INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600
Characterization of a Virulence Determinant From Group A Streptococcus: Identification of a Novel Chromosomal Region Responsible for Streptolysin S Production in Streptococcus pyogenes

Sergio M. Borgia

A thesis submitted in conformity with the requirements for the degree of Master of Science
Graduate Department of Medical Genetics and Microbiology
University of Toronto

© Copyright by Sergio Marcelo Borgia, 1997
The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author’s permission.

L’auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L’auteur conserve la propriété du droit d’auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.
Characterization of a Virulence Determinant From Group A Streptococcus: Identification of a Novel Chromosomal Region Responsible for Streptolysin S Production in *Streptococcus pyogenes*

Sergio M. Borgia
Master of Science, 1997
Department of Medical Genetics and Microbiology
University of Toronto

ABSTRACT

Streptolysin S (SLS) is a poorly understood, non-immunogenic and oxygen-stable protein produced by Group A Streptococci which is also produced by strains of Lancefield groups C, G, E, H, L. SLS results in very potent membrane-damaging effects on target cells including red blood cells, lymphocytes, polymorphonuclear leukocytes, platelets, tissue culture and tumor cells. Nothing is known of its genetic makeup. Previously in this laboratory, a *S. pyogenes* MGAS166 SLS' mutant, designated SBNH5, was constructed using the conjugative transposon Tn916. A 3.8 kb region flanking Tn916, containing the region associated with SLS production, was cloned into plasmid pACYC184. It was shown that this SLS' mutant was less virulent in a mouse model. The purpose of this investigation was to characterize this fragment.

Sequencing of the 3.8 kb insert revealed a novel sequence, with some homology to the insertion sequence IS199 from *Streptococcus mutans*, a transposase gene from *Lactobacillus helviticus* and a transposase protein from *Streptococcus thermophilus*. To date, only one other insertion sequence has been reported for *S. pyogenes*; this sequence may represent a novel insertion element in GAS. Tn916 was precisely mapped to base 1367 (5'-3') on the 3.8 kb fragment. Several putative open reading frames (ORFs) were identified; five ORFs were
selected for further characterization. One of these, ORF 5 harbours the insertion point of Tn916 within the promoter region.

Northern blot analyses revealed transcript production for ORF 5 in the wild-type (WT). However, no transcripts were produced by the mutant SBNH5 and other Tn916 generated hemolytic deficient mutants. Interestingly, Tn916 has also integrated into ORF 5 of another hemolytic deficient mutant, CS91-23. ORF 5 demonstrates nearly all the conserved elements of a functional transcript. It encodes a 53 amino acid peptide with no obvious leader sequence as well as several consecutive cysteine residues. The protein sequence of ORF 5 is inconsistent with previous biochemical studies of SLS including amino acid composition.

There are similarities between a class of molecules known as lantibiotics and the ORF 5 peptide. Lantibiotics are a sub-group of bacterial antibiotic peptides containing post-translationally modified amino acids. They belong to a class of ribosomally-synthesized, low molecular weight, heat stable, bacteriocin-like compounds that have only been found in gram positive organisms.

The protein sequence of ORF 5 was compared to 13 known lantibiotic sequences. The highest degree of homology was observed with epidemin and pep5 (from Staphylococcus epidermidis) matching 44% and 40% similarity and 22% and 20% identity respectively. In support of a functional relationship between lantibiotics and the ORF 5 peptide, cytolysin LL/LS from Enterococcus faecalis was the first lantibiotic shown to be a hemolysin. These data suggest that ORF 5 encodes either a regulatory gene for SLS production or the structural gene itself.
ACKNOWLEDGMENTS

Of course, first thanks to my parents and family.

Greatest thanks to my supervisor, mentor and good friend, Dr. Joyce de Azavedo, a real source of inspiration for true scientific scholarship. My sincerest appreciation to my co-supervisor, Dr. Donald Low, an outstanding role model and teacher, for his guidance, wit, and support. Many thanks to my committee members, Drs. Jim Brunton and Richard Ellen, and all the professors who have advised me in the project, Dr. Barry Kreiswirth, Dr. Andreas Podbielski, Dr. Michael Caparon, and Dr. Patrick Cleary. I would also like to thank my professors and mentors in microbiology especially those who have really helped and supported me since I entered the department in 1992, Dr. Andrew Bognar, Dr. Cliff Lingwood, Dr. Mario Huesca and Dr. Eleanor Fish, as well as members of the thesis examination committee. Best of luck to the new Microbiology department.
TABLE OF CONTENTS

LIST OF FIGURES .................................................................................................................. 7
LIST OF TABLES .................................................................................................................... 8

1.0 INTRODUCTION .................................................................................................................. 9
1.1 GENERAL CLASSIFICATION ............................................................................................. 9

1.2 GROUP A STREPTOCOCCAL DISEASES ............................................................................. 9
1.2.1 Diseases and Clinical Manifestations of *Streptococcus pyogenes* ................................ 9
1.2.2 Epidemiology of Group A Streptococcal Infections .................................................... 10

1.3 VIRULENCE FACTORS OF GROUP A STREPTOCOCCUS ............................................. 11
1.3.1 Factors Involved in Adherence and Colonization ....................................................... 12
     A) M Protein Mediated Adherence ................................................................................. 12
     B) Hyaluronic Acid Capsule Mediated Adherence ....................................................... 13
     C) Serum Opacity Factor Mediated Adherence ............................................................ 14
     D) Fibronectin Binding Protein Mediated Adherence ................................................... 14
1.3.2 Factors Involved in Replication and Dissemination .................................................... 17
     A) M Protein and M-like Mediated Replication and Dissemination ............................ 18
     B) Cysteine Protease Mediated Replication and Dissemination ................................ 18
     C) Streptokinase Mediated Replication and Dissemination ........................................ 19
1.3.3 Factors Involved in Avoidance of Host Defense Mechanisms .................................... 20
     A) Antiphagocytic and Anticomplimentary Properties Mediated by M Protein and
       Hyaluronic Acid, Protein SIC and C5a Peptidase ...................................................... 20
1.3.4 Factors Associated with Damage to the Host .............................................................. 22
     A) Superantigens and the Role of Streptococcal Pyrogenic Exotoxins ......................... 22
     B) Cytotoxins ............................................................................................................... 25

1.4 GENETIC CHARACTERIZATION OF GAS VIRULENCE FACTORS ............................... 25
     A) The Group A Streptococcal *Vir* Regulon ................................................................. 25
     B) Genetic Characterization of the Streptococcal Pyrogenic Exotoxins ....................... 28

1.5 STREPTOLYSIN O (SLO) .................................................................................................. 30

1.6 STREPTOLYSIN S (SLS) .................................................................................................. 33
     A) General Properties and Factors Needed for Streptolysin S Formation ................... 33
     B) Protein Chemistry of Streptolysin S ......................................................................... 36
     C) Evidence for Streptolysin S Serving as a Virulence Factor in GAS ........................ 38

2.0 OBJECTIVES ..................................................................................................................... 40
3.0 MATERIALS AND METHODS

3.1 Bacterial Strains, Plasmids and Culture Conditions ............................................. 42
3.2 DNA Manipulations .................................................................................................. 42
3.3 Southern Hybridization Analysis .............................................................................. 44
3.4 DNA Mapping .......................................................................................................... 45
3.5 DNA Sequencing ...................................................................................................... 45
3.6 Sequence Analysis ................................................................................................... 47
3.7 Determination of the Precise Insertion Point of Tn916 ............................................. 47
3.8 RNA Analyses of ORFs Within the 3.8 kb Fragment .................................................. 48
3.9 Subcloning the 3.8 kb Fragment Into DH5α for Complementation Studies .............. 49
3.10 Electroporation of pLZ12-S1 into CS91-23 ............................................................. 51
3.11 Restriction Enzyme Analysis (REA) ......................................................................... 51

4.0 RESULTS

4.1 Confirmation of Tn916 Insertion in SBNH5 ............................................................... 54
4.2 DNA Mapping .......................................................................................................... 54
4.3 DNA Sequence Analysis .......................................................................................... 56
4.4 Analysis of Insertion-like Elements Within the 3.8 kb Fragment ............................ 58
4.5 Determination of the Precise Insertion Point of Tn916 .............................................. 59
4.6 Analysis of Putative ORFs Within the 3.8 kb fragment ........................................... 63
4.7 RNA Analyses of ORFs Within the 3.8 kb Fragment ................................................ 66
4.8 Construction of pLZ12-S1 and Complementation studies ...................................... 67
4.9 Location of 3.8 kb Fragment Within 20 Clinical Isolates of GAS ............................. 73
4.10 Probing Other GAS Strains for the 3.8 kb Insert and ORF 5 ................................. 77

5.0 DISCUSSION AND FUTURE STUDIES .................................................................. 81

REFERENCES ................................................................................................................. 94
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1</td>
<td>Genetic organization of the VIR Regulon</td>
</tr>
<tr>
<td>Figure 2</td>
<td>Hemolytic pattern of MGAS166 and the isogenic hemolytic deficient mutant, SBNH5, after Tn916 mutagenesis</td>
</tr>
<tr>
<td>Figure 3</td>
<td>Confirmation of Tn916 insertion into MGAS166 166 to generate SBNH5</td>
</tr>
<tr>
<td>Figure 4</td>
<td>Sequence and amino acid homology of IS905 from Lactococcus lactis with the 3.8 kb region</td>
</tr>
<tr>
<td>Figure 5</td>
<td>Schematic representation of putative open reading frames within the 3.8 kb region</td>
</tr>
<tr>
<td>Figure 6</td>
<td>Generation of amplicons representing five non-overlapping open reading frames with the 3.8 kb region</td>
</tr>
<tr>
<td>Figure 7</td>
<td>Nucleotide sequence and protein translation of ORF 5 within the region responsible for SLS production in GAS</td>
</tr>
<tr>
<td>Figure 8</td>
<td>Strategy for determining the exact insertion point of Tn916 in SBNH5</td>
</tr>
<tr>
<td>Figure 9</td>
<td>RNA analysis of MGAS166 and SBNH5 probed with the 3.8 kb fragment responsible for SLS production</td>
</tr>
<tr>
<td>Figure 10</td>
<td>RNA analysis of MGAS166 and SBNH5 probed with ORF 5</td>
</tr>
<tr>
<td>Figure 11</td>
<td>RNA analysis of GAS strains probed with ORF 5 and 16S rRNA</td>
</tr>
<tr>
<td>Figure 12</td>
<td>Subcloning 3.8 kb insert to generate construct pLZ12-S1</td>
</tr>
<tr>
<td>Figure 13</td>
<td>Analysis by PFGE of 9 unrelated clinical isolates of GAS</td>
</tr>
<tr>
<td>Figure 14</td>
<td>Analysis of DNA from 9 clinical isolates (by Southern blotting) for the presence of 3.8 kb fragment after PFGE</td>
</tr>
<tr>
<td>Figure 15</td>
<td>Analysis of genomic DNA from 11 clinical isolates (by Southern blotting) for the presence of 3.8 kb fragment after restriction enzyme digestion</td>
</tr>
<tr>
<td>Figure 16</td>
<td>Various GAS wildtype and mutant strains probed with 3.8 kb fragment</td>
</tr>
<tr>
<td>Figure 17</td>
<td>Various GAS wildtype and mutant strains probed with ORF 5</td>
</tr>
<tr>
<td>Figure 18</td>
<td>Demonstration of ORF 5 in ATCC 27762 by PCR</td>
</tr>
</tbody>
</table>
LIST OF TABLES

Table 1  Virulence factors of GAS: Determinants mediating adherence and colonization, replication and dissemination and avoidance of host defense mechanisms ................................................................. 24

Table 2  Stains and plasmids used in this investigation ........................................ 43

Table 3  Sequence of primers used for sequencing and PCR ............................. 46

Table 4  GAS clinical strains used in identifying presence of 3.8 kb fragment responsible for SLS production ........................................................................................................... 53

Table 5  Homology results of 3.8 kb fragment using BLAST search ............. 57

Table 6  Location of 3.8 kb fragment in GAS clinical strains after PFGE and Southern blotting ..................................................................................................................... 77
INTRODUCTION

1.1) GENERAL CLASSIFICATION

*Streptococcus pyogenes* is a Gram positive coccus that is pathogenic for humans. It measures less than two μm in diameter and grows in pairs and chains. This facultatively anaerobic, catalase negative organism is β-hemolytic when grown on blood agar. Streptococci are divided into immunological groups according to cell wall carbohydrate composition. The cell wall of *S. pyogenes* contains N-acetylglucosamine and rhamnose. These streptococci are designated as belonging to the Group A. Hence, Group A streptococci (GAS) are synonymous with the species *Streptococcus pyogenes*.

GROUP A STREPTOCOCCAL DISEASES

1.2.1) Diseases and Clinical Manifestations of GAS

GAS have the ability to cause a wide range of diseases and clinical manifestations. In addition to causing relatively minor and self-limiting suppurative skin and soft tissue infections such as cellulitis, impetigo, erysipelas and pharyngitis (Bisno and Stevens 1995), GAS cause the non-suppurative, post-streptococcal sequelae of glomerulonephritis and rheumatic fever, and may cause life threatening invasive disease.

Streptococcal toxic-shock syndrome (STSS) is one of the manifestations of invasive GAS disease. STSS affects persons of all ages, but is found mostly in the elderly. It is characterized by systemic multi-organ involvement or failure (Schlievert et. al. 1996, reviewed by Stevens 1995). Eighty percent of patients
have clinical manifestations of soft tissue infection which, in as high as 70% of the cases, may result in necrotizing fasciitis (NF) or myositis (NM) requiring surgical debridement or amputation (Stevens 1992). Whether or not STSS, NF, and NM are varying degrees of the same clinical course is not clear. STSS appears to be clinically distinct from NF/NM.

The portal of entry of streptococci remains elusive in half of the cases, however a possible source may be the mucous membranes of the throat or the skin. Usually, infection begins after some form of minor trauma with or without hematoma or muscle strain and are not necessarily associated with breach of the skin integrity.

1.2.2) Epidemiology of Group A Streptococcal Infections

Infections consistent with the clinical presentation of streptococcal diseases were recognized as far back as the 5th century B.C. Epidemic erysipelas and descriptions of a disease that would be clinically referred to as NF today, have been noted in the literature (Descamps et al. 1994). It is interesting to note that the clinical presentations such as the speed with which NF progresses, the severity of pain, and the broad host range were similar then as is reported today (Weinbren and Perinpanayagam 1992).

As early as 1664 Sydenham described a disease as ‘febris scarlatina’ (Reviewed by Katz and Morens 1992). Scarlet fever has served as a predictable marker of streptococcal disease throughout the centuries. Katz and Morens (1992) have summarized streptococcal scarlet fever to have occurred in roughly three stages. The first stage occurred from ancient times to the late 1700s. The
second stage, spanned approximately 1825-1885 and was associated with a high mortality rate. In the third stage, covering the last one hundred and ten years, scarlet fever was seen as a milder disease in the developed world.

Low (1997) has reviewed the history of streptococcal disease and reports that in the latter half of this century severe GAS infections decreased markedly with mortality rates decreasing from 72% to between 7 and 27%. However, in the 1980s a significant resurgence of severe invasive GAS disease has been reported in the literature. This has been due to the increase in the number of cases of STSS and NF with or without STSS.

In one study of invasive GAS infections in Ontario, Canada, the annual incidence of NF increased fourfold from 0.85 per million population to 3.5 per million (p<0.001) from 1992 to 1995 (Kaul et. al. 1997).

