Molecular Cloning, Expression, and Characterization of the Authentic Hyaluronan Synthase from Group C Streptococcus equisimilis*

Kshama Kumari and Paul H. Weigel‡
From the Department of Biochemistry and Molecular Biology, University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma 73190

We previously reported the first cloning of a functional glycosaminoglycan synthase, the hyaluronan synthase (HAS) from Group A Streptococcus pyogenes (spHAS) (DeAngelis, P. L., Papaconstantinou, J., and Weigel, P. H. (1993) J. Biol. Chem. 268, 19181–19184). Group A spHAS was unrelated to a putative Group CHA synthase reported by others (Lansing, M., Lellig, S., Mausolf, A., Martini, I., Crescenzini, F., Oregan, M., and Prehm, P. (1993) Biochem. J. 289, 179–184). Here we report the isolation of a bona fide HA synthase gene from a highly encapsulated strain of Group C Streptococcus equisimilis. The encoded protein, designated seHAS, is 417 amino acids long (calculated molecular weight, 47,778; calculated pI, 9.1) and is the smallest member of the HAS family identified thus far. The enzyme migrates anomalously fast in SDS-polyacrylamide gel electrophoresis (∼42,000 Da). The seHAS protein shows no similarity (<2% identity) to the previously reported Group C gene, which is not an HA synthase. The seHAS and spHAS protein and coding sequences are 72 and 70% identical, respectively. seHAS is also similar to eukaryotic HAS1 (∼31% identical), HAS2 (∼28% identical), and HAS3 (28% identical). The deduced protein sequence of seHAS was confirmed by reactivity with a synthetic peptide antibody. Recombinant seHAS expressed in Escherichia coli was recovered in membranes as a major protein (∼10% of the total protein) and synthesized very large HA (M₆ > 7 × 10⁶) in the presence of UDP-GlcNAc and UDP-GlcA. The product contained equimolar amounts of both sugars and was degraded by the specific Streptomyces hyaluronidase. Comparison of the two recombinant streptococcal enzymes in isolated membranes showed that seHAS and spHAS are essentially identical in the steady-state size distribution of HA chains they synthesize, but seHAS has an intrinsic 2-fold faster rate of chain elongation (V_max) than spHAS. seHAS is the most active HA synthase identified thus far; it polymerizes HA at an average rate of 160 monosaccharides/s. The two bacterial HA synthase genes may have arisen from a common ancient gene shared with the early evolving vertebrates.

HA is an important extracellular molecule in vertebrate development, wound healing, inflammation, and tumorigenesis (1–4). It is essentially a ubiquitous, general structural component of extracellular matrices; it has specialized functions in synovial fluid, cartilage, dermis, and the vitreous; and it can be a modulator of eukaryotic cell behavior. In addition, the HA capsule produced by many Group A and Group C Streptococcus strains is an integral part of how these bacteria have developed into mammalian pathogens (5, 6). Loss of the HA capsule typically decreases virulence by >90%. Progress in studying the enzymes responsible for HA biosynthesis has accelerated greatly since the Group A spHAS was reported in 1993 (7, 8). In 1996, several groups isolated cDNAs for putative mouse and human HASs (4) and identified two distinct genes designated HAS1 and HAS2 (9–14). A third gene, HAS3, has recently been found in human and mouse (15).

In 1993, Lansing et al. (16) reported the cloning of a putative streptococcal Group C HAS. The predicted protein was 522 amino acids long and was unexpectedly different than spHAS, which is 419 amino acids long (8). In fact, the two proteins are unrelated (<2% identical) and do not score against one another in a BLAST search. However, the encoded protein identified by Lansing et al. (16) is 23–32% identical to a family of bacterial peptide transporters (17). Although the cloned Group C gene was isolated using an antibody that partially inhibited HA synthesis activity, the encoded protein was not expressed by these investigators, and it was, therefore, never shown to possess the ability to make HA.

Because spHAS shows high similarity (~30% identity) to all of the subsequently identified HAS proteins (4), it seemed unlikely that the reported Group C protein could be a bona fide HA synthase. To clarify this confusion in the literature and to pursue our interest in elucidating details of HA biosynthesis, we undertook an independent effort to clone the Group C HAS. Here we report that the active seHAS is 417 amino acids long (molecular weight 47,778), 72% identical to spHAS, and ~30% identical to the eukaryotic members of the HAS family.

EXPERIMENTAL PROCEDURES

Materials—Oligonucleotide primers were synthesized by Genosys Biotechnologies, Inc. (The Woodlands, TX) and the Great American Gene Co. (Ransom Hill Bioscience, Inc., CA). Cy-5 fluorescent sequencing primers were synthesized by the Molecular Biology Resource Facility, Oklahoma University Health Sciences Center. Restriction and DNA modifying enzymes were from Promega unless noted otherwise. UDP-

* This research was supported by National Institutes of Health Grant GM35978. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank®/EBI Data Bank with accession number(s) AF023876.

