Topological Organization of the Hyaluronan Synthase from Streptococcus pyogenes*

Coy Heldermon, Paul L. DeAngelis, and Paul H. Weigel‡

From the Department of Biochemistry & Molecular Biology, University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma 73190

Since we first reported (DeAngelis, P. L., Papaconstantinou, J., and Weigel, P. H. (1993) J. Biol. Chem. 268, 19181–19184) the cloning of the hyaluronan (HA) synthase from Streptococcus pyogenes (spHAS), numerous membrane-bound HA synthases have been discovered in both prokaryotes and eukaryotes. The HA synthases are unique among enzymes studied to date because they mediate 6–7 discrete functions in order to assemble a polysaccharide containing hetero-disaccharide units and simultaneously effect translocation of the growing HA chain through the plasma membrane. To understand how the relatively small spHAS performs these various functions, we investigated the topological organization of the protein utilizing fusion analysis with two reporter enzymes, alkaline phosphatase and β-galactosidase, as well as several other approaches. From these studies, we conclude that the NH₂ terminus and the COOH terminus, as well as the major portion of a large central domain are localized intracellularly. The first two predicted membrane domains were confirmed to be transmembrane domains and give rise to a very small extracellular loop that is inaccessible to proteases. Several regions of the large internal central domain appear to be associated with, but do not traverse, the membrane. Following the central domain, there are two additional transmembrane domains connected by a second small extracellular loop that also is inaccessible to proteases. The COOH-terminal ~25% of spHAS also contains a membrane domain that does not traverse the membrane and may contain extensive re-entrant loops or amphipathic helices. Numerous membrane associations of this latter COOH-terminal region and the central domain may be required to create a pore-like structure through which a growing HA chain can be extruded to the cell exterior. Based on the high degree of similarity among Class I HAS family members, these enzymes may have a similar topological organization for their spHAS-related domains.

HA³ is a linear polysaccharide that contains repeating disaccharide units of GlcUAβ1,3GlcNAcβ1,4 and is widely distributed throughout the animal kingdom (1). In eukaryotes, HA serves a viscoelastic structural role and functions, together with proteoglycans, as part of the extracellular matrix scaffold that creates tissue integrity (2, 3). HA also mediates the localization and movement of cells (3–5), and can alter or regulate cell behavior by acting as a signaling molecule (6, 7). HA in prokaryotes is found as a capsule that serves as a virulence factor, presumably by mimicking the eukaryotic host and thus being immuno-protective (8–10). The enzyme responsible for HA synthesis, HA synthase, was cloned originally from Group A streptococcus (11). Structurally related Class I HA synthases were then cloned from human (12–14), mouse (15–17), chicken (17), frog (17–21), Group C streptococcus (22), and chlorella virus PBCV-1 (23). A structurally distinct Pasteurella multocida HA synthase was also cloned (24). All the HA synthases reported to date produce a broad distribution of high molecular weight (e.g., ~10⁵–10⁷) HA chains from the substrates, UDP-GlcNAc and UDP-GlcUA.

With the exception of pmHAS, all of the HA synthases share regions of significant homology at the amino acid level and extensive similarity in their hydropathy profiles (25, 26). Based on these similarities the HASs have been grouped into two classes (26). Class I HASs include all the eukaryotic and streptococcal enzymes. The pmHAS is presently the only Class II member. The two streptococcal enzymes, spHAS and seHAS, are the smallest members within the HAS family containing, respectively, 419 and 417 amino acids. These two proteins are 72% identical at the amino acid level. The streptococcal enzymes have also been purified to homogeneity and are the most extensively characterized HASs biochemically (27–29). These two HASs have been characterized in membranes or as purified enzymes with respect to the Km values for the two substrate sugars, Vmax, and average product length. Tlapak-Simmons et al. (27) recently demonstrated that spHAS and seHAS function as monomeric proteins dependent on the association of ~16 cardiolipin (CL) molecules. A single HAS protein appears to be solely responsible for HA chain initiation and elongation and probably also for extrusion of the growing HA chain from the cell.

The sizes of the HAS enzymes are similar to that of other transporters such as the Lac permease (30), mannose transporter IIC (31), the porins from Eschericia coli and Rhodobacter capsulatus (32, 33), the facilitative glucose transporter (34), and the melibiose carrier protein (35). Unlike these transporters, however, spHAS and seHAS, as well as the other HAS family members, lack the number of predicted hydrophobic transmembrane domains necessary for the formation of a pore sufficiently large to transport the growing HA chain across the membrane. To understand the enzymatic and transport functions of the HAS family, it is important to determine the top-
ological organization of the active site(s) and membrane-spanning or membrane-associated regions.

Previously (11, 36) we predicted a model for spHAS organization in which the ends of the protein and a large central domain were inside the cell (e.g. similar to Fig. 3E). The eukaryotic HASs, which are ~35% larger than the streptococcal enzymes, were proposed to have this same topological organization with an additional COOH-terminal region containing several MDs (25). Since the streptococcal HASs are structurally the smallest, simplest members of this family, information on the topology of the spHAS and seHAS enzymes could help elucidate the basic topology of the remaining HAS family members.

In this study, the topology of spHAS was characterized by protease and substrate accessibility and by fusion protein analysis using alkaline phosphatase and β-galactosidase as reporter enzymes, whose correct folding and activity is dependent on cellular localization. The results support a topology of HAS that is similar to the model proposed earlier, but that entails more potentially complex interactions with the membrane and delineates the membrane-associated regions near the COOH terminus.

