

Hyaluronate synthase: cloning and sequencing of the gene from *Streptococcus* sp.

Manfred LANSING,*|| Sabine LELLIG,* Andreas MAUSOLF,† Irene MARTINI,‡ Fabiana CRESCENZI,‡ Michael O'REGAN‡ and Peter PREHM*§

*Institut für Physiologische Chemie und Pathobiochemie, Waldeyerstrasse 15, D-44 Münster, Federal Republic of Germany,

†Max Planck Institut für Biochemie, D-8033 Martinsried, Federal Republic of Germany,

and ‡Advanced Technology Division, FIDIA s.p.a., Via Ponte della Fabbrica 3/A, I-35031 Abano Terme PD, Italy

The complete nucleotide sequence of hyaluronate synthase from *Streptococcus* sp. and its flanking regions is presented. The gene locus was designated *has*. Southern-blotting results suggested that the gene was conserved in hyaluronate-producing streptococci. A putative translation-initiation codon was identified and the open reading frame consists of 1566 bp,

specifying a protein of 56 kDa. Sequences resembling the promoter and ribosome-binding site of Gram-positive organisms are found upstream of the synthase. The predicted amino-acid sequence reveals the presence of a 35-residue signal peptide. The sequence has some similarity to bacterial peptide-binding proteins.

INTRODUCTION

Hyaluronate is a linear high-molecular-mass glycosaminoglycan which has been shown to be involved in many biological interactions (Laurent and Fraser, 1992). Group A and C streptococci produce a hyaluronate capsule which amplifies infectious virulence, because it protects streptococci against immunological attacks of the infected host (Todd and Lancefield, 1928; Seastone, 1939; Hirst, 1941; Whitnack et al., 1981; Wessels et al., 1991). Hyaluronate is produced by a synthase which resides in protoblast membranes of streptococci (Markovitz et al., 1958; Markovitz and Dorfman, 1962; Stoolmiller and Dorfman, 1969; Sugahara et al., 1979) and in plasma membranes of eukaryotic cells (Prehm, 1984). We have previously identified the streptococcal synthase as a 52 kDa protein (Prehm and Mausolf, 1986). The bacterial enzyme cross-reacts immunologically with proteins in plasma membranes from eukaryotic cells (Prehm 1989; Klewes et al., 1993) and is shed from growing streptococci into the culture media (Mausolf et al., 1990). In order to gain more information on the hyaluronate synthase itself and on the genetics of hyaluronate synthesis, we have undertaken the cloning and sequencing of the gene encoding the hyaluronate synthase.

MATERIALS AND METHODS

Materials

DNA-modifying enzymes were obtained from Boehringer, NEB Gibco/BRL or Pharmacia. Radiochemicals were obtained from Amersham International, and other reagents from Sigma Chemical Co. λ gt-11 and pBS-SK were from Stratagene, pGEM-3Z from Promega, and pUC-18 from Boehringer. *Eco*RI-linker: 5'-d(CCGGAATCCGG)-3' was from NEB. Oligonucleotides were synthesized on an Applied Biosystems DNA Synthesizer 380B. The primers used were as follows: primer 1, 5'-d(GGTGGCGACGACTCCTGGAGCCCG)-3'; primer 2, 5'-d(TTGACACCAGACCAACTGGTAATG)-3'; primer 3,

5'-d(AGGAATGTCACAACCTT)-3'; primer 4, 5'-d(CCCC-TAGAGAGTCTAGA)-3'. DNA manipulations were performed as described by Sambrook et al. (1989).

Bacterial strains

Streptococcus equisimilis strain D181 was from the Rockefeller University collection. *Streptococcus equi* 68222 and *Streptococcus zooepidemicus* 68270 were from the Belfanti Serotherapeutic Institute, Milan. *Streptococcus sanguis* (Challis) was obtained from Dr. Laplace, Jena, F.R.G. *Escherichia coli* SMR 10 was from Dr. Meinhard, Münster. *E. coli* Y1089 and Y1090 were from Stratagene.

Isolation of streptococcal DNA

Genomic DNA was isolated from streptococci by the method of Nida and Cleary (1983) with slight modifications. An overnight culture of streptococci (5 ml) growing in Todd–Hewitt broth containing 40 mM DL-threonine and 30 mM glucose at 37 °C was used to inoculate 100 ml of Todd–Hewitt broth containing 40 mM DL-threonine. The culture was incubated at 37 °C. On reaching a D_{600} of 0.03, cysteine was added to give a final concentration of 6 mM, and at a D_{600} of 0.1 testicular hyaluronidase was added to give a final concentration of 3 units/ml. The bacteria were harvested, at a D_{600} value of 0.5, by centrifugation at 5000 g for 10 min, washed twice with TE buffer (10 mM Tris/HCl/1 mM EDTA) at 0 °C, resuspended in 8.75 ml of TE and incubated with lysozyme (5 mg/ml) and hyaluronidase (1.5 units/ml) for 1 h at 37 °C. Proteinase K and Pronase E (250 μ g/ml of each) were added and incubation was continued for 30 min at 37 °C. SDS was added to a final concentration of 0.5% and incubation was continued, until cell lysis was observed. TE was added to a final volume of 10 ml. NaCl and cetyltrimethylammonium bromide (CTAB) were added to final concentrations of 0.7 M and 1.13% (w/v) respectively. The solution was heated to 65 °C for 20 min and extracted with 1 vol.

Abbreviations used: FE, 10 mM Tris/HCl, 1 mM EDTA (pH 7.5); CTAB, cetyltrimethylammonium bromide.

§ To whom correspondence should be addressed.

|| Present address: Advanced Technology Division, FIDIA s.p.a., Via Ponte della Fabbrica 3/A, I-35031 Abano Terme PD, Italy.

chloroform/isoamyl alcohol (24:1). The phases were separated by centrifugation at 9000 *g* for 10 min at room temperature. The upper, aqueous phase was withdrawn and mixed with 0.6 vol. propan-2-ol to precipitate DNA.

