Iandolo

June 1, 1988

Supported by

U. S. ARMY MEDICAL RESEARCH AND DEVELOPMENT COMMAND
Fort Detrick, Frederick, Maryland 21701-5012

Contract No. DAMD17-86-C-6055

Kansas State University
Manhattan, Kansas 66506



Approved for public release: distribution unlimited.

SECURITY CLASSIFICATION OF THIS PAGE

# REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

| | |
|---|---|
| 1a. REPORT SECURITY CLASSIFICATION<br>Unclassified | 1b. RESTRICTIVE MARKINGS |
| 2a. SECURITY CLASSIFICATION AUTHORITY | 3 DISTRIBUTION / AVAILABILITY OF REPORT<br>Approved for public release; distribution unlimited |
| 2b. DECLASSIFICATION / DOWNGRADING SCHEDULE | |
| 4. PERFORMING ORGANIZATION REPORT NUMBER(S) | 5. MONITORING ORGANIZATION REPORT NUMBER(S) |
| 6a. NAME OF PERFORMING ORGANIZATION<br>Division of Biology<br>Kansas State University<br>6b. OFFICE SYMBOL (If applicable) | 7a. NAME OF MONITORING ORGANIZATION |
| 6c. ADDRESS (City, State, and ZIP Code)<br>Ackert Hall<br>Manhattan, Kansas 66506 | 7b. ADDRESS (City, State, and ZIP Code) |
| 8a. NAME OF FUNDING / SPONSORING ORGANIZATION US Army Medical Research & Development Command<br>8b. OFFICE SYMBOL (If applicable) | 9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER<br>Contract No. DAMD17-86-C-6055 |
| 8c. ADDRESS (City, State, and ZIP Code)<br>Fort Detrick<br>Frederick, Maryland 21701-5012 | 10. SOURCE OF FUNDING NUMBERS<br>PROGRAM ELEMENT NO. 61102A; PROJECT NO. 3M16 1102BS12; TASK NO. AA; WORK UNIT ACCESSION NO. 119 |

11. TITLE (Include Security Classification)
Cloning Sequencing and Structural Manipulation of the Enterotoxin D and E Genes from Staphylococcus aureus

12. PERSONAL AUTHOR(S)
John J. Iandolo

| 13a. TYPE OF REPORT | 13b. TIME COVERED | 14. DATE OF REPORT (Year, Month, Day) | 15. PAGE COUNT |
|---|---|---|---|
| Annual | FROM 11/1/86 TO 10/30/87 | 1988 June 1 | 23 |

16. SUPPLEMENTARY NOTATION

| 17. COSATI CODES | | | 18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number) |
|---|---|---|---|
| FIELD | GROUP | SUB-GROUP | |
| 06 | 13 | | Biotechnology, BW, Cloning, Enterotoxins, Sequencing, Staphylococcus aureus, Toxins, Gene cloning |
| 06 | 16 | | |

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

We cloned the gene for exfoliative toxin A from the chromosome of S. aureus expression vector λgt11. The DNA fragment expressing exfoliative toxin A (ETA) was subcloned in E. coli and unlike the etb gene was fully expressed. Both eta and etb were completely sequenced and the coding regions for both toxins were compared. Significant stretches of both DNA and protein sequence similarities were found.

We also report the sequencing of the enterotoxin D gene (entD) of S. aureus 485. The protein sequence of SED is highly similar to the other enterotoxins (especially SEA) and streptococcal pyrogenic exotoxin A.

The phage attachment sites of the lipase inactivating bacteriophage (L54a) have been identified, cloned and sequenced. A core sequence of homology within the lipase gene and the phage genome serve as sites for reciprocal recombination. The resultant integrated phage is flanked by this core sequence. Lastly, we were also able to map the bacteriophage functions necessary and sufficient for integration and excision of the phage.

| | |
|---|---|
| 20. DISTRIBUTION / AVAILABILITY OF ABSTRACT<br>☐ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS | 21. ABSTRACT SECURITY CLASSIFICATION<br>Unclassified |
| 22a. NAME OF RESPONSIBLE INDIVIDUAL<br>Virginia M. Miller | 22b TELEPHONE (Include Area Code) 301/663-7325<br>22c. OFFICE SYMBOL SGRD-RMI-S |

DD Form 1473, JUN 86 *Previous editions are obsolete.* SECURITY CLASSIFICATION OF THIS PAGE

TABLE OF CONTENTS

FOREWORD

In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and Use of Laboratory Animals of the Institute of Laboratory Animal Resources, National Research Council (NIH Publication No. 86-23, Revised 1985).

Citations of commercial organizations and trade names in this report do not constitute an official Department of the Army endorsement or approval of the products or services of these organizations.

During the past year work continued on the three projects reported in the 1985-1986 annual report. We have made considerable progress on the major goals of the original proposal which included the enterotoxins D and E, the exfoliative toxins and studies of the lipase determinant as a model for $\beta$-toxin regulation. A summary of progress over the past year is presented below.

Exfoliative Toxin

A 1.7 kb HindIII restriction fragment of DNA was isolated from plasmid pIJ002 and cloned into the replicating form of bacteriophage M13mp18 and mp19 DNA and transformed into E. coli JM103. Single stranded DNA was isolated from the bacteriophage produced by the transfected cells. This DNA was then sequenced by the dideoxy chain terminator method of Sanger using [$\alpha$-$^{35}$S]dATP instead of $^{32}$P. The sequence (Fig. 1) was analyzed by computer and shown to contain an open reading frame (ORF) that compared favorably with the predicted amino acid analysis of exfoliative toxin B (ETB). A likely methionine initiation codon was found at position 181 within range of a suitable ribosome binding site. The ORF which begins at 181 is 822 bases in length and terminates at position 1000. Translation of this ORF identified 22 of the 26 N-terminal amino acid residues (N-terminus = Lysine) as well as the C-terminal residue (Lysine) of ETB which were determined by chemical sequencing methods. A 31 amino acid signal peptide precedes the toxin molecule with an alanine residue at the proposed cleavage site, where processing of the precursor occurs to yield the mature protein.

