Genomic Location and Variation of the Gene for CRS, a Complement Binding Protein in the M57 Strains of *Streptococcus pyogenes*

Michael Binks, David McMillan and Kadaba S. Sriprakash


Updated information and services can be found at: http://iai.asm.org/content/71/12/6701

These include:

**REFERENCES**

This article cites 21 articles, 14 of which can be accessed free at: http://iai.asm.org/content/71/12/6701#ref-list-1

**CONTENT ALERTS**

Receive: RSS Feeds, eTOCs, free email alerts (when new articles cite this article), more»
Genomic Location and Variation of the Gene for CRS, a Complement Binding Protein in the M57 Strains of Staphylococcus pyogenes

Michael Binks,1 David McMillan,1 and Kadaba S. Sriprakash1,2,3*

Queensland Institute of Medical Research1 and Australian Centre for International & Tropical Health & Nutrition,2 Brisbane, and Menzies School of Health Research, Darwin,3 Australia

Received 14 January 2003/Returned for modification 5 June 2003/Accepted 1 September 2003

All isolates of serotype M1 of group A streptococci possess a gene for streptococcal inhibitor of complement (SIC) in the mga regulon, which harbors genes for other virulence factors, such as M and M-like proteins, C5a peptidase, and a regulator. In serotype M57 the gene for a protein that is closely related to SIC (crs57) is located outside the mga regulon. We mapped the location of the crs57 gene in six strains of emm57 (gene encoding the M57 protein) sequence types to an intergenic region between the ABC transporter gene (SPy0778) and the gene for a small ribosomal protein (rpsU). The noncoding sequences on both sides of crs57 exhibited high degrees of identity to the corresponding regions of sic from M1 strains. This included one of the inverted repeat sequences of IS1562 but not the insertion element itself. These observations suggest that crs57 was recently acquired by serotype M57 or its progenitor via horizontal acquisition from serotype M1. The six emm57 sequence type isolates analyzed in this study belong to two distinct molecular types (vir types VT8 and VT101). Although the crs57 sequences from VT8 strains had very few substitution mutations, the VT101 crs57 sequence had a large number of such mutations. The CRS57 proteins from these strains are secretory products and have the ability to bind to complement proteins. All these proteins contain several tryptophan-rich repeats designated DWS motifs and internal repeat sequences. In all of these structural and biochemical characteristics CRS57 resembles SIC from M1 strains. Hence, CRS57 has a functional role similar to that of SIC in an M1 strain.

Group A streptococcus (GAS) (Streptococcus pyogenes) is a human-specific pathogen that is responsible for a wide range of diseases, including immune-mediated postinfective sequelae, such as acute rheumatic fever and post-streptococcal glomerulonephritis. This pathogen has evolved diverse mechanisms to overcome host defenses against infection. The surface-exposed major protective antigen, M protein, binds to factor H and thereby may inhibit deposition of opsonin C3b on the GAS surface (9). M protein also binds to a 570-kDa human plasma protein designated C4b-binding protein that inhibits complement activation via the classical pathway (3, 15). All S. pyogenes strains express a specific protease that cleaves chemotactic complement C5a protein (4) to an inactive molecule, retarding phagocyte recruitment at the site of infection. Additionally, isolates of S. pyogenes with M1-type specificity express a secretory protein designated SIC (streptococcal inhibitor of complement function) that inhibits complement-mediated cell lysis (1). While the biological significance of SIC as an inhibitor of complement function is not clear, the presence of the gene encoding it in all serotype M1 strains (1), the extreme divergence (20), and the rapid emergence of variants (14) indicate that this protein has an important biological role. In fact, recent studies of SIC showed that it also inhibits the functions of innate immune proteins, such as secretory leukocyte proteinase inhibitor and lysozyme (5).

Biochemical studies have suggested that SIC interferes with the function of the membrane attack complex by possibly binding to one or more protein components associated with the complex (1). Fernie-King et al. (6) showed that M1 SIC binds to the C6 and C7 complement proteins, preventing their incorporation into the membrane attack complex. Despite considerable sequence diversity, all SIC variants from M1 strains have the complement-inhibiting activity (14).

