INFECTION AND IMMUNITY, Feb. 1989, p. 533–539
0019-9567/89/020533-07$02.00/0

Vol. 57, No. 2

Sequence Analysis and Expression in *Escherichia coli* of the Hyaluronidase Gene of *Streptococcus pyogenes* Bacteriophage H4489A

WAYNE L. HYNES AND JOSEPH J. FERRETTI*

Department of Microbiology and Immunology, University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma 73190

Received 15 July 1988/Accepted 3 November 1988

The hyaluronidase gene (*hylP*) from *Streptococcus pyogenes* bacteriophage H4489A was previously cloned into *Escherichia coli* plasmid pUC8 as a 3.1-kilobase *Tha*I fragment. Southern hybridization experiments confirmed the origin of this fragment in bacteriophage H4489A before determination of the nucleotide sequence of the entire fragment. Two open reading frames (ORFs) were found, the first of which specified a 39,515-molecular-weight protein identified as the bacteriophage hyaluronidase. The second ORF encoded a 65,159-molecular-weight protein of unknown function. Putative transcription and translation control sequences for each ORF were identified by using a plasmid containing a promoterless chloramphenicol acetyltransferase gene. Controlled exclusive expression of the *hylP* gene via the T7 polymerase-promoter system in *E. coli* resulted in a 40,000-dalton protein, a result consistent with the coding capacity of the *hylP* gene.

Bacteriophages isolated from lysogenic strains of group A streptococci have the ability to infect encapsulated strains of *Streptococcus pyogenes* (11, 16). In contrast, virulent phages are unable to infect these strains unless the hyaluronic acid capsule has been removed either spontaneously, as occurs in aged cultures, or through the action of the enzyme hyaluronidase (17). Benchetrit et al. (2) showed that both temperate and virulent phages possess hyaluronidase activity; however, the enzyme levels associated with the temperate phages were several orders of magnitude greater than in virulent phages. The function of the phage hyaluronidase has been suggested to be mediation of penetration of the hyaluronic acid capsule, a prerequisite for adsorption and establishment of a productive infection in the group A streptococcus (11, 16).

A phage-associated hyaluronidase was purified by Benchetrit et al. (1) from a temperate bacteriophage isolated from an M-type 49 strain of *Streptococcus pyogenes* (GT8760) after propagation on group A streptococcus strain K56. This enzyme was reported to be a 71,000-dalton glycoprotein as determined by sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis. Kjems (12) had previously suggested that the streptococcal bacteriophage DNA contained all the genetic information required for production of the hyaluronidase enzyme. Recently we cloned the gene which specified hyaluronidase production from the temperate bacteriophage H4489A isolated from an M-type 49 strain of *Streptococcus pyogenes* (10). The phage hyaluronidase gene, *hylP*, was contained on a 3.1-kilobase (kb) fragment and appeared to contain its own promoter region, as determined by expression in *Escherichia coli*. In the present communication, we report further on the expression of the *hylP* gene in *E. coli* as well as the nucleotide sequence of the *hylP* gene and the deduced amino acid sequence of the bacteriophage hyaluronidase.

* Corresponding author.

MATERIALS AND METHODS

Enzymes and chemicals. Restriction enzymes, T4 DNA ligase, DNA polymerase I, Klenow fragment of DNA polymerase I, M13 vector DNA and the M13 17-base primer were purchased from Bethesda Research Laboratories (Gaithersburg, Md.) and used according to the instructions of the manufacturer. Deoxy- and dideoxynucleotide triphosphates were from P-L Biochemicals, Inc. (Milwaukee, Wis.), and [α-^{32}P]dATP was obtained from Dupont, NEN Research Products (Boston, Mass.). Isopropyl-β-D-thiogalactopyranoside (IPTG) and 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-Gal) were purchased from Sigma Chemical Co. (St. Louis, Mo.). Plasmid pKK232-8 (5) used for the detection of promoter fragments was obtained from Pharmacia, Inc., Molecular Biology Div. (Piscataway, N.J.). Also from Pharmacia were the multifunctional vectors pT7T318U and pT7T319U. [^{35}S]methionine was obtained from ICN Radiochemicals (Irvine, Calif.).

Subcloning for nucleotide sequencing. The 3.1-kb *Tha*I fragment obtained by digestion of bacteriophage H4489A was previously ligated into pUC8 to create the chimeric plasmid pSF49 (10). Subsequent subcloning into the M13 vectors mp18 and mp19 was by insertion into the *Bam*HI-*Eco*RI sites of the polylinker.

The *Bam*HI-*Eco*RI fragment was unidirectionally degraded with nuclease *Bal* 31 by a modification of the technique of Gilmore et al. (7), with subsequent subcloning of the processed DNA into M13 phage mp18 or mp19. Any additional clones required were prepared by forced cloning of fragments into the appropriate vector. Colorless plaques from the cloning experiments were reinfected into *E. coli* JM109 and screened for the replicative form of the phage to ascertain the presence and size of the DNA insert.

Nucleotide sequencing. Sequencing reactions were done by the Sanger dideoxy-chain termination method (19) by the procedure described in the Amersham sequencing handbook. DNAs from several isolated phage clones were sequenced to obtain a completely overlapping set of sequences for both strands, so that sequences were confirmed on both strands. The nucleotide sequence was analyzed by the James

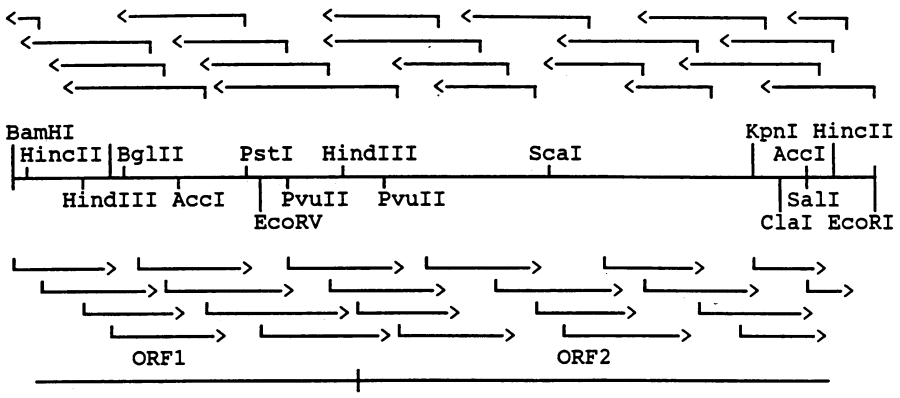


FIG. 1. Partial restriction map and sequencing strategy for the hyaluronidase gene, *hylP*, which contained the 3,127-bp *Bam*HI-*Eco*RI fragment isolated from pSF49 (10). Arrows indicate the direction and extent of sequence derived from each independent clone prepared by *Bal* 31 digestion.

