JOURNAL OF BACTERIOLOGY, May 1986, p. 385–391
0021-9193/86/050385-07$02.00/0
Copyright © 1986, American Society for Microbiology
Vol. 166, No. 2

Lysogenic Conversion of Staphylococcal Lipase Is Caused by Insertion of the Bacteriophage L54a Genome into the Lipase Structural Gene†

CHIA YEN LEE AND JOHN J. IANDOLO*
Microbiology Group, Division of Biology, Kansas State University, Manhattan, Kansas 66506

Received 2 December 1985/Accepted 17 February 1986

Staphylococcus aureus **PS54 manifests no lipase (*geh*) activity. This is due to the insertion of bacteriophage L54a DNA into the *geh* structural gene. The nucleotide sequence of this 2,968-base-pair DNA fragment was determined. Lipase deduced from the nucleotide sequence is a polypeptide of 690 amino acids which extends from nucleotide 706 to 2776.**

Many strains of staphylococci produce a true lipase or glycerol ester hydrolase (EC 3.1.1.3). The activity of the staphylococcal lipase gene is negatively regulated by bacteriophage lysogenization, also known as lysogenic conversion (3, 21). We have cloned the lipase gene (*geh*, for glycerol ester hydrolase) and shown that the mechanism of conversion is most likely due to interruption of the gene by insertion of the phage DNA (11). To characterize this unique regulation of gene expression, we sequenced the lipase gene and determined the phage insertion site by Southern hybridization analyses. Data reported in this paper indicate that lipase is a 76,000-dalton protein and that the bacteriophage insertion site lies between nucleotides 2608 and 2698, which corresponds to amino acids 635 to 644 of the lipase enzyme.

MATERIALS AND METHODS

Bacterial strains and phages. *Staphylococcus aureus* strains and phages were described in our previous paper (11). *S. aureus* PS54 harbors two temperate phages, L54a and L54b. Bacteria lysogenic for L54a are lipase negative. Both bacteriophages were eliminated by UV irradiation, and the cured strain (now lipase positive) was designated PS54C. Phage preparation, phage DNA extraction, and bulk chromosomal DNA purification were also as previously described (11).

Media and chemicals. Media for routine cultivation and detection were purchased from Sigma Chemical Co. Restriction enzymes, BAL 31 exonuclease, bacteriophage T4 DNA ligase, and nick-translation reagents were obtained from New England BioLabs, Inc., and Bethesda Research Laboratories, Inc. The large fragment of DNA polymerase (Klenow) was purchased from Boehringer Mannheim Biochemicals. Restriction enzyme digestions and other routine enzyme treatments were carried out by the procedures recommended by the suppliers. [α-^{32}P]dATP and [α-^{32}P]dCTP were purchased from New England Nuclear Corp.

Hybridization. The transfer of DNA to nitrocellulose membranes was by the method of Southern (24). The hybridization analysis was previously described by our laboratory (4). The DNA probes were made with [α-^{32}P]dCTP. After hybridization, the nitrocellulose was washed twice for 30 min each time at room temperature in 100 ml of 0.1× SSC (1× SSC is 0.15 M NaCl and 15 mM sodium citrate [pH 7.0]) containing 0.1% sodium dodecyl sulfate. The membrane was dried at 80°C for 10 min and then autoradiographed as we described previously (4).

DNA sequence analysis. Various restriction endonuclease fragments of the *geh* element were cloned into the M13 bacteriophage derivatives mp18 or mp19 (33) and propagated in *Escherichia coli* JM103. DNA sequencing was carried out by the dideoxy chain termination method of Sanger et al. (22). A computer-assisted sequence analysis was carried out with Seqaid (19), a software package kindly provided by Donald J. Roufa, Kansas State University.

Deletion mutagenesis. Plasmids pLI210 and pLI211 containing the 2.9-kilobase (kb) insert with the lipase gene from *S. aureus* PS54C (11) were linearized by endonuclease digestion at the unique *Bam*HI site and further digested with BAL 31 exonuclease to obtain various-length deletions from either end of the 2.9-kb *geh* insert. The digests were then phenol extracted, ethanol precipitated, ligated with T4 DNA ligase, and transformed into competent *E. coli* LE392 (11). Transformants were selected on L-broth plates containing 10 μg of chloramphenicol per ml. A panel of plasmids with various size deletions in the *geh* fragment was obtained from the transformants. Size estimates were made by agarose gel electrophoresis of minilysates of clones after linearization of the plasmids by restriction enzyme digestion. Alternatively, plasmids pLI210 or pLI211 were digested with restriction enzymes to delete specific sections of DNA and then religated.

RESULTS

Deletion mutagenesis. We reported earlier (11) that plasmids pLI210 and pLI211 carry a 2,968-base-pair (bp) DNA fragment containing the lipase gene (*geh*) of *S. aureus* PS54C which expressed lipase activity both in *E. coli* and *S. aureus*. To further localize the *geh* gene, various plasmids containing deletions at either end of the 2.9-kb fragment were generated. These deletions are schematically shown in Fig. 1 along with an indication of the effect of the deletion of lipase activity. Up to 500 bp could be removed from the left end of the fragment and up to 80 bp could be removed from the right end without influencing activity. Larger deletions at either end of the insert resulted in loss of enzymatic activity.

DNA sequence of the *geh* gene. The strategy used for nucleotide sequencing is shown in Fig. 2. Each restriction fragment was subcloned into bacteriophages M13 mp18 or

* Corresponding author.
† Contribution number 86-228-J from the Kansas Agricultural Experiment Station.

