

Identification and Immunogenicity of Group A *Streptococcus* Culture Supernatant Proteins

BENFANG LEI,¹ STACY MACKIE,¹ SLAWOMIR LUKOMSKI,² AND JAMES M. MUSSER^{1*}

Laboratory of Human Bacterial Pathogenesis, Rocky Mountain Laboratories, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Hamilton, Montana 59840,¹ and Department of Pathology, Baylor College of Medicine, Houston, Texas 77030²

Received 2 August 2000/Returned for modification 6 September 2000/Accepted 15 September 2000

Extracellular proteins made by group A *Streptococcus* (GAS) play critical roles in the pathogenesis of human infections caused by this bacterium. Although many extracellular GAS proteins have been identified and characterized, there has been no systematic analysis of culture supernatant proteins. Proteins present in the culture supernatant of strains of serotype M1 (MGAS 5005) and M3 (MGAS 315) mutants lacking production of the major extracellular cysteine protease were separated by two-dimensional gel electrophoresis and identified by amino-terminal amino acid sequencing and interrogation of available databases, including a serotype M1 genome sequence. In the aggregate, amino-terminal amino acid sequence data for 66 protein spots were generated, 53 unique sequences were obtained, and 44 distinct proteins were identified. Sixteen of the 44 proteins had apparent secretion signal sequences and 27 proteins did not. Eight of the 16 proteins with apparent secretion signal sequences have not been previously described for GAS. Antibodies against most of the apparently secreted proteins were present in sera from mice infected subcutaneously with MGAS 5005 or MGAS 315. Humans with documented GAS infections (pharyngitis, acute rheumatic fever, and severe invasive disease) also had serum antibodies reacting with many of the apparently secreted proteins, indicating that they were synthesized in the course of GAS-human interaction. The genes encoding four of the eight previously undescribed and apparently secreted culture supernatant proteins were cloned, and the proteins were over-expressed in *Escherichia coli*. Western blot analysis with these recombinant proteins and sera from GAS-infected mice and humans confirmed the immunogenicity of these proteins. Taken together, the data provide new information about the molecular aspects of GAS-host interactions.

The human bacterial pathogen group A *Streptococcus* (GAS) causes a wide variety of diseases globally (35). GAS is the primary cause of pharyngitis, an infection that causes substantial morbidity and economic loss and can lead to acute rheumatic fever and rheumatic heart disease, the main cause of preventable pediatric heart disease globally. Although largely controlled in the United States and other western countries, the disease persists in certain regions, for example, Salt Lake City, Utah (49). GAS can also cause serious invasive infections such as septicemia, streptococcal toxic shock syndrome, and necrotizing fasciitis. The fatality rate for patients with invasive infections has exceeded 50% in some case series. The organism can cause other postinfection sequelae such as glomerulonephritis, a disease that is also a major public health problem in developing countries. There is currently no licensed vaccine to prevent human GAS infection.

Research conducted over many decades has shown that much of the host pathology caused by GAS is mediated by extracellular proteins (8, 12, 35). Three general categories of extracellular proteins have been recognized (12). First, proteins that are actively secreted into the extracellular environment and are largely or solely found free in the culture supernatant have been described. These proteins have secretion signal sequences and are presumed to be actively exported. Examples of these proteins include potent superantigens such as streptococcal pyrogenic exotoxins A and C. A second cate-

gory of GAS extracellular proteins includes molecules with a secretion signal sequence and a conserved hexapeptide sequence (Leu-Pro-X-Thr-Gly [LPXTG]) located at the carboxy terminus that anchors the protein to the bacterial cell membrane (12). Some of these proteins or fragments thereof can also be present free in the culture supernatant, especially in the stationary phase of growth. These cell surface-anchored proteins include known virulence factors such as M protein and C5a peptidase. In recent years, evidence has been presented that a third category of GAS extracellular proteins exists (37–39, 44, 51). These proteins lack apparent secretion signal sequences and the LPXTG membrane anchor motif. Although their presence in culture supernatants may be related in part to passive release from the intracellular compartment due to cell wall autolysis, the observation that some of these proteins are found in relatively high concentrations in the logarithmic phase of growth suggests that an unknown transport mechanism participates. Interestingly, two proteins that are assigned to this class (glyceraldehyde-3-phosphate dehydrogenase and α -enolase) are enzymes involved in the glycolytic pathway of metabolism (37–39, 51).