1.3) VIRULENCE FACTORS OF GROUP A STREPTOCOCCUS

Group A streptococci produce an array of virulence factors that have been ascribed to play a role in infection. These may be grouped according to function and the contribution they make to the infectious process. Specifically, these determinants can be categorized into four functional classes: A) adherence and colonization, B) replication and dissemination; C) avoidance of host defenses, and D) factors associated with damage to the host (Table 1). It is necessary, however, to be aware that streptococcal virulence factors often play several dynamic and integrated roles in pathogenesis. Concomitantly, the role of host factors including immunity and defense mechanisms play a major role in the outcome of infection.
1.3.1) Factors Involved in Adherence and Colonization

Attachment of GAS to host cells is a necessary prerequisite for infection. Several factors produced by GAS have been shown to be associated with adherence (Table 1). These include the M protein, hyaluronic acid capsule, serum opacity factor (SOF), and several fibronectin binding molecules.

A) M-protein

The M protein of GAS was first described in 1928 by Rebecca Lancefield (Lancefield 1962). Recently, analysis of the genes encoding M-protein have demonstrated that M protein genes (emm) are members of a larger emm-like gene family and that many GAS express more than one M-like protein (mp, enn) (Bessen and Fischetti 1990, 1992; Bessen and Hollingshed 1994). M-like proteins have been found to adhere to several cell types such as keratinocytes and epithelial cells (Beachey and Courtney 1987, Caparon et. al. 1991, Courtney et. al. 1994, Ellen and Gibbons 1972, Hollingshed et. al. 1993, Tylewska et. al. 1988, Wang and Stinson 1994). In keratinocytes, M and related proteins bind membrane cofactor protein (MCP) (Accardo et. al. 1996, Okada et. al. 1995), which is localized on the surface of keratinocytes thus mediating streptococcal binding to these cells and facilitating invasion through the epidermis (Fischetti 1991).

M proteins are variable in size. With the exception of the transmembrane/wall region and a short random coil sequence at the extreme N-terminus, M proteins contain a seven-residue periodicity that results in an α-helical coiled secondary structure. The coiled structure of M proteins is rare in
eubacteria, but surprisingly, common in eukaryotic cells. Significant levels of homology are observed with myosin heavy chain, for example.

The proximal C-terminus is highly conserved between different M proteins and is associated with binding. This conserved repetitive element is comprised of at least two copies of the approximately 23-amino-acid C-repeat sequences separated by short non-repetitive spacer regions. The evolutionary conservation of this C-repeat region implies that it is important (Fischetti et. al. 1990) and this region has been shown to facilitate adherence to keratinocytes in the cutaneous epithelium and to bind factor H (Fischetti 1991, Perez-Casal et. al. 1995). A considerable degree of primary sequence variation is permissive within the N-terminal half without compromising secondary structure integrity. This diversity forms the basis for serological typing and bacterial strains produce a unique M protein profile. Immunity to one serotype is not cross-protective in humans.

B) Hyaluronic Acid Capsule

The mucoid hyaluronic capsule has been demonstrated to participate in pharyngeal colonization and infection in mice (Wessels and Bronze 1994). In a murine throat colonization model, hyaluronic acid capsule was shown to confer a selective advantage for bacterial survival in the upper respiratory tract. A transposon generated hasA mutant showed significantly reduced throat colonization as compared to the wild type (Husmann et. al. 1997). The capsule is composed of at least three genes; hasA, hasB, and hasC, which form an operon. The first gene, hasA, encodes hyaluronate synthesis (DeAngelis et. al. 1993, Dougherty and van de Rijn 1992); hasB encodes UDP-glucose
dehydrogenase (Dougherty and van de Rijn 1993); and hasC possesses sequence homology to UDP-glucose pyrophosphorylase (Crater and van de Rijn 1995, Hynes et. al. 1995). Both encapsulated and non-encapsulated strains contain the intact has loci and expression of capsule appears to be transcriptionally regulated (Crater and van de Rijn 1995).

C) Serum Opacity Factor

Serum opacity factor (SOF) has also been implicated in cellular attachment and is a apoproteinase that cleaves high-density lipoprotein in sera and is produced by certain M serotypes (Kreikemeyer et. al. 1995). The molecule consists of three domains including one responsible for the opacity reaction, a fibronectin-binding domain, and cell attachment domain.

D) Fibronectin-Binding Protein

Fibronectin is a glycoprotein that can be found in a soluble form in most body fluids or associated with cellular surfaces or as a constituent of the extracellular matrix. Interestingly, fibronectin itself serves to bind indiscriminately to various cellular components including fibrin, collagens, integrins, DNA, heparin and other proteinaceous and non-protein substances (Hynes and Yamada 1982). Its ubiquity and broad spectrum of substrates makes fibronectin an ideal microbial receptor. Once fibronectin is sequestered by GAS, colonization and dissemination are facilitated. It has been shown that several streptococcal surface molecules bind fibronectin and that such an association plays a role in the ability of the organism to adhere to epithelial cells (Hasty et. al. 1992).
The cell surface liptoteichoic acid molecule (LTA) is able to bind to fibronectin and may serve as a ligand for bacterial attachment to a variety of substrata (Courtney et al. 1992). A two step model has been proposed whereby a bacterial LTA-binding protein anchors LTA oriented with its fatty acid moiety toward the streptococcal surface which then interacts with the fatty acid binding domains that are present on the membranes of many different cell types (Hasty et al. 1992). This constitutes the first step and is thought to be a reversible, weak hydrophobic interaction. The successful completion of the first step is thought to be necessary and facilitates the second step, which would involve a more specific and irreversible interaction between the M protein and a putative M protein receptor.

A separate factor has been described by Hanski and Caparon (1992) to bind fibronectin. This surface molecule called protein F is encoded by the prfF (prfF1) gene and adheres to respiratory cells (Hanski and Caparon 1992). The regulation of expression of protein F has been examined and was found to be environmentally regulated in response to alterations in atmosphere. Expression of protein F was repressed during growth under reduced oxygen concentrations. Expression in an anaerobic environment was induced by superoxide-generating and redox-altering reagents (Gibson and Caparon 1996). These data support the theory that environmental factors play a role in the initial success of GAS to colonize a particular host and that subcutaneous reduced oxygen tension contributes to colonization.
A novel prfF2 has been identified by Jaffe et. al. (1996). Insertional inactivation of prfF2 abolished its high-affinity fibronectin binding. Most strains that lack prfF1 (encoding protein F1), but still bind fibronectin with high affinity, possess prfF2-related genes suggesting that protein F2 is also a major fibronectin-binding protein in GAS. Protein F2 is highly homologous to fibronectin-binding proteins from Streptococcus dysgalactiae and Streptococcus equisimilis, particularly in its carboxy-terminal portion.

Still other fibronectin binding proteins have been described in GAS which may play a lesser but not-trivial role in attachment of streptococci to host targets. These include the streptococcal fibronectin binding (Sfbl) protein (Molinari et. al. 1997), serum opacity factor (SOF) protein (Raconjac et. al. 1995), and FBP54 (Courtney et. al. 1992 and 1994), a 28-kDa antigen (Courtney et. al. 1992), glyceraldehyde-3-phosphate dehydrogenase (G3PD) (Pancholi and Fischetti 1992), M3 protein (Schmidt et. al. 1993), and the protein-F independent ZOP (zinc, oxygen and pH) binding pathway (Lee and Caparon 1996). In this pathway oxygen is required for the modification of a protease-resistant surface structure in a pH-sensitive reaction which permits fibronectin binding in a zinc dependent manner (Lee and Caparon 1996).

The fibronectin binding domains of these proteins are highly similar to each other and consist of repeated motifs. In prfF, for example, one binding domain recognizes the N-terminal domain of the fibronectin molecule and consists of a 27-amino acid motif that is tandemly repeated up to six times (Sela et. al. 1993). The other binding domain is located immediately upstream to the
repeat binding domain and recognizes a region on the fibronectin molecule distinct from the repeat binding target (Sela et. al. 1993). Two domains are responsible for fibronectin binding by prF2. One domain (FBRD) consists of three consecutive repeats, whereas the other domain (UFBD) resides on a non-repeated stretch of approximately 100 amino acids and is located 100 amino acids upstream of FBRD. Each of these domains is capable of binding fibronectin when expressed as a separate protein (Jaffe et. al. 1996).

As evidenced, different functions may be ascribed to different adhesins and this may vary in different strains or with regulation influenced by environmental factors. For example, M protein mediates adhesion to keratinocytes but not to Langerhans’ cells (Okada et. al. 1994), whereas fibronectin-binding proteins are necessary for adherence to Langerhans’ cells (Okada et. al. 1994) but not keratinocytes (Okada et. al. 1995, Pancholi and Fischetti 1992). One common theme, however, is that environmental regulation may optimize colonization and infection by 'stimulating' certain adhesins while 'repressing' others, depending on the site of infection and the conditions prevailing at that site.

1.3.2) Factors involved in Replication and Dissemination

Several virulence factors enhance the ability of GAS to multiply, survive, and spread within the host (Table 1). Once colonization has occurred, systemic dissemination of GAS is a major contributing factor to the clinical manifestations of STSS and NF.
A) M and M-like Protein Mediated Replication and Dissemination

The family of M proteins bind immunoglobulins, specifically IgG, and organisms with this characteristic have been associated with the ability to invade tissue from a cutaneous site of infection in a mouse model (Akesson et. al. 1994, Podbielski et. al. 1996-B, Raeder and Boyle 1993-A and 1993-B). Further evidence of this comes from strains causing impetigo studied by Bessen and Fischetti that demonstrated IgG-binding properties (Bessen and Fischetti 1990).

M and M-like proteins also bind kininogen (Barry et. al. 1992), plasmin (Berge Sjobring 1993, Lottenberg et. al. 1994), and albumin (Accardo et. al., 1996, Ben Nasr et. al. 1995). Kininogens are classified as high molecular weight (H-) kininogen and low molecular weight (L-)kininogen. H-kininogen is part of the coagulation cascade and is activated in the presence of negatively charged surfaces, including bacterial surfaces (Lottenberg 1996). This may result in the release of kinins, strong pro-inflammatory agents, that mediate physiological manifestations such as fever, vasodilation, and increased vascular permeability. These changes could promote the dissemination of GAS to other host sites.

B) Cysteine Protease Mediated Replication and Dissemination

In addition, the streptococcal cysteine protease (SpeB) [section 1.3.4 A] has also been shown to release kinins from H-kininogen (Herwald et. al. 1996). Once at a target tissue, GAS may have greater chances of survival depending on the array of virulence factors expressed for a particular host target location. In this way, foci of survival ‘niches’ may account for the localized damage seen in some GAS infections.
C) Streptokinase Mediated Replication and Dissemination

The streptokinase molecule has been cloned and sequenced and is expressed by nearly all strains of GAS (Ball et al. 1995). Streptokinase is a 47 kDa single chain molecule consisting of 441 amino acids (1323 bp open reading frame) (Ball et al. 1995). The protein cleaves plasminogen to form plasmin, a serine protease, which digests fibrin and mediates clot dissolution (Lottenberg et al. 1992). Plasmin has also been reported to degrade the extracellular matrix and its production may serve to aid GAS in causing invasive infections (Lottenberg et al. 1992). Further, GAS must compete for plasmin with host regulators like $\alpha_2$-antiplasmin. Plasmin may be sequestered in one of two ways. The first is through direct binding with the bacterial cell surface either directly or through M proteins. The second involves a complex interaction between plasminogen, fibrinogen and streptokinase (Berge and Sjobring 1993, Kuusela et al. 1992, Wang et al. 1995-A). In this system, fibrinogen binds to the bacterial cell surface and provides a site for future binding of streptokinase-plasminogen complexes. Thus, the organism acquires a surface method of hydrolyzing fibrin clots and provides a plasminogen activator that is not inhibited by $\alpha_2$-antiplasmin (Wang et al. 1995-B). Using plasmin to mitigate clot formation at the site of entry or infection undoubtedly confers GAS with a tremendous advantage for spread to other sites.
1.3.3) Avoidance of Host Defense Mechanisms

The first line of host defense factors against invading pathogens are physical barriers such as the skin and mucous membranes. Once infection has occurred, the host’s non-specific immune response confers the next level of protection. This includes activated macrophages, mast cells, polymorphonuclear cells (PML) and their activating molecules, cytokines and complement factors. The complement system is non-specifically responsible for mediating phagocytosis and in turn bacterial eradication. Its efficacy depends, to a large extent, on a set of regulatory molecules which control complement activation. The ability of streptococci to escape phagocytosis is important in spreading infection to distal sites and GAS have evolved specialized mechanisms for avoiding these host lines of defense.

A) Antiphagocytic and Anticomplimentary Properties Mediated by M Protein, Hyaluronic Acid, Protein SIC and C5a Peptidase

M types 1, 3, 12, and 28 have been the most common types associated with STSS and contribute to invasiveness through the ability of M protein to hinder phagocytosis by human PMLs (Schlievert et. al. 1979). The antiphagocytic properties of M and M-like proteins have been demonstrated in numerous studies (Akesson et. al. 1994, Frick et. al. 1994, Podbielski et. al. 1996-B, Raeder and Boyle 1993-A and 1993-B).

The mechanism of M protein efficacy remains unclear, but there is evidence to suggest that M protein binds fibrinogen (Poirier et. al. 1989), or the complement control protein factor H (Horstmann et. al. 1988). The interaction of M protein with plasma fibrinogen masks C3b-binding sites on the bacterial
surface (Poirier et al. 1989). Furthermore, it has been proposed that when the streptococcus contacts serum, factor H binds to the M molecule and inhibits or reverses the formation of C3b,Bb complexes and helps to convert C3b to the inactive form (iC3b) on the bacterial surface, preventing C3b-dependent phagocytosis (Fischetti et al. 1995, Horstmann et al. 1988, Sharma and Pangburn 1997).

The mucoid hyaluronic acid capsule prevents recognition by phagocytes and may aid in survival within the host (Wessels et al. 1991). Acapsular mutants of GAS have been shown to be more readily phagocytosed by granulocytes and were less virulent in a mouse model than the wild type strain (Wessels et al. 1991 and 1994).

Three complement regulatory molecules are sequestered by the M and M-like proteins of GAS to interfere with the complement pathway: factor H; C4b-binding protein (C4BP); and MCP (Accardo et al. 1996, Okada et al. 1995, Them et al. 1995).

Recently, the newly described extracellular protein, SIC, produced by strains of serotype M1 and M57, was shown to confer GAS with the ability to resist complement-mediated lysis (Akesson et al. 1996).

C5a is one of the primary mediators of the alternate complement pathway which recruits neutrophils to foci of infection. The C5a peptidase (scpA) expressed by GAS is highly specific for C5a and cleaves the molecule at the neutrophil binding site (Cleary et al. 1992). The expression of C5a peptidase may also play a role in the pathway of dissemination of GAS because Ji et al.
(1996) found that scpA mutants localized to the lymph nodes, while wild type organisms spread to the spleen. This implicates C5a peptidase as a virulence factor which helps colonize the host by inhibiting the influx of PMLs, thereby impeding partial clearance of the streptococci.

1.3.4) Virulence Factors Associated With Damage to the Host

Several streptococcal virulence factors may be associated with damage to host tissues and systems either directly or indirectly. Some, such as the streptococcal superantigens [sections 1.3.4 A, 1.4 B] mediate host damage in both ways, while the cytotoxins [sections 1.3.4 B, 1.5 and 1.6] predominantly induce direct damage.