**EXPERIMENTAL PROCEDURES**

**Materials and Strains—**The starting vector for expression of PCR-mediated fusion proteins was pALTER-Ex1 from Promega. The vector pKK223-3 was from Amersham Pharmacia Biotech. The vector pSKλLpHspA (37) was graciously provided by Marleen van Geest and Julie Joekema. Restriction and DNA modifying enzymes were from Promega unless noted otherwise. Oligonucleotide primers were synthesized by The Great American Gene Co. or Genosys. DNA plasmid purification was done with Qiaquick Spin Miniprep Kits from Qiagen. T7 Version sequence 2.0 sequencing kit and 5'-S-DATP were from Amersham Pharmacia Biotech. Media were from Fisher and the remaining reagents were of the highest purity available from Sigma unless otherwise noted. E. coli strain DHB9 was provided courtesy of The Beckwith Collection. E. coli CC170 cells containing TnLacZ (38) were provided by Dr. Colin Manoil. E. coli JM109 was used for LacZ fusion assays. CC118 for LacZ fusion assessment was provided courtesy of Dr. Colin Manoil. TnPhoA provided by Manoil and Beckwith (29) was used for infection of DHB9 cells and integration of the AP coding region into pKK3K. Antibody to ToadA (40) was a generous gift from Dr. Robert Webster.

**Construction of pKK3K—**Custom oligonucleotides, 5'-CCCAGGATCTTGGACT-CTTATTTAAGATTTAAGGTTGAT-3' (forward primer) and 5'-GAATTCGATATGCTTACGTTTCTAGATAGTTGTGTTTAAATAGTGGAC (reverse primer, containing a His tag sequence), were used to amplify the 600-bp sequence of spHAS from Streptococcus pyogenes chromosomal DNA by PCR with Tag polymerase (Fisher). The forward primer contains an N for allowing for either a Met or Val as the NH2-terminal amino acid; the primer is suitable for either codon, although the KK3K coding region contains an N to allow for either a Met or Val as the NH2-terminal amino acid. The forward primer has been shown to be effective for PCR with PFU polymerase (Stratagene). The PCR product and pAlter-Ex1 were digested with BamHI and KpnI and ligated into the pALTER-Ex1 restriction digest with T4 DNA ligase (New England Biolabs) and transformed into DHB9 competent cells. The resulting plasmid was designated pAP6X-1.

**Construction of AP-His3—**To create alkaline phosphatase containing a COOH-terminal His tag, the AP gene in pSKλLpHspA was amplified by 20 cycles of PCR with primer APS (5'-GGGTACTGGCGACCACTAG-3') and primer APHIS6-3 (5'-CTAGTCTAGATTTGTTGTTGGTGTGGGTTTTCGCCCCGAGCAG3'-5') using PFU polymerase (Stratagene). The PCR product and pALTER-Ex1 were digested with BamHI and XhoI and separated from the primers by a Qiagen PCR purification column. The PCR product was ligated to the pALTER-Ex1 restriction digest vector with T4 DNA ligase (New England Biolabs) and transformed into DHB9 competent cells. Colonies were screened by plasmid preparation and restriction digestion for the proper size and orientation of insert. The desired plasmid was designated pAP6X-1.

**PCR-mediated HAS Fusion Proteins—**Fig. 1 shows schematically how the spHAS-AP and spHAS-LacZ fusion protein vectors were constructed. Primers were designed to amplify different lengths of the spHAS ORF. HAS5 (5'-GGGTACTGGCGACCACTAG-3') was used for PCR with PfU polymerase and position Thr (5'-GGTTTTTTAAAAATGACGACCCCGCGG3'-5'). Primer has lambda (λ) was used with the reverse primers: Gln37 (5'-GGGTACCTATTTAATTAGTTTT3'-5'), or Pro (5'-GGGTATAGAAAGATTTTTT3'-5') from a pKK3K template. PCR products were ligated to KpnI linker from New England Biolabs. Ligation reactions and pAP6X-1 were digested with BamHI and XhoI and ligated together. PfU PCR was also used to generate spHAS fragments with a restriction site at the 3' end of the products to facilitate insertion into the vector. For these PCR reactions, HAS5′ was used with the reverse primers: Gly27 (5'-GGGTACTGGCGACCACTAG-3'), Val39 (5'-GGGTATAGAAAGATTTTTT3'-5'), Asp349 (5'-GGGTACCTATTTAATTAGTTTT3'-5'), and Lys419 (5'-GGGTATAGAAAGATTTTTT3'-5'). PCR products were digested with BamHI and KpnI and ligated into pAP6X-1. All ligations were transformed into DHB9 cells and screened by colony PCR using primers HAS5′ and PhoA. Those colonies giving the expected band size were grown in overnight culture and plasmids were purified. Plasmids were screened by restriction digest then sequenced with the PhoA primer to confirm location and frame of fusions.

**Transfer of Transposon Fusion Genes to the pALTER-Ex1 Vector—**The ORFs for the transposon-created fusion proteins were PCR-amplified with primers HAS5′ and APHIS6′. PCR products were digested with BamHI and XhoI and ligated into the pALTER-Ex1 vector. Colonies were isolated and assessed for correct construction by sequencing through the junction site.