M. Pustell DNA-protein sequencing program from International Biotechnologies Inc. (New Haven, Conn.).

Southern hybridization. DNA transfer, nick translation, and in situ hybridization were performed by the method of Southern (22) as described by Maniatis et al. (15).

Promoter fragment detection. Fragments for insertion into the promoter detection plasmid pKK232-8 (5) were obtained by digestion with appropriate enzymes and then separated from the remainder of the DNA by electrophoresis in low-melting-point agarose. The DNA fragments used were the *Bam*HI-*Bgl*II fragment and the *Pvu*II fragment (Fig. 1) for open reading frames 1 and 2 (ORF1 and ORF2), respectively.

Hyaluronidase assay. The assay of hyaluronidase activity during the purification of the enzyme was performed with brain heart infusion–bovine albumin fraction V–hyaluronic acid agar (BHB) plates. These plates were prepared essentially by the method described by Smith and Willett (21), modified by the addition of 0.1% sodium azide to the agar before pouring the plates not requiring organism growth. Other detection systems also used for hyaluronidase activity at various times throughout this study were the hyaluronic acid agarose plates described previously (10) and the dye-binding method of Benchetrit et al. (1) for the assay of partially purified solutions.

Hyaluronidase purification. The cloned hyaluronidase was purified by a procedure similar to that described by Benchetrit et al. (1) for the isolation of the native bacteriophage enzyme. An overnight culture of *E. coli* containing the *hylP* gene was centrifuged, and the cell pellet was washed twice with 100 mM phosphate buffer, pH 6.0, before being suspended in 50 mM phosphate buffer, pH 6.0, at 10 times the initial cell concentration. Lysozyme was added to a final concentration of 5 mg/ml, and the cells were incubated on ice for 30 min before being frozen at −70°C. When cells were thawed, alumina powder was added and the cells were lysed by sonification; the preparation was then clarified by centrifugation to remove the cell debris. The clarified supernatant was stored frozen and recentrifuged before use to remove any insoluble material. The clarified lysate was applied to a DEAE-cellulose (DE52; Whatman, Inc., Clifton, N.J.) column equilibrated with 50 mM sodium phosphate buffer (pH 7.2). After a washing with buffer, the material adsorbed to the column was eluted with 0.5 M NaCl in equilibrating buffer. Hyaluronidase activity was determined by using the BHB assay medium. Enzymatically active fractions were pooled and dialyzed against 50 volumes of the buffer used for the carboxymethyl cellulose (CM-cellulose) column. The dialyzed sample was applied to a CM-cellulose (Sigma Chemical Co.) column equilibrated with 50 mM sodium phosphate buffer (pH 6.0). The column was washed with buffer, and then the adsorbed material was eluted with 0.5 M NaCl in the equilibrating buffer. Fractions were collected and tested for hyaluronidase activity both with BHB agar and by use of the dye-binding method of Benchetrit et al. (3). Enzymatically active fractions were pooled and then dialyzed against distilled water prior to concentration of the sample. Concentration was achieved either by lyophilization or by placing the dialysis tube (8,000-dalton cutoff) containing the sample into polyethylene glycol. Concentrated material was applied to a SDS-polyacrylamide gel prepared essentially as described by Laemmli (13). The proteins were stained with Coomassie blue. Molecular weight markers were prestained and were obtained from Diversified Biotechnologies (Newton Centre, Mass.).

Use of the T7 RNA polymerase-promoter system. A *Bal* 31 deletion which contained the upper reading frame of the *Tha*I fragment plus the first 166 base pairs (bp) of ORF2 was used for expression of hyaluronidase from the bacteriophage T7 promoter. The fragment was isolated from low-melting-point agarose before ligation into the multifunctional vectors pT7T318U and pT7T319U (Pharmacia, Inc.), resulting in the plasmids pSF18U-1 and pSF19U-1, respectively. After it was confirmed that the appropriate fragment was inserted into the vector, *E. coli* JM109 was transformed with an insert-containing plasmid along with the plasmid pGP1-2 (obtained courtesy of S. Tabor), which contained the gene for the T7 polymerase (23). Clones containing both plasmids were selected after growth on media containing the antibiotics ampicillin and kanamycin. A clone containing pT7T319U with no insert and the RNA polymerase gene-containing plasmid pGP1-2 was used as a control.

Once appropriate clones were obtained, proteins were labeled with [^{35}S]Met essentially by the method described by Tabor and Richardson (23), modified so that the reactions were all carried out in microcentrifuge tubes (M. Gilmore, personal communication). Cells were pulse-labeled for 30, 75, and 105 min after induction of the polymerase. After labeling, the samples were subjected to electrophoresis on SDS-polyacrylamide gels, and the gels were dried and exposed to X-ray film.

RESULTS

Subcloning and Southern hybridization. The 3.1-kb *Tha*I *hylP*-containing fragment from pSF49 was obtained after digestion with *Bam*HI and *Eco*RI restriction enzymes and was subsequently ligated into similarly cut *E. coli* vectors M13mp18 and M13mp19. Expression of hyaluronidase activity by *E. coli* clones containing the *Tha*I fragment was observed in both orientations with regard to the *lac* operator-promoter region.

Southern hybridization by the nick translation procedure to label pSF49 indicated that the cloned fragment containing the *hylP* gene was present in bacteriophage H4489A DNA. The probe (pSF49) was able to hybridize to a 3.1-kb fragment obtained by digestion of the streptococcal bacteriophage H4489A DNA with *Tha*I (data not shown). The hybridization patterns observed with other restriction enzymes were in agreement with the restriction map of the cloned insert (Fig. 1).