‡ To whom correspondence should be addressed. Tel.: 405-271-1288; Fax: 405-271-3092; E-mail: paul-weigel@ouhsc.edu.

1 The abbreviations used are: HA, hyaluronan (hyaluronic acid); HAS, hyaluronan synthase; UDP-GlcA, UDP-β-D-glucuronic acid; UDP-GlcNAc, UDP-N-acetyl-D-glucosamine; spHAS, S. pyogenes HAS; seHAS, S. equisimilis HAS; muHAS, mouse HAS; ORF, open reading frame; PAGE, polyacrylamide gel electrophoresis; bp, base pair(s); kb, kilobase(s); PCR, polymerase chain reaction.
Corresponds to the region 259DRCLTNYAIDL (spHAS). Oligonucleotide The digested DNA was precipitated with ethanol, washed, and ligated (18), was partially digested with Sau3A1 to an average size of 2–12 kb.

FIG. 1. Nucleotide and deduced protein sequence of the S. equisimilis HAS ORF. The DNA sequence (top row) of the ORF encoding sehAS was cloned from strain D181 genomic DNA, as assembled (23) and described under Experimental Procedures. The encoded amino acid sequence is shown on the bottom row. Underlined regions denote transmembrane or membrane-associated domains predicted by the TMpred program (Swiss Institute for Experimental Cancer Research). Cys residues are circled, and possible HA binding motifs (B-X-B; where B is a basic residue; Ref. 42) are shaded. The sequence has been submitted to GenBank™ under accession number AF023876.

GlcNac and UDP-GlcA were from Fluka, and all other chemicals were of the highest grade available, from Sigma, unless otherwise noted. UDP-[3H]GlcNAc (29.2 Ci/mmol) was from NEN (Boston, MA) containing adsorbed phage were prehybridized at 60 °C and hybridized with the 5'-end-labeled oligonucleotide D181.5 in QuikHyb hybridization solution (Stratagene) at 80 °C. The membranes were washed with 2× SSC buffer and 0.1% (w/v) SDS at room temperature for 15 min, washed with 0.1× SSC buffer and 0.1% (w/v) SDS at 60 °C for 30 min, dried, and then exposed to Bio-Max MS film overnight at −70 °C. Positive plaques were replated and rescreened twice. Pure positive plaques were saved in 0.1 M NaCl, 0.4 M MgSO4, 20 μl Tris-HCl, pH 7.5, and 0.1% (w/v) gelatin with chloroform. PCR on these plagues using vector primers revealed three different insert sizes. PCR with a combination of vector primers and primers from different regions of the ORF, or with undigested genomic DNA as a template (23) was performed from Experimental Procedures. The 1.2-kb PCR product from D181 genomic DNA, as well as from the 1.2-kb PCR product from genomic DNA amplified D181 genomic DNA gave PCR products of the expected size. The four PCR products were cloned and sequenced using the same strategy as above. For each PCR product, sequences obtained from different clones were compared to derive a consensus sequence. Thus, we obtained a 1042-bp sequence with a continuous ORF with high homology to sehHAS.

Library Screening—Two molecular tools were used to screen the library, the cloned 459-bp PCR product and the antisense oligonucleotide D181.5 (5'-GCTGTAGATTGCACCAGTTAGC-3') derived from the 1042-bp sequence. The 459-bp PCR product was radiolabeled using the Prime-It II random primer labeling kit (Stratagene) according to the manufacturer's instructions. Oligonucleotides were immobilized on the Kinace-It kinasing kit (Stratagene) using [γ-32P]ATP. Radiolabeled oligonucleotides were isolated from unlabeled material on NucTrap Push columns (Stratagene). The 459-bp probe gave a high level of nonspecific hybridization with the λ phage vector. The oligoprobe, however, hybridized specifically with a D181 genomic digest on Southern blots. To screen the phage library, XL1-Blue MRF™ was used as a host (3000 plaques/plate). Nitrocellulose membranes (0.45 μm; Micron Separations, Inc., Westboro, MA) containing adsorbed phage were prehybridized at 60 °C and hybridized with the 5'-end-labeled oligonucleotide D181.5 in QuikHyb hybridization solution (Stratagene) at 80 °C. The membranes were washed with 2× SSC buffer and 0.1% (w/v) SDS at room temperature for 15 min, washed with 1× SSC buffer and 0.1% (w/v) SDS at 60 °C for 30 min, dried, and then exposed to Bio-Max MS film overnight at −70 °C. Positive plaques were replated and rescreened twice. Pure positive plaques were saved in 0.1 M NaCl, 14 mM MgSO4, 20 μl Tris-HCl, pH 7.5, and 0.1% (w/v) gelatin with chloroform. PCR on these plagues using vector primers revealed three different insert sizes. PCR with a combination of vector primers and primers from different regions of the ORF, or with undigested genomic DNA as a template (23) was performed from Experimental Procedures. The 1.2-kb PCR product from D181 genomic DNA, as well as from the 1.2-kb PCR product from genomic DNA amplified D181 genomic DNA gave PCR products of the expected size. The four PCR products were cloned and sequenced using the same strategy as above. For each PCR product, sequences obtained from different clones were compared to derive a consensus sequence. Thus, we obtained a 1042-bp sequence with a continuous ORF with high homology to sehHAS.