**LacZ Fusion Construction—**Primer LacZZ5′ (5'-CCCCGAGGATACTTATTTAAGATTTAAGGTTGAT-3') and LacZZ6′ (5'-CTAGTCTAGATTTGTTGTTGGTGTGGGTTTTCGCCCCGAGCAG3'-5') were used to amplify the β-galactosidase ORF from strain CC170 with PfU polymerase as above. This product and the K419AP6 construct were digested with BamHI and XhoI, gel purified, and the LacZ fragment was ligated into the Lys319 construct. The spHAS PCR amplifications used the AP6 fusion proteins described above were also ligated into the Lys319 construct. The spHAS fusion products described above were also ligated into the Lys319 construct after digestion with BamHI and KpnI to remove the Lys319 insert. Fusions were confirmed by sequencing the spHAS insert sequence through the junction site. LacZ fusions were grown in JM109 cells.
AP and LacZ Fusion Protein Assays—Clones with confirmed fusion proteins were grown overnight in LB/Amp. The overnight cultures were diluted 1:25 to start new cultures in LB/Amp grown at 30 °C with aeration for 3–4 h. Cultures were induced with isopropyl-1-thio-β-D-galactopyranoside for 3 h at 30 °C. To determine the relative activity of the various fusion proteins, we measured cell number (by A600), the amount of fusion protein (by quantitative Western analysis), and enzyme activity (by colorimetric assays). One ml of cells from each culture was transferred to a 1.7-ml microfuge tube and centrifuged at 10,000 × g for 5 min. Cell pellets were resuspended in 1 ml of ice-cold 10 mM Tris-HCl, pH 8, or phosphate-buffered saline containing 1 mM MgSO4, and 50 mM β-mercaptoethanol. A portion (100 μl) of the suspension was diluted into 0.9 ml of cold 10 mM Tris-HCl, pH 8, or phosphate-buffered saline and A600 was determined. A second 100 μl of the suspension was centrifuged at 10,000 × g for 5 min, the pellet was resuspended in 100 μl of 1 × Laemmli SDS-PAGE buffer with 10 mM dithiothreitol, heated at 95 °C for 2 min, and used for Western blots. Another 100 μl of the suspension was diluted into 0.9 ml of 1 × Tris, 0.1 mM ZnCl2, or phosphate-buffered saline containing 1 mM MgSO4, and 50 mM β-mercaptoethanol. Fifty μl of 0.1% SDS and 50 μl of chloroform were then added and the tube was vortexed and incubated at 37 °C for 5 min. At time 0, 100 μl of 0.4% p-nitrophenyl phosphate in 1 ml Tris or 5 μg/ml o-nitrophenyl-β-D-galactoside was added to each reaction tube. When a yellow reaction color was observed, 120 μl of 8.3 M Na2SO4, 83 μl of KH2PO4 was added to stop the AP assays, or 0.5 ml of 1 M Na2CO3 was added to stop the LacZ assays. The time noted, the tube was transferred to ice and A650 and A420 were obtained. Activities were determined using the formula: 1000[A420 − (1.75 × A650)] × (time in min × nmol product) as described by others (42, 43). Both the LacZ and AP assays were also done in a multwell plate reader format with 20 and 40 μl of induced resuspension and 20 μl of substrate in a total volume of 200 μl/well. Plate reader assays were scanned every 30 s over 1 h with an enzyme-linked immunosorbent assay plate reader (Spectramax 340, Molecular Devices). Activities were determined over the linear range of substrate conversion.

Assessment of Fusion Protein Localization—Based on prior published cutoff values used by numerous other investigators (31, 44–50), the various fusion proteins were divided into categories of high (>20% of maximum), intermediate (10–20% of maximum), or low (<10% of maximum) activity, relative to the highest activity for each assay method used (defined as 100%). The average activity of each fusion protein, using the same assay method was compared as a percentage to the fusion protein construct with the highest activity. For example, for the spHAS-AP fusion proteins, all the plate-based assays were evaluated as a percent of the Asep135-AP fusion protein, which gave the highest specific activity in this assay. For AP or LacZ fusion proteins with low activity, the reporter protein was assessed as inside or outside of the cell, respectively. For AP or LacZ fusion proteins with high activity, the reporter protein was assessed as outside or inside of the cell, respectively. If fusion protein pairs (i.e. both AP and LacZ fusion proteins at the same spHAS position) had intermediate activities, neither was assigned an orientation because of this ambiguity or the fusion site was considered to be part of a possible MD.

Western Blots—Samples (5–10 μl) in SDS-PAGE buffer were loaded onto polyacrylamide gels and electrophoresed at 200 V for 50 min. Proteins were transferred to nitrocellulose at 100 V for 2 h in a Bio-Rad Trans-Blot Cell. Blots were stained with 0.05% copper(II)phthalocyanine-3,4,4’,4”-tetrasulfonic acid, tetrasodium salt (Alldrich) and imaged of the blots were captured with an Alpha-Innotech IS1000 imaging system. Blots were destained and blocked with 1% bovine serum albumin in TBST (with 0.05% NaN3) for ≥1 h. Blots were washed in TBST and exposed to a 1:1000 dilution of anti-penta-His monoclonal Ab from Qiagen or anti-β-galactosidase monoclonal Ab from Calbiochem, in TBST containing 0.1% bovine serum albumin for 1 h. Antibodies were biotinylated with NHS-biotin from Pierce. Secondary development was for 1 h with 125I-streptavidin and the blots were analyzed using a phosphor screen on a Molecular Dynamics Personal Phosphorimagager. Expression levels were measured as integrated density values and converted to nanomole of fusion protein using an internal standard curve of purified spHAS-H125I. This was followed by detection with rabbit anti-mouse IgG Ab conjugated to alkaline phosphatase (Sigma). Blots were developed with 0.3 mg/ml p-nitro blue tetrazolium chloride and 0.15 mg/ml 5-bromo-4-chloro-3-indolyl phosphate p-toluidine salt in 1× Tris-HCl, pH 9.5, 0.5 mM MgCl2.