A) Superantigens and the Role of Streptococcal Pyrogenic Exotoxins

A family of proteins referred to as superantigens comprise a large group of virulence factors. These include the streptococcal pyrogenic exotoxins (SPEs) of which there are at least four known, SPE A, SPE B, SPE C, SPE F (MF), the streptococcal superantigen (SSA) and Streptococcus pyogenes mitogen (SPM). Although separate and distinct, these toxins may be functionally grouped into the superantigen family, that is, they are able to hyperstimulate T-cell responses through their ability to bind to both the outer groove of Class II major histocompatibility (MHC) molecules of antigen presenting cells (APCs) and the Vβ region of the T-cell receptor (TCR). This results in an immune disregulation leading to an enhanced production of proinflammatory cytokines, such as interferon γ (INF-γ), interleukin-6 (IL-6), and tumor necrosis factor α (TNF-α), capable of mediating shock and resulting in tissue injury (Kotb 1995, Mollick et al.)
Streptococcal toxic shock syndrome may be mediated via the ability of the SPEs or some as yet unidentified toxin/antigen to act as superantigens. SpeA and/or B is present in most cases of severe infection (Musser et. al. 1991) and SpeA has been shown to be expressed at roughly a fourfold higher level in cells grown at 37°C when compared to cells grown at 26°C. This suggests that toxin production may be favoured once colonization has reached the soft-tissues or bloodstream (Xu and Collins 1996). These exotoxins, with the exception of SPM, have been well characterized genetically (Weeks and Ferretti 1986, Hauser and Schlievert 1990, Norrby-Teglund et. al. 1994-A, Iwasaki 1993, Reda et. al. 1994).

The streptococcal pyrogenic exotoxin B (speB) is chromosomally encoded and conserved in all strains of GAS examined to date (Yu and Ferretti 1991). It serves as a cysteine protease and is also known as interleukin 1B convertase. The protease cleaves human fibronectin and vitronectin, two abundant cellular matrix proteins (Kapur et. al. 1993-A). It also cleaves human interleukin-1B precursor to generate mature interleukin-1B suggesting a critical role in inflammation and shock (Kapur et. al. 1993-B). Much evidence indicates that SpeB is a virulence factor in GAS infections (Musser et. al. 1991, Kapur et. al. 1994, Shanley et. al. 1996, Wolf et. al. 1994). SpeB also releases biologically active fragments of streptococcal surface proteins that suppress the chemotaxis of phagocytes by blocking C5a-mediated granulocyte migration (Berge and Bjorck 1995).
Table 1: Virulence factors of GAS: Determinants mediating adherence and colonization, replication and dissemination and avoidance of host defense mechanisms.

<table>
<thead>
<tr>
<th>Determinant</th>
<th>Characteristic</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Adherence and colonization</strong></td>
<td></td>
</tr>
<tr>
<td>M and M-like proteins</td>
<td>Binds to MCP on keratinocytes and fibronectin; attachment to different host cells.</td>
</tr>
<tr>
<td>Hyaluronic acid capsule</td>
<td>Pharyngeal colonization.</td>
</tr>
<tr>
<td>FBP54</td>
<td>Binds to Fibronectin/fibrinogen.</td>
</tr>
<tr>
<td>Serum opacity factor</td>
<td>Binds to Fibronectin.</td>
</tr>
<tr>
<td>28-kDa antigen</td>
<td>Binds to Fibronectin.</td>
</tr>
<tr>
<td>G3PD</td>
<td>Binds to Fibronectin.</td>
</tr>
<tr>
<td>ZOP binding pathway</td>
<td>Binds to Fibronectin.</td>
</tr>
<tr>
<td>Protein F, F2, and SFBI</td>
<td>Binds to Fibronectin.</td>
</tr>
<tr>
<td><strong>Replication and Dissemination</strong></td>
<td></td>
</tr>
<tr>
<td>M-protein</td>
<td>Binds to MCP on keratinocytes; releases kinins.</td>
</tr>
<tr>
<td>M and M-like proteins</td>
<td>Binds to Immunoglobulins.</td>
</tr>
<tr>
<td>Streptokinase</td>
<td>Converts plasminogen to plasmin.</td>
</tr>
<tr>
<td>SPE B (cysteine protease)</td>
<td>Releases kinins; cleaves fibronectin and degrades vitronectin.</td>
</tr>
<tr>
<td>Stfb1</td>
<td>Binds fibronectin on epithelial cells.</td>
</tr>
<tr>
<td><strong>Avoidance of Host Defense Mechanisms</strong></td>
<td></td>
</tr>
<tr>
<td>M protein</td>
<td>Antiphagocytic due to binding of fibrinogen; binds kininogens with release of kinins.</td>
</tr>
<tr>
<td>M and M-like proteins</td>
<td>Binds complement regulators: factor H; C4b binding protein; and membrane cofactor protein.</td>
</tr>
<tr>
<td>C5a peptidase</td>
<td>Inactivates C5a which inhibits chemotaxis of inflammatory cells.</td>
</tr>
<tr>
<td>SIC</td>
<td>Inhibits complement mediated lysis.</td>
</tr>
<tr>
<td>Hyaluronic acid capsule</td>
<td>Antiphagocytic.</td>
</tr>
<tr>
<td><strong>Factors Associated With Damage to the Host</strong></td>
<td></td>
</tr>
<tr>
<td>SPE B</td>
<td>Activates IL-1β, cleaves surface proteins from GAS which blocks neutrophil recruitment; releases kinins from H-kininogen.</td>
</tr>
<tr>
<td>Streptolysins S and O</td>
<td>Destruction of tissue and immune cells.</td>
</tr>
<tr>
<td>SPE A, B, C, and F</td>
<td>Superantigen mediated release of cytokines.</td>
</tr>
<tr>
<td>SSA, SPM</td>
<td>Superantigen mediated release of cytokines.</td>
</tr>
</tbody>
</table>
B) Cytotoxins

Toxins that damage cell membranes to such a degree that extensive leakage of intracellular constituents results in lysis and death of the cell are called cytolytic toxins or cytotoxins.

*Streptococcus pyogenes* elaborates two hemolysins, streptolysin O (SLO) and streptolysin S (SLS) which will be described in detail in sections 1.5 and 1.6. These cytotoxins are likely responsible for direct tissue damage to host cells through their pore forming mechanisms. The cytolytic properties of SLO and SLS may be involved in local tissue destruction seen in GAS infections. Moreover, GAS appear devoid of iron chelators or siderophores as shown by Eichenbaum *et al.* (1996) by their inability to acquire iron from human transferrin and lactoferrin. Cellular lysis mediated by SLO and SLS may contribute to the release of intracellular iron containing compounds such as heme, hemoglobin, and ferritin which GAS are able to utilize as an iron source.

1.4) GENETIC CHARACTERIZATION OF GAS VIRULENCE FACTORS

Having reviewed how the individual virulence factors fall within each functional category, it is appropriate, here, to introduce the genetic organization and control of expression of those which have been best characterized, such as the VIR regulon (*Vir*) (Figure 1) and the pyrogenic exotoxins.

A) The Group A Streptococcal Vir Regulon

The *vir* regulon (Figure 1) is a conserved region of chromosomal DNA encoding several homologous genes responsible for such virulence factors as the *mrp* (*fcrA*) gene which codes for an IgG-binding protein (Cleary *et al.* 1991-
B, Boyle 1995). This locus is followed downstream by the *emm* gene which codes for the coiled helical M protein with variable N-terminus.

Not shown are two loci downstream to the *emm* locus; the protein H (*sph*) gene which also binds fibronectin (Boyle 1995) followed by the SIC protein (*sic*) which interferes with complement-mediated cell lysis (Akesson *et. al.* 1996). Downstream of this is the complement inactivating C5a peptidase (*scpA*) gene (Chen and Cleary 1990). C5a peptidase was shown to be highly specific for the complement peptide C5a which is cleaved between the His-67 and Lys-68 residues, a region of the chemotaxin which binds to receptors located on the surface of polymorphonuclear leukocytes (Takahashi *et. al.* 1989). A similar locus and protein has been discovered in group B streptococci and named *scpB* and SCPB respectively (Cleary *et. al.* 1991-A).

The proteins and corresponding gene loci comprising the *vir* regulon discussed above were members of what was previously referred to as the M-like protein family that most likely arose by gene duplication and divergence (Boyle 1995). It was recently agreed that the gene located upstream of *fcrA* and previously called *mry* (for mRNA yield) or *VirR*, should be designated *mga* (multigene regulator of group A Streptococci) (Scott *et. al.* 1995). The product of this gene is a 62 kDa protein (Mga) which is a *trans*-acting positive regulator coordinately controlling the expression of the genes within the *vir* regulon (Caparon and Scott 1987, Kihlberg *et. al.* 1995, McLandsborough and Cleary 1995). This protein has the characteristics of a response regulator of a two-
Figure 1: Genetic organization of the VIR Regulon.

Figure 1: emm and emm-like genes are encoded next to each other at a chromosomal locus called the vir regulon. They are flanked by the C5a peptidase (scpA) and the regulatory gene mga. (Reprinted from Trends in Microbiology). OF = opacity factor
component transducing system (Chen et. al. 1993, Perez-Casal et. al. 1991, Podbielski et. al. 1995). To date, the sensory component has yet to be discovered. Environmental and biological factors that appear to influence Mga include pH, temperature, pCO$_2$, and [Fe$^{2+}$], and late growth phase (McIver et. al. 1995, Okada et. al. 1993, Podbielski et. al. 1992).

Other genes which are under the direct or indirect control of Mga include the fibronectin-binding serum opacity factor (sof) (Kreikemeyer et. al. 1995, Rakonjac et. al. 1995). The sof gene lies approximately 15 kb from the core VIR regulon. Genetically, OF$^+$ strains have diverged from OF$^-$ serotypes in that the vir regulon of most OF$^+$ strains contain a triplet of emm-like genes, unlike the structure of OF$^-$ strains (including strains of similar M-type), with some containing a single or double copy of emm or emm-like elements (Kehoe et. al. 1996). The streptococcal pyrogenic exotoxin B (speB), a precursor of an extracellular cysteine protease with many virulence mechanisms is also under the control of Mga (Berge and Bjork 1995, Kapur et. al. 1993-A and 1993-B, Podbielski et. al. 1996-A, Shanley et. al. 1996).

B) Genetic Characterization of the Streptococcal Pyrogenic Exotoxins

$SpeA$ is bacteriophage T12 encoded in $S. pyogenes$. It consists of 753 base pairs and codes for a 29 kDa protein that is excreted after leader-sequence cleavage as a 25 kDa peptide (Weeks and Ferretti 1986).

$SpeB$ is encoded for by a 1194 bp open reading frame which translates to 398 amino acids. Removal of the putative leader sequence results in a mature protein of 371 amino acids (Hauser and Schlievert 1990). The protein is
released extracellularly as an inactive zymogen of 40 kDa, which is converted to a 28 kDa (253 residue) active form by proteolysis or autocatalytic truncation in reducing conditions (Elliott and Dole 1947, Gerlach et. al. 1983).

SpeC, like SpeA, is phage encoded and hence mobile. It consists of 208 amino acids, with a calculated molecular weight of 24 kDa and shares homology with SpeA (Goshorn and Schlievert 1988).

SSA is encoded for by a 783 bp open reading frame that translates to a predicted 260 amino acid protein (Reda et. al. 1994). The deduced sequence is 60% identical to staphylococcal enterotoxin B (SEB) and 49% identical to SpeA. Not all strains possess the ssa locus and evidence suggests that it may recently have been acquired through horizontal gene transfer in GAS (Reda et. al. 1994).

The SpeF (MF) gene is 813 bp long encoding a protein of 271 amino acids that becomes truncated to 228 residues with a molecular weight of 25 kDa and shows no significant homology with any of the other SPEs (Iwasaki and Igarashi 1993).

It has only been in the last ten years that much of the genetic work on pathogenic streptococci has been carried out. It is fair to state that knowledge of the genetic mechanisms underlying pathogenicity in Gram positive organisms is lagging behind that of the Gram negative group. The difficulties lay, in part, with the fact that Gram positive organisms, and streptococci in particular, are difficult to manipulate genetically. Methods of introducing foreign DNA into GAS were laborious until the technique of electrotransformation became readily applicable in the mid 1980s (Harmandayan 1990). Electroporation itself is not well
standardized in streptococci because of the various technical parameters that affect transformation efficiency (Harmandayan 1990). Lastly, adequate protocols for extracting plasmids from GAS are lacking. Hence, there is an inherent and appreciable difficulty when working in streptococcal genetics, a predicament that is not trivial.

1.5) STREPTOLYSIN O (SLO)

Streptolysin O is an extracellular, cytolytic and immunogenic single chain protein which has been very well characterized biochemically. It serves as the prototype of a large group of cross-reactive hemolysins known as oxygen-labile or thiol-activated cytolysins which share mechanistic and physical properties.

In general, thiol-activated cytolysins are lethal and cardiotoxic. They are activated by thiol-reducing agents, are inactivated by small amounts of cholesterol (sterols) and are sensitive to oxygen. These toxins exhibit cross-neutralization by hyperimmune serum and share similar pH and temperature optima. The family also includes: θ-toxin (C. perfringens), cereolysin (B. cereus), pneumolysin (S. pneumoniae), tetanolysin (C. tetani), lysteriolysin (L. monocytogenes) (For reviews see Alouf and Geoffrey 1988, Halbert 1970 and Alouf 1980).

SLO is highly lytic for eukaryotic cells, especially erythrocytes. It has potent cardiotoxic effects and is lethal to laboratory animals. It does not produce β-hemolysis on the surface of blood agar plates, but does so if GAS are inoculated below the surface. Also, it is known that two different forms of the native SLO (nSLO) toxin exist, a high and low molecular weight, of which the low
Mr nSLO may be cleaved between the N-terminal Lys-77 and Leu-78. This site corresponds to an extremely sensitive cleavage site for the speB-derived streptococcal cysteine protease (Pinkney et al. 1995).

The membrane binding properties appear to be primarily cholesterol mediated and the toxin may have binding sites with different affinities (Alouf and Geoffrey 1988). Cell binding to the toxin is not saturable and the toxin is rapidly and quantitatively incorporated into target cell plasma membranes (Bhakdi et al. 1996). Initially the cytotoxin exhibits reversible monomeric binding at low and ambient temperatures and it has been shown that the C-terminus is critical for attachment as truncation of only two residues abrogates monomer binding to cell membranes (Bhakdi et al. 1996, Weller, unpublished results).

The actual mechanism of pore formation remains to be elucidated conclusively. However, it is known that aqueous transmembrane pores necessitate lateral lipid displacement which is always an ATP-independent process. The energy needed for such reactions is derived from conformational changes in the toxin molecules themselves. The general model proposed is as follows. The first phase involves the oligomerization of toxin protomers to form circularized protein complexes. Monomers bind to the membrane most likely followed by a rate limiting nucleation step involving the formation of dimers. The dimers act as crystallization points for the attachment of other monomers. SLO complexes grow into arcs and rings that represent the transmembrane pores (Palmer et al. 1995). Most likely the toxin molecules interact with cholesterol at the membrane surface to form primary toxin/cholesterol complexes. As these complexes form arc-ring structures at 37°C, it is believed that the toxin effectively sequesters cholesterol from certain regions
permitting a lipid phase transition to occur. This event leads to increased mobility of the hydrocarbon chains of the membrane phospholipids above their phase transition temperatures. A decrease in permeability and molecular membrane cohesion at the site leads to disruption and membrane fragmentation. SLO pores are the largest known measuring 30-35 nm in diameter with an inner diameter of 24 nm (Buckingham and Duncan 1983, Bhakdi et. al. 1985, Sekiya et. al. 1996).

The slo gene has been cloned and sequenced and consists of a 541 amino acid peptide with a molecular weight of 61 kDa (Kehoe and Timmis 1984, Kehoe et. al. 1987). At the C-terminal region there is a highly conserved stretch of hydrophobic residues involved in membrane binding and harbouring a single cysteine residue which confers amphiphilicity, but the toxin does not form large aggregates in solution. The N-terminal 70 amino acids are completely superfluous for pore-forming activity because they can be cleaved without loss of activity (Weller, unpublished observations).