Despite the reasonably detailed understanding of GAS extracellular proteins, there has been no systematic analysis of culture supernatant proteins with contemporary investigative methods. Importantly, this general line of investigation has provided critical new information about extracellular proteins made by diverse pathogens, such as *Mycobacterium tuberculosis*, *Haemophilus influenzae*, *Listeria monocytogenes*, and *Helicobacter pylori* (5, 6, 19–21, 25, 41, 54). For example, new vaccine candidates and drug and diagnostic targets have been identified.

The goal of the present study was to identify and characterize proteins found in relative abundance in culture superna-

* Corresponding author. Mailing address: Laboratory of Human Bacterial Pathogenesis, Rocky Mountain Laboratories, National Institute of Allergy and Infectious Diseases, National Institutes of Health, 903 S. 4th St., Hamilton, MT 59840. Phone: (406) 363-9315. Fax: (406) 363-9427. E-mail: jmusser@niaid.nih.gov.

tants of GAS strains grown *in vitro*. A comprehensive understanding of GAS pathogenesis requires molecular dissection of the pathogen proteome, especially proteins found in culture supernatants. As a first step toward this end, we analyzed culture supernatant proteins made by strains of serotype M1 and M3 GAS by two-dimensional electrophoresis and amino-terminal amino acid sequencing. Isolates expressing these M proteins are the two more abundant types causing human invasive infections such as septicemia and necrotizing fasciitis (33). To identify supernatant proteins expressed in the course of host-pathogen interactions, we used two-dimensional Western blot analyses with sera obtained from mice with experimentally induced soft tissue infection and from humans with pharyngitis, acute rheumatic fever, and severe invasive disease. The genes encoding four of eight apparently secreted and previously undescribed culture supernatant proteins were cloned, and recombinant proteins were overexpressed in *Escherichia coli*. Western blot analysis with these recombinant proteins and sera from GAS-infected mice and humans confirmed the immunogenicity of the four proteins.

MATERIALS AND METHODS

Bacterial strains and routine growth. Strains MGAS 5005 (serotype M1) and MGAS 315 (serotype M3) have been described previously and characterized extensively (16, 17, 28–30, 33). Briefly, MGAS 5005 contains the gene (*speA*) encoding streptococcal pyrogenic exotoxin A, expresses the most common variant of M1 and streptococcal inhibitor of complement (1, 17), and is representative of the most abundant serotype M1 strains recovered from humans with invasive disease globally. Similarly, MGAS 315 has the *speA* gene and all genetic features typical of serotype M3 strains commonly cultured from patients with invasive disease.

E. coli XL-1 Blue (Stratagene, La Jolla, Calif.) or *E. coli* NovaBlue and BL21(DE3) (Novagen, Madison, Wis.) were used for gene cloning and protein expression. GAS strains were grown routinely in Todd-Hewitt broth (Difco Laboratories, Detroit, Mich.) supplemented with 0.2% yeast extract (THY). Brain heart infusion (BHI) agar (Difco Laboratories) or tryptose agar with 5% sheep blood (Becton Dickinson, Cockeysville, Md.) was used as a solid medium. For mutant selection, BHI agar supplemented with spectinomycin (150 µg/ml) was used. The GAS strains were grown at 37°C in a 5% CO₂–20% O₂ atmosphere.