Accessibility to Proteases and Chemical Labeling Reagents—To assess which portions of spHAS reside in the periplasm, spheroplasts of cells expressing spHAS-H125I (i.e. the full-length protein; not a fusion protein) were subjected to several modifying agents. Overnight cultures were seeded 1:15 into 13 ml of LB/ampicillin and grown at 30 °C, 250 rpm until late log phase. Cells were induced for 3 h with 1.5 mM isopropyl-1-thio-β-D-galactopyranoside. Cells were collected by centrifugation, resuspended in 13 ml of 20 mM Tris, pH 8.5, 20% sucrose, and 0.5 ml of suspension was removed as a “no treatment” control. EDTA was added to 2 mM and lysozyme was added to 10 μg/ml. Cells were incubated on ice for 20 min and >90% cell wall removal was observed by light microscopy verifying the production of spheroplasts (51). The MgSO4 concentration was brought to 15 mM and a lysis was prepared by sonication of a 6-ml portion of the spheroplast suspension, using a
These possible MDs are at residues 7–29 (PMD1), 33–54 (PMD2), 207–232 (PMD3), 317–343 (PMD4), 350–370 (PMD5), and 379–396 (PMD6). The shortest \( \alpha \)-helical domain that could theoretically span a membrane bilayer is 20 residues and the shortest \( \beta \)-sheet domain would be 11 residues. The majority of membrane proteins that have been crystallized and whose structures have been successfully solved contain \( \alpha \)-helical membrane-spanning regions. Therefore, in determining the topology of a protein for the first time, it is important to consider the possibility that both \( \alpha \)-helical and \( \beta \)-sheet MD structures may be present. Hydropathy analysis (53) of spHAS reveals at least six hydrophobic stretches that could represent PDMs, predicted (or putative) membrane domains (Fig. 2).

Based on the three prediction algorithms of Garnier et al. (55), Chou and Fasman (56, 57), and Rose (58) as well as the two prediction programs TMPRED (59) and HMMTOP (60), a large number of topologies are possible for spHAS (Fig. 3). These models vary as to which PDM is predicted to be a MD or a TMD and the orientation of a given PDM.

To determine which, if any, of the various predicted topology models summarized in Fig. 3 might be correct, topological studies were undertaken using a combination of fusion protein analysis (31, 42, 44–50, 61), chemical labeling, protease accessibility, and substrate accessibility. Using Tn-mediated and PCR-mediated approaches, protein fusions were created with varied lengths of spHAS linked to the reporter enzymes AP or LacZ. AP is only folded correctly and active when it is in the nonreducing environment outside the cell, whereas LacZ, in contrast, is only active when it is made within the intracellular reducing environment (31, 37, 42, 44–50, 61). As the spHAS fusion protein inserts into the membrane, the reporter enzyme is localized according to the \( N \)-terminal topological signals (62–64). The reporter enzyme is then active or inactive depending on the local reducing environment in which the AP or LacZ is localized.

Unfortunately, Tn-generated spHAS-AP fusion proteins were not random; they were distributed in the ORF predominantly around residue Asp \(^{49} \) and were not adequately dispersed throughout the entire sequence. Therefore, specific PCR-generated fusion proteins were created on either side of the six PDMs. This fusion protein approach has been used successfully to determine the topology of a large number of
bacterial proteins (61) and has been shown to be accurate when the results could be compared with a crystal structure, such as for the photosynthetic reaction center of *Rhodopseudomonas sphaeroides* (65).

**Analysis of Fusion Protein Constructs**—AP fusion proteins were generated in 17 locations within spHAS (Table I) through a combination of Tn-mediated and PCR-mediated approaches (38). LacZ fusion proteins were also generated by PCR at nine of these 17 positions. The constructs were checked for expression of fusion proteins by Western blot analysis and the sequence of the fusion junctions were confirmed by DNA sequencing. The spHAS fusion proteins generated with AP or LacZ had a broad range of expression levels, as determined by Western blot analysis, and varied greatly in the degree to

**Table I**

*Activities and deduced localization of spHAS-AP and spHAS-LacZ fusions*

The amino acid position of fusion protein junctions in spHAS is indicated for both AP and LacZ fusions. All specific activities are the mean of results from all assays (one to six independent determinations) at a given fusion location. Measurements were in the linear range for time and protein. Duplicate fusion sites indicate independently derived fusions of Tn (T) and/or PCR (P) origin. Only AP assays, in which fusion protein expression could be assessed within the linear range of the standard curve, were used to determine normalized activities. Specific activities of AP fusion proteins are indicated as units/min/nmol of protein. Specific activities of LacZ fusion proteins are indicated as units/min relative to the expression level of the fusion protein at position T7 (which was included in all assays). Cells with the pKK3K vector alone, with the complete spHAS ORF, or with no vector were AP-negative. Positive controls with exogenous bacterial alkaline phosphatase gave the same specific enzyme activity as specified by the manufacturer. The conclusion about orientation with respect to the membrane is indicated for each assay method and the final deduced locations were assessed as noted in the text. MJ indicates fusion sites at or near membrane junctions. The symbol "?" indicates no clear conclusion about orientation.