Molecular studies (12) have shown that like serotype M1 strains, all serotype M57 strains possess a gene which encodes a protein closely related to the M1 SIC, designated CRS57 (the gene encoding the M51 protein). While most studies have been carried out with SIC variants of M1 strains, little is known about CRS57. In this study, we determined the exact location of crs57 in the M57 genome, examined the diversity of CRS57 from six M7 isolates belonging to two distinct molecular types (vir types VT8 and VT101), and tested the ability of the molecules to bind to the complement proteins. CRS57 is highly conserved in the major vir type (VT8) of M57 strains, whereas in the minor vir type (VT101) the protein is more diverse than VT8 CRS57. We show here that the CRS57 proteins are excretory products and that they have the ability to bind to the C6 and C7 complement proteins. Taken together, our results are consistent with single lateral acquisition of the sic gene from an emm1 strain by M57 or its progenitor.

MATERIALS AND METHODS

Strains, culture, and genomic DNA. Table 1 summarizes information about the S. pyogenes strains used in this study. S. pyogenes serotype M1 and M7 reference strains (strains 2031 and 2077, respectively) were obtained from Public Health Laboratory Services, Prague, Czech Republic. All other strains were isolates ob-
tained in the Northern Territory (NT). The GAS isolates were typed by vir typing (Fig. 1). Vir typing is based on restriction fragment length polymorphism of the mrg regulon (8, 11, 12). Growth of GAS and genomic DNA extraction were carried out as described previously (10).

**Sequencing of** crs57. Based on a partial sequence for crs57 (accession numbers AF060764 and AF060765) (12), a specific reverse primer (primer crs57R; GAG ACAAAACACTCCAGACCGG) corresponding to this gene was designed. To determine the sequence upstream of crs57, a bubble PCR was employed (17). Genomic DNA was digested with SacI and, to 15 µg of the digest, 10 µg of preammeled linker oligonucleotides (GCACGCTCTGAATCTGACAGTACAG ACTGGAGTCACAGCTGC and GATCGCAGCTGTGACCTAGTCACTCC AATGATCTGACAGCTGC) containing a SacI overhang and an internal mismatch (bubble) was ligated. Three units of T4 DNA ligase in 10% buffer containing ATP (Amersham) was used. The ligation product was then used as a template in a PCR with Tag DNA polymerase (Amersham) by using the bubble-specific primer (GAGACCTTCTAGACAGTACAGTAGTCACTCC AATGATCTGACAGCTGC) containing a SacI overhang and a primer encoding a small ribosomal protein subunit (rpsU) data (accession number AE004092), we designed primers ABCF2 and rpsUR1, a bubble PCR was employed (17).

Primer ABCF2 and primer crs57R; GAGACCTTCTAGACAGTACAGTAGTCACTCC AATGATCTGACAGCTGC containing a SacI overhang and a primer encoding a small ribosomal protein subunit (rpsU).

**TABLE 1. S. pyogenes strains used in this study**

<table>
<thead>
<tr>
<th>Strain</th>
<th>M57 vir type</th>
<th>emm sequence type</th>
<th>Source</th>
<th>Presence of sic gene a</th>
</tr>
</thead>
<tbody>
<tr>
<td>2077</td>
<td>VT8</td>
<td>emm57</td>
<td>Prague</td>
<td>Present</td>
</tr>
<tr>
<td>NS1140</td>
<td>VT101</td>
<td>emm57</td>
<td>NT</td>
<td>Present</td>
</tr>
<tr>
<td>NS38</td>
<td>VT8</td>
<td>emm57</td>
<td>NT</td>
<td>Present</td>
</tr>
<tr>
<td>NS48</td>
<td>VT8</td>
<td>emm57</td>
<td>NT</td>
<td>Present</td>
</tr>
<tr>
<td>BSA16</td>
<td>VT8</td>
<td>emm57</td>
<td>NT</td>
<td>Present</td>
</tr>
<tr>
<td>BSA5</td>
<td>VT8</td>
<td>emm57</td>
<td>NT</td>
<td>Present</td>
</tr>
<tr>
<td>2031</td>
<td>emm1</td>
<td>Prague</td>
<td>Present</td>
<td></td>
</tr>
<tr>
<td>NS44</td>
<td>emm1</td>
<td>NT</td>
<td>Present</td>
<td></td>
</tr>
<tr>
<td>NS27</td>
<td>ND b</td>
<td>NT</td>
<td>Absent</td>
<td></td>
</tr>
</tbody>
</table>

a The presence of the sic gene was determined by PCR by using CRS-specific primers (12).

b ND, not determined.