Construction of *speB*-inactivated isogenic mutant strains. A promoterless nonpolar spectinomycin resistance gene (*aad* gene) cassette was initially used to construct the *speB*-negative isogenic mutant (26, 28). However, the desired mutants of MGAS 5005 and MGAS 315 could not be obtained, presumably due to the failure of the mutant to survive antibiotic selection. The *speB* gene is expressed in the stationary phase of growth, well after antibiotic selective pressure has been supplied. Hence, when the *speB* gene is regulated by its native promoter, it is unlikely that transcription of the gene and the contiguous antibiotic resistance-conferring gene are expressed early during selection. To overcome this problem, a new nonpolar cassette with the spectinomycin resistance gene linked to a constitutive promoter was constructed. The *aad* gene and its promoter (26) were amplified using plasmid pEU904 (obtained from J. R. Scott, Emory University) DNA as the template and primers 5'-GGACCCGGAATA CATGTTATAATAACTATAAC-3' and 5'-CCTCCGGGCATGTGATTTTC CTCCTTTTATAATTTTTTAATCTGTTA-3'. The PCR product was digested with *Sma*I and inserted into pLEX5B (10) to generate recombinant plasmid pSL81. Plasmid pSEB1719 harboring the *speB* gene (36) was digested with *Sty*I and blunt ended with Klenow fragment DNA polymerase I. The *aad*-containing *Sma*I fragment of pSL81 was inserted between the *Sty*I sites to replace the majority of the *speB* gene. The suicide plasmid obtained (pSL82) was introduced into MGAS 5005 and MGAS 315 by electroporation. The recombinants were selected on medium with 150 µg of spectinomycin per ml. Presumed mutants were identified and confirmed by PCR, Southern blotting, and DNA sequencing. The *speB* mutants lacked proteolytic activity, as assessed by the casein plate assay (36).

Isolation of culture supernatant proteins. The *speB* mutants of MGAS 315 and MGAS 5005 were cultured in THY lacking proteins with a molecular mass greater than 10 kDa (protein-reduced THY [PR-THY]). To prepare PR-THY, THY was passed through a 0.22-µm-pore-size filter and the resulting material was filtered with a membrane cartridge with a 10-kDa molecular mass cutoff (S3Y10; Millipore Corporation, Bedford, Mass.) installed in a ProFlux M12 tangential-flow filtration system (Millipore Corporation). Proteins with a molecular mass greater than 10 kDa were retained by the cartridge and discarded. To isolate proteins produced in the exponential phase, MGAS 315 *speB* or MGAS 5005 *speB* was grown overnight on BHI agar and inoculated into 400 ml of PR-THY. The organisms were grown to exponential phase (optical density at 600

nm [OD₆₀₀] of ~0.4) in 5% CO₂ at 37°C. Forty milliliters of the culture was added to 900 ml of PR-THY in each of eight 1-liter bottles. The bottles were incubated in 5% CO₂ at 37°C and harvested at an OD₆₀₀ of 0.4 to 0.5. To isolate proteins produced in the stationary phase, 1 ml of exponential-phase (OD₆₀₀ ~0.4) GAS was inoculated into 900 ml of PR-THY in each of eight 1-liter bottles. The bacteria were grown in 5% CO₂ at 37°C for 16 to 18 h and pelleted by centrifugation at 12,000 × *g* for 10 min. For all growth conditions, the supernatant was concentrated to 350 ml by passage through a 0.22-µm-pore-size filter and the Millipore filtration system equipped with cartridge S3Y10. Proteins in the concentrated supernatant were precipitated by addition of (NH₄)₂SO₄ to 90% saturation and recovered by centrifugation at 18,000 × *g* for 20 min. The pellet was dissolved in a small amount of water. The proteins were dialyzed against 3 liters of water in Slide-A-Lyzer dialysis cassettes with a 10-kDa molecular mass cutoff (Pierce, Rockford, Ill.) at 4°C for 20 h with one change of water. The sample was stored at –20°C in 1.5-ml aliquots.