| Fusion site, amino acid number | Fusion method | spHAS-AP fusions | spHAS-LacZ fusions, Plate assays | Final deduced location |
|-------------------------------|--------------|-----------------|-------------------------------|----------------------|
|                               |              | Plate assays    | Spec assays                   |                      |
|                               |              | Activity | Location | Activity | Location | Activity | Location | Spec assays | Location | Final deduced location |
| Thr7                          | P            | 0            | In       | 0        | In       | 29       | In       |             |          |                      |
| Thr32                         | T            | 2            | In       | 17       | Out      | 6        | Out      |             |          |                      |
| Gly167                        | P            | 7            | Out      | 11       | ~         | 5        | Out      |             |          |                      |
| Asn60                         | P            | 5            | Out      | 14       | Out      | 52       | In       |             |          |                      |
| Phe134                        | P            | 7            | Out      | 20       | Out      |          |          |             |          |                      |
| Glu169                        | T            | 16           | Out      | 7        | In        |          |          |             |          |                      |
| Asn207                        | T            | 20           | Out      | 17       | Out      |          |          |             |          |                      |
| Ala215                        | T            | 20           | Out      | 15       | Out      |          |          |             |          |                      |
| Gly270                        | P            | 6            | ~         | 12       | ~         | 9        | Out      |             |          |                      |
| Pro217                        | P            | 6            | ~         | 8        | In        | 98       | In       |             |          |                      |
| Asp349                        | P            | 20           | Out      | 50       | Out      | 13       | Out      |             |          |                      |
| Asp349                        | T            | 19           | Out      | 67       | Out      |          |          |             |          |                      |
| Asp349                        | T            | 31           | Out      |          | Out      |          |          |             |          |                      |
| Leu353                        | T            | 27           | Out      |          | Out      |          |          |             |          |                      |
| Leu366                        | T            | 6            | In       |          | In        |          |          |             |          |                      |
| Pro376                        | P            | 18           | Out      | 38       | In        |          |          |             |          |                      |
| Lys387                        | T            | 15           | Out      |          | Out      |          |          |             |          |                      |
| Lys3819                       | P            | 2            | In       | 0        | In        | 128      | In       |             |          |                      |
| Lys3819                       | P            | 89           | In       |          | In        |          |          |             |          |                      |
which breakdown products were observed (Fig. 4A). The intact fusion proteins demonstrated an electrophoretic mobility shift proportional to the length of the spHAS portion incorporated (Fig. 4A). Fusion protein expression levels, determined by Western blot analysis with an anti-His5 Ab or anti-LacZ Ab (Fig. 4B), varied over nearly 2 orders of magnitude. For spHAS-AP fusions, the expression level generally varied inversely with the length of the fusion protein (e.g. fusion constructs at Gly37 and Asn60 had 4-fold greater expression levels compared with any other constructs). The only exception to this was for the fusion at position Thr7, which had an extremely low expression level for the AP fusion. Based on SDS-PAGE and Western analysis comparing isolated membranes and whole cell lysates, at least 95% of the various fusion proteins were localized, as expected, to the cell membrane (not shown).

In a large number of studies, the location of reporter enzyme in AP or LacZ fusion proteins has typically been judged to be intracellular or extracellular based on about a 5–10-fold difference in activity (31, 44–50). For spHAS-AP fusion proteins, an external orientation was scored as a high activity above a cutoff of 6 units/min/nM for plate reader assays and 13 units/min/nM for spectrophotometric assays. For spHAS-LacZ fusion proteins, an intracellular localization was assessed based on high activities of >25 units/min/% position Thr7 expression.

Activity assays for fusion proteins performed by the plate reader method have the benefit of needing no correction in absorbance measurements for interference from cell debris. More importantly, the plate reader assays enable determination of the linearity of the assay over multiple time points. The bacterial AP activities of fusion proteins were assessed by both the plate assay and the traditional single time point spectrophotometric assay. Results were consistent between the two assay methods, with the exception of fusion proteins produced at positions Glu169 and Asn207, for which the two assays gave conflicting results (Table I). The spectrophotometric assay usually gave higher AP specific activities. Assessments of AP cellular localization were then made as either inside, outside, or unclear, based on the activity category for each fusion protein.

The HAS-AP fusion protein at position Asp349 demonstrated a high AP activity and a low LacZ activity, indicating a periplasmic location for that region of spHAS. High-to-intermediate AP activity with low LacZ activity was also observed for fusion proteins at Gly37 and Asn60. These were complemented by fusion proteins at position Thr7, Pro317, and Lys419, which showed a high activity for the LacZ fusion and a low-to-intermediate activity for the AP fusion, indicating intracellular locations for these sites. Thus, we conclude that the NH2 and COOH termini of spHAS, as well as the region immediately prior to PDM4, are intracellular and that the two loops between PDM1 and PDM2 and between PDM4 and PDM5 are extracellular. In E. coli these two loops are exposed to the periplasm.