**Production of recombinant CRS57.** The CRS57 gene was amplified by using oligonucleotides CTACTAGGACGTACAGTAGTACAG ACTGGAGTCACAGCTGC and GATCGCAGCTGTGACCTAGTCACTCC AATGATCTGACAGCTGC containing a SacI overhang and a primer encoding a small ribosomal protein subunit (rpsU).

**FIG. 1.** Vir typing of emm57 strains was performed as described previously (11). The amplified products of the mrg regulon were digested with HaeIII (A) or HinII (B). A marker (1-kb ladder; New England Biolabs) and strain 2031 (an M1 strain) were included (lanes M and 7, respectively). Lanes 1 to 6, isolates 2077 (a reference strain from Prague), NS1140, BSA16, BSA5, NS38, and NS48, respectively.

Rad followed by immunoblotting with anti-SIC (this study) and anti-His (Invitrogen) antibodies was used to identify the purified proteins.

**Antibodies.** CRS-specific rabbit antisera (IMVS, Gilles Plains, South Australia, Australia) was obtained by immunizing rabbits with recombinant M1 SIC (pQE30; Qiagen) from reference strain 2031.

**Detection of** crs57 **in culture supernatants.** Overnight cultures (10 ml) of S. pyogenes 2077, NS38, BSA16, BSA5, NS44, and NS27 in Todd-Hewitt broth (Oxoid) were centrifuged at 12,000 × g for 10 min, and 1 ml of each culture supernatant was transferred to a new tube. The culture supernatants were concentrated by using trichloroacetic acid (final concentration, 10%) at −20°C for approximately 20 min to induce precipitation. To retrieve the precipitated proteins, each mixture was centrifuged at 16,000 × g for 20 min. The supernatant was discarded, and the pellet was resuspended in 100 µl of 0.1 M NaOH. SIC was detected by separation of the sample by PAGE (Gradipore; Minirotein II; Bio-Rad), followed by Western blotting and detection with anti-SIC antibody.

**Binding of** CRS57 **to complement.** Assays for binding of CRS57 to the complement proteins C6 and C7 were performed by using an enzyme-linked immunosorbent assay. Ninety-six-well plates (TITERTEK) were coated with recombinant proteins (50 µg/ml; CRS57 or control proteins) in PBS at 4°C overnight. After blocking...
with 5% skim milk in PBS, complement from human sera (diluted 1/100 in PBS) or commercial C6 and C7 (diluted 1/1,000 in PBS; Sigma) were added and incubated for 1 hour at 37°C in a 100-μl (final volume) mixture. The wells were then washed three times with PBS containing 0.5% Tween. Horseradish peroxidase-conjugated secondary antibodies to the complement components C6 and C7 (diluted 1:1,000; ICN Biomedical) were used to determine the extent of binding. The reaction mixture was developed with 4-chloro-1-naphthol (Sigma), and the absorbance at 450 nm was determined with a Bio-Rad benchmark microplate reader.

Nucleotide sequence accession numbers. Nucleotide sequences of the PCR products obtained from the six independent isolates of the emm57 sequence type have been deposited in the GenBank database under accession numbers AY229856, AY229857, AY229858, AY229859, AY229860, and AF060764.