Nucleotide sequence of the 3.1-kb *Tha*I fragment. The complete nucleotide sequence of the 3.1-kb *Tha*I fragment obtained from bacteriophage H4489A was determined by the dideoxy-chain termination method of Sanger et al. (19) (Fig. 2). The fragment was 3,127 bp long, with an overall A+T content of 59.3%. The restriction map generated from the nucleotide sequence was in agreement with that previously determined (10) by restriction enzyme digest analysis (Fig. 1). Examination of the sequence revealed two ORFs (ORF1 and ORF2) separated by 14 nucleotides. The first reading frame (ORF1) possesses an initiating ATG at positions 122 to 124, 6 bp from a putative ribosome-binding site (GGAGG). A possible −35 region (TTATAA) is located 16 bp upstream from a proposed −10 region (AATAATG). ORF1 proceeds until a stop codon occurs at positions 1235 to 1237. The protein specified by this sequence consists of 371 amino acids, corresponding to a molecular weight of 39,515. The second ORF contains 593 codons and could encode a protein with a molecular weight of 65,159. It starts with an ATG at positions 1252 to 1254 and terminates with a TAG at positions 3031 to 3033. The translation initiation codon of ORF2 is preceded by a possible ribosome-binding site (GGAGG) only 4 bp upstream. Putative −10 (TTAAAA) and −35 (TTACGT) regions separated by 19 bp are located within the sequence of ORF1. Possible stem loop termination structures occur at the ends of both ORFs. The first, at the end of ORF1 (bp 1274 to 1299), occurs within the nucleotide sequence of ORF2. The other, at the end of ORF2, encompasses bp 3047 to 3074. *Bal* 31-generated deletions, which removed one or the other of the ORFs, indicated that hyaluronidase activity was associated with ORF1; removal of ORF2 did not affect the enzymatic activity specified by ORF1. The function of the protein encoded by ORF2 is as yet undetermined.

Plasmid pKK232-8, which contains a promoterless chloramphenicol acetyltransferase gene (*cat*), was used to show that fragments containing the putative promoters are in fact able to function as promoters in *E. coli*. This activity was ascertained by the ability of cells containing the appropriate fragments to express resistance to chloramphenicol.

Purification of the cloned hyaluronidase. Hyaluronidase produced by *E. coli* strains carrying the *hylP* gene was partially purified by using ion exchange columns. The intracellular lysate obtained after sonification of *E. coli* cells containing the *hylP* gene was passed over a DEAE-cellulose column. Hyaluronidase activity was present in the material which passed through the DEAE column. The eluant fraction was dialyzed before being applied to a CM-cellulose column. Most of the protein appeared to pass through this column, as determined by A_{280}. Hyaluronidase activity was located in the fraction which was eluted with sodium chloride. Active fractions were pooled, dialyzed, and concentrated before being subjected to SDS-polyacrylamide gel electrophoresis. A number of bands were stained by Coomassie blue, and a major band was found corresponding to a molecular weight of 42,000 (Fig. 3).

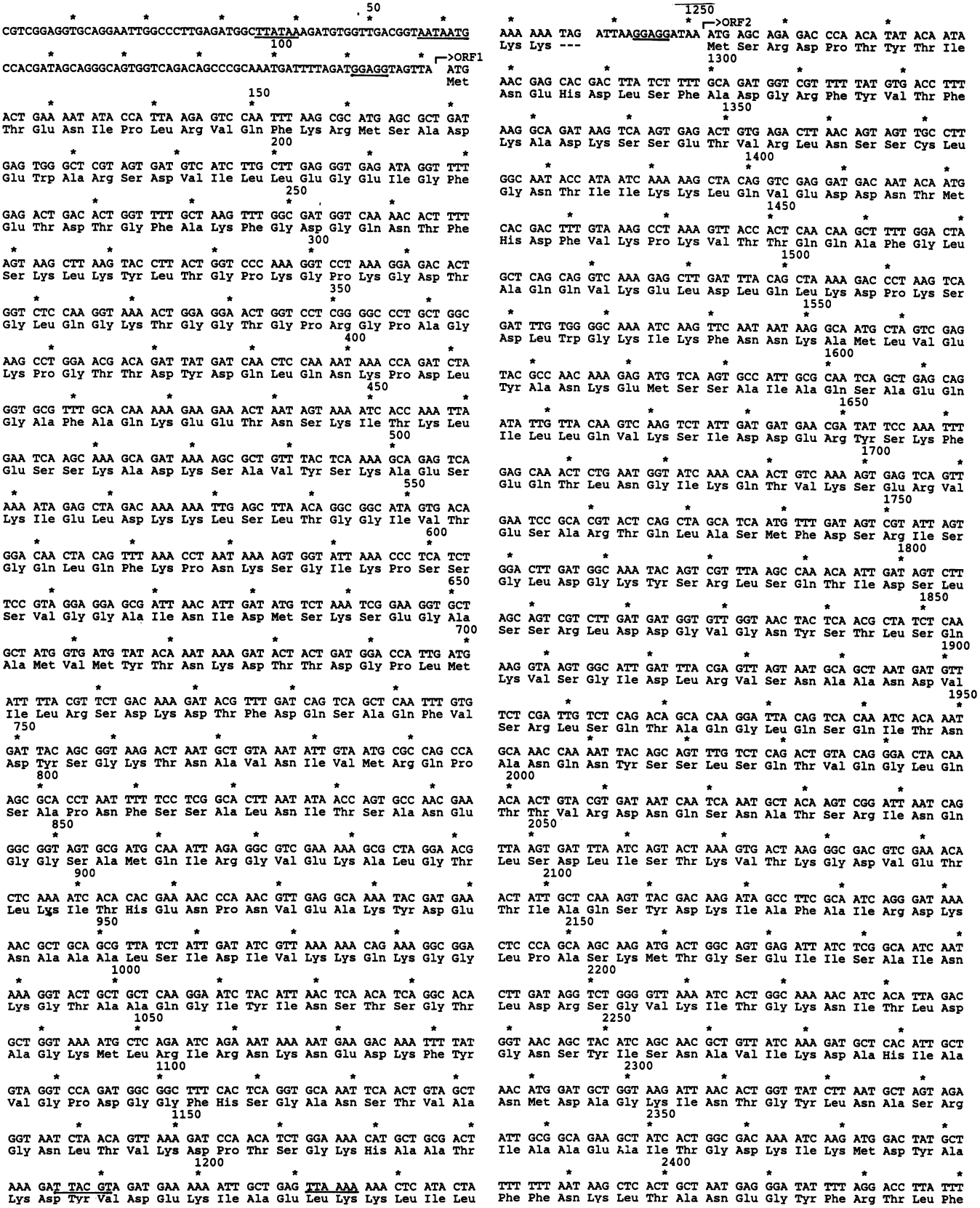
Controlled expression of the hyaluronidase by T7 RNA polymerase. Use of the T7 RNA polymerase-promoter system for the controlled expression of genes inserted into the plasmid containing the T7 promoter enabled us to identify the size of the protein encoded by the *hylP* gene. Subcloning of the *hylP* gene into vector pT7T319U created the plasmid pSF19U-1. This plasmid specified a protein of approximately 40,000 daltons, as determined by SDS-polyacrylamide gel electrophoresis. Other proteins were also labeled during this procedure; however, the 40,000-dalton protein was present only in the *hylP*-containing clone. The remaining proteins were also present in clones lacking any insert after the T7 promoter of the plasmid pT7T319U (Fig. 4). These preparations were also assayed for hyaluronidase activity immediately before addition of [^{35}S]Met. Cell preparations containing the plasmid pSF19U-1 were hyaluronidase positive, whereas those with pSF18U-1 or pT7T319U without an insert did not produce any enzymatic activity.