Paradoxically, high activities for both spHAS-AP and spHAS-LacZ fusion proteins were observed at Val134 and Pro776. In these cases, the stronger positive activity for LacZ was taken to be the better indicator of membrane orientation and they were assigned as internal. One set of spHAS fusion proteins at position Gly270 demonstrated intermediate-to-low activities for both AP and LacZ constructs and thus no orientation was indicated. The remaining fusion locations were examined using AP fusion proteins only and thus, the lack of strong AP activities is not complemented with LacZ activities to give both positive and negative indications of membrane orientation. Fusion locations at Ala208, Ala215, Leu353, and Lys397 had high AP activities, indicating an external localization. The fusion at Leu365 had low activity indicating an internal orientation of the AP at this HAS position. The fusions at Thr32, Glu169, and Asn207 give conflicting results between the plate reader and the spectrophotometric assays and were not assigned a membrane orientation.

Sensitivity of spHAS to Proteases and Chemical Labeling—To assess the presence and accessibility of any portions of spHAS that might be localized to the periplasm, spheroplasts, either intact or disrupted by sonication, were exposed to a chemical modifying reagent or a variety of proteases. In these experiments the full-length spHAS protein, not a truncated or fusion protein, was examined. As expected, all the sonicated samples showed multiple sites of accessibility to the proteases and to a biotinylating reagent reactive with primary amines (Fig. 5), demonstrating that under the conditions used, spHAS was susceptible to proteolytic and chemical modifications. However, intact spheroplasts showed no demonstrable cleavage or modification of spHAS, indicating that the protein has no extracellular regions sufficiently exposed to allow modifica-

---

![Figure 4](http://www.jbc.org/)
As a positive control, the same digestion conditions were shown to cleave the TolA protein completely in both spheroplasts and lysates (not shown). TolA, which is involved in colicin uptake, is localized to the inner membrane (40). Thus, it is very likely that any extra-cytoplasmic loops of the spHAS protein are very small and close to or intimately associated with the membrane.

Accessibility of the spHAS Active Site to Substrates—In order to determine whether the sugar nucleotide substrates are accessible to the active site of spHAS from an extracellular, periplasmic, or intracellular location, we incubated intact cells, intact spheroplasts, or sonicated spheroplasts with these substrates. Only the lysed samples demonstrated the expected formation of HA (Table II), indicating that the two substrate-binding sites and glycosyltransferase active sites are only accessible from an intracellular location. The protein regions within the HAS family that are most highly conserved, and are thus most likely to be involved in substrate binding and catalysis, are located between PMD2 and PMD4 (25, 26). The conclusion from the above substrate accessibility result is that this region is intracellular.

DISCUSSION

Stoeltmiller and Dorfman (66) first established the location of spHAS at the bacterial cell membrane in 1969. In the intervening three decades, the orientation of the protein and the location of its active site with respect to the membrane had not been established experimentally. Understanding how a membrane enzyme functions requires a basic working model of its topology and the orientation of its substrate-binding sites and active sites with respect to the membrane. This is particularly true for the HAS enzymes because they also have a putative transport function. The spHAS enzyme is the only protein required in vitro for the initiation and elongation of HA chains (11, 36) and probably in vivo for the extrusion of the growing HA chain through the membrane bilayer. Consistent with this conclusion, human HAS1 is active when purified away from other proteins, indicating that this HAS enzyme also initiates and elongates HA chains using only the HAS1 protein (67). As the first, and one of the smallest, membrane-bound glycosaminoglycan synthases to be cloned, elucidation of the mechanism...
of spHAS action could serve as a model for other members of the HAS family. The relatively small size of the HAS enzymes and their small number of predicted membrane-spanning regions, make a topological model even more important in order to comprehend their mode of action.

For the purpose of minimizing bias toward \( \alpha \)-helices, predictions from the TMPRED (59) and HMMTOP (60) programs were evaluated only as potential MDs (the PMDs shown in Fig. 1) and the structural nature of these regions was ignored. All predictive methods agree on the location of the first two hydrophobic stretches (PMD1 and PMD2) as probable TMD regions, which are predicted to be \( \beta \)-sheets. The two programs, however, predicted opposite orientations for each TMD. The stretch between residues Asn\(^{207} \)-Tyr\(^{232} \) is predicted by the three algorithms to have an amphipathic helix joined to a \( \beta \)-sheet by a turn or coil. Subjectively, PMD3 appears to be polar and this is conceptually difficult to reconcile with a typical TMD. The same location is predicted for the last three PMDs by both programs. However, the algorithms disagree as to the structures predicted for these three domains. All three algorithms predict some variation of a \( \beta \)-sheet for PMD6, although two of the three predict a break in the structure by a turn or coil. PMD4 and PMD5 are predicted to be \( \beta \)-sheets by the methods of Chou and Fasman (56, 57) and Rose (58), but to be \( \alpha \)-helices by the method of Garnier \textit{et al.} (55). The TMPRED and HMMTOP programs predicted 5 or 6 TMDs, with the NH\(_2\) terminus out or in, respectively, and disagreed as to whether PMD2 is utilized, whether PMD3 exists, and if the COOH terminus is inside. In prokaryotes, there is a strong proclivity for positive charges to be localized intracellularly and switching the charge densities of extramembranous loops can switch the preferential orientation of associated MDs (68, 69). Based on the positive-inside rule (62–64), the predicted orientations for the NH\(_2\) and COOH termini of all the HASs should be intracellular (25).