Two-dimensional electrophoresis. The first-dimension isoelectric focusing was performed with the IPGphor IEF system and Immobiline DryStrip gel strips (7 cm) with a linear immobilized pH gradient of 3 to 10 (Amersham Pharmacia Biotech, Inc., Piscataway, N.J.). The sample was prepared by adding a saturating amount of solid urea to the starting protein solution and mixing this solution with an equal volume of rehydration stock solution (8 M urea, 4% 3-[(3-cholamidopropyl)-dimethylammonio]-1-propanesulfonate [CHAPS], 4% IPG buffer, 50 mM dithiothreitol). The sample (130 µl), gel strip, and mineral oil were sequentially added to a 7-cm strip holder. The holder was loaded into the IPGphor IEF system. The strip was rehydrated with the sample solution for 11 h, and the proteins were focused in three steps consisting of 500 V for 1 h, 1,000 V for 2 h, and 8,000 V for 7 h at 20°C.

The second dimension (sodium dodecyl sulfate-polyacrylamide gel electrophoresis [SDS-PAGE]) was carried out with a Tris-HCl Ready Gel (Bio-Rad Laboratories, Hercules, Calif.). After isoelectric focusing, the gel strip was rocked gently in SDS equilibration buffer (50 mM Tris-Cl [pH 8.8], 6 M urea, 30% glycerol, 2% SDS, 10 mM mercaptoethanol, bromophenol blue) for 15 min. SDS-PAGE was conducted at 150 V for 1 h, and the gel was stained with GelCode Blue Stain Reagent (Pierce). Alternatively, the proteins were transferred to a membrane for amino-terminal amino acid sequencing or Western blotting.

Amino-terminal amino acid sequencing. Proteins separated by two-dimensional electrophoresis were transferred to a polyvinylidene difluoride membrane (Bio-Rad Laboratories) with Towbin transfer buffer using a Trans-Blot SD semidry transfer cell (Bio-Rad Laboratories) at 15 V for 40 min. The membrane was washed three times with water, stained with 0.1% Coomassie blue in 40% methanol and 1% acetic acid, and destained with 50% methanol aqueous solution. Amino-terminal amino acid sequences of the stained spots in the polyvinylidene difluoride membrane were determined by Edman degradation with an ABI/Perkin-Elmer 494 cLC protein sequencer.

Mouse sera. Mouse sera were obtained from adult (18- to 20-g) male outbred CD-1 Swiss mice (strain ICR; Harlan-Sprague-Dawley Inc., Houston, Tex.) inoculated subcutaneously with wild-type MGAS 5005 (serotype M1) and MGAS 315 (serotype M3). The animals were anesthetized by Metofane (Mallinckrodt Veterinary, Mundelein, Ill.) inhalation before all experimental procedures. Control sera were obtained prior to mouse inoculation. Inocula were prepared from GAS cultures grown to mid-logarithmic phase (OD₆₀₀ of ~0.5). The bacterial cells were harvested and washed once with sterile ice-cold and pyrogen-free phosphate-buffered saline (PBS). The cells were resuspended in an appropriate volume of PBS to give the required inoculum, which was verified by colony counts with tryptose agar plates containing 5% sheep blood.

Groups of ~115 mice were injected subcutaneously in the right flank with wild-type MGAS 5005 (2.3 × 10⁷ CFU) or MGAS 315 (1.7 × 10⁷ CFU). Preliminary experiments found that about 70 to 80% of mice survive and develop prominent gross cutaneous lesions with these inocula. Cutaneous lesions caused by these inocula healed over time in the surviving mice. The animals were sacrificed 20 to 28 days after inoculation. Serum was obtained by cardiac puncture and stored at –20°C. Sera obtained from cohorts of animals inoculated with each strain were pooled and used for Western blot analyses.