It is interesting to note that the mechanism of SLO activation may not be consistent with toxins in the same family. When the codon for a single cysteine residue in SLO was changed to encode for either an alanine or serine residue, no significant difference in its hemolytic activity or sedimentation profiles was noted, suggesting that the mutations did not affect the ability of SLO to form oligomers in membranes (Pinkney et. al. 1989). The alanine-substituted SLO was no longer prone to inactivation by sulfhydryl-reactive agents or by oxidation. These data challenge the widely held assumption that an essential cysteine residue (thiol group) is necessary for in vitro activation.
1.6) STREPTOLYSIN S (SLS)

A) General Properties of SLS and Factors Needed for Formation

SLS is a poorly-understood, non-immunogenic and oxygen-stable protein elaborated by GAS which is also produced by strains of Lancefield groups C, G, E, H, L, but not strains of groups B and D. SLS results in very potent membrane-damaging effects on target cells including red blood cells, lymphocytes, polymorphonuclear leukocytes, platelets, tissue culture and tumor cells (For reviews see Alouf and Geoffrey 1988, Alouf 1980, Bernheimer 1972, and Ginsburg 1970). By weight it is one of the most potent cytotoxins known (Koyama 1963). Intracellular membranes are also damaged by SLS (Ginsburg 1970). Phospholipids inhibit the lytic efficacy of SLS suggesting that they are involved in membrane lysis. Interestingly, cholesterol which inhibits SLO is ineffective against SLS. Inhibitors of SLS include papain, chymotrypsin and trypan blue (Ginsburg et. al. 1965).

SLS can exist in many different forms including intracellular, extracellular, and cell bound (Ginsburg 1970). It is synthesized de novo and the majority of the toxin remains cell bound and can only be detected if a carrier molecule (inducer) is added to the culture. The inducers reported so far include the RNase-resistant fraction of yeast RNA (RNA core) and several non-ionic detergents such as Tween and Triton-X as well as bovine serum albumin (BSA). In addition, lipoteichoic acid has been implicated as a carrier (Theodore and Calandra 1981). Hemolytic activity can be transferred from one carrier to
another. Surprisingly, the carriers need not be chemically related (Duncan and Mason 1976, Theodore and Calandra 1981).

Bernheimer (1949) and Bernheimer and Rodbart (1948) showed that the formation of streptolysin S occurred at the resting stage in simple chemically defined phosphate buffer medium containing maltose and MgSO₄, in addition to the aforementioned carrier molecule. Taketo and Taketo (1964 and 1965) demonstrated that an extract of *S. pyogenes* after grinding, that is, a cell-free extract, yielded SLS with stabilization of the hemolysin by the carrier molecule. Those investigations also revealed that maltose was required and sufficient for the intracellular formation of SLS, in the absence of a carrier. Taken together these results suggest that maltose is necessary for intracellular SLS synthesis and that release into the medium is accomplished by the carrier. Calandra and coworkers have provided evidence that this intracellular hemolysin, which they named 'intracellular potential streptolysin S', was mainly localized to the cellular membrane, and that a carrier molecule facilitated its extraction (Calandra and Roger 1981, Calandra 1980, Calandra *et. al.* 1976, Calandra and Oginski 1975). Hence, the two known factors that were necessary for SLS production were the carrier molecule and maltose supplemented in the growth medium.

In addition a third factor was shown to be critical for production of SLS. Akao *et. al.* (1988) showed that neither the supernatant nor the precipitate alone obtained from sonicated cells yielded SLS; however, a mixture of supernatant and precipitate showed SLS activity. The precipitate could not be replaced, for example, from a different bacterial culture, but the supernatant from other
cultures could. The medium used was brain heart infusion broth (BHI) and a broth constituent was suspected of being another factor necessary for SLS formation (Akao et. al. 1988). This essential component was shown not to be lipid, was heat stable and filterable. Protein digests of BHI components had a stimulatory effect for SLS production and these peptides could not be replaced by carriers and maltose for SLS production of the precipitate. Thus, three inducer factors are necessary for maximal SLS formation: carrier, maltose and peptides. The last two are needed for intracellular production, the carrier is needed for possible stabilization and extracellular release (Akao et. al. 1988).

The essential peptide has been partially characterized and has disulfide bridges as essential structures (Akao et. al. 1992). Briefly, the SLS-inducing peptide was purified 185-fold from the pronase digest of BSA, and the streptococcal sonicate produced 61,900 hemolytic units of SLS after addition of 1µmol of the peptide in the presence of maltose and carrier. The molecular weight of the peptide was estimated to be approximately 1000 Da. The presence of disulfide bridges were inferred from the amino acid composition of the protein. The purified peptide showed the presence of aspartic acid, glutamic acid and leucine in a 1:1:3 molar ratio, in addition to 4 mol of cysteine after oxidation. Further, under reducing conditions, the 1000 Da peak broke down into smaller peaks thus suggesting that the peptide actually consists of three proteins linked with two disulfide bonds. In addition, the presence of the disulfide bonds is necessary for SLS formation as treatment of the peptide with reducing agents abrogated SLS production (Akao et. al. 1992).
B) Protein Chemistry of Streptolysin S

Evidence that the active principle is protein in nature as well as the amino acid composition has been reported by Koyama (1963). Confirmation, using proteases and the incorporation of radiolabeled amino acids, was demonstrated by Lai et. al. who also published an amino acid determination of the toxin (Lai et. al. 1978).

In the past, the purification of carrier-free SLS has never been successful as the peptide-inducer complex appear to form a tight interaction preventing the liberation of the carrier oligonucleotide from the SLS hemolytically active moiety without undergoing decay or denaturation. Lai et. al. (1978) have been able to improve the purification of SLS after inducing its production with yeast-soluble RNA. The estimated molecular weight of the protein was approximately 15 kDa containing 32 amino acid residues and nearly 7.1 kDa for the associated carrier molecule. Koyama partially purified an SLS-oligonucleotide mixture and found 12 different amino acids present in varying molar ratios (Koyama 1963). Bernheimer obtained an S20w of 2.4 suggesting MW of approximately 20 kDa (Bernheimer 1972). The ratio of peptide to nucleic acid was estimated at 0.3 by Koyama giving a MW of near 2800-3000 Da or approximately 28 amino acids (100-110 Da/residue). The sizes (of SLS-carrier complexes) as well as the numbers of amino acid residues reported by Lai are somewhat in agreement with those reported by Bernheimer based on Koyama's data and molecular weight determination.
More recently, Alouf et. al. successfully separated a hemolytically active peptide of about 1800 Da. which seems more consistent with the expected amino acid composition (Alouf and Geoffroy 1988). This molecule was believed to be SLS in its pure form and exhibited a pI of 9.2 (3.6 for SLS-RNA core complex). A preliminary attempt to sequence purified SLS failed as did amino terminal determination by Edman degradation, suggesting that the amino terminal may have been blocked, or that the peptide moiety was cyclic (Alouf, personal communication).

Notwithstanding the consensus among Lai, Koyama and Alouf regarding the lack of cysteine in the SLS protein and the proportions of certain residues, there appears to be a discrepancy in either the proportions, or even the presence, of others. Care should be taken in the interpretation of such data as earlier work by Lai et. al. (1978) and Koyama (1963) revealed the presence of contaminating protein(s) in RNA-core used as a carrier. The definitive amino acid composition and sequence will be revealed when the structural gene for SLS is identified.

The mechanism of action of SLS remains to be understood. However, the fact that certain phospholipids can inhibit SLS action suggests that SLS interacts with phospholipids in the cellular membrane. Moreover, SLS does not appear to possess phospholipase activity (Elias et. al. 1966, Okazaki 1971). A slightly different mechanism of action has been reviewed by Bernheimer (1972) that postulated that SLS may alter lipids in membranes such that the cell could no longer regulate a proper ion exchange across the membrane. The consequence
of such a perturbed osmotic balance would cause the cell to expand with inevitable membrane rupture. To this end Duncan and Mason (1976) have shown evidence that supported this model. Rabbit erythrocytes pre-labeled with $^{86}\text{Rb}^+$ that were co-incubated with active SLS showed that the loss of intracellular $^{86}\text{Rb}^+$ preceded the escape of intracellular hemoglobin by several minutes so that by the time 4% of hemoglobin was released, over 50% of $^{86}\text{Rb}^+$ was released. In addition, further experiments revealed that the majority of SLS toxin did not irreversibly bind to cellular targets, but that active supernatant fractions could be recycled several times (Duncan and Mason 1976). This is in contrast to Elias' et. al. (1966) findings that 75% of SLS' hemolytic activity was abrogated after a ten minute incubation with red blood cell ghosts. The reason for this discrepancy is not known, however, once the gene for SLS is identified and characterized, studies using the purified recombinant toxin should serve to resolve discrepancies in the biochemical data.

C) Evidence for Streptolysin S Serving a Virulence Factor in GAS

Among the virulence factors of GAS, the role of streptolysin S in pathogenesis has not been elucidated fully. This may be due to the fact that although SLS's effect in vitro is known (albeit generally) in laboratory assays, its effect in vivo remains to be determined. Furthermore, the SLS protein is non-immunogenic, as antibodies neither have been experimentally generated nor detected in sera from patients with GAS infections and the lack of successful antibody production against the protein prevents immunochemical investigation.
Most clinical isolates of GAS produce SLS. Occasionally SLS\textsuperscript{-} strains have been isolated from patients with pharyngitis. Naturally occurring strains of SLS\textsuperscript{-} GAS are rare and although such clinical isolates occur, they often are not recognized because non-hemolytic mutants from throat swabs are ignored in the clinical laboratory and dismissed as \textit{viridans} streptococci (Nida and Cleary 1983).

It is likely that SLS does have a role as a virulence factor as demonstrated in studies using SLS deficient mutants. Owens \textit{et al.} (1978) used non-specific chemical mutagenesis to create SLS deficient strains. These SLS deficient mutants showed reduced virulence when injected intraperitoneally into mice. However, the SLS mutant demonstrated growth differences in relation to the wildtype parent suggesting that the expression of other genes may also have been altered (Owens \textit{et al.} 1978).

Previous studies in this laboratory provide stronger evidence implicating SLS as a virulence factor, namely that Tn916 generated hemolytic deficient mutants (SBNH5) showed reduced virulence by changes in weight and fewer necrotic lesions when inoculated subcutaneously in a murine model than did the wildtype MGAS166 [see section 1.7.]. Mice challenged with the wildtype strain depicted, on average, a 1.3 gram decrease in weight versus a 1.0-1.1 gram increase (p<0.05) in weight when challenged by the mutant strain over a 24 hour period (Betschel \textit{et al.} 1997).

Although there is evidence that SLS is likely a virulence factor associated with necrosis, the paucity of accurate biochemical data and the lack of genetic characterization necessitate further work on SLS. Thus, the purpose of the
current investigation was to better characterize this virulence factor at the molecular level of gene expression by identifying the gene(s) responsible for SLS and/or its production in GAS.

2.0) OBJECTIVES

Previously in our laboratory, Stephen Betschel generated an *S. pyogenes* strain, MGAS166, devoid of SLS production by insertional inactivation using the tetracycline encoding conjugative transposon Tn916. Briefly, Tn916 was transferred from *Enterococcus faecalis* CG110 into MGAS166 by mating and transconjugants were screened phenotypically for abrogation of the SLS phenotype (Figure 2). One SLS deficient transconjugant (SBNH5) containing a single copy of the transposon was identified. Subsequently, a 3.8 kb region flanking the Tn916 insertion was excised and cloned into the low copy number plasmid pACYC184 (Betschel et. al. 1993).

Essentially, the specific aims are as follows: 1) to obtain and examine the sequence of the 3.8 kb insert and identify the area of Tn916 transposition; 2) to identify putative open reading frames (ORFs) and analyze them in the wild-type (WT) and mutant strains; and 3) to perform complementation studies on the hemolytic deficient mutant to observe if phenotypic restoration of the hemolytic trait occurs.
Figure 2: Hemolytic pattern of MGAS166 and the isogenic hemolytic deficient mutant, SBNH5, after Tn916 mutagenesis.

Panel A: WT MGAS166 showing clear zones of β-hemolysis.

Panel B: Mutant SBNH5 devoid of β-hemolysis after Tn916 insertion.
MATERIALS AND METHODS

3.1) Bacterial Strains, Plasmids and Culture Conditions

Strains used in this investigation are listed in Table 2. Strains were grown in Todd-Hewitt broth (Oxoid, Basingstoke, England) or on Columbia agar (Oxoid) plates containing 5% defibrinated sheep blood (Quelab Laboratories, Quebec, PQ). When antibiotic selection was required, 2000 μg/ml streptomycin (Sigma Laboratories, St. Louis, MO) and 5 μg/ml tetracycline (Sigma) were added to the appropriate media and filter sterilized. Strains T18P, MGAS166, and CS91-23 were kindly provided by Drs. Patrick Schlievert, (University of Minnesota, Minneapolis, MN), James Musser (Baylor College of Medicine, Houston, TX) and Patrick Cleary (University of Minnesota, Minneapolis, MN) respectively. Plasmid pLZ12-Km was kindly provided by Dr. Michael Caparon (Washington University School of Medicine, St. Louis, MO).

3.2) DNA Manipulations (Extraction, Restriction, Electrophoresis)

To confirm the insertional inactivation of SBNH5 by Tn916, genomic DNA was isolated as follows from mutant SBNH5 and wildtype MGAS166 strains. Essentially, 1.4 mL of overnight culture were centrifuged at 14,000 rpm in a microfuge (Eppendorf) to pellet cells. Cells were washed with 1.0 mL lysis buffer (50mM glucose; 25 mM Tris, pH 8.0; 10 mM EDTA, pH 8.0; 150 mM NaCl) and incubated in 0.5 mL of lysis mixture (lysis buffer plus 10 μL DNase-free RNAse [10 mg/mL] (Sigma); 50μL mutanolysin [1 mg/mL] (Sigma); 50 μL lysozyme...
### Table 2: Plasmids and strains used in this investigation:

<table>
<thead>
<tr>
<th>Strain or plasmid</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>MGAS166</td>
<td>M1T1, SLS&lt;sup&gt;+&lt;/sup&gt;, resistant to streptomycin at 2000 μg/mL. (Musser et. al. 1993)</td>
</tr>
<tr>
<td>SBNH5</td>
<td>Tn916 generated mutant of MGAS166 devoid of SLS production. (Betschel et. al. 1993)</td>
</tr>
<tr>
<td>ATCC 27762</td>
<td>ATCC strain producing SLO, but not SLS. (Ginsburg 1970)</td>
</tr>
<tr>
<td>CS91-23</td>
<td>Tn916 generated hemolytic deficient mutant. (Nida and Cleary 1983)</td>
</tr>
<tr>
<td>T18-P</td>
<td>Clinical M18, SLS&lt;sup&gt;+&lt;/sup&gt;, resistant to streptomycin at 2000 μg/mL. (Schlievert et. al. 1977)</td>
</tr>
<tr>
<td>SB30-2</td>
<td>Tn916 generated hemolytic deficient mutant of T18-P (Betschel et. al. 1993)</td>
</tr>
<tr>
<td>GAS clinical isolates</td>
<td>20 strains of varying M-types, 4 from NF, 3 from STSS cases. See Table 6 (Mt Sinai Hospital clinical strains collection, Toronto, Canada)</td>
</tr>
<tr>
<td>pLZ12-S1</td>
<td>pLZ12-Km containing a 3.8 kb region responsible for SLS production subcloned into the Km resistance gene at the HindIII site. (Hanski et. al. 1992-B)</td>
</tr>
<tr>
<td>pKmobsacB</td>
<td>Conjugative suicide vector (Schafer et. al. 1994).</td>
</tr>
<tr>
<td>Escherichia coli DH5α</td>
<td>Laboratory strain used to propagate pLZ12-S1.</td>
</tr>
</tbody>
</table>
[10mg/mL] (Sigma) at 37°C for at least one hour. This was followed by addition of 30 μL of 10% sodium-dodecyl-sulfate (SDS) (Sigma) and the lysate was mixed by inversion and kept at room temperature for 10 minutes. After the addition of 50 μL of proteinase K (Sigma), the lysate was incubated at 45°C for one hour and 100μL of 5M NaCl was added followed by vigorous shaking. Finally, 80 μL of 10% cetyltrimethylammonium bromide (CTAB) (Sigma) was added and after thorough mixing, the mixture was incubated at 65°C for 10 minutes. The aqueous phase (approximately 450 μL) was then extracted twice with phenol/chloroform (25:24) and once with chloroform. DNA was precipitated with 1.0 mL 95% ethanol and 1/10 volume 3M sodium acetate (Sigma) at -20°C for 30 minutes. DNA was pelleted by centrifugation at 14,000 rpm washed with 1.0 mL 70% ethanol, air dried, and resuspended in 50 μL sterile distilled water. DNA was restricted with HindIII (Boeringer Mannheim, Laval, PQ) for at least 5 hours at 37°C, subjected to 0.8% agarose gel electrophoresis (with 5 μg/ml ethidium bromide) (Bio-Rad, Mississauga, Ontario) and visualized on a UV transiluminator.