Expression Cloning of the Putative sehHAS—PCR primers were designed at the start and stop codon regions of sehHAS to contain an EcoRI restriction site in the sense oligonucleotide (5'-AGGATCCTGAGGTTTCTGGATCC-3') and a MluI site in the antisense oligonucleotide (5'-AGAATTCTCGAGTTATATTTTTTTATTTTATTCTTTG-3') and a EcoRI site in the antisense oligonucleotide (5'-AGAATTCTGAGGTTTCTGGATCC-3'). Using Pfu DNA polymerase (Stratagene), these primers amplified a 1.2-kb PCR product from D181 genomic DNA, as well as from the purified hybridization-positive phage. The 1.2-kb product from genomic DNA was purified by agarose gel electrophoresis, digested with PstI and EcoRI, and cloned directionally into PstI- and EcoRI-digested pKK223 vector. The ligated vector was transformed into E. coli SURE cells that were grown at 30 °C. Colonies were isolated, and their plasmid DNA was purified. Of six colonies, five had the correct size insert, and one had no insert. All five clones were verified to have sehHAS activity, and the inserts were sequenced.

HA Synthase Activity—HA synthase activity was assayed in membranes prepared from fresh log phase cells. Cells were harvested at 3000 × g and washed at 4 °C with phosphate-buffered saline, and membranes were isolated by modification of the protoplast method of Ito et al. (21) as described earlier (22). Membranes from Streptococcus pyogenes and S. equisimilis were obtained by a modification (22) of the protoplast procedure of Van de Rijn and Drake (23). Membranes were incubated at 37 °C in 50 mM sodium and potassium phosphate, pH 7.0,
with 20 mM MgCl₂, 1 mM dithioerythritol, 120 μM UDP-GlcA, and 300 μM UDP-GlcNAc. Incorporation of sugar was monitored by using UDP-[14C]GlcA and/or UDP-[3H]GlcNAc. Reactions were terminated by addition of SDS to a final concentration of 2% (w/v) and heating at 90 °C for 1 min. Product HA was separated from precursors by descending paper chromatography and measured by determining incorporated radioactivity remaining at the origin (22, 24).

**Gel Filtration Analysis—** Radiolabeled HA produced in vitro by membranes containing recombinant seHAS or spHAS was analyzed by gel filtration chromatography on a column (1.0 × 50 cm) of Sephacryl S500HR or S400HR (Pharmacia). Samples (0.2 ml) were eluted with 5 mM Tris, pH 8.0, 0.2 M NaCl, and 1.0 ml fractions were collected and assessed for 14C and/or 3H radioactivity. Authenticity of the HA polysaccharide was assessed by treatment of product with the HA-specific hyaluronate lyase of *Streptomyces hyalurolyticus* (EC 4.2.2.1; Ref. 18) at 37 °C for 3 h. The digest was then subjected to gel filtration.

**SDS-PAGE and Western Blotting—** SDS-PAGE was performed according to the method of Laemmli (25). Electrotransfers to nitrocellulose (0.1 μm; Schleicher & Schuell) were performed in 20% methanol using a Bio-Rad mini Transblot device. The blots were blocked with 2% bovine serum albumin in TBS. Protein A/G alkaline phosphatase conjugate (Pierce) and p-nitroblue tetrazolium/5-bromo-4-chloro-3-indolyl p-toluidine salt (Bio-Rad) were used for detection.

**DNA Sequence and Analysis—** Plasmids were sequenced on both strands by the Oklahoma University Health Sciences Center Sequencing Facility using fluorescent vector primers. Sequencing reactions were performed using a Thermosequenase kit (Amersham Corp.) for fluorescent-labeled primers (with 7-deazaG). Samples were electrophoresed on a Pharmacia ALFExpress DNA sequencer, and data were analyzed by the ALF Manager program, version 3.02. Internal regions of inserts were sequenced with internal primers using the ABI Prism 377 program, version 2.1.1. Ambiguous regions were sequenced manually using Sequenase™ 7-deaza-DNA polymerase (United States Biochemical) and α-35S-labeled dATP (Amersham Life Sciences). The se-
sequences obtained were compiled and analyzed using DNASIS, version 2.1 (Hitachi Software Engineering Co., Ltd.). The nucleotide and amino acid sequences were compared with other sequences in GenBank™ and other data bases.

General—Radioactivity was determined using Ultima Gold scintillation fluid (Packard) in a Packard Tri-Carb model 2300TR liquid scintillation spectrometer. Protein was quantitated by the method of Bradford (26) using bovine serum albumin as a standard.