The present HAS-reporter-protein fusion results confirm internal orientations of the NH\(_2\) and COOH termini of spHAS, as well as the regions after PMDs 2, 5, and 6, and before PMD 4. Fusion results indicate an external orientation for the region after PMD 4 and within PMDs 2, 3, 5, and 6. Fusions at a junction immediately following an out-to-in TMD frequently interfere with stop-transfer signals in the fusion protein, which can result in anomalous extracellular localization and, therefore, give activities that are high for AP and low for LacZ (39, 60, 61, 65). This behavior could explain the high activities of Gly\(^{37} \) and Asn\(^{60} \) in PMD2, and Asp\(^{249} \) and Leu\(^{353} \) in PMD5, indicating these PMDs are out-to-in membrane spanning regions. Similarly, fusions positioned after a start-transfer signal for a domain that normally inserts into the membrane without traversing the bilayer may yield fusion proteins that orient the reporter enzyme extracellularly (44, 62–64). Thus, the external indication from fusion proteins within PMD 3 and 6, which are flanked by fusions showing internal orientation of both sides of that PMD, most likely indicate a membrane domain that does not completely span the bilayer.

The results with the fusions at Thr\(^{32} \), Glu\(^{169} \), Asn\(^{207} \), and Gly\(^{270} \) gave ambiguous assignments. The external activity indicated by AP fusion at Asn\(^{207} \) (as well as Ala\(^{208} \)) may suggest a need to expand the length of PMD3 NH\(_2\)-terminal beyond what is indicated by hydropathy plots alone. This NH\(_2\)-terminal region may be capable of forming an amphipathic helix with a hydrophobic face, so that PMD3 would actually be longer than indicated in Fig. 6. We believe that another cause of ambiguity in these four fusions is because \( \beta \)-sheets, rather than \( \alpha \)-helices, constitute a significant proportion of the preceding MDs. So far, the fusion protein approach has been primarily utilized to study membrane proteins that are predominantly \( \alpha \)-helical (60). Theoretically, \( \alpha \)-helices are energetically stable even if they are present as single domains that do not interact with other domains, whereas a single \( \beta \)-sheet should be far less stable due to unpartnered polar hydrogen bond donors and acceptors in the hydrophobic environment. In the latter case, multiple interacting \( \beta \)-sheets would need to be present in order to create an energetically stable situation. Alternatively, an energetically favorable situation might also be created by the formation of specific interactions between a \( \beta \)-sheet and several CL molecules, which are required for streptococcal HAS activity. The enzymatic activities of both streptococcal HASs are very dependent on phospholipids, with CL being the most effective (28). The human HAS1 does not appear to have a requirement for CL (67).

If the reporter enzyme replaces a membrane localization determinant or is within a membrane-inserted \( \beta \)-sheet, the result could be the localization of the reporter proteins in both extracellular and intracellular compartments (61, 63, 64). There would be a distribution of final orientations for the population of spHAS fusion proteins, some will locate the reporter enzyme internally, some externally, and some may be unstable and likely to be degraded. Thus, the fusion proteins at junctions following the second NH\(_2\)-terminal \( \beta \)-sheet (PMD2) might be unstable in their orientation and would give ambiguous results as we observed. Since PMD4 and PMD5 are predicted to be helical and to include multiple membrane localiza-
tion signals (39), fusions in this region are more likely to be stable in their conformation and give clear-cut in or out results (38, 61–64).

To characterize further the topology of the ambiguous regions in spHAS, chemical labeling and protease accessibility experiments were carried out on intact cells, spheroplasts, and cell lysates. If any regions of spHAS have significant periplasmic exposure, then they should be labeled by the N-hydroxysuccinimidyl ester reagent at exposed lysine residues and/or digested by one or more of the proteases. Intact cells (not shown) or spheroplasts demonstrated no labeling or proteolysis, although positive results were obtained with lysates and TolA controls in all cases. These data support the conclusion that the majority of the spHAS protein is intracellular and that any extracellular protein loops are either very small or intimately associated with the membrane. These results also indicate that PMD3 is not likely to be a bona fide TMD because any extracellular loops it would form with an adjacent TMD should be sufficiently long and hydrophilic to have been detected in these experiments.

The various HAS family members share significant homology over the region corresponding to residues Gly137–Glu305 in spHAS (25). This region contains sequence motifs that are also conserved in the chitin and cellulose synthases and that have been shown to be important for their glycosyltransferase activities. This region, logically then, would be the location of the enzyme active sites. The present report provides the first experimental demonstration, based on the inaccessibility of substrates to spHAS in intact spheroplasts, that the active site of HAS is located intracellularly.

Based on the present results from several different approaches, we propose a model for the topology of spHAS (Fig. 6), in which the protein contains only four TMDs (PMDs 1, 2, 4, and 5 of Fig. 2). The amino and carboxyl termini and the large central region of the molecule are intracellular and the only extracellular regions of spHAS are very short protein loops between PIM1 and PIM2 and between PIM4 and PIM5. Based on their hydrophobicity and the turns in the middle of the predicted β-sheet structures that they form, we conclude that PIM3 and PIM6 are membrane-associated regions, but not TMDs. These β-sheet MDs could help form a pore for the growing HA chain (28), in a similar fashion to that observed for the K⁺ channel (69) or the FhuA siderophore transporter (70). If these regions are helical, they could assist in the formation of an acceptor site that sits within the plane of the membrane as observed for the Lac permease (71).