Genomic DNA library

High-molecular-mass DNA from *S. equisimilis* D181 was partially digested with *AluI*, and fragments of 2–5 kb were isolated by sucrose gradient centrifugation (Weis, 1989). The isolated *AluI* fragments were dephosphorylated, methylated with *EcoRI* methylase and ligated to *EcoRI* linkers (Wu et al., 1987). The linker-containing DNA was ligated to *EcoRI*-digested λ gt-11 arms and packaged into infectious virions using the 'one-strain-packaging' system *E. coli* SMR10 according to the method of Rosenberg et al. (1985). The genomic library was established using *E. coli* Y1089 and amplified with *E. coli* Y1090.

Library screening

The library was screened as described by Huynh et al. (1986) with polyclonal antibodies raised against the synthase which had previously been isolated in our laboratory (Prehm and Mausolf, 1986). Cross-reactive antibodies to *E. coli*/ λ gt-11 lysates were removed by pseudoscreening as described by Sambrook et al. (1989).

Clone characterization

Plate lysates were prepared from positive λ clones. The plates were washed with 5 ml of phosphate-buffered saline and proteins were precipitated by addition of polyethylene glycol 6000 to a final concentration of 20% (w/v). The samples were incubated at 4 °C for 2 h. After centrifugation at 35000 *g* for 30 min at 4 °C the pellet was resuspended in 0.5 ml of phosphate-buffered saline. Proteins were precipitated by the method of Wessel and Flügge (1984), separated on a 10% SDS/polyacrylamide gel and analysed by Western blotting with antibodies raised against the hyaluronate synthase.

Inserts of the λ clones were amplified using PCR by the method of Saiki et al. (1988) using primers 1 and 2 adjacent to the *EcoRI* cloning site. PCR was carried out in a Bio-med thermocycler 60. The thermocycling profile consisted of an initial denaturation at 95 °C for 2 min, followed by 35 cycles at 95 °C, for 1 min; at 45 °C, for 1 min; and at 72 °C, for 5 min. PCR products were purified according to Both et al. (1991), digested with different restriction enzymes to map the inserts and analysed by agarose-gel electrophoresis.

Southern blotting was carried out essentially as described by Sambrook et al. (1989). Probe preparation, hybridization and detection were carried out with the non-radioactive DIG-DNA labelling kit from Boehringer (Mannheim).

DNA-sequence analysis

Nucleotide sequences were determined by the dideoxy-chain-termination method according to Sanger et al. (1977) using the T7 DNA polymerase sequencing kit (Pharmacia) or the fmol sequencing kit (Promega). Double-stranded plasmid DNA was purified using the plasmid preparation kit from Diagen and sequenced according to Chen and Seeburg (1985).

λ DNA was sequenced using the protocol of Snyder et al. (1987) with slight modifications. DNA (5 μ g) was double digested with *KpnI* and *SacI*, phenol extracted, ethanol precipitated, and resuspended in 20 μ l of water. The DNA was denatured by

addition of 2 μ l of a solution containing 2 M NaOH and 2 mM EDTA for 10 min at 37 °C. Sodium acetate (13 μ l; 1.5 M) was added and the DNA was precipitated with 2.5 vol. of ethanol at –20 °C, and the sample was placed in a methanol/solid CO₂ bath for 15 min. After centrifugation the DNA was resuspended in 15 μ l of 7.5 mM Tris/HCl/7.5 mM dithiothreitol/5 mM MgCl₂, pH 7.5, containing 0.7 ng/ μ l ³²P-labelled primer 1 or 2 and incubated for 15 min at 55 °C. Sequencing reactions were carried out as for plasmid DNA sequencing.

PCR-amplified DNA from λ gt-11/2LK-D was sequenced as described by Both et al. (1991) with primers 3 and 4 which were deduced from the hyaluronate-synthase-gene sequence.

λ gt-11/2LK-D DNA was isolated as described above and sequenced using the fmol sequencing kit (Promega) with ³²P-end-labelled primer 2.

DNA sequences were analysed with DNASTAR software (DNASTAR Inc. Madison, U.S.A.), GENEPRO software (Riverside Scientific, Washington, U.S.A.) and CLUSTAL software (Higgins and Sharp, 1988).

Isolation of peptides and amino-acid-sequence analysis

Purified hyaluronate synthase (200 μ g) was dissolved in 200 μ l of 6 M urea and diluted with 400 μ l of 20 mM Tris/HCl, pH 7.8, and digested with 5% (w/v) trypsin or V8-protease for 24 h. The mixture was acidified with trifluoroacetic acid and peptides were centrifuged at 10000 *g* for 3 min. The supernatant was applied to a C₁₈ reverse-phase h.p.l.c. column (VYDAC 218 TP54). Peptides were eluted with a linear gradient of 0.1% trifluoroacetic acid in water to 0.1% trifluoroacetic acid in 70% acetonitrile in 30 min with a flow rate of 1 ml/min. The elution was monitored at 220 nm. The fractions were dried in a speedvac and analysed in a gas-phase sequencer (Applied Biosystems 120A).

RESULTS

Isolation of λ gt-11 clones expressing hyaluronate synthase

A λ gt-11 genomic library was established from *S. equisimilis* D181 and screened with antibodies against the hyaluronate synthase. Positive plaques were found on plates which had not been induced by isopropylthiogalactoside, indicating that those recombinants contained a promoter which was functional in *E. coli*. Ten positive clones were identified from 50000 plaques. These clones were purified and amplified.