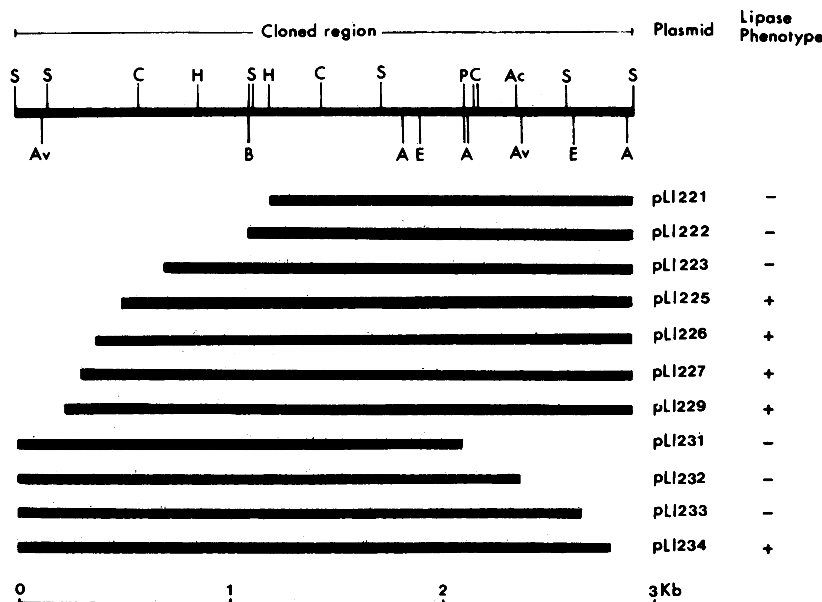

FIG. 1. Restriction endonuclease map of the 2,968-kb lipase-containing DNA fragment and the deleted fragments used to map the location of the *geh* gene. Only the cloned regions are shown. A, *Alu*I; Ac, *Acc*I; Av, *Ava*II; B, *Bcl*I; C, *Cla*I; E, *Eco*RII; H, *Hpa*II; P, *Pvu*II; S, *Sau3*AI. Restriction sites were determined from the DNA sequence. Lipase activity is indicated by + and −.

mp19 and sequenced on one strand. However, 98% of the 2,968-bp fragment, including the entire *geh* reading frame, was sequenced on both strands (Fig. 3). Computer analysis of both strands showed that there was only one large open reading frame, extending from residue 706 to 2776, that could code for a polypeptide similar in length to that reported for lipase. Within this open reading frame, there are 13 ATG potential initiation codons. However, since our molecular size estimate, based on sodium dodecyl sulfate-polyacrylamide gel electrophoresis, was about 70,000 daltons and the smallest published molecular size of purified lipase from *S. aureus* is 100,000 daltons (28), it is likely that the codon at position 706 is the actual initiation codon. Furthermore, it is preceded 4 bp upstream by a potential Shine-Dalgarno sequence (GAGGTGAT) which matches exactly with the 3′ terminus of *E. coli* 16S rRNA (23, 25).

DNA sequence features. The guanine-plus-cytosine (G+C) content of the proposed lipase DNA sequence was 37.5%, which is typical of the *S. aureus* genome (30 to 38%) (13). However, the region from nucleotide 300 to the start codon (ATG) at position 706 was very low in G+C content (22%). This region was therefore thought to contain potential binding sites for RNA polymerase to initiate transcription. A search of consensus sequences for *E. coli* promoters (8, 20) upstream from the translation start site showed numerous potential transcription initiation sites. However, results from the deletion mutagenesis experiments indicated that the sequence up to 500 bp from the left end of the DNA fragment was not essential for expression. A search of the interval from bp 500 to the start codon at bp 706 did reveal a hexanucleotide string beginning at bp 625 having the sequence TAATAT that we consider to be the actual promoter.

Codon usage in the lipase gene is shown in Table 1. Two of the codons, AGA (arginine) and GGA (glycine) are rarely used in highly expressed *E. coli* genes (6, 7) but a high percentage is found in the lipase gene. Since lipase is highly expressed in *Staphylococcus* spp. (ca. 1 mg/ml of stationary-phase culture), the high percentage of these codons may suggest that the tRNA species recognizing the AGA and