This topological model for spHAS still leaves unanswered the question of how the streptococcal HASs, which contain so few MDs (only six) can translocate the large HA polymer through the bilayer, especially since the functional enzyme is a protein monomer (27). The other sugar transporters whose topologies have been characterized have 12 membrane-spanning helices (34, 35, 45, 70, 71). An HA tetrasaccharide would occupy a volume of ~225 Å³ (~20 × 2.5 × 4.5 Å) excluding water coordinated by H-bonds or hydration. If 12 helical TMDs are necessary for the formation of a sufficiently large pore to transport a mono- or disaccharide, then spHAS is deficient. It should be noted, however, that in most of these transporters there are actually two pores, a sugar and a cation pore (72). Between 4 and 8 of the helices are usually involved in the transport of the sugar.

We demonstrated recently that spHAS and seHAS require phospholipid, particularly CL, for activity (28). Potentially, this phospholipid component could create a portion of the pore/channel and help stabilize the protein, particularly the β-sheets. In addition, we call attention to the lengths of TMD 1 and TMD 2, which are ~23 and ~22 residues, respectively. If these domains are indeed β-sheets, then each sheet is sufficiently long to span the membrane twice or to cross the bilayer at a 45° angle with respect to the plane of the membrane. This feature would increase the capability of spHAS to create a larger pore by greatly expanding the volume of a pore made with participation of these two domains. Consistent with an important role for this region, we find that deletion of residues Val1-Thr24 gives an inactive enzyme, whether in membranes or lysates (not shown).

The two MDs shown in Fig. 6 (PMD3 and PMD6 in Fig. 2) may also be long enough in a β-sheet conformation to actually span the membrane twice, thus increasing the number of membrane spanning domains to 10. Furthermore, there are several predicted β-sheets, which are also long enough to span a membrane, in the central domain between residues 70 and 310. These putative central domain β-sheets are neither very hydrophobic nor hydrophilic, but may be stable in a β-barrel type structure and able to help form a pore in a membrane by interaction with other more hydrophobic membrane domains and associated phospholipids. If these β-sheets are indeed partially embedded in the membrane, then another explanation for the ambiguous results obtained for fusion proteins at Val134, Glu169, and Asn207 may be perturbation of the formation of such a β-barrel, or localization of the reporter enzyme in an altered reducing environment, which might enable some molecules to fold correctly and be active. However, no experimental results currently indicate the presence of these additional membrane-associated sheets and more in depth structural analyses will be necessary to unravel the formation of the HA pore in HAS.

In summary, we have identified the basic topology for the streptococcal HASs, and the likely topology for the conserved regions within the larger HAS family, although the topology of the unique COOH-terminal portions of the eucaryotic HAS enzymes cannot be inferred from the present study. We have verified the earlier suggestion (25) that the orientation of the active site is intracellular. Further studies are in progress to determine what proportion of the MDs are, in fact, β-sheets versus α-helices, and how these domains contribute to pore formation and HA translocation.

Acknowledgments—We thank Janet A. Weigel for performing the iodinations and Dr. Valerie Tapak-Simmons for helpful discussions throughout this study. We are very grateful to Drs. Robert E. Webster and Rafael Tosado-Acevedo for providing the TolA antibody.

REFERENCES

1. Meyer, K., and Palmer, J. W. (1934) J. Biol. Chem. 107, 629–634
2. Evered, D., and Whelan, J. (1989) Ciba Found. Symp. 143, 1–288
3. Toole, B. P. (1991) Cell Biology of the Extracellular Matrix (Hay, E. D., ed) pp. 305–341, Plenum Press, New York
4. Knudson, C. B., and Knudson, W. (1993) FASEB J. 7, 1233–1242
5. Laurent, T. C. (1987) Acta Otolaryngol. 112, 7–24
6. West, D. C., Hampson, I. N., Arnold, F., and Kumas, S. (1985) Science 226, 1324–1326
7. Rockey, D. C., Chung, J. J., Mckee, C. M., and Noble, P. W. (1996) Hepatology 25, 86–90
8. Whitnack, E., Bisco, A. L., and Beachey, E. H. (1981) Infect. Immun. 31, 985–991
9. Wessels, M. R., Goldberg, J. B., Moses, A. E., and Dicesare, T. J. (1994) Infect. Immun. 62, 433–441
10. Wessels, M. R., Moses, A. E., Goldberg, J. B., and Dicesare, T. J. (1991) Proc. Natl. Acad. Sci. U. S. A. 88, 8317–8321
11. DeAngelis, P. L., Papanastassios, J., and Weigel, P. H. (1993) J. Biol. Chem. 268, 19181–19184
12. Itano, N., and Kimata, K. (1996) Biochem. Biophys. Res. Commun. 222, 816–820
13. Shiyaj, A. M., Heldin, P., Butter, E. C., Yoshino, T., and Briskin, M. J. (1996) J. Biol. Chem. 271, 23395–23399
14. Watanabe, K., and Yamaguchi, Y. (1996) J. Biol. Chem. 271, 22945–22948
15. Itano, N., and Kimata, K. (1996) J. Biol. Chem. 271, 9875–9878
16. Spicer, A. P., Augustine, M. L., and McDonald, J. A. (1996) J. Biol. Chem. 271, 23400–23406
17. Spicer, A. P., and McDonald, J. A. (1998) J. Biol. Chem. 273, 1923–1932