Proteins from lysates of positive clones were separated by SDS/PAGE and analysed after Western blotting. One of the clones produced two proteins of about 52 and 56 kDa, which were recognized by the antibodies (Figure 1, lane B). This clone was designated λ gt-11/2LK. A similar pattern of proteins was found in protoblast membranes extracted from streptococci (Figure 1, lane A). Proteins from *E. coli*/ λ gt-11 lysates (Figure 1, lane C) were not stained with the antibodies. Hyaluronate synthase was purified from lysates of λ gt-11/2LK and from membrane preparations of *S. equisimilis* D181 as described (Prehm and Mausolf, 1986). Proteolytic fragments were isolated and the amino-acid sequences were determined. An identical sequence was found in both preparations (VAVVLWGGD rP), indicating that the recombinant protein was indeed the synthase.

λ gt-11/2LK contained a 3.1 kb insert (Figure 2) which could be cleaved by *EcoRI* into 2.1 kb, 0.33 kb and 0.67 kb fragments. The 2.1 kb fragment was subcloned in λ gt-11 to generate λ gt-11/2LK-D (Figure 2). This clone also expressed the entire hyaluronate synthase (results not shown) and was used for sequence analysis.

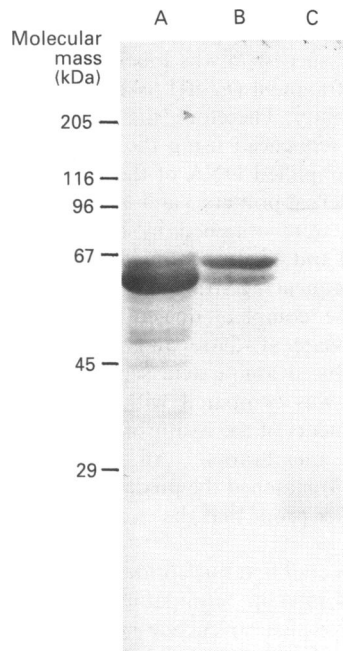


Figure 1 Comparison of the streptococcal and recombinant hyaluronate synthase

Streptococcal hyaluronate synthase was isolated from streptococcal membranes as described by Prehm and Mausolf (1986) and proteins isolated from λ gt-11/2LK and λ gt-11 lysates were compared by Western-blotting experiments using anti-(hyaluronate synthase) antibodies. Lane A, streptococcal hyaluronate synthase; lane B, λ gt-11/2LK lysate; lane C, λ gt-11 lysate. Protein molecular mass standards are shown on the left.

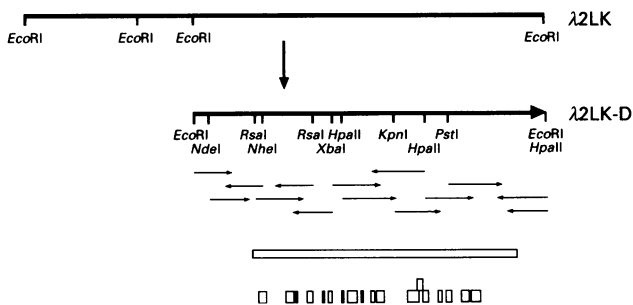


Figure 2 Sequencing strategy for the hyaluronate synthase

The original λ gt-11/2LK clone was digested with *EcoRI* and subcloned to yield an insert of 2.1 kb which also expressed the entire synthase. The extent and direction of individual sequence analysis are indicated by arrows. The synthase coding region is represented by the large open box below the 2.1 kb *EcoRI* fragment (thick arrow). Small boxes represent segments verified by peptide-sequence analysis of the streptococcal hyaluronate synthase.

Conservation of the hyaluronate synthase gene

To determine if the gene encoding the hyaluronate synthase was conserved in other hyaluronate-synthesizing streptococcal species, Southern-blotting experiments were carried out. The results are shown in Figure 3. When the 3.1 kb insert of λ gt-11/2LK was used as probe, hybridization to DNA from all three hyaluronate-producing streptococci (*S. equi*, *S. zooepidemicus* and *S. equisimilis*) was observed, while no hybridization to DNA of *S. sanguis* (a non hyaluronate producer) was