DISCUSSION

The gene specifying hyaluronidase production (*hylP*) from an M-type 49 group A streptococcal bacteriophage was previously cloned into *E. coli* (10). Properties of this cloned enzyme appeared similar to those of the native bacteriophage hyaluronidase, with Southern hybridization experiments confirming that the DNA fragment containing the *hylP* gene was derived from bacteriophage H4489A.

Nucleotide sequence analysis of the 3.1-kb fragment showed the presence of two ORFs. Both reading frames contained possible promoter regions immediately preceding the suggested ribosome-binding site. The putative −10 and −35 regions of the *hylP* gene (ORF1) differ by only one base from that reported as the possible promoter sequence for the streptokinase gene (*skc*) from *Streptococcus equisimilis* (14). Located within the first 406 bp of the 3.1-kb fragment is a sequence capable of acting as a promoter in *E. coli*, as determined by the use of the chloramphenicol acetyltransferase promoter detection plasmid pKK232-8. Although expression of chloramphenicol acetyltransferase in this system does not confirm that the putative promoter is the actual promoter for ORF1, it does show that the *Bam*HI-*Bgl*II fragment contains a sequence able to function in *E. coli* to promote expression of the *cat* gene. A similar observation was made with the promoter region of ORF2 which is contained within the 596-bp *Pvu*II fragment.

The genome of a bacteriophage would be expected to be highly organized and lack extraneous DNA, which may explain why the two genes present on the 3.1-kb *Tha*I fragment are separated by only 14 bp. This high degree of organization may also account for the presence of promoter and termination structures within the nucleotide sequence of another gene product. The presence of a promoter or termination sequence within the coding sequence of another gene may also be a way of regulating expression of gene products, possibly as a result of structural or conformational changes (18). It is possible that the translation of the two reading