RESULTS

Initial attempts to clone the seHAS gene by screening Group C genomic DNA or a library for the ability to hybridize with nucleic acid probes based on the complete spHAS coding sequence (8) were unsuccessful. In retrospect, this result may be due to the fact that the longest identical region between the two genes is only 20 bp. There are numerous mismatches throughout the two coding regions that might have compromised efficient hybridization with the full-length probe. We then turned to a PCR approach using oligonucleotide primers based on several regions of high identity among spHAS, DG42 (now known to be a developmentally regulated X. laevis HAS and designated xlHAS; Ref. 4), and NodC (a Rhizobium β-GlcNAc transferase; Ref. 20). The xlHAS and NodC proteins are 50 and 10% identical to spHAS, respectively. This strategy yielded a 459-bp PCR product, the sequence of which was 66.4% identical to spHAS, indicating that a Group C homologue of the Group A HA synthase gene had been identified. The complete coding region of the gene was then reconstructed using a similar PCR-based strategy as outlined under “Experimental Procedures” (Fig. 1). A final set of PCR primers was then used to amplify the complete ORF from genomic DNA. When this 1.2-kb PCR fragment was incorporated into the expression vector pKK223 and transformed into E. coli SURE cells, HA synthase activity was demonstrated in isolated membranes from 5 of the 5 colonies tested.

The ORF of the reconstructed gene encodes a novel predicted protein of 417 amino acids that was not in the data base (Fig. 1); it is 2 amino acids shorter than spHAS (4, 8). The two bacterial proteins are 72% identical (Fig. 2), and the nucleic acid sequences are 70% identical. The predicted molecular weight of the seHAS protein is 47,778, and the predicted pI is at pH 9.1. The three recently identified mammalian HASs are clearly related to the bacterial proteins (Fig. 3). The overall

The ORF of the reconstructed gene encodes a novel predicted protein of 417 amino acids that was not in the data base (Fig. 1); it is 2 amino acids shorter than spHAS (4, 8). The two bacterial proteins are 72% identical (Fig. 2), and the nucleic acid sequences are 70% identical. The predicted molecular weight of the seHAS protein is 47,778, and the predicted pI is at pH 9.1. The three recently identified mammalian HASs are clearly related to the bacterial proteins (Fig. 3). The overall

FIG. 4. Hydropathy plots for seHAS and predicted membrane topology. Hydropathicity plots for the two streptococcal HASs were generated by the method of Kyte and Doolittle (46) using DNASIS (A). The proposed topology for the two proteins (B) conforms to the charge-in rule and puts the large central domain inside. This domain is likely to contain most of the substrate binding and catalytic functions of the enzymes. Cys226, which is conserved in all HAS family members, is shown in the central domain (circled C).

FIG. 5. Overexpression of recombinant seHAS and spHAS in E. coli. Membrane proteins (5 µg/lane) were fractionated by SDS-PAGE using a 10% (w/v) gel under reducing conditions. The gel was stained with Coomassie Blue R-250, photographed, scanned, and quantitated using a Molecular Dynamics Personal Densitometer (model PDS1P60). The position of HA synthase is marked by the arrow. Lane A, native spHAS (Group A); Lane C, native seHAS (Group C); Lane E, recombinant seHAS; Lane P, recombinant spHAS; Lane V, vector alone. Molecular weights of standards (Bio-Rad, low Mr) are shown in thousands.
Lys405 in spHAS). We have proposed (4) that predicted MD5 is predicted for authentic HA. Membranes from cells transformed with fore, the molar ratios of the sugars incorporated into product is 1:1, as into the total amount of HA fractionated on the column is 1.4, which is V

bers are more divergent (Fig. 4). The enzyme reaction was stopped by addition of EDTA to a final concentration of 25 mM. Half of the reaction mix was treated with Streptomyces hyaluronidase at 37 °C for 3 h. SDS 2% (w/v) was added to hyaluronidase-treated (C) and untreated (A and B) samples, which were heated at 90 °C for 1 min. The samples were diluted to 500 μl with column buffer (5 mM Tris, 0.2 M NaCl, pH 8.0) and clarified by centrifugation, and 200 μl was injected onto a Sephacryl S500HR column. Fractions (1 ml) were collected and radioactivity was determined. BD is the peak elution position of blue dextran ( –2 × 10⁶ Da; Pharmacia). Vc marks the excluded volume and Vv the included volume. The ratio of [14C]GlcA/[^3]H]GlcNAc incorporated into the total amount of HA fractionated on the column is 1.4, which is identical to the ratio of specific activities of the two substrates. Therefore, the molar ratios of the sugars incorporated into product is 1:1, as predicted for authentic HA. Membranes from cells transformed with vector alone did not synthesize HA (not shown).

identity between the two groups is ~28–31%, and in addition, many amino acids in seHAS are highly conserved with those of the eukaryotic HASs (e.g. Lys/Arg or Asp/Glu substitutions). Within mammalian species, the same family members are almost completely identical (e.g. muHAS1 and huHAS1 are 95% identical; muHAS2 and huHAS2 are 98% identical). However, even within the same species, the different HAS family members are more divergent (e.g. muHAS1 and muHAS2 are 53% identical; muHAS1 and muHAS3 are 57% identical; muHAS2 and muHAS3 are 71% identical).