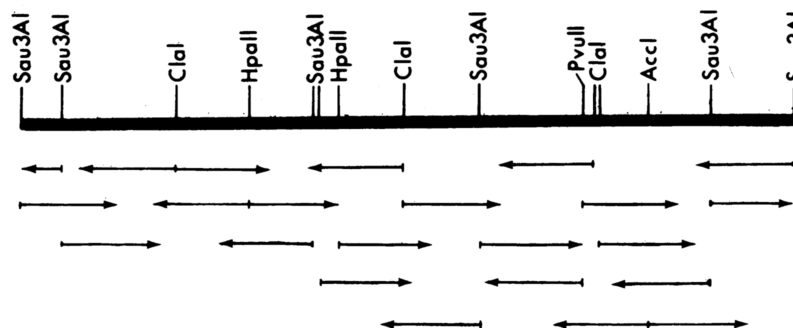

FIG. 2. Sequencing strategy for the *geh* gene insert. The arrows indicate the direction of sequencing. The length of the arrows indicates the start and extent of the sequence determined. Restriction enzyme sites used to generate fragments suitable for cloning are also shown.

```
Sau3A
GATCATTGCACCTGCAACTTTAAATAATGTAATGACCAAGTTTGTGATTGAAGTCATAAAAGTACGTCTCACTATTGGTAACGCTGTACCAATTAAAATCATAATCAGTGCAACATAGGG  120
    AvaII              Sau3A
CATAAGTGGACCTATGATTGAGCGAATGGCTAGATGAACATGATCGACGAAAATAGTGTTGTTACCATTAATCGTAAAAGGAATAAAGAAACATAGTATGCCCACTAAACTATAGACAAA  240

AAAACGCCATGCACTTGGTTGTTGTGCATTAGAATGATATTGATTCATTAAAGCAACCCCTTTGTTTAAATGAATACACAAAACTGTATGATGCATCTTCCCCTTAATGAGATGAATCAT  360

TATTTTAATTTAGAAAAATCTGAAAACTTACTATAATTGTATAGTTTGAATTATTTTCATACCAATACAAATTAACTAATTATATATAGATTGAAACTATATTACTTAATAAAATATTTA  480
                                                                                          ClaI
TCTTAAATGTTGTTGTGTTGATTCAACACCACAACTAAAAGTGTTTATAAATTATTTGGAAATACACATATTTGTAAATGATTAGTATCGATTTAATATCGTATTATTAAATTTTTATTA  600
                                                                       -35                     pribnow     *
ATTTTGTAGTCTTAATCAAAAAATAATATATGTCATGTTATATTGAAGGTGCAGTTGTTTTTCATTCTCAAGAGGGGGTCAAAAAAATACTTTTGAGGTGATTATATGTTAAGAGGACAA  720
                                                                                                 S.D.    MetLeuArgGlyGln  5

GAAGAAAGAAAGTATAGTATTAGAAAGTATTCAATAGGCGTGGTGTCAGTGTTAGCGGCTACAATGTTTGTTGTGTCATCACATGAAGCACAAGCCTCGGAAAAAACATCAACTAATGCA  840
GluGluArgLysTyrSerIleArgLysTyrSerIleGlyValValSerValLeuAlaAlaThrMetPheValValSerSerHisGluAlaGlnAlaSerGluLysThrSerThrAsnAla  45
                             HpaII
GCGGCACAAAAAGAAACACTAAATCAACCGGGAGAACAAGGGAATGCGATAACGTCACATCAAATGCAGTCAGGAAAGCAATTAGACGATATGCATAAAGAGAATGGTAAAAGTGGAACA  960
AlaAlaGlnLysGluThrLeuAsnGlnProGlyGluGlnGlyAsnAlaIleThrSerHisGlnMetGlnSerGlyLysGlnLeuAspAspMetHisLysGluAsnGlyLysSerGlyThr  85

GTGACAGAAGGTAAAGATACGCTTCAATCATCGAAGCATCAATCAACACAAAATAGTAAAACAATCAGAACGCAAAATGATAATCAAGTAAAGCAAGATTCTGAACGACAAGGTTCTAAA 1080
ValThrGluGlyLysAspThrLeuGlnSerSerLysHisGlnSerThrGlnAsnSerLysThrIleArgThrGlnAsnAspAsnGlnValLysGlnAspSerGluArgGlnGlySerLys  125
                                              BclI(Sau3A)                        Sau3A
CAGTCACACCAAAATAATGCGACTAATAATACTGAACGTCAAAATGATCAGGTTCAAAATACCCATCATGCTGAACGTAATGGATCACAATCGACAACGTCACAATCGAATGATGTTGAT 1200
GlnSerHisGlnAsnAsnAlaThrAsnAsnThrGluArgGlnAsnAspGlnValGlnAspThrHisHisAlaGluArgAsnGlySerGlnSerThrThrSerGlnSerAsnAspValAsp  165
                    HpaII
AAATCACAACCATCCATTCCGGCACAAAAGGTAATACCCAATCATGATAAAGCAGCACCAACTTCAACTACACCCCCGTCTAATGATAAAACTGCACCTAAATCAACAAAAGCACAAGAT 1320
LysSerGlnProSerIleProAlaGlnLysValIleProAsnHisAspLysAlaAlaProThrSerThrThrProProSerAsnAspLysThrAlaProLysSerThrLysAlaGlnAsp  205

GCAACCACGGACAAACATCCAAATCAACAAGATACACATCAACCTGCGCATCAAATCATAGATGCAAAGCAAGATGATACTGTTCGCCAAAGTGAACAGAAACCACAAGTTGGCGATTTA 1440
AlaThrThrAspLysHisProAsnGlnGlnAspThrHisGlnProAlaHisGlnIleIleAspAlaLysGlnAspAspThrValArgGlnSerGluGlnLysProGlnValGlyAspLeu  245
          ClaI
AGTAAACATATCGATGGTCAAAATTCCCCAGAGAAACCGACAGATAAAAATACTGATAATAAACAACTAATCAAAGATGCGCTTCAAGCGCCTAAAACACGTTCGACTACAAATGCAGCA 1560
SerLysHisIleAspGlyGlnAsnSerProGluLysProThrAspLysAsnThrAspAsnLysGlnLeuIleLysAspAlaLeuGlnAlaProLysThrArgSerThrThrAsnAlaAla  285

GCAGATGCTAAAAAGGTTCGACCACTTAAAGCGAATCAAGTACAACCACTTAACAAATATCCAGTTGTTTTTGTACATGGATTTTTAGGATTAGTAGGCGATAATGCACCTGCTTTATAT 1680