RESULTS

Divergence of crs57 sequence. PCR performed with primers flanking the 3’ end of the ABC transporter gene (SPy0778) and the 5’ end of rpsU suggested that in all emm57 isolates the two genes are farther apart than they are in serotype M1 strains or 39 other strains tested. Data for the six isolates of the emm57 sequence type and strains of representative non-emm57 sequence types are shown in Fig. 2. The six isolates of the emm57 sequence type belong to two distinct vir types, VT8 and VT101. The results of a comparison of the SPy0778-rpsU regions in emm1 and emm57 strains are presented schematically in Fig. 3. In all six emm57 isolates tested, the region between the SPy0778 and rpsU genes contained crs57. Remnants of the emm1 mga regulon were present immediately upstream and downstream of the crs57 gene. In emm1, there is an IS1562 sequence in reverse orientation between the sic gene and the scpA gene. The IS1562 sequence was not found within the SPy0778-rpsU region in the emm57 strains. However, the intervening sequence between the sic gene and the IS1562 in M1, including a copy of the terminal inverted repeat of the insertion element, was conserved in the SPy0778-rpsU region in all of the emm57 strains and occurred in the same position in relation to the crs57 gene (Fig. 3). These findings suggest that a crs57 gene was recently acquired by M57 or its progenitor, possibly by horizontal transfer from M1.

The emm57 strains used in this study include one strain from Prague (a reference strain) and five strains from the NT that belong to two distinct vir types (Fig. 1). Interestingly, the vir type profile of strain 2077 is identical to the VT8 profile. The derived CRS57 sequences of all of the VT8 isolates are highly homologous except for a few insertions or deletions of residues (including short and long repeats) (Fig. 4). A comparison of the corresponding DNA sequences also showed that there is a silent mutation in addition to the mutations that account for the amino acid changes. By contrast, the crs57 gene in NS1140, which is a VT101 strain, had several replacement mutations. Interestingly however, examination of the C-proximal half of the NS1140 CRS57 revealed a greater resemblance to the corresponding region of SIC from strain AP1 than to the corresponding region of VT8 CRS57. A comparison of the sic...

FIG. 2. PCR amplification of the SPy0778-rpsU intergenic region in isolates of the emm57 sequence type (lanes 1 to 6), in an isolate of the emm1 sequence type (lane 7), and in NS27 (lane 8). Lanes 1 to 6, amplification products from 2077, NS1140, BSA16, BSA5, NS38, and NS48, respectively. Lane M contained HindIII-digested lambda phage DNA as a size marker.

FIG. 3. Comparison of part of the mga locus in M1 and the SPy0778-rpsU region in M57 and other strains. The shaded areas in the SPy0778-rpsU region represent homologous regions. The mottled area between the regions surrounding the sic gene in M1 and the regions surrounding the crs57 gene in M57 shows a high level of similarity. This area also includes one of the inverted repeats of IS1562. The downward arrowheads beneath the mottled rectangle suggest that there was horizontal acquisition of this region from M1. The horizontal arrows show the relative positions of the ABC-F2 and rpsU-R1 primers used to amplify SPy0778-rpsU regions (the sequences are given in the text). The figure is not to scale.
sequence from strain AP1 (accession number X92968) and the VT8 crs57 sequence showed that there was a relatively high number of replacement mutations in the C-proximal third of the molecule (26 replacements in the C-proximal third and 12 replacements in the rest of the protein). Of particular interest is the conservation of tryptophan-containing short repeats. The frequencies of the five amino acids surrounding the tryptophan residues were determined (Fig. 5). The most conserved amino acid residues in the first three of the five positions are (D/G)-W-(S/G). We therefore refer to this as the DWS motif. The M57 isolates tested in this study each had 15 to 21 DWS motifs.

CRRS57 is an excretory product. The results described above showed that the crs57 gene is at the end of the ABC transporter immediately upstream of the rpsU gene in all of the emm57 strains. Since all of the crs57 sequences analyzed have a signal sequence, it is reasonable to expect that, like the SIC protein, the CRS57 protein in these strains is an excretory product. We confirmed this by analyzing culture supernatants and whole-cell extracts from overnight cultures of all six emm57 isolates by PAGE and Western blotting (Fig. 6; data for NS1140 not shown). Anti-SIC antibodies reacted with 40- to 50-kDa bands in culture supernatants of strain 2031 (a positive control strain) and the emm57 strains, as reported previously (6). The observed size differences between the CRS57 proteins of different strains in Fig. 6 is consistent with derived crs57 sequence data.