50
CGTCGGAGGTGCAGGAATTGGCCCTTGAGATGGCTTTATAAAGATGTGGTTGACGGTAATAATG
100
CCACGATAGCAGGGCAGTGGTCAGACAGCCCGCAAATGATTTTAGATGGAGGTAGTTA
>ORF1
ATG
Met
150
ACT GAA AAT ATA CCA TTA AGA GTC CAA TTT AAG CGC ATG AGC GCT GAT
Thr Glu Asn Ile Pro Leu Arg Val Gln Phe Lys Arg Met Ser Ala Asp
200
GAG TGG GCT CGT AGT GAT GTC ATC TTG CTT GAG GGT GAG ATA GGT TTT
Glu Trp Ala Arg Ser Asp Val Ile Leu Leu Glu Gly Glu Ile Gly Phe
250
GAG ACT GAC ACT GGT TTT GCT AAG TTT GGC GAT GGT CAA AAC ACT TTT
Glu Thr Asp Thr Gly Phe Ala Lys Phe Gly Asp Gly Gln Asn Thr Phe
300
AGT AAG CTT AAG TAC CTT ACT GGT CCC AAA GGT CCT AAA GGA GAC ACT
Ser Lys Leu Lys Tyr Leu Thr Gly Pro Lys Gly Pro Lys Gly Asp Thr
350
GGT CTC CAA GGT AAA ACT GGA GGA ACT GGT CCT CGG GGC CCT GCT GGC
Gly Leu Gln Gly Lys Thr Gly Gly Thr Gly Pro Arg Gly Pro Ala Gly
400
AAG CCT GGA ACG ACA GAT TAT GAT CAA CTC CAA AAT AAA CCA GAT CTA
Lys Pro Gly Thr Thr Asp Tyr Asp Gln Leu Gln Asn Lys Pro Asp Leu
450
GGT GCG TTT GCA CAA AAA GAA GAA ACT AAT AGT AAA ATC ACC AAA TTA
Gly Ala Phe Ala Gln Lys Glu Glu Thr Asn Ser Lys Ile Thr Lys Leu
500
GAA TCA AGC AAA GCA GAT AAA AGC GCT GTT TAC TCA AAA GCA GAG TCA
Glu Ser Ser Lys Ala Asp Lys Ser Ala Val Tyr Ser Lys Ala Glu Ser
550
AAA ATA GAG CTA GAC AAA AAA TTG AGC TTA ACA GGC GGC ATA GTG ACA
Lys Ile Glu Leu Asp Lys Lys Leu Ser Leu Thr Gly Gly Ile Val Thr
600
GGA CAA CTA CAG TTT AAA CCT AAT AAA AGT GGT ATT AAA CCC TCA TCT
Gly Gln Leu Gln Phe Lys Pro Asn Lys Ser Gly Ile Lys Pro Ser Ser
650
TCC GTA GGA GGA GCG ATT AAC ATT GAT ATG TCT AAA TCG GAA GGT GCT
Ser Val Gly Gly Ala Ile Asn Ile Asp Met Ser Lys Ser Glu Gly Ala
700
GCT ATG GTG ATG TAT ACA AAT AAA GAT ACT ACT GAT GGA CCA TTG ATG
Ala Met Val Met Tyr Thr Asn Lys Asp Thr Thr Asp Gly Pro Leu Met
ATT TTA CGT TCT GAC AAA GAT ACG TTT GAT CAG TCA GCT CAA TTT GTG
Ile Leu Arg Ser Asp Lys Asp Thr Phe Asp Gln Ser Ala Gln Phe Val
750
GAT TAC AGC GGT AAG ACT AAT GCT GTA AAT ATT GTA ATG CGC CAG CCA
Asp Tyr Ser Gly Lys Thr Asn Ala Val Asn Ile Val Met Arg Gln Pro
800
AGC GCA CCT AAT TTT TCC TCG GCA CTT AAT ATA ACC AGT GCC AAC GAA
Ser Ala Pro Asn Phe Ser Ser Ala Leu Asn Ile Thr Ser Ala Asn Glu
850
GGC GGT AGT GCG ATG CAA ATT AGA GGC GTC GAA AAA GCG CTA GGA ACG
Gly Gly Ser Ala Met Gln Ile Arg Gly Val Glu Lys Ala Leu Gly Thr
900
CTC AAA ATC ACA CAC GAA AAC CCA AAC GTT GAG GCA AAA TAC GAT GAA
Leu Lys Ile Thr His Glu Asn Pro Asn Val Glu Ala Lys Tyr Asp Glu
950
AAC GCT GCA GCG TTA TCT ATT GAT ATC GTT AAA AAA CAG AAA GGC GGA
Asn Ala Ala Ala Leu Ser Ile Asp Ile Val Lys Lys Gln Lys Gly Gly
1000
AAA GGT ACT GCT GCT CAA GGA ATC TAC ATT AAC TCA ACA TCA GGC ACA
Lys Gly Thr Ala Ala Gln Gly Ile Tyr Ile Asn Ser Thr Ser Gly Thr
1050
GCT GGT AAA ATG CTC AGA ATC AGA AAT AAA AAT GAA GAC AAA TTT TAT
Ala Gly Lys Met Leu Arg Ile Arg Asn Lys Asn Glu Asp Lys Phe Tyr
1100
GTA GGT CCA GAT GGC GGC TTT CAC TCA GGT GCA AAT TCA ACT GTA GCT
Val Gly Pro Asp Gly Gly Phe His Ser Gly Ala Asn Ser Thr Val Ala
1150
GGT AAT CTA ACA GTT AAA GAT CCA ACA TCT GGA AAA CAT GCT GCG ACT
Gly Asn Leu Thr Val Lys Asp Pro Thr Ser Gly Lys His Ala Ala Thr
1200
AAA GAT TAC GTA GAT GAA AAA ATT GCT GAG TTA AAA AAA CTC ATA CTA
Lys Asp Tyr Val Asp Glu Lys Ile Ala Glu Leu Lys Lys Leu Ile Leu
1250
>ORF2
AAA AAA TAG ATTAAGGAGGATAA ATG AGC AGA GAC CCA ACA TAT ACA ATA
Lys Lys --- Met Ser Arg Asp Pro Thr Tyr Thr Ile
1300
AAC GAG CAC GAC TTA TCT TTT GCA GAT GGT CGT TTT TAT GTG ACC TTT
Asn Glu His Asp Leu Ser Phe Ala Asp Gly Arg Phe Tyr Val Thr Phe
1350
AAG GCA GAT AAG TCA AGT GAG ACT GTG AGA CTT AAC AGT AGT TGC CTT
Lys Ala Asp Lys Ser Ser Glu Thr Val Arg Leu Asn Ser Ser Cys Leu
1400
GGC AAT ACC ATA ATC AAA AAG CTA CAG GTC GAG GAT GAC AAT ACA ATG
Gly Asn Thr Ile Ile Lys Lys Leu Gln Val Glu Asp Asp Asn Thr Met
1450
CAC GAC TTT GTA AAG CCT AAA GTT ACC ACT CAA CAA GCT TTT GGA CTA
His Asp Phe Val Lys Pro Lys Val Thr Thr Gln Gln Ala Phe Gly Leu
1500
GCT CAG CAG GTC AAA GAG CTT GAT TTA CAG CTA AAA GAC CCT AAG TCA
Ala Gln Gln Val Lys Glu Leu Asp Leu Gln Leu Lys Asp Pro Lys Ser
1550
GAT TTG TGG GGC AAA ATC AAG TTC AAT AAT AAG GCA ATG CTA GTC GAG
Asp Leu Trp Gly Lys Ile Lys Phe Asn Asn Lys Ala Met Leu Val Glu
1600
TAC GCC AAC AAA GAG ATG TCA AGT GCC ATT GCG CAA TCA GCT GAG CAG
Tyr Ala Asn Lys Glu Met Ser Ser Ala Ile Ala Gln Ser Ala Glu Gln
1650
ATA TTG TTA CAA GTC AAG TCT ATT GAT GAT GAA CGA TAT TCC AAA TTT
Ile Leu Leu Gln Val Lys Ser Ile Asp Asp Glu Arg Tyr Ser Lys Phe
1700
GAG CAA ACT CTG AAT GGT ATC AAA CAA ACT GTC AAA AGT GAG TCA GTT
Glu Gln Thr Leu Asn Gly Ile Lys Gln Thr Val Lys Ser Glu Arg Val
1750
GAA TCC GCA CGT ACT CAG CTA GCA TCA ATG TTT GAT AGT CGT ATT AGT
Glu Ser Ala Arg Thr Gln Leu Ala Ser Met Phe Asp Ser Arg Ile Ser
1800
GGA CTT GAT GGC AAA TAC AGT CGT TTA AGC CAA ACA ATT GAT AGT CTT