3.3) Southern Hybridization Analysis

DNA was transferred to Hybond N+ nylon membranes (Amersham, Oakville, ON) overnight and probed with the appropriate probes labeled using the enhanced chemiluminescence (ECL) direct labeling system (Amersham) according to the instructions of the manufacturer.

3.4) DNA mapping
The 3.8 kb fragment cloned into pACYC184 was excised with HindIII and digested with four restriction enzymes, Clal, BamHI, Xbal and EcoRV (Boeringer Mannheim) for at least 1.5 hours at 37°C and subjected to 0.8% agarose gel electrophoresis in order to determine the orientation of the insert.

3.5) DNA Sequencing

The entire 3.8 kb was sequenced commercially (Mobix Inc., Hamilton, Ont.) using the Applied Biosystems automated sequencer according to the Sanger dideoxy method. Primers were constructed (Mount Sinai Hospital Research Institute) from the flanking regions of the insert on pACYC184 and sequencing was carried out from both ends; new primers were constructed internally, whenever necessary, to continue sequencing (Table 3). One region near an open reading frame (ORF), named ORF 5 (see below), was sequenced twice to confirm the nucleotide sequence in that region.
<table>
<thead>
<tr>
<th>Primer name</th>
<th>Sequence (5' - 3')</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>AB6912</td>
<td>TGTAGCACCTGAAGTCA</td>
<td>Sequencing left from pACYC184 into 3.8 kb sequence</td>
</tr>
<tr>
<td>AB6913</td>
<td>AGATTTCTACACGGTG</td>
<td>Sequencing right from pACYC184 into 3.8 kb sequence</td>
</tr>
<tr>
<td>s341</td>
<td>GATCAAGCAACAGCTGC</td>
<td>Continuation of 3.8 kb sequence</td>
</tr>
<tr>
<td>s324</td>
<td>GGCCTTAGGAGTTATGTC</td>
<td>Continuation of 3.8 kb sequence</td>
</tr>
<tr>
<td>s413</td>
<td>TTTAAATCAGTCAGCAGCG</td>
<td>Continuation of 3.8 kb sequence</td>
</tr>
<tr>
<td>s1244</td>
<td>TGATGGTTTACACATAG</td>
<td>Continuation of 3.8 kb sequence Amplification of ORF 5 (539 bp)</td>
</tr>
<tr>
<td>sr1440</td>
<td>ATGTGCTAGGCTTTCGTC</td>
<td>Continuation of 3.8 kb sequence</td>
</tr>
<tr>
<td>a393</td>
<td>GCAGACAGATGTTAGTT</td>
<td>Continuation of 3.8 kb sequence</td>
</tr>
<tr>
<td>a317</td>
<td>GGTTAAGGGAGTAGCCA</td>
<td>Continuation of 3.8 kb sequence</td>
</tr>
<tr>
<td>a348</td>
<td>GAATCAGGGCGACTGCCT</td>
<td>Continuation of 3.8 kb sequence</td>
</tr>
<tr>
<td>a405</td>
<td>CCAAGATATAAGCTGC</td>
<td>Continuation of 3.8 kb sequence</td>
</tr>
<tr>
<td>a474</td>
<td>GGCATAAGGTGTTAGA</td>
<td>Continuation of 3.8 kb sequence Amplification of ORF 5</td>
</tr>
<tr>
<td>s1652</td>
<td>GTAGCGGAAGTTATACG</td>
<td>Amplification of ORF 1 (260 bp)</td>
</tr>
<tr>
<td>a1912</td>
<td>ACTACCTTCTCAGTTCCTC</td>
<td>Amplification of ORF 1</td>
</tr>
<tr>
<td>s1855</td>
<td>TACTAGATGTACCTGC</td>
<td>Amplification of ORF 2 (469 bp)</td>
</tr>
<tr>
<td>a2324</td>
<td>CTATGTGAAAATTGACG</td>
<td>Amplification of ORF 2</td>
</tr>
<tr>
<td>s2322</td>
<td>TAGACAAATGCGCTCTTC</td>
<td>Amplification of ORF 3 (369 bp)</td>
</tr>
<tr>
<td>a2691</td>
<td>AAAGCATAGGCGACTGC</td>
<td>Amplification of ORF 3</td>
</tr>
<tr>
<td>s1111</td>
<td>ATGTTATCGTTGAAACG</td>
<td>Amplification of ORF 4 (333 bp)</td>
</tr>
<tr>
<td>a1444</td>
<td>ACAAGGACAAAGCTAGC</td>
<td>Amplification of ORF 4</td>
</tr>
</tbody>
</table>
### 3.6) Sequence Analyses

Since the fragment was cloned into pACYC184 using only one restriction enzyme, its orientation in the plasmid may be either 'coding' or 'non-coding'. Thus, the 'sense' and 'antisense' strands were sequenced and database search analyses were done in both directions with both strands. Analysis of sequence data was carried out using the Wisconsin GCG sequence analysis program as well as the FASTA algorithm and BLAST search engines available at the National Centre for Biotechnology Institute (NCBI). The DNA sequence was also compared to the GAS Streptococcal genome sequencing project at the University of Oklahoma (McShan et. al. 1997). Open reading frame (ORF) analysis was carried out using the Wisconsin GCG program.

### 3.7) Determination of the Precise Insertion Point of Tn916

Four PCR reactions were set up using known left and right outward primers from Tn916 and left and right outward primers from the 3.8 kb fragment. Chromosomal DNA (20 ng) from SBNH5 extracted as already described was used as a template. PCR reactions were set up containing 200 μM dNTPs (Pharmacia Biotech, Montreal, PQ), 1U Taq polymerase (Perkin Elmer), 1X Taq reaction buffer containing Mg²⁺ (Perkin Elmer) and 20 pmol of each primer per reaction in a total volume of 50 μL for a total of 40 cycles. Cycling conditions were as follows: 1 cycle of 94°C for 4 minutes, 40 cycles of 94°C for 30 seconds, 45°C for 30 seconds, 72°C for 1 minute and a final cycle of 72°C for 10 minutes. Amplicons were subjected to 0.8% agarose gel electrophoresis and were purified from the gel using the Qiaex II Gel Extraction Kit (Qiagen,
Chatsworth, CA) according to the manufacturer's instructions. The purified product was sequenced as described and sequence data was compared by 'best fit' analysis against SBNH5 genomic DNA.

3.8) RNA Analyses of ORFs Within the 3.8 kb Fragment

To investigate the coding capacity of the putative ORFs within the 3.8 kb region, probes based on the five selected ORFs were generated by PCR using appropriate primers (Table 3) and used in Northern blotting experiments. Total RNA from strains MGAS166, SBNH5, ATCC 27762, CS91-23, T18P and SB30-2 was extracted using Trizol (Gibco BRL) according to the recommendations of the manufacturer.

Briefly, a 100 μL volume of bacteria was mixed with an equal volume of diethylpyrocarbonate (DEPC)-treated water containing 0.5 mg/mL mutanolysin and 1 mg/mL lysozyme. The mixture was incubated at 37°C and after 30 minutes, 1.0 mL Trizol reagent was added followed by a 5 minute incubation at room temperature. Chloroform (200μL) was then added to each tube and the mixture was vortexed for 1 minute before centrifugation at 12,000 g for 15 minutes at 4°C. After centrifugation, the aqueous phase was removed from the tubes, 0.5 mL isopropanol was added and the tubes were incubated at room temperature for 10 minutes before centrifugation at 12,000 g at 4°C. The supernatant was discarded, the pellet containing RNA was washed with 75% ethanol, and centrifuged at 7,500 g for 5 minutes at 4°C. After removal of the ethanol, the RNA pellet was air dried and dissolved in 100 μL of DEPC-treated water. All solutions and equipment were rendered RNase free according to
accepted recommendations ('Current Protocols' 1997). RNA was isolated at mid-log in the growth phase and every two hours thereafter up to a maximum of ten hours, quantified spectrophotometrically, standardized and loaded onto 1.9% formaldehyde/agarose gels.

RNA electrophoresis and Northern blot transfer were performed as described in 'Current Protocols'. Gels were run at 80 mA for 2-3 hours and transferred by capillary action onto Hybond N+ nylon membranes (Amersham) overnight and probed with either the entire 3.8 kb fragment or individual ORFs that were α32PdCTP-labeled (NCI, Costa Mesa, CA) using the 'Ready-to-Go DNA labeling beads' (Pharmacia Biotech) according to the recommendations of the manufacturer. The integrity of RNA was checked by simultaneously probing all strains with a conserved 16S rRNA sequence.

3.9) Subcloning the 3.8 kb Fragment Into DH5α for Complementation Studies

The 3.8 kb fragment was excised from pACYC184 using HindIII, purified as described above, and subcloned into the HindIII site within the kanamycin resistance gene of the shuttle vector pLZ12-Km. Plasmid pLZ12-Km is approximately 5.6 kb in length and also contains a chloramphenicol and kanamycin resistance gene. Prior to ligation the vector (usually 5 μg) was dephosphorylated by incubating with five units of calf intestinal alkaline phosphatase (CIAP) to prevent self-ligation and 1X buffer (Boeringer Mannheim) at 37°C for half an hour. Ligation was carried out in a 20 μL reaction volume
containing 20 ng vector DNA; 5X molar excess of insert DNA; 2μL of 10X T4 DNA ligase buffer (660 mM Tris-HCl, 50 mM MgCl₂, 10 mM dithioerythritol (DTT), 10 mM ATP pH 7.5); and 1 unit of T4 ligase (Boeringer Mannheim). The ligation reactions were incubated at 16°C for 14 hours and stored at -20°C until used for transformation.

Transformation was carried out using CaCl₂ competent E. coli DH5α cells. Briefly, E. coli were made competent as follows: 25 mL of cells were grown to mid-log phase (OD₆₅₀ = 0.4), shaken at 37°C, and centrifuged at 5,000 rpm for 5 minutes. The supernatant was discarded and the cells were gently resuspended in 12.5 mL ice cold 50 mM CaCl₂ and left on ice for 40 minutes. Cells were respun as above and resuspended in 1.2 mL of 50 mM CaCl₂.

Transformation was carried out in polypropylene tubes (Falcon) as follows: 200 μL of chilled competent cells were aliquoted into the tubes and half the reaction volume of the ligation mixtures above were added and gently mixed by pipetting. The cells were then incubated for 30 minutes on ice followed by a brief 45 second heat shock in a 45°C water bath. Cells were then chilled on ice for 2 minutes and rescued in 1.0 mL of room temperature SOC medium for 1.5 hours with shaking. After rescue the cells were diluted and selected for growth on chloramphenicol plates (15 μg/ml). Confirmation of construct uptake was done on kanamycin plates (25 μg/ml). Clones that were positive phenotypically were checked for the presence of the 3.8 kb insert by plasmid extraction as previously described (Birnboim and Doly 1979), agarose electrophoresis and
Southern blotting. A clone containing the 3.8 kb insert was designated pLZ12-S1 and chosen for further work.

3.10) Electroporation of pLZ12-S1 into GAS CS91-23

Electroporation was carried out essentially as described by Caparon et. al. (1991). Briefly, GAS CS91-23 was grown overnight in Todd-Hewitt Broth (Oxoid) supplemented with 0.5% yeast extract (Difco) and 10mM glycine (Sigma). Cells were washed three times in cold 15% glycerol and 200 μL of cells were placed in a 0.2 cm electroporation cuvette (BioRad). A minimum of 1 μg of pLZ12-S1 was added and cells were subjected to 2.5 kV, 100Ω using the BioRad Gene Pulser. The mixture was rescued in non-selective Todd-Hewitt Broth for 2 hours at 37°C and plated on Columbia Base agar supplemented with 5% sheep blood, tetracycline (5 μg/ml) and chloramphenicol (1 μg/ml). Transconjugants were identified by selecting for colonies that had a restored β-hemolytic phenotype.

3.11) Restriction Enzyme Analysis (REA)

To examine the prevalence and location of the 3.8 kb fragment in a representative sample of GAS clinical specimens, a total of 20 unrelated GAS isolates (Table 4), 17 from Ontario and 3 from out of province were analyzed by restriction digestion as follows. Nine strains were cut with Smal and resolved by pulsed field gel electrophoresis (PFGE) essentially as described by Murray et. al. (1990). Chromosomal DNA from the remaining 11 strains including MGAS166 was extracted and restricted using HindIII as described above. DNA was transferred to nylon membranes and probed with the labeled 3.8 kb fragment.
containing the area responsible for SLS production by Southern blotting as described previously.
Table 4: GAS clinical strains used in identifying presence of 3.8 kb fragment responsible for SLS production.

<table>
<thead>
<tr>
<th>Code #</th>
<th>M</th>
<th>Case Definition</th>
<th>NF Class</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Analyzed by conventional electrophoresis</td>
<td></td>
</tr>
<tr>
<td>5640</td>
<td>3</td>
<td>Invasive</td>
<td>No</td>
</tr>
<tr>
<td>5811</td>
<td>4</td>
<td>Invasive</td>
<td>No</td>
</tr>
<tr>
<td>5657</td>
<td>6</td>
<td>Invasive</td>
<td>No</td>
</tr>
<tr>
<td>5454</td>
<td>NT</td>
<td>STSS</td>
<td>No</td>
</tr>
<tr>
<td>5619</td>
<td>1</td>
<td>STSS</td>
<td>No</td>
</tr>
<tr>
<td>5385</td>
<td>28</td>
<td>Invasive</td>
<td>No</td>
</tr>
<tr>
<td>5711</td>
<td>NT</td>
<td>Invasive</td>
<td>No</td>
</tr>
<tr>
<td>5367</td>
<td>11</td>
<td>Invasive</td>
<td>No</td>
</tr>
<tr>
<td>5661</td>
<td>12</td>
<td>Invasive</td>
<td>No</td>
</tr>
<tr>
<td>6079 (BC)</td>
<td>1</td>
<td>STSS</td>
<td>No</td>
</tr>
<tr>
<td>6079 (W)</td>
<td>1</td>
<td>STSS</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Analyzed by Pulsed Field Gel Electrophoresis</td>
<td></td>
</tr>
<tr>
<td>5809</td>
<td>12</td>
<td>STSS</td>
<td>Yes</td>
</tr>
<tr>
<td>5809-1</td>
<td></td>
<td>Health care worker contact to 5809, NA</td>
<td></td>
</tr>
<tr>
<td>5786</td>
<td>1</td>
<td>Invasive</td>
<td>Yes</td>
</tr>
<tr>
<td>5787</td>
<td>1</td>
<td>Non-Invasive</td>
<td>Yes</td>
</tr>
<tr>
<td>5823</td>
<td>1</td>
<td>Invasive</td>
<td>No</td>
</tr>
<tr>
<td>5823-A</td>
<td></td>
<td>Hospital roommate to 5823, no information available</td>
<td></td>
</tr>
<tr>
<td>5835</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>6060</td>
<td>3</td>
<td>Invasive</td>
<td>No</td>
</tr>
<tr>
<td>6078</td>
<td>4</td>
<td>Invasive</td>
<td>Yes</td>
</tr>
</tbody>
</table>

STSS: Toxic Shock by definition
NF: Necrotizing Fasciitis
NA = Data not available
RESULTS

4.1) Confirmation of Tn916 Insertion in SBNH5

Tn916 contains a single HindIII restriction site. Cleavage at this site divides the transposon into two fragments of 6 kb and 12 kb (Clewell and Gawron-Burke 1986). After HindIII digestion of mutant SBNH5 chromosomal DNA, each copy of Tn916 which has integrated into the chromosome should yield two bands that hybridize with the 3.8 kb probe. To confirm that the Tn916 insertion into the SBNH5 chromosome was in the same locus as the 3.8 kb fragment cloned from the MGAS166 parental strain, SBNH5 genomic DNA restricted with HindIII revealed the presence of two bands (7.7 and >12 kb) when probed with the 3.8 kb insert. The parent strain, MGAS166, revealed only one band at 3.8 kb (Figure 3).