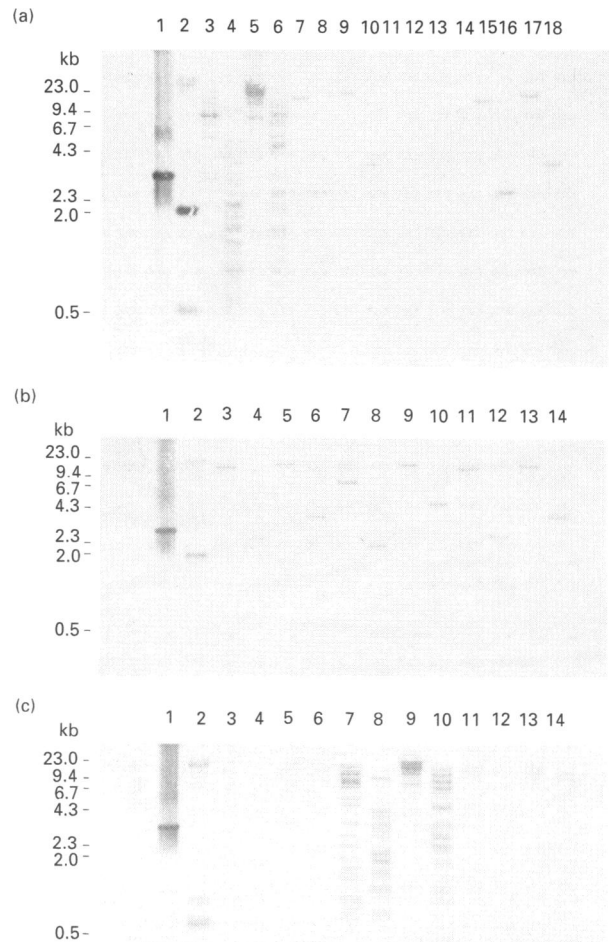


Figure 3 Hybridization studies with the hyaluronate synthase gene

(a) Southern blot probed with the 3.1 kb fragment of λ gt-11/2LK amplified by PCR. Lanes: 1, PCR product; 2, 2LK *EcoRI* EI; 3–6, *S. equisimilis* D181 DNA digested with *EcoRI*, *HindIII*, *BamHI* and *EcoRV* respectively; 7–10, *Strep. equi* DNA digested with the same enzymes; 11–14, *Strep. sanguis* DNA digested with the same enzymes; 15–18, *Strep. zooepidemicus* DNA digested with the same enzymes. (b) Probe: *XbaI*–*PstI* internal *has* fragment. Lanes 1 and 2 are as in (a); lanes 3–6, *S. equi* DNA digested with enzymes as in (a); lanes 7–10, D181 DNA digested with the same enzymes; lanes 11–14, *S. zooepidemicus* DNA digested with the same enzymes. (c) Probe: flanking *EcoRI* fragment. Sample order is as in (b). Molecular masses are indicated on the left-hand side.

seen (Figure 3a). The hybridization pattern observed with *S. equisimilis* was unexpected and indicated the presence of multiple copies of the DNA sequence used as probe. However, blotting experiments carried out with an internal fragment of the synthase gene and a flanking DNA fragment (Figures 3b and 3c) showed that this repetitive pattern was not linked to the gene itself. The similarity of the hybridization pattern obtained with various enzyme digests suggests that the synthase locus is conserved in these strains. The repetitive pattern observed for *S. equisimilis* D181 DNA indicates that the synthase gene is flanked by a DNA segment which is unique to this organism and which is present in 5–10 copies in the chromosome, suggesting that it may be an insertion sequence or a transposon.

DNA sequences

It was not possible to subclone the 2.1 kb fragment into plasmids such as pUC-18, pGEM-3Z and pBS-SK for conventional DNA

GAA TTC CCT TGT TTT AGG AAT GTC ACA ACC TTA TGG ATG GTT TAC 45
 GGC TTC GTT TAA GTG ATT GCG CAA TTT TGA GAA CAG AAA TGC CTT 90
 CTT CAA CAT ATG CCT CTA TCA TTA CAA GTT CGT TTG TGG TAA GAT 135
 GGG TAT AGG TCA TCT ATG TTA TTC CTT TCA GAC AAA TGT GGT GTT 180
 TAT CTG AGC CTA ACA TAG ATG GCT TTT TCT GTC TAA CTT AAT TAT 225
 -35 ACA AAT TGG GAT TTA TAA AAA TAA AAT CAA AAA TTT TCA GAT AAA 270
 TCT ATT TCC TTT TTT CAG AAT TAT GGT ATA ATC TAA TTT AAT TAT 315
 TTT AAC TAA CTA TTT TGT TAG GAG GAA GGT TAT ATG ACA GTA CTA 360
 S/D
 M T V L
 GGA ACA AAA GCA TGT AAG CST CTG GGC TTA GCG ACA GTT ACG CTA 405
 G T K A C K R L G L A T V T L
 GCC TCT GTT GCT GCC TTG ATG GCT TGT CCA AAT AAG CAA TCA CGG 450
 A S V A A L M A C P N K Q S A
 TCA ACA GAC AAA AAG AGT GAG ATT AAT TGG TAT ACG CCA ACT GAG 495
 S T D K K S E I N W Y T P T E
 ATC ATC ACA CTT GAT ATT TCA AAA AAT ACA GAT ACC TAT TCT GCT 540
 I I T L D I S K N T D T Y S A
 TTG GCT ATT GGT AAT TCT GGC AGT AAC CTT TTG CGT GCT GAT GCT 585
 L A I G N S G S N L L R A D A
 AAA GGG AAA TTA CAG CCT GAT TTA GCT GAA AAG GTT GAT GTG TCA 630
 K G K L Q P D L A T V D V S
 GAG GAT GGC TTG ACC TAT ACA GCA ACC CTG CGA CAT GGC TTG AAA 675
 E D G L T A T A T L R H G L K
 TGG TCT GAT GGT AGT GAT CTA ACA GCA GAG GAC TTT GAG TAC AGT 720
 H S D G S D L T A E D F E Y S
 TGG CAG CGA ATG GTC GAT CCT AAG ACA GCC TCA GAG TAC GCT TAC 765
 W Q R M V D P K T A S E Y A Y
 TTG GCA ACT GAG TCA CAT CTG AAA AAC GCA GAG GAC ATT AAT AGC 810
 L A T E S H V K N A E D I N S
 GGG AAA AAT CCT GAT CTA GAC TCT CTA GGG GTA AAG GCT GAT GGG 855
 G K N P D L D S L G V K A D G
 AAT AAG GTT ATT TTT ACC TTA ACG GTG CCG GCA CCA CAA TTT AAG 900
 N K V I F T L T V P A P Q F K
 AGC TTG CTA TCC TTC TCT AAC TTT GTG CPT CAA AAA GAA TCC TTT 945
 S L L S F S N F V C Q K E S F
 GTC AAG GAC GCT GGC AAG GAC TAT GGG ACA ACA TCA GAA AAA CAA 990
 V K D A G K D Y G T T S E K Q
 ATT TAT TCT GGT CCT TAT ATT GTC AAG GAC TGG AAT GGC ACT AGC 1035
 I Y S G P Y I V K D W N G T S
 GGA ACC TTT AAG CTA GTA AAG AAT AAA AAC TAT TGG GAC GCC AAA 1080
 G T F K L V K N M Y H D A K
 AAC GTC AAA ACT GAG ACA GTT AAT GTT CAA ACG GTT AAA AAG CCA 1125
 N V K T E T V N T V K K P
 GAT ACA GCT GTT CAA ATG TAC AAG CAA GGT AAG CTA GAC TTT GCA 1170
 D T A V Q M Y K G A T A L V A
 AAT ATT TCT GGT ACC TCA GCT ATT TAC AAT GCT AAT AAA AAG CAT 1215
 N I S G T S A I Y N A N K K H
 AAG GAC GTT GTT CCA GTT CTT GAG ACA ACA GCC TAT ATC GTA 1260
 K D V V P V L E A T T A Y I V
 TAT AAC CAG ACA GGA GCT ATT GAA GGC TTG AAC AGT CTT AAA ATT 1305
 Y N Q T G A L N S L K I
 CGT CAA GCC TTG AAT TTG GCA ACA GAC CGT AAG GGA ATT GTA TCT 1350
 R Q A L N L A T A C A K S A V
 GCG GCT CTT GAT ACA GGA TCA AAG CCG GCT ACA GCG CTT GTT CCT 1395
 A A V D T G S K P A T A L V F
 ACA GGT CTT GCT AAA TTA TCT GAC GGA ACA GAT CTA ACA GAG CAT 1440
 T G L A K L S D G T D L T E H
 GTA GCA CCT GGC TAT AAA TAC GAT AAG GAG GCA GCA AAG CTC 1485
 V A P G Y K Y D D K E A A K L
 TTC AAG GAA GGC TTA GCA GAG CTG GGC AAG GAT GCC TTG ACA ATC 1530
 F K E G L A E L G K D A L T I
 ACC ATC ACA GCT GAT GCT GAT GCC GCC TGC GCC AAG TCT GCA GTG 1575
 I T A D A D A C A C A K S A V
 GAT TAC ATC AAG GAA ACC TGG GAA ACA GCT CTT CCT GGC TTG ACC 1620
 D Y I K E T W E T A L P G L T
 GTT GAA GAG AAA TTT GTT CCT TTC AGC CAA CGT CTT GAG GAT ACT 1665
 V E E K F V P F S Q R L E D T
 AAA AAC CAA AAC TTT GAG GTT GCA GTT CTT TGG GGT GGT GAT 1710
 K N Q N F E V A V V L W G G D
 TAT CCA GAA GGG TCT ACC TTC TAT GGC TTG TTT AAA TCA GGT TCT 1755
 Y P E G S T F Y K F S G S
 GCT TAT AAC TAT GGC AAA TTT ACG AAT GCA GAC TAT GAT GCC GCC 1800
 A Y N Y G K F T N A D Y D A A
 TAC AAT AAG GCA CTA ACA CAA GAT GCC TTG AAT ATT GAT CGC GCT 1845
 Y N K A L T T D A L I D A A
 GCT GAT GAC TAT AAG GCA GCA GAA AAA GCC CTT TAT GAC AAT GCT 1890
 A D D Y K A A E K A L Y D N A
 CTT TAC AAT CCC CTT TAC CTT CCG TAG TGG TGA GGC CTT GCA AAA 1935
 L Y N P L Y L P *
 TGG AAG CAT CAA AGG TCT TAT AAG TAA TTC AAC TGG TCT AAA TGA 1980
 TGT CTT TAC TTA CGC TTA TAA GGA ATA ACA ATA ATA TTA TCG TAA 2025
 GCT TAT ACT GAC CTA GCA AGG ATA GCT CCT TGG TGG GAA GTG TTT 2070
 AAG AGA AGC TTA GGG CTA GCC GGA ATT C 2098