Human sera. Inasmuch as GAS causes several very different types of infections, sera obtained from humans with three distinct diseases were used for Western blot analysis. Convalescent-phase serum was obtained from a child with uncomplicated GAS pharyngitis. The convalescent-phase serum was drawn 18 days after the pharyngitis diagnosis. The M protein type of the infecting strain was not known, and the organism was not available for analysis. Paired acute- and convalescent-phase sera obtained from a patient with an invasive GAS infection were also studied. The convalescent-phase serum was drawn 48 days after the invasive-disease diagnosis. The M type of the infecting strain (serotype M1) was inferred by sequence analysis of the region of the *emm* gene encoding the hypervariable amino terminus of this protein. We also used serum drawn from a patient in the acute phase of an initial rheumatic fever episode. The M protein type of the infecting strain was not known, and the organism was not available for analysis. However, serologic analysis indicated infection in an M1 organism.

Western blot analysis. Proteins separated by two-dimensional electrophoresis or SDS-PAGE were transferred to nitrocellulose membranes (Immobilon-NC; Millipore Corporation) with Towbin transfer buffer using a Trans-Blot SD semi-

TABLE 1. Primers, vectors, cloning sites, and recombinant plasmids used in gene cloning

Gene ^a	Primers	Vector/cloning sites	Plasmid
<i>sp13</i>	5'-GTGGCTAGCCAAGATGTAAATGCTTGAGACG-3' 5'-CGGATCCTTACTCTTGCCAATAACATATGTTAG-3'	pET21b/ <i>NheI</i> , <i>Bam</i> HI	pSP13
<i>sp22</i>	5'-GTGTTTCATATGGATAGTTTTTCTGCTAATCAAG-3' 5'-AGGATCCTTAATTGGTCTGATTCCAAC-3'	pET21b/ <i>NdeI</i> , <i>Bam</i> HI	pSP22
<i>sp24-f1</i>	5'-GCGCAACATATGACCAATACAGAGTTGAGC-3' 5'-AGGATCCTAATGCATACCGTGTGACTCG-3'	pET21b/ <i>NdeI</i> , <i>Bam</i> HI	pSP24-f1
<i>sp24-f2</i>	5'-ACCATGGATCCTGAAACAAACCGATTC-3' 5'-CGAATTCTAACCATCTCTTTTGGAAATGGT-3'	pET21d/ <i>NcoI</i> , <i>EcoRI</i>	pSP24-f2
<i>sp35</i>	5'-ACCATGGAGGATTTAAGTACTAAG ATTG-3' 5'-CGGATCCTTAGTGTGGATAAATAATGTAAC-3'	pET21d/ <i>NcoI</i> , <i>Bam</i> HI	pSP35

^a Fragment 1 (f1) represents amino acid residues 57 to 282. Fragment 2 (f2) represents amino acid residues 1029 to 1366.

dry transfer cell (Bio-Rad Laboratories) at 15 V for 40 min. The membrane was treated with 10 ml of block solution (1:2,000 Amersham Liquid Block in 150 mM NaCl and 100 mM Tris-HCl, pH 7.4) for 1 h and incubated for 1 h with primary antibodies added to the block solution. The membrane was then rinsed twice and washed three times for 15 min each with 0.1% Tween 20 in PBS. The membrane was incubated with goat anti-mouse or anti-human immunoglobulin G (heavy plus light chains) horseradish peroxidase-conjugated secondary antibodies (Sig-

ma, St. Louis, Mo.) for 1 h and rinsed and washed as described above. Antigen-antibody reactivity was visualized by enhanced chemiluminescence.