Shortly after these sequence data were published, several errors were discovered which required redetermination of the sequence. These data are shown in Figure 2 and reflect substantial changes. Additional chemical sequence data obtained from Drs. Schmidt, Spero and Johnson-Winegar at

USAMRIID agrees totally with the protein derived from the corrected DNA sequence and provides an important internal control of the sequence data. The new derived protein sequence data indicates that the ETB molecule is 277 amino acid residues and has a 31 residue signal peptide which when cleaved leaves a mature protein of 246 residues. The molecular weights of the species are 30,769 for the precursor and 27,318 for the mature protein. Furthermore, there is complete agreement between the DNA derived sequence and the chemically determined sequence for the first 40 amino acids (Fig. 2, underlined) and between the first 48 amino acids of a cyanogen bromide derived peptide of ETB beginning at residue 172 and continuing through residue 219 (Fig. 2, underlined). The composition of the protein is interesting in that there is neither tryptophane or cysteine present and the molecule lacks the cysteine loop found in the enterotoxins. The transcription signals originally reported (Publication #1) remain unchanged and show -35 and -10 promoter sequences, a ribosome binding site and a transcription stop signal that closely approximates the canonical sequences. Following these studies we began cloning of the eta gene. This element is chromosomal so a different strategy was utilized to isolate it. Bulk chromosomal DNA prepared from S. aureus UT0002 was used to construct a bacteriophage lambda gtll library. The plaques (40,000) were screened with rabbit antiserum that had been extensively adsorbed with lysates of the E. coli host strain. Plaques binding antibody were identified with $^{125}$I-labeled Protein A and exfoliative toxin A (ETA) production was confirmed by Western blot. Five clones were found that reacted positively with the antiserum. One of these phage was randomly selected for further study. The DNA insert isolated from this phage was approximately 3.2 kilobases and was recloned into the shuttle vector pLI50 and transformed into

E. coli LE392. Immunoblots confirmed that this plasmid contained the eta structural gene which was expressed in and biologically active in E. coli.

Deletion analysis further localized the gene to a 1391 bp fragment. This fragment was sequenced by the Sanger dideoxy method and is presented in Fig. 3. The G+C content of eta is 31% and is typical of the S. aureus genome. However, the G+C content of the 150 bp sequence upstream from the methionine start codon (nucleotide 313) was even lower in G+C content (19%), suggesting that the region could serve as the potential binding site for RNA polymerase to initiate transcription. A potential -35 sequence and a -10 sequence that could serve as promoter regions were identified (Fig. 3). Furthermore, the probable ATG translation start site is preceded by the sequence GGATGA, which qualifies as a potential ribosome binding site. A potential transcription stop codon is located at position 1154 and is followed 79 bp downstream by a stem-loop structure at positions 1232-1259.

Translation of the ORF (Fig. 3) yielded a 280 amino acid polypeptide that corresponded to the published properties of ETA. A 38 amino acid signal peptide precedes the N-terminus of the mature ETA protein which is cleaved immediately after the sequence Ala-Lys-Ala of residues 36-38. Removal of the signal peptide results in a mature ETA protein containing 242 amino acid residues with a molecular weight of 26,950. The sequence was identical to that of peptides of ETA that were determined by automated Edman degradation. A contrasting finding was that the C-terminal amino acid is glutamic acid rather than lysine. This result was substantiated independently by O'Toole and Foster (J. Bacteriol. 169: 3910-3915. 1987).

The amino acid composition of the DNA derived mature protein sequences of ETA and ETB were compared to the chemically derived values (Table 1). The

data are quite similar and confirm that ETA has a single tryptophane and methionine residue and the lack of cysteine. ETB was found to lack cysteine and tryptophane as already indicated. Comparison of the amino acid composition of both ETA and ETB indicates that they are reasonable similar proteins that are rich in polar amino acids.

Direct comparison of the protein sequences of ETA and ETB is shown in Fig. 4. Three prominent regions of similarity are evident in which the match was extensive. The first occurred in the N-terminal portion of the molecule at positions 46 - 70 (20 of 25 residues match, 80%), the second near the middle at positions 106 - 134 (17 of 29 residues match , 58%) and the third near the C-terminus at positions 201 through 221 (17 of 21 residues match, 81%). No other regions of significant similarity were present. The total number of amino acids matched by computer alignment was 110 (45%) out of an average of 245 residues. This extensive similarity might not have been predicted because of the lack of antigenic relationship between the toxins. However, when the relative hydropathicities are compared (Fig. 5), it is clear that much of the sequence of each toxin represents highly conserved domains in which the amino acid differences are fairly conservative. We interpret this to indicate that folding of the two proteins is similar, so that the sites of biological activity, presumably focused at the regions of sequence homology, can be similarly presented to the appropriate substrate.

Enterotoxin D

In the last annual report, I indicated that 5 E. coli clones were isolated that produced SED and contained the entD gene on a chromosomal DNA fragment cloned into pBR322. We have continued analysis of these clone and present a summary of these data below.

To determine the insert sizes of the five positive clones, minilysates of each were prepared. Agarose gel electrophoresis of the DNA from these cells showed that they contained varying insert sizes. The plasmid containing the smallest insert (3.2 Kbp) was retained and designated pIB486. Further subcloning of entD was achieved by digesting pIB486 with EcoRI and NaeI, attaching an EcoRI linker to the blunt, NaeI end, and ligating into the EcoRI site of pUC18. This plasmid, pIB488, contains a 2.0 Kbp insert including a complete entD gene as determined by Western blotting of the cellular extract from cells containing this plasmid.

Overlapping clones for sequencing were obtained by the method described by Dale et al.(Plasmid 31: 31-40, 1985) The 2.0 Kbp insert from pIB488 was cloned into the bacteriophage sequencing vector, M13mp19. Single-stranded recombinant phage DNA was harvested, extracted, and annealed to a 20 bp oligonucleotide (RD20) that hybridizes to the EcoRI site within the multiple cloning region. The DNA was then digested with EcoRI which cleaves only within the annealed portion of the molecule. The 3' to 5' exonuclease activity of T4 DNA polymerase was then used to obtain variable deletions of the insert DNA. Poly-A tails were added using TdT terminal transferase which allows recircularization of the molecule by annealing with RD20 on the other end of the fragment. Subsequent ligation was carried out with T4 DNA ligase. These ligated molecules were then transformed into E. coli JM109. Infected cells from the resulting plaques were picked and grown in L-broth. Phage DNAs from these cultures were harvested and size fractionated on a 1.0% agarose gel. DNA from the deletions selected was purified and Sequenced by the dideoxy chain termination method. The opposite strand was also sequenced by

going through the above process on an M13mp19 clone that had the 2.0 Kbp fragment inserted in the opposite orientation.