sequencing. The subcloned fragments were always shorter than the original DNA, suggesting instability of the clone at high copy number. However, the 1.27 kb fragment from the right-hand side (*Xba*I to *Eco*RI) could successfully be subcloned, indicating that the left-hand 0.83 kb fragment (*Eco*RI to *Xba*I) was responsible for the plasmid instability. Therefore, the left-hand side of the 2.1 kb fragment was sequenced using the fmol sequencing kit with primer 2. PCR-amplified DNA of the λ gt-11/2LK-D was sequenced with the internal primers 3 and 4. Fragments from the λ gt-11/2LK-D insert were obtained by digestion with *Nde*I, *Nhe*I, *Pst*I, *Kpn*I, *Rsa*I and *Xba*I, subcloned into plasmid vectors and sequenced. The sequencing strategy is shown in Figure 2.

Figure 4 shows the complete nucleotide sequence of the streptococcal hyaluronate synthase and its flanking regions, together with the deduced amino-acid sequence. The deduced amino-acid sequence was compared with sequences obtained from proteolytic fragments of the hyaluronate synthase isolated from streptococcal membranes. All peptide sequences (underlined in Figure 4) matched the predicted protein sequence perfectly. This is further proof that the cloned gene encodes the hyaluronate synthase.

Starting at the ATG codon at nucleotides 349–351, there is an open reading frame of 1566 bp, terminating at a TAG codon at nucleotides 1915–1917, which can encode a protein of 522 amino-acid residues, with a molecular mass of 56624 Da. Although the ATG codon at position 424–426 could also be an initiation codon, that at position 349–351 is a more likely candidate, because a possible Shine–Delgarno sequence, GAGG, that is very similar to the ribosome-binding sequences in other Gram-positive bacteria (McLaughlin et al., 1981; Moran et al., 1982) is located 8 bp upstream of the ATG codon.

The transcriptional start site has not been determined. However, putative –35 and –10 promoter sequences, TATAACA and TAAAT, for a Gram-positive promoter with a nearly ideal spacing of 18 bp (Moran et al., 1982) were identified (boxed in Figure 4). The region immediately upstream of the coding region is AT-rich (74% compared with 57% within the coding region), a feature that is characteristic of Gram-positive promoter regions (Moran et al., 1982; Uhlen et al., 1984; Fahnstock et al., 1986).