AlaAspAlaLysLysValArgProLeuLysAlaAsnGlnValGlnProLeuAsnLysTyrProValValPheValHisGlyPheLeuGlyLeuValGlyAspAsnAlaProAlaLeuTyr  325
                                                                                                  Sau3A
CCAAATTATTGGGGTGGAAATAAATTTAAAGTTATCGAAGAATTGAGAAAGCAAGGCTATAATGTACATCAAGCAAGTGTAAGTGCATTTGGTAGTAACTATGATCGCGCTGTAGAACTT 1800
ProAsnTyrTrpGlyGlyAsnLysPheLysValIleGluGluLeuArgLysGlnGlyTyrAsnValHisGlnAlaSerValSerAlaPheGlySerAsnTyrAspArgAlaValGluLeu  365
                                                     AluI                                                    EcoRII
TATTATTACATTAAAGGTGGTCGCGTAGATTATGGCGCAGCACATGCAGCTAAATACGGACATGAGCGCTATGGTAAGACTTATAAAGGAATCATGCCTAATTGGGAACCTGGTAAAAAG 1920
TyrTyrTyrIleLysGlyGlyArgValAspTyrGlyAlaAlaHisAlaAlaLysTyrGlyHisGluArgTyrGlyLysThrTyrLysGlyIleMetProAsnTrpGluProGlyLysLys  405

GTACATCTTGTAGGGCATAGTATGGGTGGTCAAACAATTCGTTTAATGGAAGAGTTTTTAAGAAATGGTAACAAAGAAGAAATTGCCTATCATAAAGCGCATGGTGGAGAAATATCACCA 2040
ValHisLeuValGlyHisSerMetGlyGlyGlnThrIleArgLeuMetGluGluPheLeuArgAsnGlyAsnLysGluGluIleAlaTyrHisLysAlaHisGlyGlyGluIleSerPro  445
                                                                         PvuII(AluI)          AluI
TTATTCACTGGTGGTCATAACAATATGGTTGCATCAATCACAACATTAGCAACACCACATAATGGTTCACAAGCAGCTGATAAGTTTGGAAATACAGAAGCTGTTAGAAAAATCATGTTC 2160
LeuPheThrGlyGlyHisAsnAsnMetValAlaSerIleThrThrLeuAlaThrProHisAsnGlySerGlnAlaAlaAspLysPheGlyAsnThrGluAlaValArgLysIleMetPhe  485
        ClaI               ClaI
GCTTTAAATCGATTTATGGGTAACAAGTATTCGAATATCGATTTAGGATTAACGCAATGGGGCTTTAAACAATTACCAAATGAGAGTTACATTGACTATATAAAACGCGTTAGTAAAAGC 2280
AlaLeuAsnArgPheMetGlyAsnLysTyrSerAsnIleAspLeuGlyLeuThrGlnTrpGlyPheLysGlnLeuProAsnGluSerTyrIleAspTyrIleLysArgValSerLysSer  525
                                                                                                    AccI
AAAATTTGGACATCAGACGACAATGCTGCTATGATTTAACGTTAGATGGCTCTGCAAAATTGAACAACATGACAAGTATGAATCCTAATATTACGTATACGACTTATACAGGTGTATCA 2400
LysIleTrpThrSerAspAspAsnAlaAlaTyrAspLeuThrLeuAspGlySerAlaLysLeuAsnAsnMetThrSerMetAsnProAsnIleThrTyrThrThrTyrThrGlyValSer  565
        AvaII
TCTCATACTGGTCCATTAGGTTATGAAAATCCTGATTTAGGTACATTTTTCTTAATGGCTACAACGAGTAGAATTATTGGTCATGATGCAAGAGAAGGATGGCGTAAAAATGATGGTGTC 2520
SerHisThrGlyProLeuGlyTyrGluAsnProAspLeuGlyThrPhePheLeuMetAlaThrThrSerArgIleIleGlyHisAspAlaArgGluGlyTrpArgLysAsnAspGlyVal  605
                                                                                                                  Sau3A
GTACCAGTGATTTCGTCATTACATCCGTCCAATCAACCATTTGTTAATGTTACGAATGATGAACCTGCCACACGCAGAGGTATCTGGCAAGTTAAACCAATCATACAAGGATGGGATCAT 2640
ValProValIleSerSerLeuHisProSerAsnGlnProPheValAsnValThrAsnAspGluProAlaThrArgArgGlyIleTrpGlnValLysProIleIleGlnGlyTrpAspHis  645
                          EcoRII
GTCGATTTTATCGGTGTGGACTTCCTGGATTTCAAACGTAAAGGTGCAGAACTTGCCAACTTCTATACAGGTATTATAAATGACTTGTTGCGTGTTGAAGCGACTGAAAGTAAAGGAACA 2760
ValAspPheIleGlyValAspPheLeuAspPheLysArgLysGlyAlaGluLeuAlaAsnPheTyrThrGlyIleIleAsnAspLeuLeuArgValGluAlaThrGluSerLysGlyThr  685

CAATTGAAAGCAAGTTAAATTCATCTTCTGAATTTAATATGCTATGTAAATCGTGCTGTTATCATGGCACATCAGATATAAGTAGCATCACAGTGTTGAATTTAAAAATAGTAAAGTGAA 2880
GlnLeuLysAlaSer
                                              AluI                            Sau3A
ATAAAGCGCCTGTCTCATTAGCGAAAACTAAAGGGACAGGCGTATCTGTTTATGAGCTTAATAAATTGTATGAATAATATGGTTGATC  2968
                                            mRNA stop
```