The DNA and derived protein sequence of the 2.0 kbp fragment is presented in Fig. 6. Analysis of the fragment revealed a large open reading frame that could encode a 258 amino acid protein with a molecular weight of 29,768. Previous amino acid analysis of the termini of the mature SED protein indicates that a serine residue is at the amino terminus. Three serine residues are present near the amino terminus of the precursor protein that could mark the amino terminus of the mature protein. By comparing the amino acid composition of the three polypeptide sequences (starting with the three serine residues) to the published amino acid composition of SED, the actual mature polypeptide sequence can be predicted. The polypeptide starting with amino acid 30 provides the most consistent amino acid composition to that of previously published results. This polypeptide is 228 amino acids in length and has a molecular weight of 26,360 which is also in agreement with previously reported molecular weight of 27,300 daltons.

Sequence comparison of the deduced amino acid sequence of SED to that of SEA, SEB, SEC, or streptococcal pyrogenic exotoxin A (SPEA) show that there are 51.6%, 41.1%, 34.9%, and 39.2% similarity, respectively. The relatively high degree of similarity between SED and SEA was expected because SED and SEA contain similar cross-reactive antigenic determinants. However, it is notable that SED is also very similar to SPEA a gene carried by a bacteriophage of the genus Streptcoccus. These data are taken to imply that there is a relationship among these toxins that is not necessarily reflected in their primary biological activity (emesis for the enterotoxins and rheumatic fever for

SPEA), but is manifest in secondary biological features such as mitogenicity, enhancement of endotoxic shock, and immune suppression.

Future work will focus on the regulation of expression of the entD gene and ultimately its mode of action utilizing techniques of site-directed mutagenesis.

Regulation of the Lipase gene

We had previously used deletion analysis to approximate the location of the insertion site (attB) of the bacteriophage L54a in the lipase structural gene (geh) of S. aureus. It was located within the terminal 360 bp of the gene. In order to determine the exact site of insertion, we cloned and sequenced fragments of the S. aureus PS54 lipase gene which contained the chromosome/bacteriophage junction fragments (attL and attR). These fragments were in turn used to probe a 4.5 kb ClaI fragment of the bacteriophage L54a genome that contained the phage attachment site (attP). The restriction map and the sequencing strategy are shown in Fig. 7.

The sequence of the att sites (Fig. 8) reveals an 18 bp core sequence common to all four regions. This feature is similar to the bacteriophage lambda att sites in which the common core is a 15 bp sequence. Unlike the common core of the lambda att sites which have an 80% A+T content, the core of the L54a att sites is only 61 percent A+T. The A+T content of the DNA flanking the core region (the arms) which extends from -50 to +50 is 63% in the attP site and 55% in the attB site. In view of the fact that the percent A+T of the S. aureus genome and staphylococcal phages is 62-70%, the percent A+T found in the core sequence and the surrounding region is not untypical.

Also indicated in Fig. 8 are regions of dyad symmetry, inverted repeats and direct repeats. These probably represent possible binding sites for

proteins that mediate the recombination. No attempt was made to confirm the protein binding capacity of these regions. However, inasmuch as there are unaltered tandem direct repeats of the core sequence flanking the prophage as a result of integration, the crossover point must occur within the core sequence. Furthermore, the flanking core sequences also suggest that recombination occurs via staggered cuts, and that recombination is not only site-specific but also orientation-specific.

In every reported system of site-specific recombination, the gene encoding the enzyme which mediates the recombination reaction is located near the recombination site, so our initial approach to identify the gene or genes responsible for L54a recombination (recombinase gene) was centered on the DNA near the attP site. Two DNA fragments containing attP were cut from the bacteriophage genome. One (4.5 Kb) contained attP with DNA extending rightward from it (ClaI restriction fragment, see Fig. 9) and the other (3.5 Kb) contained attP with DNA extending leftward from it (PvuII-HindIII restriction fragment, see Fig. 9), These fragments were individually cloned into a shuttle vector pLI50 and the resultant plasmids designated pLI461 and pLI475, respectively (Fig. 9). The plasmids were transformed into protoplasts derived from S. aureus RN4220 and the presence of recombinase activity was tested by assaying for lipase activity. The results shown in Fig. 9 indicate that RN4220(pLI461) had no lipase activity, whereas RN4220(pLI475) remained lipase positive. The recombinase expressed in RN4220(pLI461) mediates the recombination between attP on plasmid pLI461 and attB on the lipase gene of the RN4220 chromosome indicating that the recombinase gene is located within the segment of DNA rightward from the attP site.

More precise mapping of the recombinase gene function was achieved by cloning and testing recombination activity of sequential deletions of the 4.5 Kb ClaI fragment. The results of this experiment are schematically shown in Fig. 9 indicating the size of the deletion along with an indication of the effect of the deletion on integration. These data indicate that one end of the recombinase gene is located to the left of the EcoRV site (about 2 Kb rightward to the attP site). In addition, since there is no promoter in the vector preceding the cloning site, the 2 Kb fragment must contain the promoter of the recombinase gene.

Our results also suggest the functioning of more than one gene and is supported by the following argument. Our assay for integration is the loss of lipase activity consequent to insertion. However, after prolonged incubation of strain PS54 lipase activity could be detected at a frequency of $10^{-4}$ to $10^{-3}$ due to loss of the prophage. The same phenomenon was observed with transformants carrying the cloned attachment sites. The plasmids pLI461, pLI462 and pLI463 would also convert these cultures to the lipase negative phenotype at a frequency of $10^{-4}$. Consequently, this implies that the excision gene is also located within the same 2.1 Kb cloned DNA fragment as the gene for integration. It is possible that in L54a, two enzymes are responsible for recombination and are located within the short 2.1 Kb DNA segment.