Figure 5 shows a hydropathy plot of the synthase. Starting at the ATG initiation codon, the first 35 amino-acid residues have the characteristics of a typical bacterial leader sequence (Vlasuk et al., 1983): a stretch of 12 amino-acid residues containing basic residues, followed by a hydrophobic core of 14 amino acids and ending with a hydrophilic region. There was a possible cleavage site for removal of the signal peptide at residue 35 which conformed closely with the –1, –3 rule defined by von Heijne (1983). The remainder of the protein sequence did not contain any significant hydrophobic stretches.

Computer searches have revealed amino-acid-sequence similarity between the hyaluronate synthase and oligopeptide transport proteins from *Bacillus subtilis* (Perego et al., 1991; Rudner et al., 1991), *Streptococcus pneumoniae* (Alloing et al., 1990), *E. coli* (Kashiwagi et al., 1990; Abouhamad et al., 1991) and *Salmonella typhimurium* (Hiles et al., 1987). These similarities

Figure 4 Nucleotide and deduced amino-acid sequence of the hyaluronate synthase

Nucleotides are numbered on the right-hand side of the sequence, and the DNA strand is listed 5' to 3'. The deduced amino-acid sequence is written below the nucleotide sequence; residues that were determined by peptide sequencing are underlined. Putative –35 and –10 promoter sequences (boxed) and the ribosome-binding site (S/D) are noted upstream of the coding region. Palindromic sequences are indicated by dotted lines. The possible cleavage site of the signal sequence is indicated by an arrow.

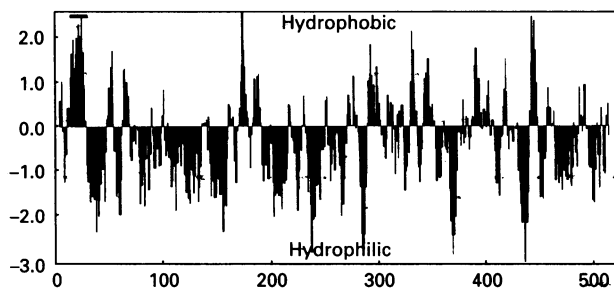


Figure 5 Hydropathy plot of the hyaluronate synthase

The graph is generated using the standard parameters of Kyte and Doolittle (1982) using a window size of 7. The y-axis indicates hydrophilicity (negative) and hydrophobicity (positive). The x-axis denotes the positions of amino-acid residues starting from the initiation codon. The black bar denotes the signal sequence.

(a)

HAS	MTVLGKACKRLGLATVTLASVAALMAC-----PNKQASATDK-KSEINWYTPTEIITL
OPPA	MK-----KRWGIIVTLMILFTLVLSACGGTGSNGEGKDKSGKTTLNINIKTEPFSL
	*. * * * * * * *
HAS	DISKNTDYSALAIAGNSGNLLRADAKGLQPLDAEKVDVSEDLTYTATLRHGLKWSDG
OPPA	HPGLANDSVSGGVIRQTFEGLTRINADGEPEEGMASKIETSKDGTKYTFPIRDGVKWSNG
	*. * * * * * * *
HAS	SDLTAEDFEYSWQRMVDPKTA SEYAVLATESHVKNAEDTNSGKNPDLDSLGVKADGNKVI
OPPA	DPVTAQDFEYAWKALDPNNEQYAYQLY--YIKGAEAAANTGKGS-LDDVAVKAVNDKTL
	*. * * * * * * *
HAS	-FTLTPVAPQFKSLLSFNFVQKESFVKDAGKDYGTTEKQIYSGPIYVKDWNVTSGTF
OPPA	KVELNNTPTPYFTELTAFTYMPINEKIAEKNNKWTNAGDDYVSNQPFKMTAWK-HSGSI
	*. * * * * * * *
HAS	KLVKNKNYDAKMKVKTETVNVQTVKPKDTAVQMYKQKGLDFANIS-GTSAIYNANKHKHD
OPPA	TLEKNDQYWDKDKVKLKKIDHVMINNNTELKFKFAGELDWAGHPLGQLPTESLPTLKKD
	*. * * * * * * *
HAS	VVPVLEATTAVIVYVNTGAI EGLNSLKIQAALMLATDRKGI VSAAVDTGSKPATALVPTG
OPPA	GSLHVEPIAGVYMYKFNTEAKPLDNVNIKALTYSLDRQSI VKNVTQGEQMPAMAAVPT
	*. * * * * * * *
HAS	LAKLSGDTLDEHVPAGY-KYDDKEAALFKE-GLAELG---KDALTTITADADAACA
OPPA	MKGFEDNKE-----GYFKNDVYKTAKEYLEKGLKEMGLSKASDLPKIKLSYNTDDAHA
	*. * * * * * * *
HAS	KSAVDYIKETWETALPGLTVEEKVFPFSORLEDTKQNF EFAVAVLGGDYPGEGSTFYGLF
OPPA	KIA-QAVQEMWKKNL-GVDVLDNSENWYIDKLSQDQVIGRMGLDGFNDPINFLFL
	*. * * * * * * *
HAS	--KGSAYNYGKFTNADYDAANKALT-----TDALNIDAAA-----
OPPA	RDKNNGNNDTG-WENPEFKLLNQSQTETDKTKRAELLKKAEGIFIDEMPVAPIYFYDT
	*. * * * * * * *
HAS	----DDYKAAEKALYDNLVNPVLYLP
OPPA	WVQDENLKGVI MPGTGEVYFRNAYFK
	*. * * * * * * *

(b)