Gly Leu Asp Gly Lys Tyr Ser Arg Leu Ser Gln Thr Ile Asp Ser Leu
1850
AGC AGT CGT CTT GAT GAT GGT GTT GGT AAC TAC TCA ACG CTA TCT CAA
Ser Ser Arg Leu Asp Asp Gly Val Gly Asn Tyr Ser Thr Leu Ser Gln
1900
AAG GTA AGT GGC ATT GAT TTA CGA GTT AGT AAT GCA GCT AAT GAT GTT
Lys Val Ser Gly Ile Asp Leu Arg Val Ser Asn Ala Ala Asn Asp Val
1950
TCT CGA TTG TCT CAG ACA GCA CAA GGA TTA CAG TCA CAA ATC ACA AAT
Ser Arg Leu Ser Gln Thr Ala Gln Gly Leu Gln Ser Gln Ile Thr Asn
GCA AAC CAA AAT TAC AGC AGT TTG TCT CAG ACT GTA CAG GGA CTA CAA
Ala Asn Gln Asn Tyr Ser Ser Leu Ser Gln Thr Val Gln Gly Leu Gln
2000
ACA ACT GTA CGT GAT AAT CAA TCA AAT GCT ACA AGT CGG ATT AAT CAG
Thr Thr Val Arg Asp Asn Gln Ser Asn Ala Thr Ser Arg Ile Asn Gln
2050
TTA AGT GAT TTA ATC AGT ACT AAA GTG ACT AAG GGC GAC GTC GAA ACA
Leu Ser Asp Leu Ile Ser Thr Lys Val Thr Lys Gly Asp Val Glu Thr
2100
ACT ATT GCT CAA AGT TAC GAC AAG ATA GCC TTC GCA ATC AGG GAT AAA
Thr Ile Ala Gln Ser Tyr Asp Lys Ile Ala Phe Ala Ile Arg Asp Lys
2150
CTC CCA GCA AGC AAG ATG ACT GGC AGT GAG ATT ATC TCG GCA ATC AAT
Leu Pro Ala Ser Lys Met Thr Gly Ser Glu Ile Ile Ser Ala Ile Asn
2200
CTT GAT AGG TCT GGG GTT AAA ATC ACT GGC AAA AAC ATC ACA TTA GAC
Leu Asp Arg Ser Gly Val Lys Ile Thr Gly Lys Asn Ile Thr Leu Asp
2250
GGT AAC AGC TAC ATC AGC AAC GCT GTT ATC AAA GAT GCT CAC ATT GCT
Gly Asn Ser Tyr Ile Ser Asn Ala Val Ile Lys Asp Ala His Ile Ala
2300
AAC ATG GAT GCT GGT AAG ATT AAC ACT GGT TAT CTT AAT GCT AGT AGA
Asn Met Asp Ala Gly Lys Ile Asn Thr Gly Tyr Leu Asn Ala Ser Arg
2350
ATT GCG GCA GAA GCT ATC ACT GGC GAC AAA ATC AAG ATG GAC TAT GCT
Ile Ala Ala Glu Ala Ile Thr Gly Asp Lys Ile Lys Met Asp Tyr Ala
2400
TTT TTT AAT AAG CTC ACT GCT AAT GAG GGA TAT TTT AGG ACC TTA TTT
Phe Phe Asn Lys Leu Thr Ala Asn Glu Gly Tyr Phe Arg Thr Leu Phe

```
                        2450
            *           *           *           *
GCT AAA AAT ATC TTT ACC ACA TCT GTA CAG GCT GTC ACT ACG TCT GCG
Ala Lys Asn Ile Phe Thr Thr Ser Val Gln Ala Val Thr Thr Ser Ala
                        2500
 *           *           *           *           *
AGT AAG ATT ACA GGT GGT GTG CTT TCT GCC ACG AAT GGT GCG AGT AGG
Ser Lys Ile Thr Gly Gly Val Leu Ser Ala Thr Asn Gly Ala Ser Arg
                      2550
    *           *           *           *           *
TGG GAT TTG AAT AGT GCA AAT ATT GAT TTT AAT CGA GAT GCC ACA ATT
Trp Asp Leu Asn Ser Ala Asn Ile Asp Phe Asn Arg Asp Ala Thr Ile
                     2600
      *           *           *           *           *
AAT TTT AAC AGC AAA AAC AAT GCG CTT GTT CGA AAA TCA GGT ACC AAC
Asn Phe Asn Ser Lys Asn Asn Ala Leu Val Arg Lys Ser Gly Thr Asn
                     2650
          *           *           *           *           *
ACT GCT TTT GTT CAC TTT AGC AAT GCG ACA CCA AAA GGC TAT AGA GGC
Thr Ala Phe Val His Phe Ser Asn Ala Thr Pro Lys Gly Tyr Arg Gly
                           2700
            *           *           *           *
TCA GCG TTG TAT GCG TCA ATC GGG ATA ACC TCA TCA GGA GAT GGC ATC
Ser Ala Leu Tyr Ala Ser Ile Gly Ile Thr Ser Ser Gly Asp Gly Ile
                           2750
 *           *           *           *           *
GAC AGC GCT TCG TCT GGA CGT TTC TGT GGA GTT AGG TTT TTC CGG TAC
Asp Ser Ala Ser Ser Gly Arg Phe Cys Gly Val Arg Phe Phe Arg Tyr
                           2800
    *           *           *           *           *
GCT GAA GGG TTA CAG CAT ACA GCA AAG GTC GAT CAA GCC GAA ATT TAT
Ala Glu Gly Leu Gln His Thr Ala Lys Val Asp Gln Ala Glu Ile Tyr
                              2850
      *           *           *           *           *
GGT GAT GAT ATT GTC TTT AGC GAC GAT TTT AAC ATC GAT CGT GGC TTT
Gly Asp Asp Ile Val Phe Ser Asp Asp Phe Asn Ile Asp Arg Gly Phe
                              2900
          *           *           *           *           *
AAG ATG CGG CCT AGC CTA ATG CCA AAA ATG GTC GAC TTA AAC AAG ATG
Lys Met Arg Pro Ser Leu Met Pro Lys Met Val Asp Leu Asn Lys Met
                                    2950
            *           *           *           *
TAC CAG GCA ATT TTG GCT CTC GGC CGC TGC TGG CTG CAT GCT AAT AAC
Tyr Gln Ala Ile Leu Ala Leu Gly Arg Cys Trp Leu His Ala Asn Asn
                                          3000
 *           *           *           *           *
ACG GCT TGG TCG TGG AAT TTT GAT ACA CGC AGC GCA ATC ATC GCA GAA
Thr Ala Trp Ser Trp Asn Phe Asp Thr Arg Ser Ala Ile Ile Ala Glu
                     3030                    3050
    *           *           *           *           *
TAT AAC GCA CAC ATT AAT AAC TTA TAG GAGAAACAATGGATTTAACGCTTAAAA
Tyr Asn Ala His Ile Asn Asn Leu ---
                              3100
          *           *           *           *           *
ACAAAGATTTAAACACACTATATAGTGTACTAGACAAAATCAAAATCACGAACATGCGAGCAA
   3127
   *
ACCG
```