The overall membrane topology predicted for seHAS was identical to that for spHAS (Fig. 4) and the eukaryotic HASs reported thus far (4). The protein has two putative transmembrane domains at the amino terminus and 2–3 membrane-associated or transmembrane domains at the carboxyl end (Fig. 4B). The hydropathy plots for the two streptococcal enzymes are virtually identical (Fig. 4A), and they illustrate the difficulty in predicting the topology of the extremely hydrophobic region of ~90 residues at Lys313–Arg406 in seHAS (Lys313–Lys405 in spHAS). We have proposed (4) that predicted MD5 is not a transmembrane domain but, rather, MD5 is an amphiphatic helix or a reentrant loop that remains on the cytoplasmic side of the membrane.

seHAS was efficiently expressed in E. coli cells. Approximately 10% of the total membrane protein was seHAS as assessed by densitometric scanning (Fig. 5). The prominent seHAS band at 42 kDa is quantitatively missing in the vector-only control (lane E versus lane V). This unusually high level of expression for a membrane protein was also found for spHAS, using the same vector in E. coli SURE cells (Fig. 5, lane P). About 8% of the membrane protein in this case is spHAS. In contrast, the amount of seHAS in Group C membranes (Fig. 5, lane C) is not more than 1% of the total membrane protein. The spHAS in Group A membranes (Fig. 5, lane A) is barely detectable.

The recombinant seHAS expressed in E. coli SURE cells does not synthesize HA in vivo because these cells lack UDP-GlcA, one of the required substrates (not shown). Membranes containing the recombinant seHAS protein, however, synthesize HA when provided with the substrates UDP-GlcNAc and UDP-GlcA (Table I). Using 120 μM UDP-GlcA and 300 μM UDP-GlcNAc, HA synthesis was linear with membrane protein (at least 1 h (not shown). Also, membranes prepared from nontransformed cells or cells transformed with vector alone have no detectable HAS activity. HA synthesis is negligible (Table I) when Mg²⁺ is chelated with EDTA or when either of the two substrates is omitted (1–3% of control). Recombinant seHAS also showed the expected specificity for sugar nucleotide substrates, being unable to copolymerize either UDP-GalA, UDP-Glc, or UDP-GalNAc with either of the two normal substrates (Table I).

Based on gel filtration analysis, the average mass of the HA synthesized by seHAS in isolated membranes is 5–10 × 10⁶ Da (Fig. 6). The distribution of HA sizes resolved in this experiment ranged from 2 to 20 × 10⁶. The product of the recombinant seHAS is judged to be authentic HA based on the equimolar incorporation of both sugars (1:1 ratio) and its sensitivity to degradation by the specific Streptomyces hyaluronidase (Fig. 6). Although the conditions for continuous total HA synthesis were not optimal in this case (because ~90% of one substrate was incorporated into product), the enzyme produced a broad
Comparison of the Streptococcal HA Synthases

**Fig. 8.** Kinetic analysis of the HA size distributions produced by seHAS and spHAS. E. coli membranes containing equal amounts of seHAS or spHAS protein were incubated at 37 °C with 1.35 mM UDP-[14C]GlcA, 1.3 × 10^6 dpm/mol, and 3.0 mM UDP-GlcNAc as described under “Experimental Procedures.” These substrate concentrations are greater than 15 times the respective K_m values. Samples taken at 0.5, 1.0, and 60 min were quenched with SDS and chromatographed over Sephacryl S400HR as described under “Experimental Procedures.” The positions of blue dextran (BD, ~2 × 10^6 Da), the void volume (V_o), and the included volume (V_i) are shown. The size distributions of HA synthesized by seHAS (○) and spHAS (■) at 0.5 (●) and 1.0 min (●) and 1.0 min (●) and (●) are shown. B, the HA profiles in the fractionating range of the column (fractions 12–24) are normalized to the percentage of total HA in each fraction. The values above the arrows are the molecular weights (in millions) of HA determined directly in a separate experiment using a Dawn multangle laser light-scattering instrument (Wyatt Technology Corp.). SeHAS (●, ■, △); spHAS (○, □, ○); 0.5 min (○, ●); 1.0 min (●, □, △); 60 min (△, △).

Distribution of HA chain lengths. The peak fraction corresponds to an HA mass of 7.5 × 10^6 Da, which is a polymer containing approximately 36,000 monosaccharides.