Confirmation that the lipase negative phenotype was due to integration of the plasmid containing the attP site and the recombinase gene was obtained from Southern hybridizaton analyses. ClaI digested bulk chromosomal DNA prepared from transformants of various plasmids was hybridized to a probe of the 770 bp ClaI-D fragment of the lipase gene which contains the attB site.

Since ClaI cleaves each of the plasmids at least once but does not cleave the ClaI-D fragment of th lipase gene, the probe would identify two bands if integration occurred, whereas it would identify only one band if there was no integration. Fig. 10 shows the results of this analysis which confirm that integration has occurred.

Lysogenization of L54a in many strains of S. aureus results in loss of lipase activity caused by insertion of the prophage genome at the 3' or carboxyl end of the lipase structural gene which is essential for catalysis. This indicates that a truncated catalytically inactive lipase protein deleted by 46 amino acids should be produced by the lysogenized strain. Indeed, preliminary immunological screening indicates that the lysogenized strain does produce a cross reactive material that lacks lipase activity. By examining the nucleotide sequence of the left junction of the bacterial chromosome and L54a DNA (i.e., attL), a stop codon TAA was found adjacent to the core sequence (Fig. 8). The sequence analyses, therefore, support the mechanism of the lysogenic conversion of the lipase.

The importance of this work is as a model for the regulation of the β-toxin of Staphylococcus which is also mediated by phage conversion. This activity however, is governed by both positive and negative conversion. There are two converting phages that mediate the expression of β-toxin. One of these is a negative regulator of expression similar to the L54a system, but the second is both a positive and negative converting phage. This second phage carries the staphylokinase gene and upon lysogenization confers staphylokinase activity and inactivates β-toxin activity in the host cell. Some of the interesting question to ask are whether the insertion sites of this phage and L54a are similar and within the structural β-toxin gene, whether these phages

are similar to or have given rise to the phages that carry the entA and spea genes and to the elements that harbor ent and the other extracellular toxins of Staphylococcus? Many other interesting questions relating not only to toxigenesis but also to the basic biology of phage conversion in this system can be asked. We are moving in these directions for the future.

Publications


1. Jackson, M.P. and J.J. Iandolo. 1986. Sequence of the exfoliative toxin B gene of Staphylococcus aureus. J. Bacteriol. 167: 726-728.

2. Lee, C.Y. and J.J. Iandolo. 1986. Integration of staphylococcal phage L54a occurs by site-specific recombination: Structural analysis of the attachment sites. Proc. Nat. Acad. Sci. USA. 83: 474-5478.

3. Lee, C.Y., J.J. Schmidt, A.D. Johnson-Winegar, L. Spero and J.J. Iandolo. 1987. Sequence determination and comparison of the exfoliative toxin A and toxin B genes from Staphylococcus aureus. J. Bac eriol. 169:3904-3909.

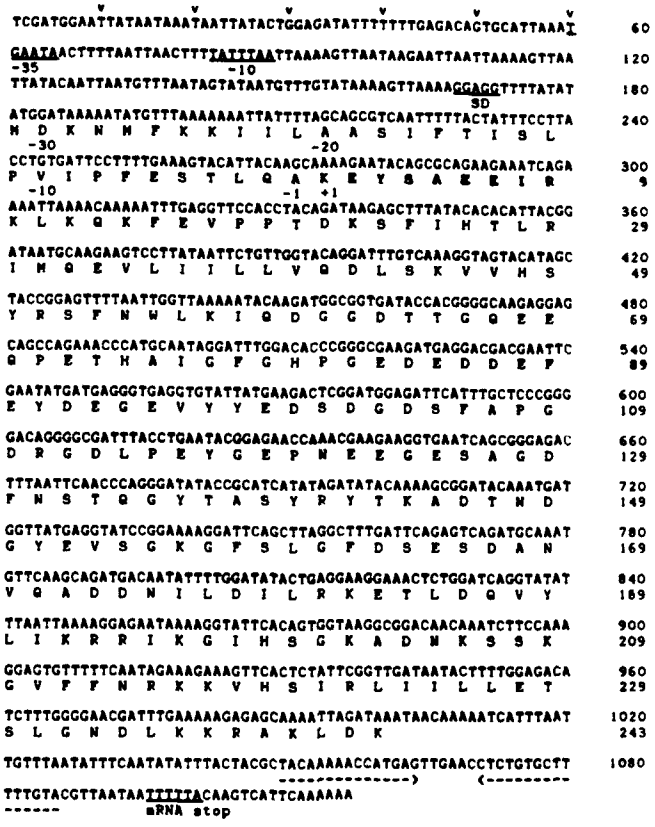



FIG. 1 Nucleotide sequence of the DNA fragment containing the ETB gene. The corresponding letter codes for the amino acids of the ETB polypeptide are: A, Ala; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; Y, Tyr. The presumptive −35, −10, SD (Shine Dalgarno ribosome binding site), −1, +1 protease processing site, the stem-loop structure (facing arrows), and the mRNA transcription stop sequence are indicated beneath the sequence.