FIG. 2. Nucleotide sequence of the *Tha*I fragment obtained from the group A streptococcus temperate bacteriophage H4489A. The deduced amino acid sequences of the two ORFs are given below the nucleotide sequence. Putative promoter and ribosome-binding site sequences are underlined. The start positions of the ORFs are indicated. ORF1 is the *hylP* gene; ORF2 codes for an unknown protein.

frames is in some way coupled or closely coordinated, as has been suggested to occur within the genes of the streptococcal transposon Tn*917* (20).

The first ORF of pSF49 encodes a 39,515-dalton protein which has hyaluronidase activity. The deduced amino acid sequence of the downstream reading frame has a size of 65,159 daltons, the function of which remains to be elucidated. A computer search of the gene bank library indicated that some similarity, albeit at low significance levels, existed between the protein encoded by ORF2 and the sigma 1 protein of reovirus type 3.

Neither of the cloned products from the H4489A bacteriophage had an N-terminal region indicative of a signal peptide. This finding was not surprising, since bacteriophages would not be expected to produce extracellular products when all the processes involved in replication and assembly of the phage particle are intracellular.

The molecular weights of hyaluronidase obtained from the deduced sequence and that observed after expression from the T7 promoter by the T7 RNA polymerase as well as by partial purification and SDS-polyacrylamide gel electrophoresis are considerably smaller than that reported by Benchetrit et al. (1) for the hyaluronidase purified from the M-type 49 bacteriophage GT8760. Comparison of the amino acid content deduced from the DNA sequence with that of the purified phage-associated sample reported by Benchetrit et al. (1) showed major differences between the two proteins.

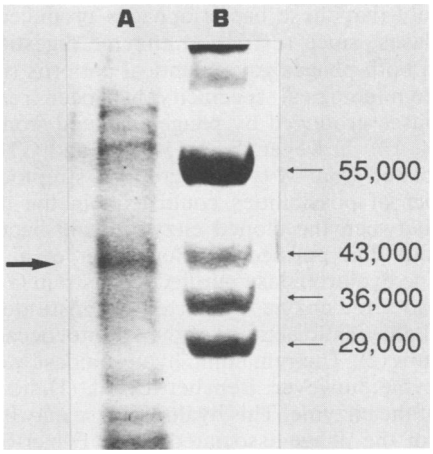


FIG. 3. Coomassie blue-stained SDS-polyacrylamide gel of partially purified hyaluronidase after ion exchange chomatography. Lane A, Partially purified protein. The position of the 42,000-dalton protein is indicated by the arrow. Lane B, Molecular weight markers.

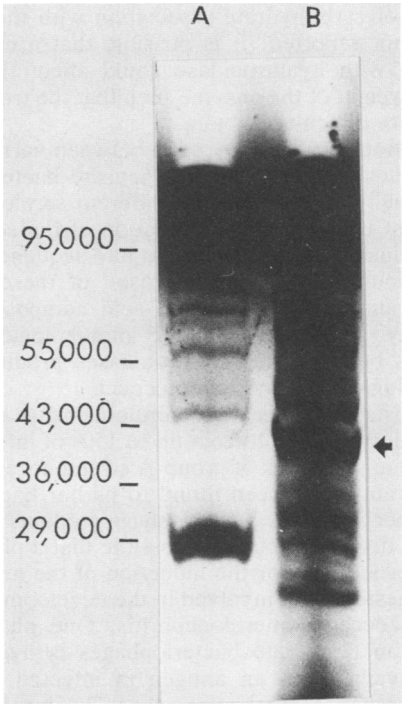


FIG. 4. Expression of the *hylP* gene by the T7 RNA polymerase-promoter system. Cultures were pulse-labeled with [^{35}S]Met 75 min after induction of the RNA polymerase. Cells were harvested, boiled for 5 min in cracking buffer, and loaded onto a 12% SDS-polyacrylamide gel. Labeled proteins were detected by autoradiography. Lane A, *E. coli* containing pT7T319U and pGP1-2; lane B, *E. coli* containing pSF19U-1 and pGP1-2. The arrow indicates the position of the 40,000-dalton band present in strains carrying ORF1 (lane B).

It is unlikely that these bacteriophages produced different hyaluronidases, since restriction enzyme digestions of the DNA from both phages gave identical patterns (results not shown). Immunological specificity has been reported for hyaluronidases produced by phages isolated from different M types (4, 12); however, both H4489A and GT8760 were initially isolated from M-type 49 group A streptococci.

A number of possibilities could explain the differences observed between the cloned enzyme from bacteriophage H4489A and that purified by Benchetrit et al. (1) from GT8760. The hyaluronidase purified from strain GT8760 was a phage-associated enzyme, reported to constitute only 25% of the total enzymatic activity of the streptococcal bacteriophage culture (2). The remaining hyaluronidase was present as free enzyme; however, Benchetrit et al. (1) did not purify this form of the enzyme. This hyaluronidase may be different from that of the phage-associated form. Fiszer-Szafarz (6) found that a hyaluronidase purified from a group A streptococcus migrated as two isoenzymes in a hyaluronic acid-containing gel, although the nature of the differences between the two forms was not investigated. Another possibility is that the enzyme preparation purified by Benchetrit et al. (1) may have contained extraneous material attached to the hyaluronidase derived from the medium or the streptococcal host strain used in the propagation of the bacteriophage.

