Review

Streptococcal superantigens: categorization and clinical associations

Robert J. Commons[1,2], Pierre R. Smeesters[2,3], Thomas Proft[4], John D. Fraser[4], Roy Robins-Browne[2,5], and Nigel Curtis[1,2,6]

[1] Department of Paediatrics, The University of Melbourne, Parkville, Australia
[2] Murdoch Childrens Research Institute, Parkville, Australia
[3] Laboratoire de Génétique et Physiologie Bactérienne, Institut de Biologie et de Médecine Moléculaires, Faculté des Sciences, Université Libre de Bruxelles, Gosselies, Belgium
[4] Department of Molecular Medicine and Pathology, Faculty of Medical and Health Sciences, University of Auckland, Auckland, New Zealand
[5] Department of Microbiology and Immunology, The University of Melbourne, Parkville, Australia
[6] Infectious Diseases Unit, Royal Children's Hospital Melbourne, Parkville, Australia

Superantigens are key virulence factors in the immunopathogenesis of invasive disease caused by group A streptococcus. These protein exotoxins have also been associated with severe group C and group G streptococcal infections. A number of novel streptococcal superantigens have recently been described with some resulting confusion in their classification. In addition to clarifying the nomenclature of streptococcal superantigens and proposing guidelines for their categorization, this review summarizes the evidence supporting their involvement in various clinical diseases including acute rheumatic fever.

Streptococcal infections

Group A streptococcus (GAS; see Glossary) or *Streptococcus pyogenes* is a Gram-positive coccus that can be part of the normal upper respiratory tract and skin microbiota. It is exclusively a human pathogen and causes a wide variety of diseases [1,2]. These range from superficial infections, such as pharyngitis and impetigo, to invasive infections, such as septicemia, myositis, scarlet fever, necrotizing fasciitis, and streptococcal toxic shock syndrome (STSS) [3], as well as post-infectious sequelae, including acute rheumatic fever and post-streptococcal glomerulonephritis. Globally, GAS is responsible for more than 500 000 deaths per year, predominantly in low income settings [1]. *Streptococcus dysgalactiae* subspecies *equisimilis* [SDSE; a group C (GCS) and G streptococcus (GGS)] is an emerging global pathogen that can also colonize and infect humans causing a similar spectrum of disease to GAS including throat carriage and pharyngitis [4], STSS, and other invasive disease [5]. Other GCS and GGS have also been associated with invasive disease in human and animal infections, such as *Streptococcus canis* [6,7].

As with most infectious diseases, the pathogenesis of invasive GAS, GCS, or GGS disease is a complex interaction between the virulence factors of the pathogen and the host response [8–10]. Streptococci possess a myriad of virulence factors, including M proteins and superantigens, which are believed to be important in invasive disease. M proteins have many functions, including assisting in cell adhesion, avoidance of phagocytosis, inducing bradykinin release, and interfering with complement binding [11]. Soluble M1 protein has been proposed to possess superantigen activity [12]. Over 200 different M types have been described and can be used to classify both GAS and SDSE [13,14]. Multiple other virulence factors that assist in evading the innate and adaptive immune systems have

Corresponding author: Smeesters, P.R. (pierre.smeesters@mcri.edu.au).
Keywords: streptococcus; toxin; nomenclature; toxic shock; invasive disease.

1471-4914/$ – see front matter
 http://dx.doi.org/10.1016/j.molmed.2013.10.004

Glossary

Acute rheumatic fever (ARF): an inflammatory disease occurring 2 to 3 weeks following GAS infection that manifests with joint, skin, heart, and/or brain involvement. Thought to be caused by T cell crossreactivity with host tissues.
Group A streptococcus (*Streptococcus pyogenes*) (GAS): β hemolytic Gram-positive coccus bacteria that can colonize the skin or mucous membranes and cause pharyngitis or invasive diseases.
Group C/G streptococcus (GCS/GGS): usually β hemolytic, these bacteria have recently been associated with invasive disease and found to produce superantigens. The most common cause of invasive disease with GCS/GGS in humans is *Streptococcus dysgalactiae* subspecies *equisimilis* (SDSE).
Invasive streptococcal disease: diseases such as septicemia, myositis, scarlet fever, necrotizing fasciitis, and streptococcal toxic shock syndrome.
Kawasaki disease (KD): an acute febrile vasculitic illness that can cause coronary artery damage in young children.
M protein: streptococcal virulence factor that has a role in cell adhesion and avoidance of phagocytosis. They also induce bradykinin release and interfere with binding by complement.
Streptococcal toxic shock syndrome (STSS): an acute toxin-mediated, life-threatening, multisystem illness characterized by fever, rash, hypotension, and multi-organ failure.
Superantigen: extracellular protein toxin that bypasses conventional antigen processing and activates large numbers of T cells.

been described, including streptolysin O, streptococcal inhibitor of complement, and streptodornase D [15].