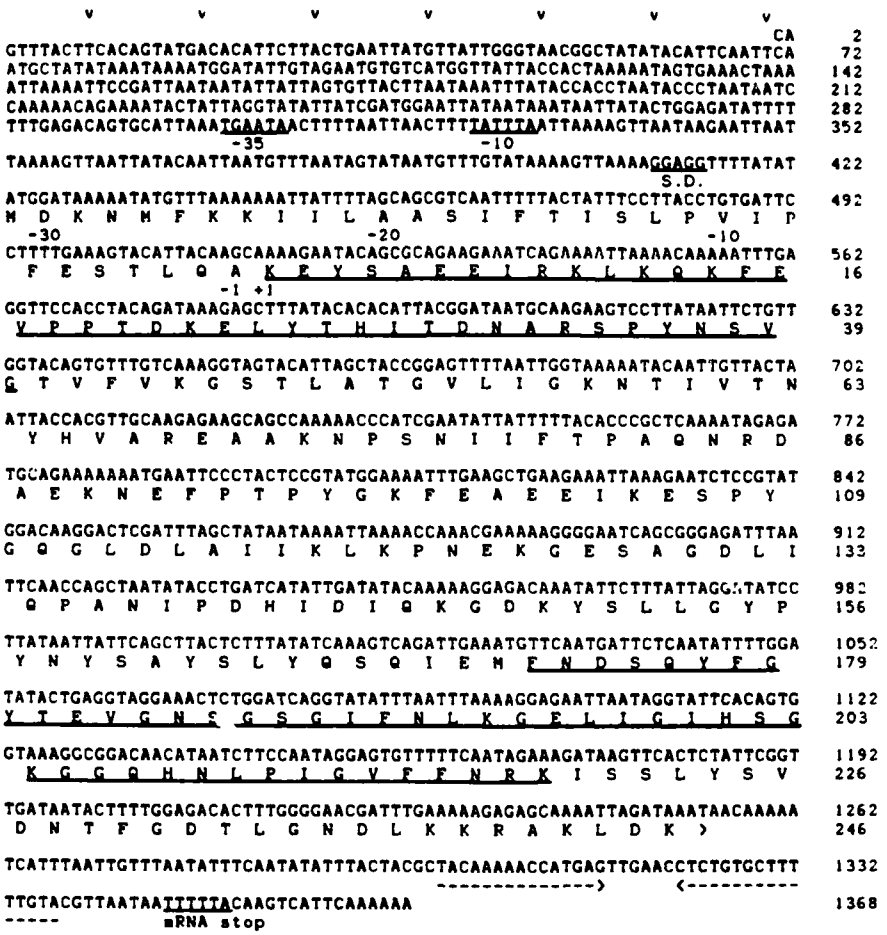


FIG. 2 Sequence of the 1,368-bp DNA fragment containing the *etb* gene and sequence of ETB derived from it. The locations of the presumptive −35, −10, Shine-Dalgarno ribosome binding site (SD), −1, +1 protease processing site, stem-loop termination structure (facing arrows), the stop codon (>), the mRNA transcription stop site, and the chemically derived peptide sequence (underlined) are indicated.

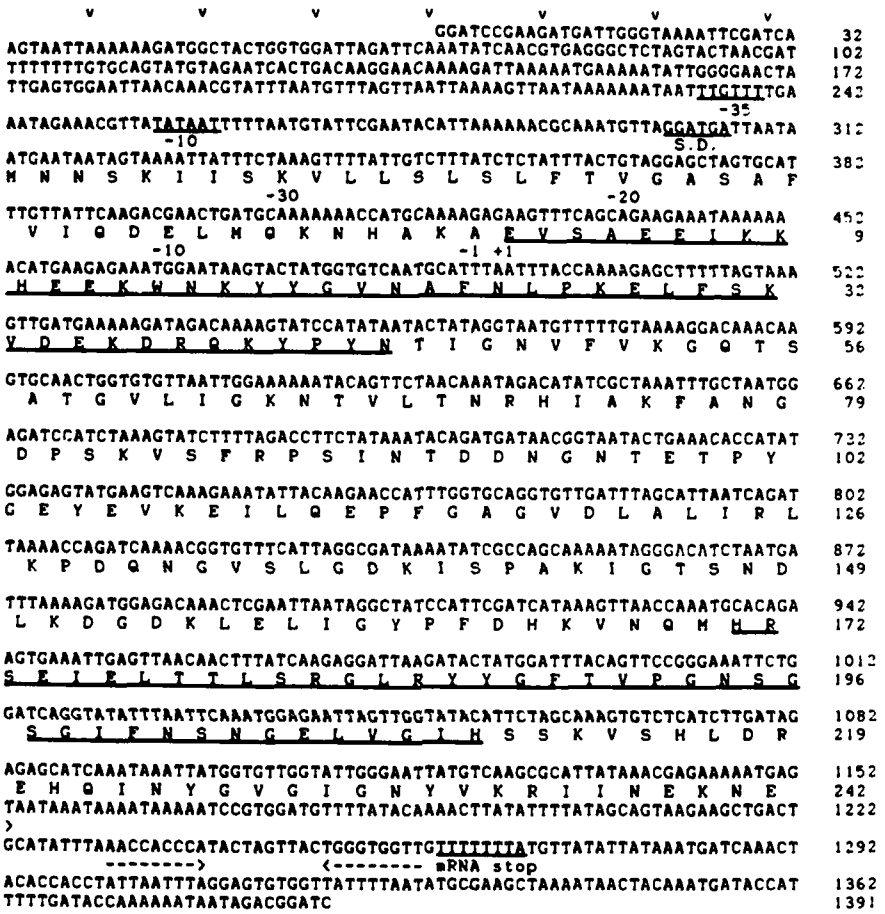


FIG. 3 Sequence of the 1,391-bp DNA fragment containing the *eta* gene and sequence of ETA derived from it. The locations of the presumptive −35, −10, Shine-Dalgarno ribosome binding site (SD), −1, +1 protease processing site, stem-loop termination structure (facing arrows), the stop codon (>), the mRNA transcription stop site, and the chemically derived peptide sequence (underlined) are indicated.

TABLE 1. Comparison of the amino acid composition of ETB predicted from the DNA sequence and that obtained from protein analysis

| Amino acid | Predicted from DNA sequence | From protein analysis[a] |
|---|---|---|
| Tryptophan | 1 | 1 |
| Lysine | 21 | 22 |
| Histidine | 6 | 5 |
| Arginine | 11 | 5 |
| Aspartic acid | 25 | 29 |
| Threonine | 11 | 12 |
| Serine | 19 | 17 |
| Glutamic acid | 24 | 26 |
| Proline | 7 | 12 |
| Glycine | 22 | 21 |
| Alanine | 10 | 13 |
| Cysteine | 0 | 0 |
| Valine | 11 | 9 |
| Methionine | 1 | 1 |
| Isoleucine | 15 | 17 |
| Leucine | 19 | 16 |
| Tyrosine | 11 | 13 |
| Phenylalanine | 11 | 9 |
| Asparagine | 9 | ND[b] |
| Glutamine | 9 | ND |

[a] The protein analysis is from Johnson et al. (9).
[b] ND, Not done.