4.2) DNA Mapping

Other than the fact that the 3.8 kb fragment from MGAS166 cloned into pACYC184 was responsible for SLS expression, nothing was known about the fragment when this work was begun. Consequently, a partial restriction map was obtained to initially determine the orientation of the 3.8 kb fragment in pACYC184. Restriction enzyme analysis was carried out using Clal and BamHI which do not cut within the fragment. Xbal and EcoRV have cut sites 2956 and 2264 base pairs from the 5’ end respectively, indicating insertion from 5’ to 3’ by our sequencing convention. These sites were subsequently confirmed by the sequence data [see section 4.3].
Figure 3: Confirmation of Tn916 insertion into MGAS166 to generate SBNH5

Lane 2 shows SBNH5 genomic DNA restricted with HindIII and probed by Southern blotting with the 3.8 kb insert. Two bands are visible at 6.7 kb and >12 kb due to the internal HindIII site within Tn916. Lane 3 shows the corresponding fragment at 3.8 kb in the MGAS166 WT. Lane 1 is the 1 kb marker.
4.3) DNA Sequence Analyses

Sequencing was carried out on the 3.8 kb fragment from both ends. Sequencing revealed that the 3.8 kb region, thought to contain the gene responsible for SLS production, was actually 3,732 bp. Searches conducted using both the FASTA and BLAST algorithms respectively, revealed no exact homologous matches, supporting the claim that the sequence is novel. However, sequence similarity with DNA or protein sequences from various species, mainly within the streptococcus and related genera, was observed using the 5'-3' sequence (Table 5). Sequence similarity (55% in 318 bp) with the epf gene, an extracellular virulence marker, in *Streptococcus suis*, and with the M protein gene of group G Streptococcus (52.4% in 328 bp) as well as an M 12 type GAS was also noted. In addition, some homology (56.5% in 124 bp) with perfringolysin O, the thiol-activated hemolysin from *Clostridium perfringens* was reported indicating overlap with a toxin of similar cytolytic properties (Shimitsu *et al.* 1991).

The entire genome of a representative M-type 1 GAS has been mapped by pulsed field gel electrophoresis (Suvarov and Ferretti 1996). Several virulence factors have been localized to specific chromosomal areas. More notably, a cluster of virulence genes can be localized on the physical map to less than 25% of the entire genome; this area consists of the 400 kb and 120 kb fragments of the *Sal* digested chromosome. In addition, the University of Oklahoma has sequenced approximately 95% of a representative M1 GAS chromosome,
Table 5: Results of homology searches using the BLAST algorithm at the NCBI.

<table>
<thead>
<tr>
<th>Organism sharing homology</th>
<th>Gene/Protein</th>
<th>Degree of Homology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Streptococcus mutans</td>
<td>IS199*</td>
<td>80.6% (in 180 bp)</td>
</tr>
<tr>
<td>Streptococcus thermophilus</td>
<td>IS1191</td>
<td>57.3% (in 213 bp)</td>
</tr>
<tr>
<td>Lactobacillus helviticus</td>
<td>transposase</td>
<td>78% (33/42 a.a.)</td>
</tr>
<tr>
<td>Lactococcus lactis</td>
<td>IS905</td>
<td>73% (in 74 bp)</td>
</tr>
<tr>
<td>S. thermophilus</td>
<td>transposase</td>
<td>67% (29/43 a.a.)</td>
</tr>
<tr>
<td>Streptococcus suis</td>
<td>epf</td>
<td>55% (in 318 bp)</td>
</tr>
<tr>
<td>S. suis</td>
<td>EF protein</td>
<td>45%-52%</td>
</tr>
<tr>
<td>Group G streptococcus</td>
<td>emm</td>
<td>52.4% (in 324 bp)</td>
</tr>
<tr>
<td>S. pyogenes</td>
<td>emm55</td>
<td>54.6% (in 174 bp)</td>
</tr>
<tr>
<td>Clostridium perfringens</td>
<td>perfringolysin</td>
<td>56.5% (in 124)</td>
</tr>
</tbody>
</table>

* IS = Insertion Sequence
date, and deposited the sequence as contiguous fragments (McShan et. al. 1997). Roughly 3500 bp from the 3.8 kb sequence described here overlap with two distinct contiguous sequences, 204 and 264, in their depository. Interestingly, a 218 bp sequence corresponding to the area near Tn916 insertion [see section 4.6] is absent from their sequences. Attempts were made to locate the 3.8 kb fragment on the map of the GAS chromosome by matching any part of the 3.8 kb sequence with a contiguous fragment that also contains a known virulence marker found on the map. However, contigs 204 and 264 failed to match with any known GAS virulence markers, and as the contigs are not in schematic order, further sequence ‘walking’ or extrapolation could not be done.

4.4) Analysis of Insertion-Like Elements Within the 3.8 kb Fragment

Bacterial insertion sequences (IS) are genetic elements capable of mobilization that contain the minimal genes necessary for transposition (Campbell et. al. 1975). They are usually less than 2 kb in length and may integrate within the host genome in several copies to produce rearrangements, deletions and genetic recombinations (Lawrence et. al. 1992). More importantly, IS are important in that they can form transposons that are able to mediate mobilization of genes that encode antibiotic resistance (Clewell and Gawron-Burke 1986 and Neu 1992). Kapur et. al. (1994) were the first to report the only IS found in S. pyogenes despite their presence in many other genera and several streptococcal species (Murphy 1989).
In the course of examining the DNA sequence of the 3.8 kb fragment for homology with reported DNA sequences some overlap was observed with insertion sequences from *Streptococcus mutans* IS199 (Macrina *et. al.* 1996), *Streptococcus thermophilus* IS1191 (Guedon *et. al.* 1995) as well as with a transposase gene and protein from *Lactococcus lactis* IS905 (Dodd *et. al.* 1994) and a transposase protein from *S. thermophilus*. The transposase encoded for by IS1191 from *S. thermophilus* is identical to the transposase encoded for by IS905 from *L. lactis*. These transposases are 391 amino acids in length, and 95 of these overlap in total (identities or positives) with the translated sequence from the 3.8 kb fragment (Figure 4). Identical amino acids are referred to as identities while synonymous residues are called positives. All of the homologous regions are located between nucleotides 792 and 1394 on the 3.8 kb map. This region may represent a second and novel IS/transposase from GAS.

4.5) Analysis of Putative Open Reading Frames Within the 3.8 kb Fragment

Translation of the 3.8 kb sequence revealed numerous putative open reading frames (ORFs) (Figure 5). The Wisconsin GCG computer program uses the ATG start codon and either UGA, UAA, UAG as stop codons to identify coding sequences. Thus, as seen in figure 5, the number of putative ORFs may be very high, not all of which are likely candidates for genes. To an extent, generation of ORFs also depends on the quality of sequence obtained and whether or not there are insertion-deletion generated frame shifts in the primary sequence. For these reasons not all ORFs are equally appealing for
Figure 4: Sequence and amino acid homology of IS905 from *Lactococcus lactis* with the 3.8 kb region from MGAS166.

**Figure 4:** FASTA and Blast searches of the 3.8 kb fragment revealed homology with known insertion sequences (IS) from several gram positive organisms. Shown here is a representative region of overlap at the DNA and amino acid level with IS905 and its corresponding transposase from *Lactococcus lactis.*
Figure 5: Schematic representation of putative open reading frames within the 3.8 kb region.

ORFs generated using the CGC Wisconsin gene program shows six coding frames. Insertion of Tn916 at bp 1367 is shown in yellow. The five ORFs amplified by PCR are highlighted in green.
Figure 6: Generation of amplicons representing five non-overlapping open reading frames with the 3.8 kb region.

Figure 6: Five non-overlapping PCR amplicons were generated using primers described in Table 3. Lanes 1 through 5 correspond to ORFs 1 through 5 respectively. The sizes of the ORFs are as follows: ORF 1 (260 bp), ORF 2 (469 bp), ORF 3 (369 bp), ORF 4 (333 bp) and ORF 5 (539 bp). Lane 6 is the 100 bp marker.
characterization. However, based on conserved genetic elements such as promoter motifs and proximity to Tn916 insertion (see section 4.6 and 4.7) five ORFs (ORF 1 through ORF 5) were selected for further characterization (highlighted ORFs in Figure 5) and amplified from pACYC184 using PCR (Figure 6).

ORF 5 demonstrates nearly all the conserved elements of a functional transcript. Namely, a consensus Shine-Dalgarno (AGGAGG) sequence is located exactly 10 bp upstream of the ATG start codon. Approximately 150 bp upstream is the characteristic Pribnow box forming the -10 promoter (TATAAT). Exactly 17 bp upstream of this lies the -35 promoter region sequence of TTTACA which is similar to the one found in E. coli. This ORF codes for a peptide of 53 amino acids which appears devoid of any signal sequence (Figure 7). There are no obvious rho-independent termination sites.

It is important to note the unusual presence of several cysteine residues 24 amino acids from the amino terminal. Seven cysteines, five consecutive, followed by two tyrosine, followed by two more cysteines, are found within nine residues from amino acids 24 to 33. The possible significance of this cysteine string will be addressed in the discussion (see section 5.0).

4.6) Determination of the Precise Insertion Point of Tn916

The rationale for determining the exact insertion point of Tn916 in SBNH5 was as follows: 1) the sequence of Tn916 is known and left and right outward primers were generated; 2) four PCR reactions were set up using left and right
Figure 7: Nucleotide sequence and protein translation of ORF 5 within the region responsible for SLS production in GAS.

```
-35
AGGTTTACATATGTTCAATTTCATCTAATAAAAGTGATAAGAATGATGTTGTT
-10 P.B.
TCCAAATGTATTAAATTAGAAAAATGATATTATTTTCATATTTGGATCTATCAACAA

RVYILILIPYYYNKSNDN*KIVV

61
CACATGTTGTATGTTAACCTCGTAAGGAAACCAACACACAAATTCAACAAAATTTT

VLQQYNASLVLVTLLFRKL

TAAGTTAAAAAAATGACGACGCTGATTCTATGATGATTGATGACTAAAC

131
ATTCATCAATTTTATTGTGAGCTGACCACAAGAATAGCTAATGATAATTTATGCTAGTC

*G*K*TTTVLISYLLDKEVN*

Start
CTTATGTTAAAAATTTACTATCAATATTTATGCATGATGATTGATGACTAAAC

191
GAAATCAATTTAAATGACGACTTATGAAATGATAATTTATGCTAGTC

LMKLFTSNILTATSVAEETQQV

GCTCCTGAGGCTGGTGCTGGTGTCTATCCTTTTCTCCTTAAATTGCTACTGGAGT

251
CGAGGACCTCCAGGACAAAGAGCGACACTGATGAAACACACAAAGGATATACGCTGATTCCA

APGCCCCTTCTTCCCFSTTATGTS-

500
GGTAATCCATGTCGATGCGCTGACTTATGACGAGTAAATGAAATTTATGCTATCC

C56
CCATTAAGGACCTCCACATCCGGCTTCAATGCGGCTCATTTATGATATGATACCGAG

GSNQGGGSGSYTPGK*SI*HL

TATGTGGATGATATATGGGATATGAGTT

391
ATACACCATTACATTTGCTTTACATTCTCAAA
```

Figure 7: The nucleotide sequence and protein translation of ORF 5 is depicted with conserved gene elements demarked as follows: minus 35 and minus 10 promoter elements are shown in yellow, ribosomal binding motif (Shine-Dalgarno) sequence shown in green, start and stop codons shown in pink. The sequence codes for 53 amino acids and the transcript is roughly 400 bp long.
Figure 8: Schematic diagram representing the strategy for determining the Tn916 insertion point.

1 kb amplicon → sequence using L.O. → best fit to Tn 916

Figure 8: Four PCR reactions set up using left/right outward primers (LO/RO) from Tn916 and left/right primers from 3.8 kb fragment. Two amplicons of the correct combination (spanning fragment and Tn916 primers) were generated, the left one was sequenced. This sequence was matched onto the left terminus of Tn916; the first diverging base indicates chromosomal SBNH5 DNA. This base matched position 1367 on the 3.8 kb insert (5'-3') and was the exact Tn916 I.P.
outward primers from Tn916 (L.O. and R.O.) and left and right outward primers from the 3.8 kb fragment (s341 and a393) with chromosomal DNA from SBNH5 serving as template; (as two primer sets were employed, four combinations for reactions were possible; two yielded amplicons corresponding to the correct primer combination that span the fragment primers and the Tn916 primers) 3) the left amplicon was sequenced and this sequence was matched with the left terminus of Tn916; 4) the first diverging base indicates the beginning of chromosomal SBNH5 DNA. This base mapped precisely to position 1367 on the 3.8 kb insert (5'-3') and is the exact insertion point of Tn916 (Figure 8). This insertion point in the 3.8 kb fragment, lies midway within the promoter region of ORF 5, eight bp downstream of the -35 element and 9 bp upstream of the -10 TATA box. Hence ORF 5 was selected for further characterization.

4.7) RNA Analyses of ORFs Within the 3.8 kb Fragment

To investigate whether or not any of the selected ORFs within the 3.8 kb fragment were involved in SLS production, Northern blotting experiments were performed. Total RNA from strains MGAS166, SBNH5, ATCC 27762, CS91-23, T18P and SB30-2 was isolated at the mid-point of the exponential phase of growth and every two hours thereafter up to ten hours and probed with either the entire 3.8 kb fragment or the PCR generated ORFs.

In the wild-type (WT), transcripts which hybridized with the entire 3.8 kb fragment (Figure 9) were produced maximally at 6 hours post log phase. Moreover, transcripts, again produced maximally post 6 hours hybridized with ORF 5 exclusively (Figure 10). However, no transcripts corresponding to the 3.8
kb fragment, or ORF 5, were produced by the SBNH5 mutant. ORF 2 did not hybridize with either strain.