FIG. 3. Nucleotide sequence of the 2,968-bp DNA fragment containing the *geh* gene and the predicted amino acid sequence of the lipase. The potential promotor region (−35 and pribnow) (8, 20), the possible start site for transcription (*), the Shine-Dalgarno sequence (S.D.) (23, 25), and the possible mRNA stop site (arrows marking the inverted repeats for the potential stem-loop) (20) are indicated.

TABLE 1. Codon usage of the *S. aureus geh* gene

Amino acid	Codon	No. of condons	Amino acid	Codon	No. of codons
Phe	TTT	12	Tyr	TAT	20
	TTC	6		TAC	3
Leu	TTA	21	Term[a]	TAA	1
	TTG	5		TAG	0
	CTT	7	His	CAT	25
	CTC	0		CAC	1
	CTA	2	Gln	CAA	44
	CTG	1		CAG	4
Ile	ATT	12	Asn	AAT	44
	ATC	12		AAC	8
	ATA	8	Lys	AAA	43
Met	ATG	13		AAG	13
Val	GTT	16	Asp	GAT	34
	GTC	2		GAC	7
	GTA	13	Glu	GAA	27
	GTG	7		GAG	5
Ser	TCT	5	Cys	TGT	0
	TCC	3		TGC	0
	TCA	21	Term	TGA	0
	TCG	7	Trp	TGG	7
Pro	CCT	9	Arg	CGT	7
	CCC	2		CGC	6
	CCA	16		CGA	3
	CCG	5		CGG	0
Thr	ACT	14	Ser	AGT	15
	ACC	2		AGC	1
	ACA	27	Arg	AGA	10
	ACG	11		AGG	0
Ala	GCT	11	Gly	GGT	28
	GCC	5		GGC	7
	GCA	26		GGA	15
	GCG	10		GGG	2

[a] Term, Termination codon.

GGA codons are in much higher concentration in *S. aureus* than in *E. coli*.

Also indicated in Fig. 3 is a potential stem-loop structure that can be formed from nucleotide 2887 to 2912 (109- to 134-bp distal to the stop codon of the lipase gene); the stem is followed by a potential mRNA stop sequence of TTTA (9, 20) at position 2919. These features may compose the mRNA terminator of the lipase gene. No attempt was made to determine whether any proposed regulatory sequences are functional either in vivo or in vitro.

Amino acid sequence features. The hydropathicity plot (10) of the predicted lipase protein (not shown) indicated that the N-terminal region of the protein contains a signallike peptide characterized by a hydrophilic region, followed by a stretch of hydrophobic amino acids. Based on rules deduced from known proteins with signal peptides (15, 31) and the calculated secondary structure of this protein (2), one would predict that the peptidase cleavage site is at the position immediately after the Ala-Gln-Ala sequence (residues 35 to 37), leaving the mature lipase protein with an N terminus at amino acid 38 (serine). The predicted amino acid sequence indicates that the lipase precursor would be a basic protein with a pI of 9.75. However, the predicted mature protein would have a pI value of 9.3, which agrees with the published value (1). The amino acid sequence (Fig. 3) also indicates that lipase contains more polar amino acids (67% polar residues) than nonpolar amino acids, but many microbial lipases also share the same characteristics (26). The protein contains 13 methionine residues but no cysteine.

Insertion site for phage L54a. Staphylococcal phage L54a inserts into the *Cla*I D fragment (corresponding to nucleotides 2197 to 2968 in Fig. 3) of the *geh* DNA insert (11). We wished to further localize the insertion site and to determine whether it lies within the structural gene of lipase. The *Cla*I D fragment was isolated and cloned into the *Acc*I site of the replicative-form DNA of bacteriophage M13 mp18. This placed the right end of the *Cla*I D fragment adjacent to the *Pst*I site of M13 mp18. After linearization of the plasmid with *Pst*I, the *Cla*I D fragment was digested with BAL 31 exonuclease to generate serial deletions from the right end of the fragment (i.e., nucleotide 2776), while the left end was protected by M13 sequence. The deleted molecules were then blunt-end ligated and transformed into *E. coli* JM103. A set of deletions were obtained that corresponded to the removal of about 80, 170, 270, and 360 bp from the right end of the fragment. These plasmids were used as probes in Southern hybridization of genomic digests of the phage-lysogenized strain *S. aureus* PS54. If the probe made with the deleted fragment contained the insertion site, it should identify two bands representing phage-chromosome junction fragments in *Cla*I- or *Pst*I-digested DNA (both enzymes cleave the bacteriophage genome at least once) of the lysogenic strain. On the other hand, if the probe made with deleted *Cla*I fragment D does not contain the insertion site due to BAL 31 digestion, it should identify only one junction fragment. Probes with either 80 or 170 bp deleted from the right end of the *Cla*I D fragment (Fig. 4A and B) identified both junction fragments. On the other hand, probes with 270 or 360 bp deleted from the right end of the fragment (Fig. 4C and D) identified only one junction (left junction) fragment. The faint band at 1.5 kb in the *Cla*I digestion shown in lane 4 of each panel was due to incomplete digestion of an

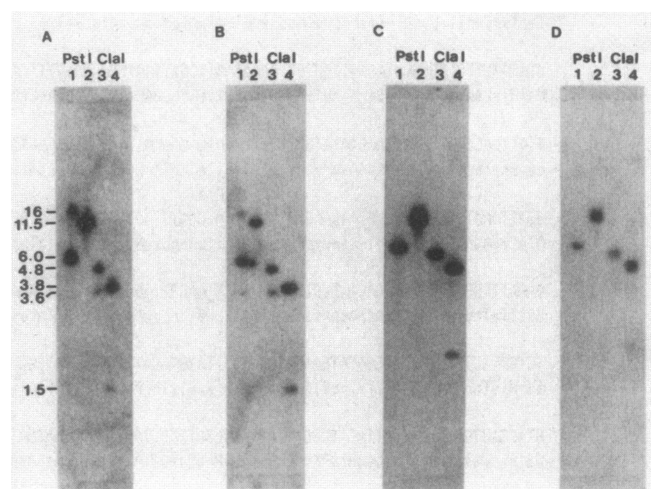

FIG. 4. Southern hybridization analyses to determine the approximate insertion site of phage L54a. The DNA from strains PS54 and PS54C was digested with *Pst*I and *Cla*I, electrophoresed in agarose gels, and blotted to nitrocellulose sheets. The blots were hybridized with ^{32}P-labeled probes obtained from the BAL 31-digested *Cla*I D fragments. Deletions of 80 (A), 170 (B), 270 (C), and 360 (D) bp from the right end of the *Cla*I D fragment were used. In each panel, lanes 1 and 3 are digested chromosomal DNA of strain PS54. Lanes 2 and 4 are DNA of strain PS54C. Lanes 1 and 2 are DNA digested with *Pst*I, and lanes 3 and 4 are DNA digested with *Cla*I. Size markers are listed at the left in kilobases.