1 (81)	GKLQPLDAEKVDVSEDLTYTATLRHGLKW-----SDGSDLTAEDFEYS
2 (81)	GEPEEGMASKIETSKDGTKYTFPIRDGVKW-----SNGDPVTAQDFEYA
3 (54)	GHPSPGVAEKWE-NKDFKVTVFLHRENKAKW-----SDGTPVTAHQDFVYS
4 (80)	GHPAPGVAESWD-NKDAKVTVFLHRENKAKW-----SDGTPVTAHQDFVYS
5 (76)	GNLPAVAEDWESVDGLTYTYKIRKGVKWFSDGEEYA---EVTAKDFVNG
6 (72)	TEVIPGLAEKWEVSDGKTYTFHLRKGKWK--HDNKEFKPTRELNADDVVS
	*. * * * * * * *

Figure 6 Sequence similarities to hyaluronate synthase

Sequence alignments were carried out using the Clustal software (Higgins and Sharp, 1988). Asterisks (*) represent exact matches and points (.) conservative changes. (a) Alignment of the hyaluronate synthase (HAS) and the OPPA protein of *B. subtilis* (Perego et al., 1991). (b) Sequence similarity between hyaluronate synthase and peptide-binding proteins from a number of different organisms. The position of the first amino-acid residue in each sequence is indicated in parentheses. Sequences: 1, hyaluronate synthase; 2, OPPA proteins from *B. subtilis* (Perego et al., 1991); 3, periplasmic oligopeptide-binding-protein precursor from *S. typhimurium* (Hiles et al., 1987); 4, polyamine-induced protein precursor from *E. coli* (Kashiwagi et al., 1990); 5, AMIA protein from *S. pneumoniae* (Alloing et al., 1990); 6, dipeptide-binding protein from *E. coli* (Abouhamad et al., 1991).

range from 23–30% over the entire sequence, when the synthase was compared with these proteins on an individual basis (Figure 6a). An alignment of all the above protein sequences reveals a stretch with significant amino-acid similarity (Figure 6b). No other significant similarity was found to other protein sequences.

DISCUSSION

In this paper we have reported the cloning and sequencing of the gene encoding the hyaluronate synthase from streptococci. To our knowledge, this is the first report of cloning of this gene from any organism and we propose the designation *has* for the gene locus. The cloning was successfully achieved in λ gt-11, but the gene could not be subcloned in its entirety in plasmid vectors, suggesting that the protein may be toxic for *E. coli*, as has been reported for other membrane-associated proteins (Gregory et al., 1990). The stability observed in λ was probably due to the short life cycle of this phage.

Southern-blotting results suggested that the *has* gene is conserved in hyaluronate-producing streptococci, indicating a uniform mechanism of synthesis. This observation, along with the fact that antibodies raised against the streptococcal synthase cross-react with proteins in the eukaryotic cell membrane (Prehm, 1989; Klewes et al., 1993), may suggest strong conservation of the synthase in all hyaluronate-synthesizing organisms.

The streptococcal synthase is found in protoblast membranes and is therefore expected to be preceded by a signal sequence. Such a signal sequence was, in fact, identified on analysis of the predicted protein sequence (Figure 4). The cleavage of this 3.5 kDa signal peptide would yield a protein of 53.1 kDa. This cleavage would account for the two forms of hyaluronate synthase observed on Western blots of proteins from λ gt-11/2LK lysates and streptococcal protoblast membranes (Figure 1). Unfortunately, attempts to determine N-terminal sequences of these two forms proved unsuccessful. However, sequence determination on peptide fragments of purified mixtures of the two forms yielded homogeneous sequences, thus supporting the Western-blot data which indicate that the proteins are two different forms of the synthase. Previous preparations always contained a majority of the processed form (Prehm and Mausolf, 1986). The processed form is easily shed from protoblast membranes of growing streptococci (Mausolf et al., 1990). The cleavage of the signal sequence could be a key factor in controlling this release. The release is also facilitated by the apparent absence of extended hydrophobic domains (Figure 6) and cell wall domains such as those of the M proteins (Hollingshead et al., 1986; Miller et al., 1988; Mouw et al., 1988; Haanes and Cleary, 1989), protein G (Fahnestock et al., 1986) and Fc-receptor (Heath and Cleary, 1989).

Significant amino-acid-sequence similarity was observed between the synthase and a number of peptide-binding proteins. All of these binding proteins form part of a larger transport complex which contains integral membrane-transport proteins and one or more ATPases (Hyde et al., 1990). Generally, the components of these transport systems are encoded in operonic structures in which the peptide-binding protein is the first gene (Alloing et al., 1990; Rudner et al., 1991; Perego et al., 1991). This may suggest that the hyaluronate synthase forms part of a complex of proteins which synthesizes and exports hyaluronate by a mechanism similar to that involved in peptide transport as has been proposed for other export systems (Blight and Holland, 1990).

The authors thank Mrs. A. Blanke for excellent technical assistance, Dr. Mersmann for his help in data retrieval from the EMBL-database and Professor V. Rizza for providing the strains and Dr. A. Negro for oligonucleotide synthesis. This work was supported by the Deutsche Forschungsgemeinschaft.