To confirm the importance of ORF 5 in the production of SLS, RNA transcripts from four different GAS strains were analyzed by Northern blotting (Figure 11). RNA extracted at six hours post mid-log phase from strain MGAS166 was visualized as the positive control. Strain CS91-23 and SB30-2 are M-12 and M-18 Tn916 generated hemolytic deficient mutants of wild-types CS91-23 and T18-P respectively (Table 2). The SLS* strains CS91-23 and SB30-2, which also harbour single Tn916 insertions, failed to produce transcripts corresponding to ORF 5 at any time points. In addition, T18P, the parental strain of SB30-2, did produce transcripts that hybridized with ORF 5. This difference was most marked at post 6 and 8 hours. Strain ATCC 27762 is a naturally occurring SLS* strain whose RNA transcripts corresponding to ORF 5 were produced at all time points. The rRNA signal used as an RNA control was observed in all lanes at approximately 1300 bp.

4.8) Construction of pLZ12-S1 and Complementation Studies

To determine if the hemolytic phenotype of the parental strain MGAS166 could be restored by complementation in the SBNH5 mutant a streptococcal shuttle vector, pLZ12-Km, was used to generate a construct harbouring the 3.8 kb fragment. This vector is a low copy plasmid that is capable of mediating transfer between Gram negative and Gram positive organisms. Briefly the
RNA analysis of MGAS166 and SBNH5 probed with the 3.8 kb fragment responsible for SLS production.

Figure 9:  
Total RNA extracted from mutant SBNH5 (lanes 2-7) and WT MGAS 166 (lanes 7-13) was quantified, standardized, blotted and probed using $\alpha^{32P}$ labeled 3.8 kb fragment. Lane 1 is an RNA standard, lane 2 is SBNH5 RNA harvested at mid-log, lanes 3-7 are SBNH5 RNA at post 2, 4, 6, 8, and 10 hours. Lane 8 is MGAS 166 RNA harvested at mid-log, lanes 9-13 are MGAS 166 RNA at post 2, 4, 6, 8, and 10 hours. The mutant strain is devoid of any transcripts within the 3.8 kb fragment at all time points, while the WT shows maximal transcript expression 4 and 6 hours post mid-log.
Figure 10: RNA analysis of MGAS166 and SBNH5 probed with ORF 5.

Lane 1 is MGAS 166 RNA harvested at 2 hours post mid-log, lanes 2-4 are WT MGAS 166 RNA at post 4, 6, and 8 hours, lanes 5 and 6 are mutant SBNH5 RNA harvested at 2 and 6 hours post mid-log. The mutant strain is devoid of any transcripts for ORF 5 at both time points, whereas the WT shows maximal transcript expression 6 hours post mid-log.
Figure 11: RNA analysis of GAS strains probed with ORF 5 and 16S rRNA.

Figure 11: Total RNA extracted from various GAS WT and hemolytic deficient mutants. The upper band represents the 16S rRNA control (c. 1300 bp), whereas the lower band is the signal seen with the ORF 5 probe (c. 400 bp). WT strains have transcripts corresponding to ORF 5, whereas mutant strains do not.

ATCC 27762: Lanes: 1, 5, 9, 13, 17
CS91-23: Lanes: 2, 6, 10, 14, 18
T18P WT: Lanes: 3, 7, 11, 15, 19
SB30-2: Lanes: 4, 8, 12, 16, 20
strategy was as follows: pLZ12-Km has two antibiotic resistance genes, one for kanamycin (Km) and one for chloramphenicol (Cm). The 3.8 kb fragment was subcloned into the HindIII site within the Km resistance gene. This construct, named pLZ12-S1, was then transformed into E. coli DH5α. Following transformation, the first 20 positive clones were screened by plasmid extraction and one clone revealed an intact pLZ12-S1 construct of 9.4 kb. When digested with HindIII the original 5.6 kb vector and 3.8 kb insert were apparent (Figure 12). Both the 9.4 kb construct and the 3.8 kb region hybridized with the 3.8 kb fragment when probed by Southern blotting.

Electroporation of pLZ12-S1 was then attempted under varying conditions on the SBNH5 mutant. Several attempts were conducted using 1.5, 1.75 and 2.5 kV/0.2 cm cuvette gap respectively. In addition, the resistance was decreased from 200Ω to 100Ω to concentrate the voltage potential resulting in a shorter time constant. Repeated transformation failed to yield any bona fide colonies, most likely due to the poor transformability of SBNH5.

To address this problem another favourable candidate and similar mutant, CS91-23, was selected for complementation studies (Table 2). Strain CS91-23 exhibits the hemolytic deficient phenotype and also has Tn916 inserted into ORF 5. Electrottransformation of pLZ12-S1 into CS91-23 yielded three transformants that revealed restoration of the hemolytic phenotype when plated on blood agar with the appropriate antibiotic selection. The presence of pLZ12-S1 could not be detected in these cells and the observed hemolysis is likely due
Figure 12: The 3.8 kb fragment was subcloned into pLZ12-Km as described in the text and transformed into E. coli DH5α. Plasmid preparations were performed on E. coli and restricted with HindIII. The vector alone is 5.6 kb. Lanes 1, 4 and 7 are 1 kb marker DNA. Lane 5 is the linearized construct pLZ12-S1 (9.4 kb). Lane 6 is pLZ12-S1 cut with HindIII; upper band is the 5.6 kb vector, lower band is the 3.8 kb insert.
to a shift in Tn916 position which would restore the function of the ORF 5 product.

4.9) Location of 3.8 kb Fragment Within 20 Clinical Isolates of GAS

To examine the prevalence of the 3.8 kb fragment responsible for SLS production, and to see if the fragment was located on the same chromosomal region in various GAS strains, nine clinical strains were restricted with SmaI and analyzed by PFGE and Southern blotting (Figure 13 and 14). The results are summarized in table 6. M-types roughly correlated with electrophoretic type pattern, however, the location of the 3.8 kb fragment did not correlate with M-type. In these strains analyzed by PFGE the 3.8 kb probe hybridized with a conserved band around 170 kb in strains 5823, 5823A, 5835, 6060; with a band c. 450 kb in strains 5786 and 5787 (isolated from husband and wife); as well as with a band c. 485 kb in strains 5809 from Midland, Ontario and a health care worker contact, 5809-1; while a unique strain, 6078, probed near 160 kb.

Eleven other clinical strains were analyzed by conventional gel electrophoreses, Southern blotted and probed with the 3.8 kb insert. A conserved band approximately 3.8 kb was present in all strains except 5619, 5367, and 6079W. In these strains the probe hybridized at approximately 2.6-2.8 kb (Figure 15). Of particular interest is that 6079W and 6079BC are from two different sites in the same patient (wound and blood culture) yet the 3.8 kb fragment probed at different loci. It is likely that the strains are genetically
Figure 13: Genomic DNA from 9 clinical strains were analyzed by PFGE as described in the text. Five distinct restriction patterns were observed, A through E, with one isolate (B1) exhibiting a one band shift difference to pattern B. Sizes ranged from 160 to 485 kb and the data is summarized in table 6. Lane 1: strain 5809, lane 2: 5809-1, lane 3: 5786, lane 4: 5787, lane 5: 5823, lane 6: 5823-A, lane 7: 5835, lane 8: 6060, lane 9: 6078.
Figure 14: Analysis of DNA from 9 clinical isolates (by Southern blotting) for the presence of 3.8 kb fragment after PFGE.

Figure 14: Genomic DNA from the 9 clinical strains resolved by PFGE (Figure 13) was transferred to nylon membranes and probed with the 3.8 kb fragment by Southern blotting. Lane order is the same as in figure 13; the last lane on right is λ DNA PFGE marker. The 3.8 kb probe hybridized with strains as follows: an approximate 450 kb fragment in strains 5786 and 5787; an approximate 485 kb fragment in strains 5809 and 5809-1 (a health care worker contact of patient colonized with 5801); an approximate 170 kb fragment in strains 5823, 5823A, 5835, and 6060; whereas a unique strain, 6078, probed with a fragment near 160 kb.
Figure 15: Analysis of genomic DNA from 11 clinical isolates (by Southern blotting) for the presence of 3.8 kb fragment after restriction enzyme digestion.

Figure 15: Genomic DNA from 10 clinical strains and the WT MGAS166 were restricted with HindIII and resolved by conventional gel electrophoresis. DNA was transferred to nylon membranes and probed with the 3.8 kb fragment by Southern hybridization. A conserved band approximately 3.8 kb hybridized with the probe in all strains except 5619, 5367, and 6079W. In these strains the probe hybridized with fragments of approximately 2.6-2.8 kb. Lane 1: MGAS, lane 2: 5640, lane 3: 5811, lane 4: 5657, lane 5: 5619, lane 6: 5454, lane 7: 5385, lane 8: 5711, lane 9: 5367, lane 10: 5661, lane 11 and 12: 6079 BC and 6079 W respectively. (First and last lanes are 1 kb marker).
different and further investigation is necessary to determine the reason for this discrepancy.

These results show that all isolates from patients with invasive and non-invasive disease possess the DNA region responsible for hemolysis. However, this region is not always found on the same area of the chromosome.

Table 6: Location of 3.8 kb fragment in GAS clinical strains after PFGE and Southern blotting.

<table>
<thead>
<tr>
<th>No. of Strains</th>
<th>Case Definition</th>
<th>M-Type</th>
<th>PFGE Pattern</th>
<th>Fragment Localization</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>STSS</td>
<td>M-12, ?</td>
<td>A, A</td>
<td>485 kb</td>
</tr>
<tr>
<td>2</td>
<td>NF</td>
<td>M-1</td>
<td>B, B</td>
<td>450 kb</td>
</tr>
<tr>
<td>1</td>
<td>Invasive</td>
<td>M-3</td>
<td>D</td>
<td>170 kb</td>
</tr>
<tr>
<td>1</td>
<td>Invasive, NF</td>
<td>M-4</td>
<td>E</td>
<td>160 kb</td>
</tr>
<tr>
<td>3</td>
<td>Invasive, ?, N/A</td>
<td>M1, ?, M-22</td>
<td>B1, C, C</td>
<td>170, 170, 170</td>
</tr>
</tbody>
</table>

4.10) Probing Hemolytic Deficient Mutants for the 3.8 kb Insert and ORF 5

To determine the location of the 3.8 kb fragment on the other hemolytic deficient mutants used in this investigation Southern blotting was performed on genomic DNA restricted with HindIII. The 3.8 kb insert and ORF 5 probes hybridized with a band of approximately 2.8 kb in the ATCC 27762 and T18-P strains. Two bands hybridizing to DNA of high molecular weight corresponding approximately to 6 kb and 12 kb were observed in the mutant strain CS91-23. This hybridization profile was identical to the pattern seen in the SBNH5 mutant and confirms the Tn916 disruption of the region responsible for SLS expression (Figure 16, 17). The ATCC strain was shown to contain ORF 5 by PCR (Figure 18).
Figure 16: Restriction enzyme digests of genomic DNA from various GAS wildtype and mutant strains probed with the 3.8 kb fragment.

Figure 16: Genomic DNA from various GAS WT and hemolytic deficient mutants was isolated, restricted with HindIII and probed by Southern blotting using the 3.8 kb fragment as a probe. Lane 3: MGAS166 shows a hybridization signal at 3.8 kb, lanes 4 and 6: SBNH5 and SB30-2 mutants respectively show bands at approximately 6 kb and >12 kb, lane 5 and 7: WT ATCC 27761 and WT T18-P respectively show a band at 2.8 kb. Lane 1 is a 1 kb marker.
Figure 17: Restriction enzyme digests of genomic DNA from various GAS wildtype and mutant strains probed with ORF 5.

Figure 17: Genomic DNA from various GAS WT and hemolytic deficient mutants was isolated, restricted with HindIII and probed by Southern blotting using ORF 5 as a probe. Lane 3: MGAS166 shows hybridization signal at 3.8 kb, lanes 4 and 6: SBNH5 and SB30-2 mutants respectively show bands at approximately 6 kb and >12 kb, lane 5: WT ATCC 27761 did not give a positive signal. Lane 7: WT T18-P shows a band at 2.8 kb. Lane 1 is a 1 kb marker.
Figure 18: Demonstration of ORF 5 in ATCC 27762 by PCR.

Figure 18: The 539 bp ORF 5 shown to be present in the hemolysin negative WT strain ATCC 27762 strain by PCR (lane 3). PCR conditions used were as described in the text. Lanes 1 and 2 show ORF 5 amplified from pACYC184 harbouring the 3.8 kb insert and MGAS166 respectively as positive controls. Lane 5 is a 1 kb marker.
DISCUSSION AND FUTURE STUDIES

The precise role of SLS in virulence has not been fully elucidated. Historically, SLS has been a difficult toxin to isolate and purify due to the fact that it appears to remain sequestered by a carrier or stabilizing molecule. In addition, the gene or genes associated with SLS production have not previously been identified. Other than generating SLS- mutants and showing that hemolytic deficient mutants were less virulent in murine models (Owens et. al. 1978, Betschel et. al. 1997), no specific details of the SLS gene or its regulation are known. One of these studies (Owens et. al. 1978) employed chemical mutagenesis to produce hemolytic deficient mutants that, however, had pleiotropic effects on the growth rate. Recently, Liu et. al. (1997) generated Tn916 insertional mutants which failed to express SLS and revealed a linkage between SLS activity and a growth requirement for riboflavin, but this mutant also exhibited pleiotropic effects on growth rate and an altered exoprotein profile. Hence, this study is the first time that any genetic work has been carried out on SLS and SLS deficient mutants that have no gross changes other than the inability to elaborate SLS compared with their corresponding WT strains.

Here, we demonstrate that the insertion point of Tn916, in our SLS- mutant strain SBNH5, lies directly in the promoter element of a novel ORF, called ORF 5. Northern analysis revealed that strain SBNH5 as well as other Tn916 generated hemolytic deficient mutants, such as CS91-23 and T18-P, lacked transcripts for ORF 5 inferring that mutants do not produce the peptide encoded for by this gene and that this peptide is necessary for SLS production.
ORF 5 contains all the essential elements necessary to encode a functional gene, namely consensus promoter elements, a ribosomal binding motif and a start codon in the correct locations. The sequence encodes a peptide of 53 amino acids with an unusually long cysteine string spanning seven out of nine consecutive residues. Such a sequence has not been previously reported in a protein this size nor for the SLS protein in the literature (Koyama 1963, Lai *et. al.* 1978). In fact, the amino acid composition as predicted by Koyama (1963) and Alouf (1988) failed to reveal the presence of any cysteine residues. However, this repeated cysteine motif is present in several metallothionein and ferredoxin/iron-sulfur binding proteins (Otaka and Ooi 1987) as well as in cysteine string proteins (van de Goor and Kelly 1996) and EGF-like domain cysteine pattern signatures (Davis 1990). It has not escaped our notice that the essential protein necessary for SLS production reported by Akao *et. al.* (1992) was composed of three peptides with cysteine mediated disulfide bridges. Whether or not there is any functional relationship between the protein encoded for by ORF 5 and the essential peptide described by Akao is not known. It is tempting to speculate that if there is, then such a protein may serve as an accessory peptide for SLS formation in the absence of an adequate exogenous essential peptide. It may be that this peptide is a regulatory protein for SLS formation.

Alternatively, ORF 5 may code for a regulatory transcript controlling DNA transcription. In fact, this is the case with δ-toxin from *Staphylococcus aureus* which is a small (26 amino acids), but potent cytotoxin. In *S. aureus*, the *agr*
locus encodes two divergent transcripts (RNAII and RNAIII), each with its own promoter P2 and P3 respectively (Kornblum et. al. 1990). The promoter P3 initiates the formation of a 514 bp transcript called RNAIII, i.e., the δ-toxin transcript (hld), which, in turn, co-regulates transcription of P2 (Janzon and Arvidson 1990, and Novick et. al. 1993). The principal role of the remainder of the agr locus, the RNAII transcript, is to direct transcription of RNAIII through the action of four genes, agrA, B, C, and D that form a two-component regulatory system in an auto-feedback loop with RNAIII (Novick et. al 1995). The agr locus also mediates control of extracellular and surface proteins such as staphylococcal α-toxin and protein-A via the RNAIII transcript (Vandenesch et. al. 1991). The RNAIII peptide, i.e., δ-toxin itself is not regulatory (Novick et. al 1993).