Demonstration of binding of CRS57 to complement proteins C6 and C7. Binding of SIC from an M1 strain to intermediate terminal complement complexes was recently demonstrated by Fernie-King et al. (6). To show that recombinant CRS57 proteins bind to C6 and C7 complement proteins in serum, we purified the fusion protein containing thioredoxin (pBAD thiocloning system). In our hands the expression of CRS57 in the pQE system gave low yields, the recombinant proteins accumulated in inclusion bodies, and the proteins were insoluble and often degraded upon purification. The thioredoxin fusion proteins...
improved the yield, solubility, and stability of the recombinant proteins considerably. Thioredoxin by itself did not bind to the complement proteins (Fig. 7). Plates coated with the recombinant proteins were reacted with human serum, and binding of the C6 and C7 complement proteins was detected by using secondary antibodies. The results demonstrated that the complement component C6 and C7 proteins were bound to CRS57 (Fig. 7A). To confirm this, we also tested binding using purified C6 and C7. In these experiments the recombinant CRS57 coated onto plates was reacted with the C6 and C7 proteins (Fig. 7B). The results confirmed that there is an interaction between the complement proteins and CRS57.

**DISCUSSION**

We reported previously (12) that in type 57 isolates the gene for CRS57 is outside the mga regulon, which contains emm, emm-like, scpA, and mga genes. In addition to confirming this, in this study we showed that in all six type 57 isolates tested the gene is located between the ABC transporter (Spy0778) and the rpsU gene. The organization of the Spy0778-rpsU region is highly conserved in the nonemm57 GAS strains tested and in three genomic sequences (M1, M3, and M18) (2, 7, 18). The CRS57 sequence of NS1140 exhibited considerable diversity compared to the other CRS57 sequences analyzed. Despite the difference, the junction points on either side of crs57 in all six emm57 strains are highly conserved. These observations suggest that horizontal acquisition of the crs gene must have been a one-off process and may have been a recent evolutionary event. Recently, Ma et al. (16) found that sic was distributed in a greater number of strains when they examined pharyngitis isolates from Japanese children. We believe that if horizontal acquisition was the mode of spread of sic among GAS strains, as these authors indicate, the source may been serotype M1 because this serotype has an intact IS1562.

Since CRS57 proteins have all of the structural characteristics of the SIC from an M1 strain, it is reasonable to expect that
CRS57 proteins also have similar biological properties. We tested this hypothesis by determining whether the CRS57 proteins, like SIC from M1 strains, bind to the C6 and C7 complement proteins. Our results show that the CRS57 proteins interact with the complement proteins. Hence, it may be inferred that CRS57 has the same properties as SIC from M1 with respect to the interaction with the complement proteins.

Recent serological observations in the NT indigenous population, in which post-streptococcus glomerulonephritis is highly endemic, revealed that 57% of the population has an-


diction of emm57 strains is high. Thus, at least the newly acquired antibodies to SIC in our study population are most likely due to emm57 strains. If so, this further supports our hypothesis that CRS57 is expressed and is antigenic during natural infection. Thus, if antibodies confer selection pressure for variants of sic from M1 strains (as proposed by Hoe et al. [14]), the same pressure could also operate to select variants of crs57 given that SIC and CRS57 have common biological and biochemical properties. Whereas a large number of mutations have been observed in the crs57 gene from VT101, the frequency of point mutations in the remaining five epidemiologically unrelated emm57 isolates was low; only short insertions or deletions, particularly of the repeat sequences, accounted for most of the mutations.

The mga regulon harbors antigenically highly variable genes. This regulon may be a mutational hot spot in the GAS genome. Since in M57 the crs57 gene is outside this region, it may be less prone to mutations than sic from M1.

All CRS proteins have short conserved DWS repeats. The tryptophan content of some CRS molecules could be as high as 6%. The role of the conserved tryptophan-containing motif is not known. In some proteins tryptophan-rich motifs may have a role in membrane binding. For instance, streptolysin O, a thiol-activated cytolysin, has a tryptophan-rich domain in the C-terminal region which is essential for membrane binding (21). Since SIC is known to enter host cells efficiently (13), it is possible that the DWS motif is responsible for promoting an interaction between SIC and the host cell membrane. Further work to test this hypothesis is in progress.

ACKNOWLEDGMENTS

This work was supported by the Australian National Health and Medical Research Council.

We thank Jon Hartas for technical help with the initial cloning experiments.

REFERENCES