The deduced protein sequence of seHAS was confirmed by the ability of antibodies to the spHAS protein to cross-react with the Group C protein (Fig. 7). Polyclonal antibodies to the whole spHAS protein or to just the central domain of spHAS also reacted very well with the seHAS protein. (Figs. 7, lanes 7–10). Antipeptide antibody to the carboxyl terminus of spHAS did not cross-react with this somewhat divergent region in the seHAS protein (Fig. 7, lanes 5 and 6). However, antipeptide antibody directed against the spHAS sequence Glu^{147}-Thr^{161} recognized the same predicted sequence in seHAS very efficiently (Fig. 7, lanes 2 and 3). This antipeptide antibody also reacts with the wild-type seHAS and spHAS proteins in streptococcal membranes and confirms that the native and wild-type enzymes from both species are of identical size (Fig. 7, lanes 1 and 4). Like the spHAS protein (7, 8), seHAS migrates anomalously fast during SDS-PAGE (Figs. 5 and 7). Although the calculated weight is 47,778, the M_r by SDS-PAGE is consistently ~42,000. Because of the sequence identity within their central domain regions and the overall identical structure predicted for the two bacterial enzymes (Fig. 4), the peptide-specific antibody against the region Glu^{147}-Thr^{161} can be used to normalize for HAS protein expression in membranes prepared from cells transformed with genes for the two different enzymes. Using this approach, membranes with essentially identical amounts of recombinant spHAS or seHAS were compared with respect to the initial rate of HA synthesis and the distribution of HA product sizes (Fig. 8).

As shown for spHAS (8), the synthesis of HA chains by seHAS is processive. The enzyme appears to stay associated with high probability to a growing HA chain until it is released as a final product. Therefore, it is possible to compare the rates of HA elongation by seHAS and spHAS by monitoring the size distribution of HA chains produced at very early times, during the first round of HA chain synthesis. At 0.5 and 1.0 min, reflecting elongation of the first HA chains, the size distribution of HA produced by seHAS was significantly shifted to larger species compared with spHAS (Fig. 8, A and B). During this interval, total HA synthesis was linear (Fig. 8A). seHAS synthesized 2.4 and 1.9 times as much HA as spHAS at 0.5 and 1.0 min, respectively (Fig. 8A). The peak fractions at 0.5 min correspond to molecular weights of 1.1 × 10^6 and 0.6 × 10^6 for the HA produced by seHAS and spHAS (Fig. 8B). The seHAS peak at 1.0 min corresponds to molecular weight ~2 × 10^6, whereas the spHAS value was lower (molecular weight ~1.1 × 10^6). At 1 min, about 22% of the HA molecules made by seHAS were already larger than ~2.8 × 10^6, whereas only 2% of HA made by spHAS was in this size range. By 60 min, however, the two distributions of HA sizes were essentially the same (Fig. 8B), with peaks at molecular weight ~7.5 × 10^6.

The seHAS enzyme, therefore, has an inherently faster V_max than spHAS in its ability to polymerize a growing HA chain. Nonetheless, the final distributions of HA sizes made by the two bacterial enzymes at steady-state are virtually identical. Based on gel filtration analysis of HA product sizes at various times, we estimate that the average rate of elongation by seHAS is ~9,000 monosaccharides/min at 37 °C, which is about 100% faster than that for spHAS. In ~5 min, the seHAS enzyme could polymerize HA chains of ~10 × 10^6.

**DISCUSSION**

The results reported here demonstrate that the cloned seHAS is an authentic HA synthase and that the 52-kDa protein
Comparison of the Streptococcal HA Synthases

Conserved in seHAS (Cys226, Cys262, Cys281, and Cys367). Only family (4). Four of the six Cys residues in spHAS are reported by Lansing et al. (16) is not the Group C HAS. Claims that the latter protein was an HA synthase or a protein needed for HA synthesis has led to confusion and controversy over the last few years (4) because an antibody to this uncharacterized protein has been widely distributed and used to study putative eukaryotic HAS proteins in a variety of systems (27–32). Because these studies were, in fact, not monitoring HAS protein levels, they unfortunately contain erroneous conclusions and must be reexamined. Unlike the Group C 52-kDa peptide transporter (16), the seHAS protein is highly homologous to the 7 reported HA synthases that now make up this rapidly growing family (4). Four of the six Cys residues in spHAS are conserved in seHAS (Cys226, Cys262, Cys281, and Cys367), Only Cys226 in seHAS (Cys225 in spHAS) is conserved in all members of the HAS family (Fig. 4B; Ref. 4). Because sulfhydryl-reactive agents, such as p-mercuribenzoate (39) or N-ethylmaleimide, greatly inhibit HAS activity, it is likely that this conserved Cys is necessary or important for enzyme activity. Initial results from site-directed mutagenesis studies indicate that a C225S mutant of spHAS has only 5–10% of wild-type activity (40).