There is evidence, however, that would negate the possibility of ORF 5 serving as a regulatory transcript. The transcript for ORF 5 does not appear to contain the specific features of a DNA binding regulator or processing enzyme. Specifically, most regulatory transcripts, with the exception of RNAIII from S. aureus, are significantly longer than ORF 5 and code for a conserved amphipathic helix-loop-helix motif of between 50 and 60 residues that is absent in the ORF 5 transcript (for review see Mitchell and Tjian, 1989). Secondly, other GAS mga mutants generated by Tn916 insertion demonstrate a decreased hemolytic phenotype, and when ORF 5 was used as a probe in Northern blotting experiments of these mutants, the ORF 5 specific message was decreased in three different serotypes (Podbielski, personal communication), whereas the
SLO message was completely unaffected. It seems unlikely that knocking-out the regulator \textit{mga} would, in turn, abrogate another regulator. It remains possible that that ORF 5 could be a co-factor, but not an inhibitor, since its expression parallels the high or low hemolysis phenotype.

The data obtained in this study indicate that ORF 5 is involved in SLS formation and imply that either ORF 5 alone or in conjunction with other regions on the GAS chromosome is necessarily involved in SLS production or regulation in GAS. If, however, ORF 5 does encode a regulatory peptide for SLS production, a point mutation in this gene would abrogate the SLS phenotype while leaving a non-functional transcript for hybridization. It is possible that the ATCC 27762 strain containing ORF 5 is SLS deficient due to such a point mutation. Alternatively, the regulatory ORF 5 may be intact, while the structural gene for SLS is affected by some other factor.

In addition, we demonstrate that the single copy insertion of Tn916 into other mutagenized SLS' strains, such as CS91-23 and T18P, lies in nearly identical genomic loci as the \( \beta \)-hemolytic GAS SLS' mutant, SBNH5. The fact that the hybridization profile of a hemolytic deficient mutant, CS91-23, generated in another study (Nida and Cleary 1983) is identical to the originally generated SBNH5 mutant (i.e., two bands at 6 and \( >12 \) kb), when probed with both the 3.8 kb insert and ORF 5 by Southern blotting, suggests that Tn916 has inserted within the ORF 5 region of this mutant as well.

DNA homology searches carried out on the genomic region flanking Tn916 insertion revealed partial matches with insertion sequence (IS) elements
and transposase genes. This may account for the band shift mobility, i.e., hybridization signals at 2.8 or 3.8 kb, observed with the clinical isolates reported previously (see section 4.9), leading us to postulate that the region responsible for SLS expression may be found on different loci of the GAS chromosome. Insertion sequences often contain strong outward reading promoters (Salyers et al. 1995). Such a promoter upstream of either the SLS gene itself, (or of a regulatory gene), may regulate SLS production and consequently virulence in vivo. These promoters may also account for the different amounts of hemolysis observed in vitro, in different strains of GAS. Only one IS element, IS1239, from GAS has so far been reported (Kapur et al. 1994) and was found to be a 981-bp ORF that potentially encodes a 326 amino acid polypeptide with substantial homology to the E. coli IS30 transposase. The present study is the only other investigation in which significant homology with known IS/transposases in GAS has been found.

The significance of IS elements within the genome of pathogenic organisms cannot be overlooked. In addition to providing the mobilizable elements for antibiotic resistance genes, they also may mediate the transposition of pathogenicity islands (PAI). PAIs have only recently been described and the definition is still developing, however, two properties appear to be common to all PAIs. First, PAIs are chromosomally encoded, and second they signify a group of virulence genes that behave as distinct and functional units (Lee 1996, Mecsas and Strauss 1996). It is clear from G+C compositions and homology analyses that the origin of PAIs are exogenous and that PAIs transfer between
organisms to confer complex virulence properties to the recipient bacteria. A number of pathogenic organisms harbour PAIs such as the LEE locus of *E. coli* and *CagA* of *Helicobacter pylori* (Knapp *et. al.* 1986 and Censini *et. al.* 1996). Thus, IS elements may serve to mediate horizontal gene transfer between non-pathogens and pathogens through PAIs and transposons. The advantage of obtaining complex virulence traits in one discrete genetic event rather than over the course of several vertical generations is of significant importance to the fitness of a potential pathogen.

Complementation studies using the entire 3.8 kb insert failed to demonstrate phenotypic restoration of the hemolytic phenotype when attempts were made to introduce this fragment into the hemolytic deficient mutants SBNH5 and CS91-23. Several future studies should be able to resolve this difficulty. First, it may be possible to introduce a smaller construct that contains only ORF 5 into one of the hemolytic deficient mutants. Ligation have already been carried out that subclone ORF 5 into the *HindIII* site of pLZ12-km. Electroporation of the shuttle vector pLZ12-Km may be favoured when introduced as a smaller vector. Thus, the size of the new construct would be 6.1 kb rather than 9.4 kb. Previous studies with pLZ12-Km have successfully employed a smaller derivative (Caparon, personal communication). Also, since electroporation into SBNH5 and CS91-23 was unsuccessful it may be possible to introduce this new construct into strain T18P, another SLS- strain, as some strains are more transformable than others (Caparon, personal communication). Transformants that have the restored hemolytic phenotype may then be
assessed in a murine model of virulence to observe if virulence has also been restored to original WT levels.

Barring this approach, recombination studies in the hemolytic deficient mutant strain ATCC 27762 may be carried out using the conjugative shuttle/suicide vector, pKmobsacB (Schafer et al. 1994, kindly provided by Dr. P. Cleary). For this purpose, a construct has already been generated called pSBSLS-1 which contains the 3.8 kb insert subcloned into the HindIII site of pKmobsacB which is able to mobilize with the help of a co-resident RP4 helper plasmid. This construct, pSBSLS-1, when transformed into E. coli S17-1, a strain with an RP4 derivative integrated into the chromosome, should be able to transfer into ATCC 27762 and through homologous recombination integrate into the altered region responsible for SLS production to demonstrate SLS activity. Alternatively, a modified 3.8 kb fragment or ORF 5 containing either insertions or deletions may be subcloned into pKmobsacB and introduced into the MGAS166 WT to abrogate SLS activity. A conclusive way of demonstrating the importance of ORF 5 in SLS production would be to use site directed mutagenesis and allelic replacement to generate ORF 5 knock-outs which could be used in virulence studies. This approach would remove any doubt as to unwanted polar effects mediated by Tn916 transposition.

Other future studies may include expression studies of ORF 5 in a suitable expression or fusion-protein vector. If the ORF 5 product is inactive, i.e. it does not serve as a functional hemolysin in vitro, it may be a regulatory protein. This possibility, however unlikely for reasons discussed above, could be
investigated by conducting footprinting studies that block the DNA binding site of the regulatory protein.

There also exists the possibility that the protein product of ORF 5 serves as an 'endogenous' carrier molecule for the release of SLS, *in vivo*, in the absence of one of the aforementioned inducers (see section 1.6A). A simple method of investigating this possibility would be to propagate a mid-log culture of MGAS166 in THB, wash the cells, and then co-incubate with the expressed product of ORF 5. It should be possible to recover a high SLS yield from the resting cells. The mutant strain, SBNH5, should serve as a negative control because of the lack of a functional internal carrier (due to the mutagenized ORF 5) to transport SLS to the cell surface. Thus, even in the presence of the external ORF 5 peptide supplement, minimal SLS yield would be obtained, unless the ORF 5 peptide could be readily internalized by the cells.

If the ORF 5 encoded polypeptide is able to produce hemolysis *in vitro*, it suggests that it is a structural gene that encodes a functional hemolysin. Another way of determining whether or not the ORF 5 product is structural would be to generate a peptide by chemical methods and raise antibodies to a specific determinant within ORF 5. Hemolysis inhibition studies could then be performed to observe if ORF 5 antibodies can abrogate SLS activity *in vitro*.

Although the ORF 5 peptide is inconsistent with previously reported amino acid compositions of SLS, it is not inconceivable that ORF 5 is the structural gene for SLS. There are many striking similarities between a class of molecules known as lantibiotics and the ORF 5 peptide, despite the lack of homology.
observed in the BLAST searches. Lantibiotics are a sub-group of bacterial antibiotic peptides containing post-translationally modified amino acids such as lanthionine and 3-methyllanthionine as well as D-forms of amino acids (Schnell et al. 1988). Their unique structural properties result from the presence of intra-molecular rings formed by thioether bonds of these modified amino acids from the originally coded cysteine and serine residues in the pre-protein. The term lantibiotic has subsequently been applied to these classes of ribosomally-synthesized, low molecular weight, heat stable, bacteriocin-like compounds that have, to date, only been found in gram positive organisms (For reviews see Nes and Tagg 1996). In addition, a cluster of accessory genes are found flanking the gene responsible for coding the structural peptide. These accessory genes are responsible for coding proteins that cleave, modify, transport and export the structural lantibiotic as well as for conferring resistance to a particular organism’s own bacteriocin (Schnell et al. 1992). The pre-protein of lantibiotics consists of the leader peptide, containing 23 to 36 amino acids, and the pro-peptide region, spanning 22 to 44 residues. Cleavage of the leader sequence is accomplished by a protease that is encoded for on the lantibiotic gene cluster.

There are three categories of lantibiotics grouped according to cleavage site. The first group is known as the minus -2 Pro compounds which contain a proline residue in the penultimate position of the leader peptide; the second group are the “double Gly group” which contain two glycine residues near the cleavage site; and the third is the “lactocin S group” which resemble lactocin from Lactobacillus sake and differ from the first two groups. Two major sub-
groups have been defined (Jung 1991). Type A lantibiotics are screw-shaped amphipathic molecules of masses 2164 to 3488 Da carrying 2 to 7 net positive charges whereas type B lantibiotics are more globular molecules ranging from 1959 to 2041 Da, having no net or a net negative charge. At least thirteen lantibiotics have been described and are well characterized. These include subtilin produced by *Bacillus subtilis*, nisin from *Lactococcus lactis*, and salivaricin A from *Streptococcus salivarius* (Gross *et. al.* 1973, Banerjee and Hansen 1988; Gross and Morall 1971, Buchman *et. al.* 1988; Ross *et. al.* 1973).

The only lantibiotic to be described in *S. pyogenes* is streptococcin A-FF22 (SA-FF22) (Tagg and Wannamaker 1976). Early studies indicated that this bacteriocin was a small extracellular protein active against other gram positive organisms, specifically streptococci (Tagg *et. al.* 1973-A, 1973-B). Purified SA-FF22 was shown to have a molecular mass of 2795 Da with one lanthionine and two met-lanthionine residues. A non-inhibitory form was isolated which was devoid of the first four amino acids, indicating that a particular sequence of amino acids at the N-terminus may be necessary for biological activity (Jack and Tagg 1991). Subsequently, the gene for SA-FF22, called *scnA* was cloned and sequenced (Hynes *et. al.* 1993). Tagg and Bannister (1979) have reported that the pores formed by SA-FF22 are relatively unstable, short-lived and approximately 0.5 nm in diameter. The protein sequence of ORF 5 was compared to 13 known lantibiotic sequences. Homology was lowest from SA-FF22 from *S. pyogenes* with only a minimum of 25% similarity and 10% identity. The highest degree of homology was observed with epidermin and pep5.
matching 44% and 40% similarity respectively, and 22% and 20% identity respectively. These two lantibiotics come from *Staphylococcus epidermidis*. Searches with other lantibiotics yielded intermediate similarity values. The lack of stronger homology is not disconcerting, however, as some comparisons within lantibiotics, themselves, yielded no stronger homology.

Additional features of the ORF 5 gene product in common with lantibiotics include the following: a) functionally the SLS hemolysin parallels the membrane integrating and pore forming ability of lantibiotics, b) they are both very similar in size, c) there is homology with lantibiotic leader sequences and the ORF 5 leader sequence, specifically the first four residues, d) the ORF 5 peptide has the ability to form a complex 3-dimensional structure because it contains many cysteine and serine residues capable of being post-translationally modified, e) ORF 5 is flanked by other ORFs possibly representing lantibiotic accessory genes. This last point is of special importance if expression studies are to be carried out on ORF 5 because cloning of ORF 5 would presumably have to include the necessary accessory genes for a functional lantibiotic to be produced. Recently, novel expression systems have been designed that allow for the successful protein engineering of lantibiotics (Kuipers et. al. 1996).

In support of a functional relationship between lantibiotics and the ORF 5 peptide, cytolysin LL/LS from *Enterococcus faecalis* was the first lantibiotic shown to be a hemolysin. It was observed that the hemolytic phenotype correlated with bacteriocin activity (Brock and Davie 1963). The genetic determinant of the *E. faecalis* hemolysin/bacteriocin was found to reside on a 60
kb conjugative plasmid, pAD1 (Dunny and Clewell 1975). To date, cytolysin LULS is the only lantibiotic shown to exert hemolytic activity. Homology of the ORF 5 peptide to cytolysin LULS was shown to be 37% similar and 22% identical. Two alternately processed cytolysin LULS bacteriocins, along with lactocin S, comprise the 'lactocin S group' of lantibiotics. The N-terminal residue of lactocin S is blocked for amino acid sequencing by Edman degradation (Skaugen 1994). Interestingly, Alouf (personal communication) was not able to sequence SLS by Edman degradation and speculated that the N-terminus may be blocked or cyclic. In addition, the post-translational modification of cysteine residues may account for the lack of cysteine content in previously reported amino acid compositions of the SLS molecule (Alouf and Geoffroy 1988, Koyama 1963), as free cysteines are never found in lantibiotics (Nes and Tagg 1996).

Examination of the ORF 5 peptide also reveals that the double glycine motif is present 22 and 23 residues from the N-terminus with the corresponding minus -2 proline residue in place for proteolytic cleavage. If this is the cleavage site of the ORF 5 protein then the aforementioned cysteine string motif would become the N-terminus of the pro-peptide. As demonstrated on hydropathy plots, Chou/Fasman and Kyte-Doolittle plots the N-terminus of the pro-peptide would be predicted to be significantly hydrophobic and could potentially serve as a membrane insertion/pore forming region.

If the predicted similarities between lantibiotics and the ORF 5 peptide are correct then the ORF 5 gene would signify a novel lantibiotic produced by S.
*Streptococcus pyogenes*. The discovery of a distinct and new bacteriocin from GAS has immediate implications in the food preservation industry where lantibiotics are used as preservatives (WHO 1969), and in the pharmaceutical industry where this class of molecules represent viable options in exploiting novel antibiotic development.

Further work is necessary to conclusively demonstrate whether ORF 5 participates in a structural or regulatory manner in the production of SLS by GAS. Future studies should further elucidate the mechanism of SLS expression and regulation in GAS, and given the importance of treatment of streptococcal infections, this work may have further impact on the development of future vaccine strategies for this organism. These studies, in conjunction with previous work begun in this laboratory, have laid a strong foundation for further work on the genetics of streptolysin S in the hope of contributing to a better understanding of the pathogenesis of *Streptococcus pyogenes* in humans.
REFERENCES


Eichenbaum, Z., E. Muller, S. A. Morse, and J. R. Scott. 1996. Acquisition of iron from host proteins by the group A streptococcus. *Infect. Immun.* 64:5428-5429.


Harmandayan, R. 1990. Transformation of *Streptococcus mutans* by electroporation. Master of Science thesis, Department of Microbiology, Faculty of Dentistry, University of Toronto.


101


McShan, M. 1997. The Streptococcal genome project has been deposited at www.dna1.uchicago.edu and the raw sequence data is available for both download and Blast searches. (Project team: B. Roe, M. McShan, and J. J. Ferretti).