REFERENCES

- Abouhamad, W. N., Manson, M., Gibson, M. M. and Higgins, C. F. (1991) Mol. Microbiol. **5**, 1035–1047

- Alloing, G., Trombe, M. C. and Claverys J. P. (1990) *Mol. Microbiol.* **4**, 633–644
- Blight, M. A. and Holland, I. B. (1990) *Mol. Microbiol.* **4**, 873–880
- Both, B., Krupp, G. and Stackebrandt, E. (1991) *Anal. Biochem.* **199**, 216–218
- Chen, E. Y. and Seeburg, P. H. (1985) *DNA* **4**, 165–170
- Fahnestock, S. R., Alexander, P., Nagle, J. and Filpula, D. (1986) *J. Bacteriol.* **167**, 870–880
- Gregory, R. J., Cheng, S. H., Rich, D. P., Marshall, J., Sucharita, P., Hehir, K., Ostegaard, L., Klinger, K. W., Welsh, M. J. and Smith, A. E. (1990) *Nature (London)* **347**, 382–386
- Haanes, E. and Cleary, P. P. (1989) *J. Bacteriol.* **171**, 6397–6408
- Heath, D. G. and Cleary, P. P. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 4741–4745
- Higgins, D. G. and Sharp, P. M. (1988) *Gene* **73**, 237–244
- Hiles, I. D., Gallagher, M. P., Jamieson, D. J. and Higgins, C. F. (1987) *J. Mol. Biol.* **195**, 125–142
- Hirst, G. K. (1941) *J. Exp. Med.* **73**, 493–506
- Hollingshead, S. K., Fishetti, V. A. and Scott, J. R. (1986) *J. Biol. Chem.* **261**, 1677–1686
- Huynh, T. V., Young, R. A. and Davis, R. W. (1986) in *DNA-Cloning: A Practical Approach*, (Glover, D. M., ed.), vol. 1, IRL Press, Oxford
- Hyde, S. C., Emsley, P., Hartshorn, M. S., Mimmack, M. M., Gileadi, M., Pearce, S. R., Gallagher, M. P., Gill, D. R., Hubbard, R. E. and Higgins, C. F. (1990) *Nature (London)* **346**, 362–365
- Kashiwagi, K., Yamaguchi, Y., Sakai, Y., Kobayashi, H. and Igarashi, K. (1990) *J. Biol. Chem.* **265**, 8387–8391
- Klewes, L., Turley, E. A. and Prehm, P. (1993) *Biochem. J.*, in the press
- Kyte, J. and Doolittle, R. F. (1982) *J. Mol. Biol.* **157**, 105–132
- Laurent, T. C. and Fraser, J. R. E. (1992) *FASEB J.* **6**, 2397–2404
- Markovitz, A. and Dorfman, A. (1962) *J. Biol. Chem.* **237**, 273–278
- Markovitz, A., Cifonelli, J. A. and Dorfman, A. (1958) *J. Biol. Chem.* **234**, 2343–2350
- Mausolf, A., Jungmann, J., Robenek, H. and Prehm, P. (1990) *Biochem. J.* **267**, 191–196
- McLaughlin, J. R., Morray, C. L. and Rabinowitz, J. C. (1981) *J. Biol. Chem.* **256**, 11283–11291
- Miller, L., Gray, L., Beachey, E. and Kehoe, M. (1988) *J. Biol. Chem.* **263**, 5668–5673
- Moran, C. P., Lang, N., LeGrice, S. F. L., Lee, G., Stephens, M., Sonenshein, P., Pero, J. and Losick, R. (1982) *Mol. Gen. Genet.* **186**, 339–346
- Mouw, A. R., Beachey, E. H. and Burdett, V. (1988) *J. Bacteriol.* **170**, 676–684
- Nida, K. and Cleary, P. (1983) *J. Bacteriol.* **155**, 1156–1161
- Perego, M., Higgins, C. F., Pearce, S. R., Gallagher, M. P. and Hoch, J. A. (1991) *Mol. Microbiol.* **5**, 173–185
- Prehm, P. (1984) *Biochem. J.* **220**, 597–600
- Prehm, P. (1989) in *CIBA Found. Symp.* **143**, 21–40
- Prehm, P. and Mausolf, A. (1986) *Biochem. J.* **235**, 887–889
- Rosenberg, S. M., Stahl, M. M., Kobayashi, I. and Stahl, F. W. (1985) *Gene* **38**, 165–175
- Rudner, D. Z., LeDeaux, J. R., Ireland, K. and Grossman, A. D. (1991) *J. Bacteriol.* **173**, 1388–1398
- Saiki, R. K., Gelfand, P. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. and Erlich, H. A. (1988) *Science* **239**, 487–491
- Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) *Molecular Cloning*, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor
- Sanger, F., Nicklen, S. and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463–5467
- Seastone, C. V. (1939) *J. Exp. Med.* **70**, 361–378
- Snyder, M., Elledge, S., Sweeter, D., Young, R. A. and Davis, R. W. (1987) *Methods Enzymol.* **154**, 107–128
- Stoolmiller, A. C. and Dorfman, A. (1969) *J. Biol. Chem.* **244**, 236–246
- Sugahara, K., Schwartz, N. B. and Dorfman, A. (1979) *J. Biol. Chem.* **254**, 6252–6261
- Todd, E. W. and Lancefield, R. C. (1928) *J. Exp. Med.* **48**, 751–767
- Uhlen, M., Guss, B., Nilsson, B., Batenback, S., Philipson, L. and Lindberg, M. (1984) *J. Biol. Chem.* **259**, 1695–1702
- Vlasuk, G. P., Inouye, S., Ito, H., Itakura, K. and Inouye, M. (1983) *J. Biol. Chem.* **258**, 7141–7148
- von Heijne, G. (1983) *Eur. J. Biochem.* **133**, 17–21
- Weis J. H. (1989) in *Current Protocols in Molecular Biology* (Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. and Struhl, K., eds.), Wiley & Sons, New York
- Wessel, D. and Flügge, U. I. (1984) *Anal. Biochem.* **138**, 141–143
- Wessels, M. R., Moses, A. E., Goldberg, J. B. and DiCesare, T. J. (1991) *Proc. Natl. Acad. Sci. U.S.A.* **88**, 8317–8321
- Whitnack, E., Bisno, A. L. and Beachey, E. H. (1981) *Infect. Immun.* **31**, 985–991
- Wu, R., Wu, T. and Ray, A. (1987) in *Guide to Molecular Cloning Techniques*, (Berger, S. and Kimmel, A. R., eds.), pp. 343–349, Academic Press, San Diego