In mammals, three genes, designated has1, has2, and has3, have been identified (9–15) and mapped to three different chromosomes in both human and mouse (33). In amphibians, the only HAS identified thus far is the developmentally regulated DG42, which was cloned in 1988 (34) and was recently shown to encode HA synthase activity by analysis of the recombinant protein in yeast membranes (19). Probably other X. laevis HAS genes will soon be identified. A plausible evolutionary relationship among the eight known HAS family members is shown in Fig. 3. One interpretation of this divergent evolution model is that a primitive bacterial HAS precursor may have been usurped early during vertebrate development. Alternatively, the bacterial pathogenic strategy of making an HA capsule could have begun when a primitive bacterium captured a primordial HAS. In any case, convergent evolution of the streptococcal and eukaryotic HAS enzymes to a common structural solution seems unlikely.

The result reported here showing different HA elongation rates for the two bacterial HA synthases, which are 70% identical, is the first comparative analysis within the HAS family (4) and will be important in guiding future studies to understand the structural basis for polymer assembly. Although more detailed kinetic analyses are required to verify that the main difference between the two enzymes is in $V_{\text{max}}$, the present studies were all performed with substrate concentrations at least 15-fold above the $K_m$ values reported for HAS enzymes. The three mammalian isozymes for HAS have not yet been characterized enzymatically or with respect to their HA product size. The finding that HAS elongation rates can differ also has important physiological implications for this family, because the differences among HAS1, HAS2, and HAS3 amino acid sequences are as great as or greater than the differences between seHAS and spHAS. It will be informative and probably biologically relevant if there are similar differences among the eukaryotic enzymes with respect to rates of HA synthesis or HA size distribution.

HA synthesis occurs at the plasma membrane (35–37), and the HA is either shed into the medium or remains cell-associated to form the bacterial capsule or a eukaryotic pericellular coat (38). The sugar nucleotide substrates in the cytoplasm are utilized to assemble HA chains that are extruded through the membrane to the external space. It may be relevant to note that the Group C HA capsule is generally much larger than the Group A capsule, which could be a direct consequence of the kinetic difference between the two HA synthases. All HAS proteins are predicted to be membrane proteins with a similar topology (4).

Knowledge of the protein topology is the very hydrophobic carboxyl portion of the HAS proteins will likely be critical for understanding how these enzymes extend the growing HA chain as it is simultaneously extruded through the membrane. Our present model for the predicted topology of the streptococcal HAS proteins contains a reentrant loop domain or an amphipathic helical domain that is membrane-associated but does not span the membrane (Fig. 4B, MD5). Reentrant loop membrane domains have been proposed in channel proteins, such as the glutamate receptor (41). Such domains may help form or modulate a channel. HAS proteins presumably also contain a channel-like region to sequester and translocate the growing HA polymer across the plasma membrane. The presence of a reentrant loop domain could facilitate this HA chain translocation in HAS and could increase the size of the channel that this small protein, with only five membrane domains, could form. The unprecedented enzymatic activity of this enzyme family may require unusual and complex interactions of the protein with the lipid bilayer.

Preliminary results based on analysis of spHAS-alkaline phosphatase fusion proteins indicate that the amino and carboxyl termini and the large central domain are all intracellular (41). The seHAS protein is also predicted to contain a large central domain (−63% of the total protein) that likely contains the two substrate binding sites and the two glycosyltransferase activities needed for HA synthesis (8). Although current programs cannot reliably predict the number or nature of membrane-associated domains within the long carboxyl-terminal hydrophobic stretch, the proposed topological arrangement (Fig. 4B) agrees with the present evidence and applies as well to the eukaryotic enzymes, which are 40% larger, primarily due to extension of the carboxyl-terminal end of the protein with two additional predicted transmembrane domains (4).

Acknowledgments—We thank Coy Heldermon and Dr. Valerie Tlapak-Simmons for helpful discussions, Debbie Blevins for help preparing the manuscript, and Dr. Paul DeAngelis for use of the S-500HR column. We especially thank Dr. DeAngelis for excellent input and help throughout this study and for construction of the spHAS construct in pKK223.