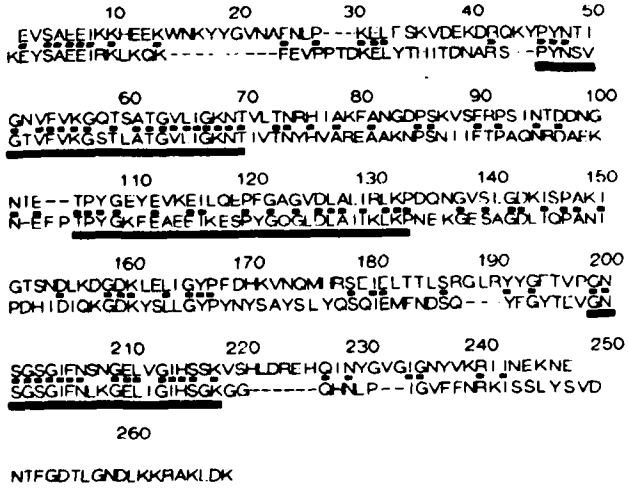


FIG. 4. Comparison of the amino acid sequences of ETA (top row) and ETB (bottom row). Sequence identities are indicated by bars, and dashed lines indicate gaps introduced to produce the optimal alignment. Numbering includes gaps and does not correspond to the residue number obtained from the DNA sequence. Alignment was constructed by computer using the algorithm of Wilbur and Lipman (30) with a K-tuple of 1, window of 20, and gap penalty of 1. Three regions of substantial homology are underlined.

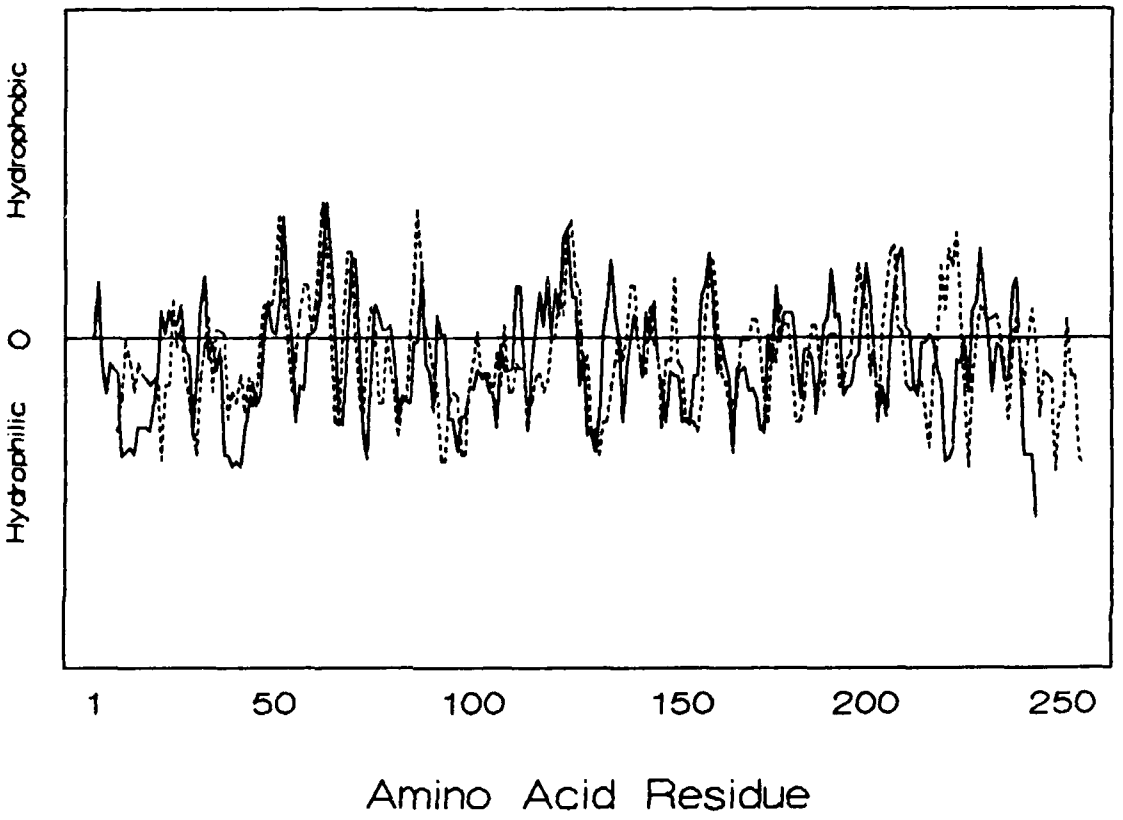



Fig. 5. Hydropathicity plot of Exfoliative toxins A (dashed line) and B (solid line). Hydropathicity was averaged over 4 amino acid residues using the algorithm of Kyte and Doolittle.