Streptococcal superantigens

Superantigens are important virulence factors produced by certain bacteria including GAS, GCS, GGS, *Staphylococcus aureus*, *Yersinia pseudotuberculosis* [16], and *Mycoplasma arthritidis* [17]. Superantigens are extracellular protein toxins whose properties include pyrogenicity, mitogenic activity for specific T cell subsets [defined by the variable region of the beta chain (Vβ) of the T cell receptor], and the ability to increase host susceptibility to endotoxic shock and suppress immunoglobulin production [16,18,19]. They bypass the conventional antigen presentation process, activating large numbers of T cells without major histocompatibility complex (MHC) class II restriction. This leads to extensive immune activation with a large release of proinflammatory cytokines, such as tumor necrosis factor α (TNFα), interleukin-2 (IL-2), and interferon γ (IFNγ), which underlies the endothelial leak, shock, and widespread organ damage that occur in toxic shock syndrome [20].

The protein structure and the immunological properties of superantigens have been comprehensively reviewed recently by Fraser and Proft [16]. Briefly, superantigen crystal structures reveal a conserved tertiary structure within all GAS superantigen subgroups, consisting of a two-domain fold that contains the common β-barrel globular domain at the N terminus and a β-grasp motif at the C terminus. Two conserved amino acid motifs (Prosite PS 00277 and PS 00278) are located at the interface between the N-terminal and C-terminal domains. Many superantigens are believed to contain a zinc-dependent C terminus binding site, which allows them to form a dimer that binds to the β-chain of the MHC class II molecule on antigen presenting cells. By contrast, some superantigens possess a generic low affinity binding site for the α-chain of the MHC class II molecule at the N terminus that can either prevent contact with the displayed peptide or bind over it [16]. Recent data show that some superantigens also bind to the co-stimulatory molecule CD28 on T cells and that this binding enhances T cell sensitivity to superantigens with increased superantigen activity and cytokine production [21,22]. The superantigen domain involved in this CD28 binding is conserved among five staphylococcal superantigens (SEA, SEC1, SEC2, SEE, and TSST-1) and the streptococcal superantigen SpeA [21,22]. Figure 1 indicates that the CD28 binding motif is also present in other streptococcal superantigens such as SSA, SpeI, SmeZ, SpeH, or SpeO.

This review summarizes the currently recognized streptococcal superantigens, collates the described streptococcal superantigen variants and alleles, and assesses their relatedness through phylogenetic analysis before proposing a structured nomenclature for future discoveries. Finally, we assess the evidence for streptococcal superantigens in disease.

Nomenclature of streptococcal superantigens

Over the past decade it has become apparent that a number of variants (alleles) exist for most superantigen genes. Some of these variants have been highlighted in publications, whereas others have only been registered with the Genome Sequence Database (GenBank). SpeA and SpeC were the first superantigen toxins identified in 1924 and 1960, respectively. A protein named SpeB was later found to be a cysteine protease rather than a superantigen, with its superantigenic activity attributable to contamination from the superantigen SmeZ [23]. Similarly, contamination with SmeZ and SpeC led to the incorrect description of SpeF (a DNase B) having superantigenic activity [24].

It was not until SSA was discovered by Mollick *et al.* in the early 1990s that the existence of other streptococcal superantigens was confirmed [25]. Subsequent screening of various genomic databases such as the '*Streptococcus pyogenes* M1 database' at Oklahoma University, the *S. equi* database at the Sanger Centre, and the more recently completed genome sequencing of M3 and M18 GAS isolates assisted in the detection of *speG*, *speH*, *speI*, *speJ*, *speK*, *speL*, and *speM*, which encode additional superantigens [26–28]. The rapid increase in the number of streptococcal superantigens and their variants that have been recognized has led to some confusion in their nomenclature (Box 1).

Allele diversity

To provide an up-to-date classification of all known superantigen variants, we searched the National Center for Biotechnology Information Nucleotide Database for all nucleotide sequences designated as known superantigens. Uniprot was also searched for all protein sequences designated as known superantigens and PubMed was searched using the keyword 'superantigen*' for publications that described superantigen variants (alleles). Relevant nucleotide sequences were combined into recognized superantigen groups and aligned using MUSCLE [29]. New variants were labeled with previously designated superantigens retaining their nomenclature where possible (Table 1). Table 1 does not consider function and includes variants such as *smeZ6*, *smeZ19*, *smeZ23*, *smeZ31*, and *smeZ59N* that are known to be non-functional [30,31]. A total of 145 unique streptococcal superantigen nucleotide sequences belonging to 14 superantigen groups were identified. The number of nucleotide alleles per group ranged from one (*szeN*, *szeP*, and *szeF*) to 58 (*smeZ*). Interestingly, 16 variants were protein duplicates and the protein sequences available in the database were truncated for 39 variants (Table 1). Therefore, thus far, 91 unique sequences encoding complete streptococcal superantigen mature proteins have been described.

Overlap between GAS and GCS/GGS superantigens

Most GAS superantigens are associated with bacteriophages except for *speG*, *speJ*, and *smeZ*, which are chromosomally encoded [16,32]. However, they can be absent, especially *speJ*, which may exist within a locus associated with a mobile genetic element [33]. There is homology between toxins in different streptococcal species as well as in *S. aureus*, with some streptococcal superantigens more closely related to staphylococcal superantigens than other streptococcal superantigens. Together with their bacteriophage location, this homology supports recent interspecies horizontal transfer [24]. The genomic context of