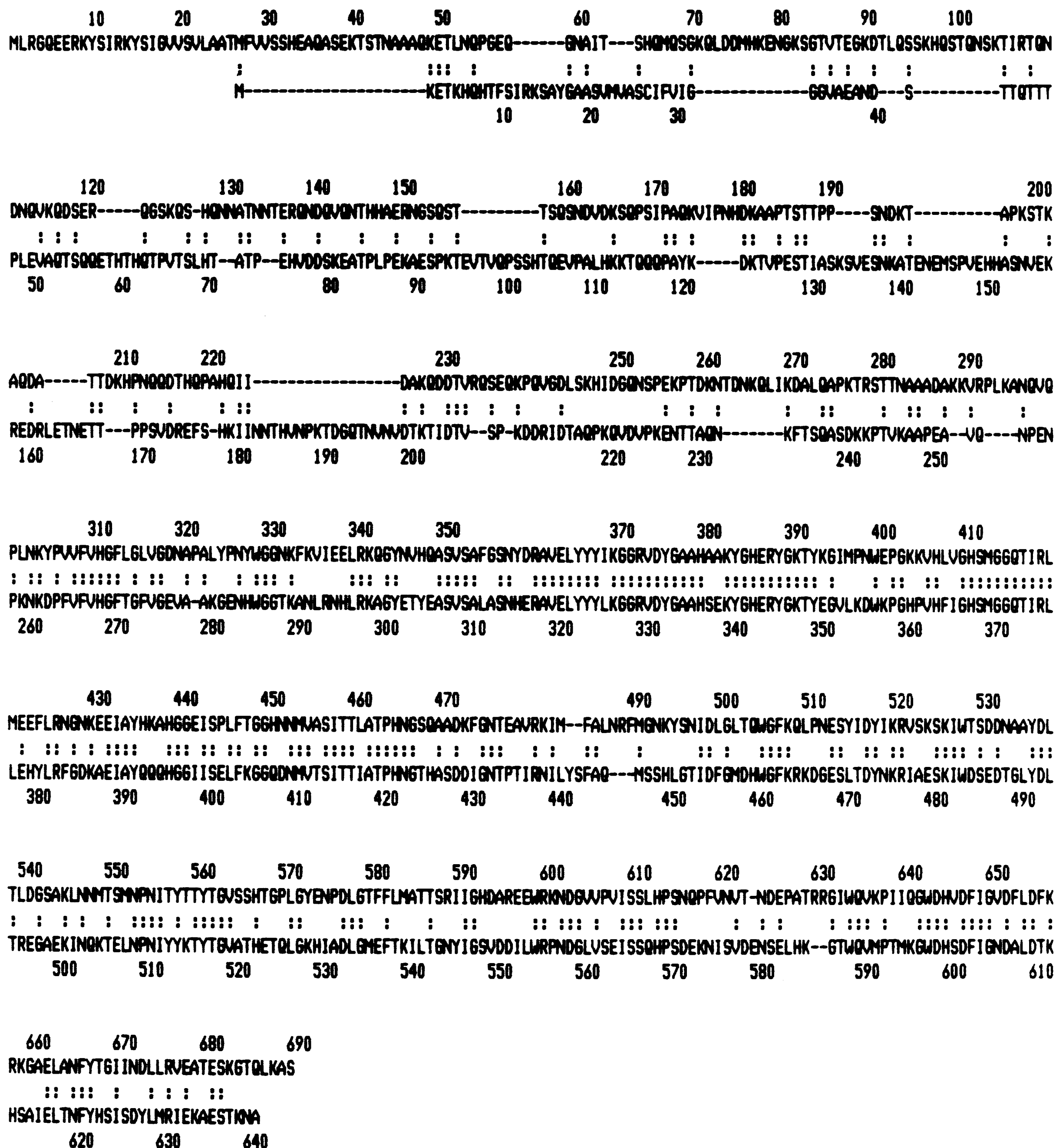

FIG. 5. Comparison of the amino acid sequence of lipases from *S. aureus* PS54C (top line) and *S. hyicus* (bottom line [5]). The sequences were aligned with a computer by the method of Wilbur and Lipman (32). Gaps were introduced to obtain maximum homology. The default parameters set were as follows: K-tuple, 1; window size, 20; gap penalty, 1. Colons between the amino acids of the two sequences indicate matches. The single-letter designation of amino acids is used.

additional *Cla*I site in the 3.8-kb fragment of the *geh* gene (11). When the blots for Fig. 4C and D were exposed longer (data not presented), a faint band corresponding to the right junction fragment was detected in Fig. 4C (the 270-bp deletion) but not in Fig. 4D (the 360-bp deletion). This result implied that bacteriophage L54a DNA inserted into the *Cla*I D fragment between 270 and 360 bp from the right end of the element. Thus, L54a inserted into the structural gene of lipase near the carboxyl end of the protein between amino acids 635 and 664.