REFERENCES

1. Toole, B. P. (1990) Curr. Opin. Cell Biol. 2, 839–844
2. Laurent, T. C., and Fraser, J. R. E. (1992) FEBS Letts. 6, 2397–2404
3. Lesley, J., and Kincade, P. R. (1993) Adv. Immunol. 54, 271–325
4. Weigel, P. H., Hascall, V. C., and Tammi, M. (1997) J. Biol. Chem. 272, 13997–14000
5. Husmann, L. K., Yung, D.-L., Hollingshead, S. K., and Scott, J. R. (1997) Infect. Immun. 65, 1422–1430
6. Schmidt, K.-H., Gunther, E., and Courtney, H. S. (1996) Med. Microbiol. Immunol. 184, 169–173
7. DeAngelis, P. L., Papacostantinou, J., and Weigel, P. H. (1993) J. Biol. Chem. 268, 14568–14571
8. DeAngelis, P. L., Papacostantinou, J., and Weigel, P. H. (1993) J. Biol. Chem. 268, 19181–19184
9. Itano, N., and Kimata, K. (1996) J. Biol. Chem. 271, 9875–9878
10. Shyjan, A. M., Heldin, P., Butcher, E. C., Yoshino, T., and Briskin, M. J. (1996) J. Biol. Chem. 271, 23385–23399
11. Itano, N., and Kimata, K. (1996) Biochem. Biophy. Res. Commun. 222, 816–820
12. Spicer, A. P., Augustine, M. L., and McDonald, J. A. (1996) J. Biol. Chem. 271, 23400–23406
13. Watanabe, K., and Yamaguchi, Y. (1996) J. Biol. Chem. 271, 22945–22948
14. Fulop, C., Salustri, A., and Hascall, V. C. (1997) Arch. Biochem. Biophys. 337, 261–265
15. Spicer, A. P., Olson, J. S., and McDonald, J. A. (1997) J. Biol. Chem. 272, 8957–8961
16. Lansing, M., Lellig, S., Mausolf, A., Martini, I., Crescenzi, F., Oregan, M., and Prehm, P. (1993) Biochem. J. 298, 179–184
17. Alonso, G., Trombe, M.-C., and Claverie, J.-P. (1990) Mol. Microbiol. 4, 633–644
18. Caparon, M. G., and Scott, J. R. (1991) Methods Enzymol. 204, 556–586
19. DeAngelis, P. L., and Achyuthan, A. M. (1996) J. Biol. Chem. 271, 23657–23660
20. Nagahashi, S., Sudo, M., Ono, N., Sawada, R., Yamaguchi, E., Uchida, Y., Mio, T., Takagi, M., Ariyawa, S., and Yamada-Okaibe, H. (1995) J. Biol.
Comparison of the Streptococcal HA Synthases

21. Ito, K., Sato, T., and Yura, T. (1977) Cell 11, 551–559
22. DeAngelis, P. L., and Weigel, P. H. (1994) Biochemistry 33, 9033–9039
23. van de Rijn, I., and Drake, R. R. (1992) J. Biol. Chem. 267, 24392–24306
24. Triscott, M. X., and van de Rijn, I. (1986) J. Biol. Chem. 261, 6004–6009
25. Laemmli, U. K. (1970) Nature 227, 680–685
26. Bradford, M. M. (1976) Anal. Biochem. 72, 248–254
27. Rittig, M., Lütjen-Drecoll, E., and Prehm, P. (1992) Exp. Eye Res. 54, 455–460
28. Klewes, L., Turley, E. A., and Prehm, P. (1993) Biochem. J. 290, 791–795
29. Rittig, M., Flugel, C., Prehm, P., and Lütjen-Drecoll, E. (1993) Graefe’s Arch. Clin. Exp. Ophthalmol. 231, 313–317
30. Klewes, L., and Prehm, P. (1994) J. Cell. Physiol. 160, 539–544
31. Pitsillides, A. A., Archer, C. W., Prehm, P., Baylis, M. T., and Edwards, J. C. (1995) J. Histoch. Cytochem. 43: 263–273
32. Prehm, S., Herrington, C., Nickel, V., Volker, W., Briko, N. I., Blinnikova, E. I., Schmiedel, A., and Prehm, P. (1995) J. Anat. 187, 271–277
33. Spicer, A. P., Seldin, M. F., Olsen, A. S., Brown, N., Wells, D. E., Doggett, N. A., Itano, N., Kimata, K., Inazawa, J., and McDonald, J. A. (1997) Genomics 41, 493–497
34. Rosa, F., Sargent, T. D., Rebbert, M. L., Michaels, G. S., Jamrich, M., Grunz, H., Jonas, E., Winkles, J. A., and Dawid, I. B. (1988) Dev. Biol. 129, 114–123
35. Markovitz, A., Cifonelli, J. A., and Dorfman, A. (1959) J. Biol. Chem. 234, 2343–2350
36. Philipson, L. H., and Schwartz, N. B. (1984) J. Biol. Chem. 259, 5017–5023
37. Prehm, P. (1984) Biochem. J. 220, 597–600
38. Knudsen, C. B., and Knudsen, W. (1993) FASEB J. 7, 1233–1241
39. Sugahara, K., Schwartz, N. B., and Dorfman, A. (1979) J. Biol. Chem. 254, 6252–6261
40. Heldermon, C., Tlapak-Simmons, V., DeAngelis, P., and Weigel, P. (1996) Mol. Biol. Cell 7, 55 (abstr.)
41. Bennett, J. A., and Dingledine, R. (1995) Neuron 14, 373–384
42. Yang, B., Yang, B. L., Savani, R. C., Turley, E. A. (1994) EMBO J. 13, 286–296
43. Higgins, D. G., and Sharp, P. M. (1988) Gene 73, 237–244
44. Sneath, P. H. A., and Sokal, R. R. (1973) Numerical Taxonomy, W. H. Freeman & Co., San Francisco
45. Kyte, J., and Doolittle, R. F. (1982) J. Mol. Biol. 157, 105–132