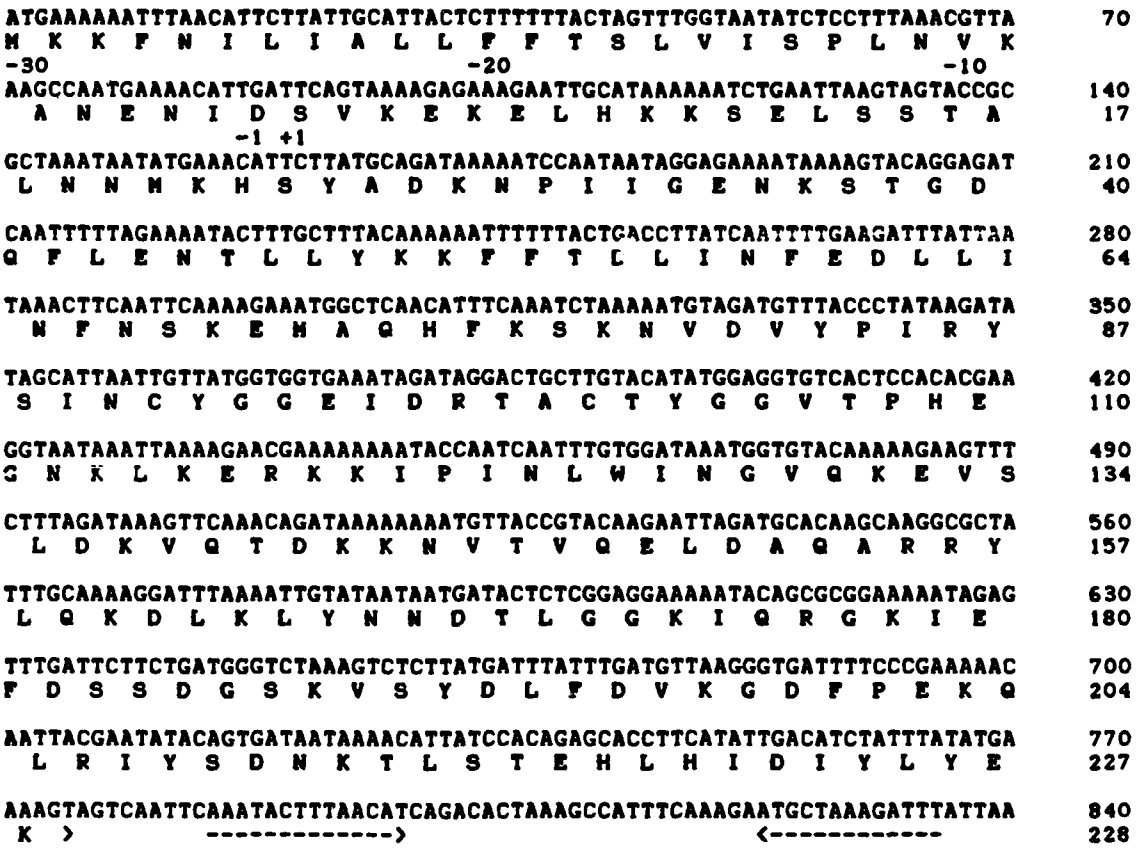



Fig. 6 Nucleotide and amino acid sequence analysis of the entD gene. The amino terminal serine residue is indicated, +1. Arrows indicate the putative transcription termination signal.

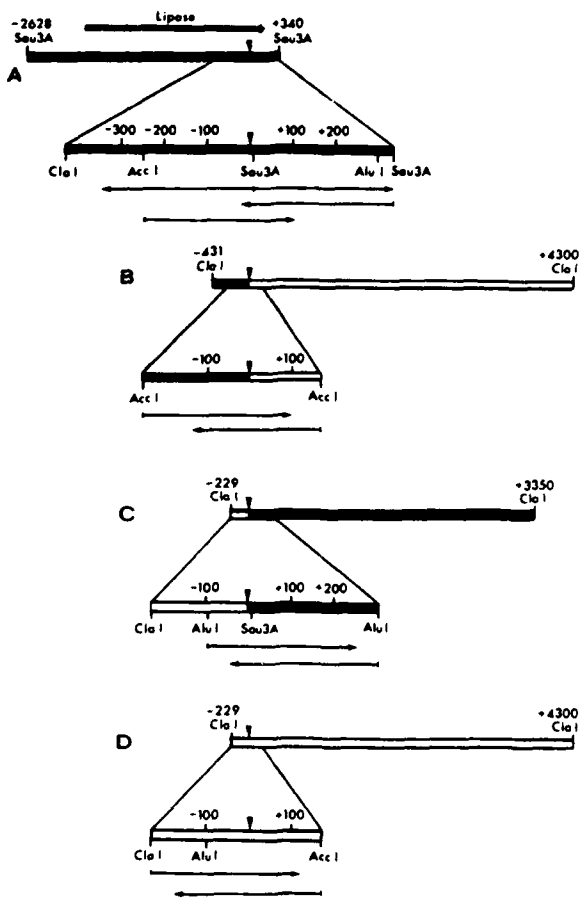



FIG. 7 Cloned primary restriction fragments of the genomes of *S. aureus* PS54 and PS54C and of the genome of bacteriophage L54a containing the attachment sites. DNA sequencing strategy is also indicated. (*A*) *attB*. (*B*) *attL*. (*C*) *attR*. (*D*) *attP*. Isolation of these fragments is described in *Materials and Methods*. Solid bars represent bacterial DNA, and open bars L54a DNA. The approximate location of the *geh* gene is indicated in *A*. An expanded restriction map of the region containing each *att* site is shown under the primary fragment. Arrows indicate the direction of sequencing. Arrowheads above the line represent the approximate location of the center of the core. Numbers indicate the distance (bp) from the center of the core sequence.