DISCUSSION

The nucleotide sequence of the entire 2,968-kb DNA sequence containing the *geh* gene has only a single large open reading frame, extending from nucleotide 706 to 2776, that represents the probable coding region for lipase. Evi-

dence supporting this conclusion is threefold. (i) A Shine-Dalgarno sequence which is complementary to the 3′ end of *E. coli* 16S rRNA was found 4 nucleotides upstream of the probable ATG initiation codon. (ii) BAL 31 deletion mutagenesis identified this region as being required for lipase activity. (iii) Data from sodium dodecyl sulfate-gel electrophoresis indicated that lipase is a protein with a molecular size of about 70,000 daltons, which agrees closely with the molecular weight of the protein deduced from the DNA sequence.

The region upstream from the proposed start codon contains a very high proportion of A+T bases, resulting in many possible binding sites for RNA polymerase. In fact, several potential sequences similar to *E. coli* promoters were found, but no attempt was made to determine which promoter was used in transcribing the lipase mRNA.

The efficiency of staphylococcal promoters in the *E. coli* background is variable and only some cloned genes are transcribed even though their proposed promoters correspond to consensus sequences. For example, a staphylococcal enterotoxin B (18) and exfoliative toxin B (unpublished data) are not transcribed from their own promoters in *E. coli* but must be positioned downstream from a strong gram-negative promoter. Other staphylococcal genes such as *geh* are expressed, perhaps indicating that additional regulatory signals are ncessary for expression.

Vadehra and Harmon (29) and Mates (12) showed that the activity of lipase could be inhibited by the thiol blocking agents iodoacetic acid and *p*-chloromercuribenzoate. These results led Arvidson (1) to conclude that disulfide bonds and free SH groups were essential to lipase activity. However, even though disulfide bonds may be important linkages in the maintenance and stability of the tertiary structure of a protein, we found no cysteine in the deduced protein sequence of lipase. Proteins lacking cysteine or with a low content of cysteine are more flexible molecules whose tertiary structure relies on weaker bonds. It has been noted that many bacterial, extracellular proteins contain low levels of cysteine (16). This may indicate that proteins with high flexibility pass more readily through the rigid cell wall. For example, Tweten and Iandolo (27) showed that staphylococcal enterotoxin B, which contains a single disulfide loop, is sequestered within the cell wall and passes slowly into the extracellular medium. Furthermore, in the case of lipase, the lack of cysteine may allow conformational changes which, as suggested by Verger (30) and Pownall et al. (17), may be necessary for enzymatic activity when a water-soluble enzyme reacts with a hydrophobic lipid.

It is interesting that hydropathicity data indicate that the region from amino acid 35 to 310 is predominantly hydrophilic, whereas more hydrophobic residues were found in the region from residue 310 to the carboxyl end. Since lipase is an enzyme whose substrate is hydrophobic, we would predict that the hydrophobic region from amino acid 310 to the C terminus is associated with the active site of the enzyme. This conclusion is supported by (i) the deletion in plasmid pLI233 (Fig. 1) which has about 20 amino acids deleted from the C-terminal end and is lipase negative, (ii) the finding that the phage L54a insertion site is near the carboxyl end of the molecule and also results in loss of activity, and (iii) as indicated below, the extensive homology which exists in this region with lipase cloned from other species of staphylococci. Therefore, the region either contains the active site or is important for the conformation of the active site.

Recently, the lipase gene has been cloned and sequenced from *Staphylococcus hyicus* (5) (denoted as *lip*; this abbreviation has already been used for the lipoic acid locus in the genus *Staphylococcus* [14], therefore *geh* is probably more appropriate). The molecular size of the deduced lipase from *lip* is 71,000 daltons, which is very close to the size of the lipase deduced from *geh*. The substrate specificity of the two enzymes appears to be identical, and therefore the proteins are likely to be evolutionarily related. A comparison of the amino acid sequences of the two proteins is shown in Fig. 5. There is a significant homology between these two enzymes, particularly at the C-terminal ends in the region from amino acid residues 360 to 467 of the *geh* gene product and residues 310 to 417 of the protein from the *lip* gene. Considering the similarities in substrate specificity, one might expect conservation of the sequence in functional regions of the enzyme, as shown, while sequence divergence or drift might be expected to occur in noncatalytic regions of the molecule. In fact, the regions conserved between the two enzymes correspond to the hydrophobic amino acid-rich region which we concluded to be associated with enzyme activity. A similar result (ca. 46% match) was observed in comparison of the nucleotide sequences.

The data presented confirm that the mechanism of negative lysogenic conversion of staphylococcal lipase is by insertional inactivation. We are cloning the attachment site of the phage to determine the precise insertion site and to investigate possible mechanisms of DNA recombination in phage-host interactions in the staphylococcal system.

ACKNOWLEDGMENTS

This work was supported by Public Health Service grant AI-17474 from the National Institute of Allergy and Infectious Diseases and U.S. Army contract DAMD17-86-C-6055.