Hill (9) purified a hyaluronidase from a type 4 group A streptococcus and estimated its molecular weight to be 50,000 by SDS-polyacylamide gel electrophoresis. The studies of both Benchetrit et al. (1) and Hill (9) indicated that the streptococcal hyaluronidases were glycoproteins, although the nature of carbohydrate association with the hyaluronidase was not reported. It is possible that a carbohydrate association with hyaluronidase could affect the apparent molecular weight of the enzyme such that the true molecular weights were difficult to obtain.

Any relationship that may exist between various streptococcal hyaluronidases, whether they are bacteriophage or chromosomally controlled from different serological types, is at present unknown. Determination of the sequences of these proteins could give an insight into sequences common to the various enzymes. Comparison of these nucleotide sequences and hence the amino acid composition of the protein may explain some of the immunological diversity that occurs between the hyaluronidases produced by different serological types of streptococci.

Acute poststreptococcal glomerulonephritis as a sequela has been shown to result from up to 15% of infections with nephritogenic serotypes of group A streptococci, and many of these strains have been found to harbor bacteriophages (24). Whether any of the bacteriophage products has a role in nephritis is unknown, but it is possible that a phage protein could be responsible for the induction of the immune complexes suggested to be involved in the development of acute poststreptococcal glomerulonephritis. One phage product expressed on temperate bacteriophages is hyaluronidase, which is available as an antigen in infected individuals. Antibody responses to bacteriophage hyaluronidase were found to be greatest in patients who developed acute poststreptococcal glomerulonephritis after an infection with an M-type 49 streptococcus strain (8). A direct relationship between lysogeny and the development of nephritis was unlikely, as patients who did not develop acute poststreptococcal glomerulonephritis after an M-type 49 streptococcal infection also showed a rise in antibody titer to the bacteriophage hyaluronidase. The hyaluronidase, however, may aid in the dispersion of the streptococcus from the initial site of infection, thereby acting as a virulence factor for the host organism.

ACKNOWLEDGMENT

This work was supported by Public Health Service grant AI19304 from the National Institutes of Health.

LITERATURE CITED

1. **Benchetrit, L. C., E. D. Gray, R. D. Edstrom, and L. W. Wannamaker.** 1978. Purification and characterization of a hyaluronidase associated with a temperate bacteriophage of group A, type 49 streptococci. J. Bacteriol. **134**:221–228.
2. **Benchetrit, L. C., E. D. Gray, and L. W. Wannamaker.** 1977. Hyaluronidase activity of bacteriophages of group A streptococci. Infect. Immun. **15**:527–532.
3. **Benchetrit, L. C., S. L. Pahuja, E. D. Gray, and R. D. Edstrom.** 1977. A sensitive method for the assay of hyaluronidase activity. Anal. Biochem. **79**:431–437.
4. **Benchetrit, L. C., L. W. Wannamaker, and E. D. Gray.** 1979. Immunological properties of hyaluronidases associated with temperate bacteriophages of group A streptococci. J. Exp. Med. **149**:73–83.
5. **Brosius, J.** 1984. Plasmid vectors for the selection of promoters. Gene **27**:151–160.
6. **Fiszer-Szafarz, B.** 1984. Hyaluronidase polymorphism detected by polyacrylamide gel electrophoresis. Application to hyaluronidases from bacteria, slime molds, bee and snake venoms, bovine testes, rat liver lysosomes and human serum. Anal. Biochem. **143**:76–81.
7. **Gilmore, M. S., K. S. Gilmore, and W. Goebel.** 1985. A new strategy for ordered DNA sequencing based on a novel method for the rapid purification of near-milligram quantities of a cloned restriction fragment. Gene Anal. Technol. **2**:108–114.
8. **Halperin, S. A., P. Ferrieri, E. D. Gray, E. L. Kaplan, and L. W. Wannamaker.** 1987. Antibody response to bacteriophage hyaluronidase in acute glomerulonephritis after group A streptococcal infection. J. Infect. Dis. **155**:253–261.
9. **Hill, J.** 1976. Purification and properties of streptococcal hyaluronate lyase. Infect. Immun. **14**:726–735.
10. **Hynes, W. L., and J. J. Ferretti.** 1987. Cloning of the hyaluronidase gene from *Streptococcus pyogenes* bacteriophage H4489A, p. 150–152. *In* J. J. Ferretti and R. Curtis III (ed.), Streptococcal genetics. American Society for Microbiology, Washington, D.C.
11. **Kjems, E.** 1958. Studies on streptococcal bacteriophages. 2. Adsorption, lysogenization and one step growth experiments. Acta Pathol. Microbiol. Scand. **42**:56–66.
12. **Kjems, E.** 1958. Studies on streptococcal bacteriophages. 3. Hyaluronidase produced by the streptococcal phage-host system. Acta Pathol. Microbiol. Scand. **44**:429–439.
13. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
14. **Malke, H., B. Roe, and J. J. Ferretti.** 1985. Nucleotide sequence of the streptokinase gene from *Streptococcus equisimilis* H46A. Gene **34**:357–362.
15. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
16. **Maxted, W. R.** 1952. Enhancement of streptococcal bacteriophage lysis by hyaluronidase. Nature (London) **170**:1020–1021.
17. **McClean, D.** 1941. The capsulation of streptococci and its relation to diffusion factor (hyaluronidase). J. Pathol. Bacteriol. **53**:13–27.
18. **McClure, W. R.** 1985. Mechanism and control of transcription initiation in prokaryotes. Annu. Rev. Biochem. **54**:171–204.
19. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
20. **Shaw, J. H., and D. B. Clewell.** 1985. Complete nucleotide

sequence of macrolide-lincosamide-streptogramin B-resistance transposon Tn*917* in *Streptococcus faecalis*. J. Bacteriol. **164:** 782–796.

21. **Smith, R. F., and N. P. Willett.** 1968. Rapid plate method for screening hyaluronidase and chondroitin sulfatase-producing microorganisms. Appl. Microbiol. **16:**1434–1436.
22. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98:**503–517.
23. **Tabor, S., and C. C. Richardson.** 1985. A bacteriophage T7 RNA polymerase/promoter system for the controlled exclusive expression of specific genes. Proc. Natl. Acad. Sci. USA **82:**1074–1078.
24. **Wannamaker, L. W., S. Skjold, and W. R. Maxted.** 1970. Characterization of bacteriophages from nephritogenic group A streptococci. J. Infect. Dis. **121:**407–418.