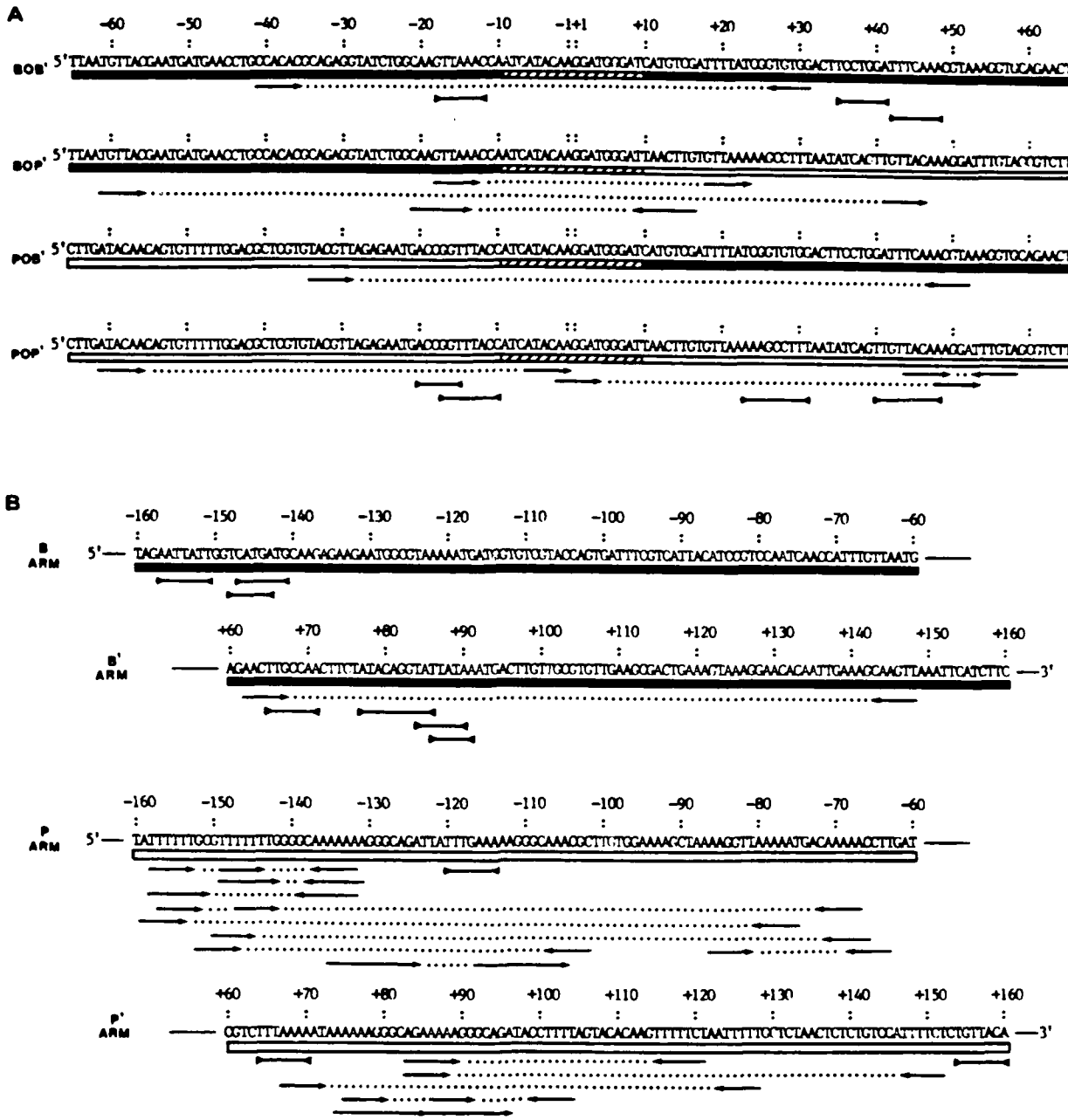
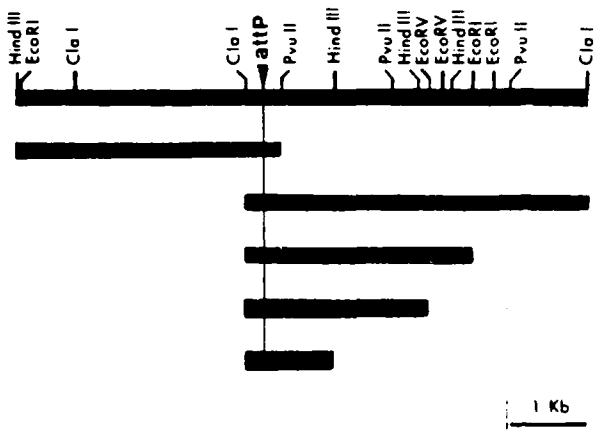



FIG. 8 Nucleotide sequences of the regions containing the *att* sites. Sequences are numbered from the center of the core; the base immediately to the right is +1 and the base immediately to the left is −1. (*A*) The central 130 bp of each of the four *att* sites that encompasses 65 bp on each side. (*B*) Distal portions of the four arms extending 100 bases leftward from −60 and 100 bases rightward from +60. Nomenclature shown at left of each region (BOB', etc.) is adapted from that of the bacteriophage λ system (24). Solid bars represent bacterial DNA; open bars, L54a DNA; hatched bars, core sequences. Molecular palindromes (►◄), inverted repeats (→ ←), and direct repeats (→ →) are indicated. Direct repeats are omitted in *B*, except one set, found in the P and P′ arm, that is of special importance and is discussed in the text. Dotted lines connect the pairs of repeats. The criteria used in marking the sequence features are as follows: minimum of 6 bp with no mismatches in inverted repeats and direct repeats; a single or no central mismatch with at least 3 bp to each side, or two central mismatches with at least 4 bp to each side, in molecular palindromes.




| Plasmid | Lipase Phenotype | Integration |
|---|---|---|
| pLI475 | + | - |
| pLI461 | - | + |
| pLI462 | - | + |
| pLI463 | - | + |
| pLI464 | + | - |

FIG. 9. Localization of the recombinase gene near the *attP* site. Relevant restriction sites are indicated. Arrowhead indicates the approximate *attP* site. Vertical line indicates the approximate location of the core sequences. Lipase phenotype of transformants generated by transforming the various deleted plasmids is indicated. Integration as determined by lipase activity is also indicated.


FIG. 10. Southern hybridization analysis of integration. DNA from the transformants containing the deleted plasmids was digested with *Cla* I, subjected to electrophoresis in agarose, blotted to nitrocellulose, and hybridized with $^{32}$P-labeled probe prepared from the *Cla* I fragment D containing the *attB* site (i.e., the DNA fragment from base pair −430 to +340 of Fig. 1A). Digested DNA was from RN4220 (lane 1), RN4220(pLI461) (lane 2), RN4220(pLI462) (lane 3), RN4220(pLI463) (lane 4), RN4220-(pLI464) (lane 5), and RN4220-(pLI475) (lane 6).

DISTRIBUTION LIST

| | |
|---|---|
| 5 copies | Commander<br>US Army Medical Research Institute of Infectious Diseases<br>ATTN: SGRD-UIZ-M<br>Fort Detrick, Frederick, MD 21701-5011 |
| 1 copy | Commander<br>US Army Medical Research and Development Command<br>ATTN: SGRD-RMI-S<br>Fort Detrick, Frederick, MD 21701-5012 |
| 12 copies | Defense Technical Information Center (DTIC)<br>ATTN: DTIC-DDAC<br>Cameron Station<br>Alexandria, VA 22304-6145 |
| 1 copy | Dean<br>School of Medicine<br>Uniformed Services University of the Health Services<br>4301 Jones Bridge Road<br>Bethesda, MD 20814-4799 |
| 1 copy | Commandant<br>Academy of Health Sciences, US Army<br>ATTN: AHS-CDM<br>Fort Sam Houston, TX 78234-6100 |