LITERATURE CITED

1. **Arvidson, S. O.** 1983. Extracellular enzymes from *Staphylococcus aureus*, p. 745–808. *In* C. S. F. Easmon and C. Adlam (ed.), Staphylococci and staphylococcal infections, vol. 2. Academic Press, Inc., New York.
2. **Chou, P. Y., and G. D. Fasman.** 1974. Prediction of protein conformation. Biochemistry **13**:222–245.
3. **Duval-Iflah, Y.** 1972. Lysogenic conversion of the lipase in *Staphylococcus pyogenes* group III strains. Can. J. Microbiol. **18**:1491–1497.
4. **Dyer, D. W., M. I. Rock, and J. J. Iandolo.** 1984. Autogenous transduction of ϕ11de in *Staphylococcus aureus*: transfer and genetic properties. J. Bacteriol. **158**:689–695.
5. **Gotz, F., F. Popp, E. Korn, and K. H. Schleifer.** 1985. Complete nucleotide sequence of the lipase gene from *Staphylococcus hyicus* cloned in *Staphylococcus carnosus*. Nucleic Acids Res. **13**:5898–5906.
6. **Grantham, R., C. Gautier, M. Gouy, M. Jacobzone, and R. Mercier.** 1981. Codon catalog usage is a genome strategy modulated for gene expressivity. Nucleic Acids Res. **9**:r43–r74.
7. **Grosjean, H., and W. Fiers.** 1982. Preferential codon usage in prokaryotic genes: the optimal codon-anticodon interaction energy and the selective codon usage in efficiently expressed genes. Gene **18**:199–209.
8. **Hawley, D. K., and W. R. McClure.** 1983. Compilation and analysis of *Escherichia coli* promoter DNA sequences. Nucleic Acids Res. **11**:2237–2255.
9. **Holmes, W. M., T. Platt, and M. Rosenberg.** 1983. Termination of transcription in *E. coli*. Cell **32**:1029–1032.
10. **Kyte, J., and R. F. Doolittle.** 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. **157**:105–132.
11. **Lee, C. Y., and J. J. Iandolo.** 1985. Mechanism of bacteriophage conversion of lipase activity in *Staphylococcus aureus*. J. Bacteriol. **164**:288–293.

12. **Mates, A.** 1974. Characterization and properties of purified *Staphylococcus aureus* lipase. Microbios **9:**61–74.
13. **Oeding, P.** 1984. Taxonomy and identification, p. 1–32. *In* C. S. F. Easmon and C. Adlam (ed.), Staphylococci and staphylococcal infections, vol. 1. Academic Press, Inc., New York.
14. **Pattee, P. A., J. M. Jones, and A. C. Yost.** 1984. Chromosome map of *Staphylococcus aureus*, p. 123–130. *In* S. J. O'Brien (ed.), Genetic maps 1984. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
15. **Perlman, D., and H. O. Halvorson.** 1983. A putative signal peptidase recognition site and sequence in eukaryotic and prokaryotic signal peptides. J. Mol. Biol. **167:**391–409.
16. **Pollock, M. R., and M. H. Richmond.** 1962. Low cyste(i)ne content of bacterial extracellular proteins: its possible physiological significance. Nature (London) **194:**446–449.
17. **Pownall, H. J., R. L. Jackson, J. D. Morrisett, and A. M. Grotto, Jr.** 1979. Structure and dynamics of reassembled plasma lipoproteins, p. 123–143. *In* A. M. Scanu, R. W. Wissler, and G. S. Getz (ed.), The biochemistry of atherosclerosis. Marcel Dekker Inc., N.Y.
18. **Ranelli, D. M., C. L. Jones, M. B. Johns, G. J. Mussey, and S. A. Kahn.** 1985. Molecular cloning of staphylococcal enterotoxin B gene in *Escherichia coli* and *Staphylococcus aureus*. Proc. Natl. Acad. Sci. USA **82:**5850–5854.
19. **Rhoads, D. D., and D. J. Roufa.** 1985. Emetine resistance of Chinese hamster cells: structures of wild-type and mutant ribosomal protein S14 mRNAs. Mol. Cell Biol. **5:**1655–1659.
20. **Rosenberg, M., and D. Court.** 1979. Regulatory sequences involved in the promotion and termination of RNA transcription. Annu. Rev. Genet. **13:**319–353.
21. **Rosendal, K., and P. Bulow.** 1965. Temperate phages influencing lipase production by *Staphylococcus aureus*. J. Gen. Microbiol. **41:**349–356.
22. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
23. **Shine, J., and L. Dalgarno.** 1974. The 3′ terminal sequence of *Escherichia coli* 16S ribosome RNA: complementarity to nonsense triplets and ribosome binding sites. Proc. Natl. Acad. Sci. USA **71:**1342–1346.
24. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98:**503–517.
25. **Stormo, G. D., T. D. Schneider, and L. M. Gold.** 1982. Characterization of translational initiation sites in *Escherichia coli*. Nucleic Acids Res. **10:**2971–2996.
26. **Sugiura, M.** 1984. Bacterial lipases, p. 505–523. *In* B. Borgstrom and H. L. Brockman (ed.), Lipases. Elsevier Science Publishing, Inc., New York.
27. **Tweten, R. K., and J. J. Iandolo.** 1983. Transport and processing of staphylococcal enterotoxin B. J. Bacteriol. **153:**297–303.
28. **Vadehra, D. V., and L. G. Harmon.** 1967. Isolation and purification of staphylococcal lipase. Appl. Microbiol. **15:**292–295.
29. **Vadehra, D. V., and L. G. Harmon.** 1967. Characterization of purified staphylococcal lipase. Appl. Microbiol. **15:**480–483.
30. **Verger, R.** 1980. Enzyme kinetics of lipolysis. Methods Enzymol. **64:**340–392.
31. **Von Heijne, G.** 1983. Patterns of amino acids near signal-sequence cleavage sites. Eur. J. Biochem. **133:**17–21.
32. **Wilbur, W. J., and D. J. Lipman.** 1983. Rapid similarity searches of nucleic acid and protein data banks. Proc. Natl. Acad. Sci. USA **80:**726–730.
33. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33:**103–119.