Cloning and expression of the genes encoding SP13, SP22, SP24, and SP35. The primers, vectors, and restriction enzymes used for PCR amplification and cloning of the *sp13*, *sp22*, and *sp35* genes and two *sp24* gene fragments are listed in Table 1. The primers were designed on the basis of gene sequences present in an available M1 GAS genome sequence (*Streptococcus pyogenes* Genome Se-

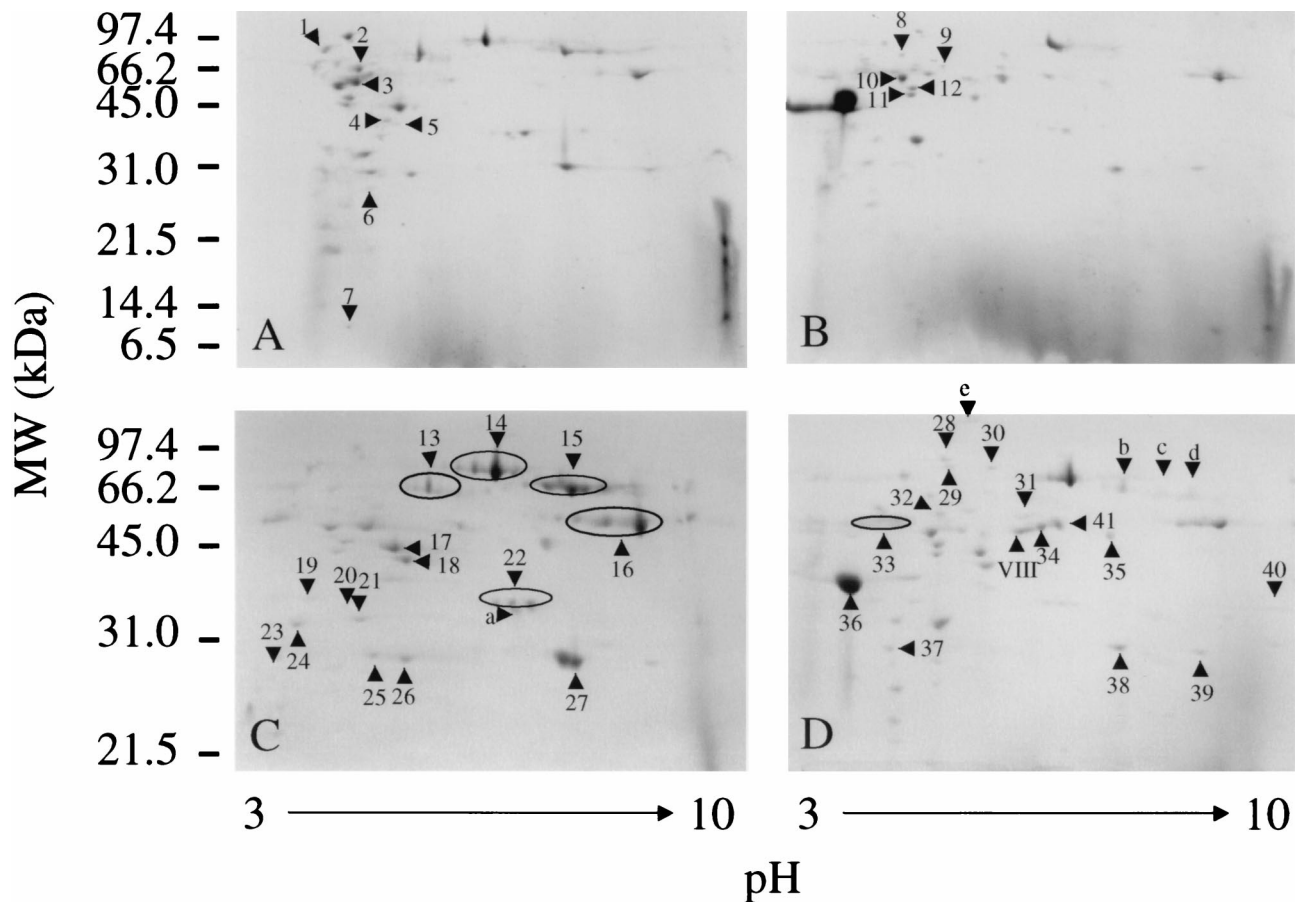


FIG. 1. Two-dimensional electrophoresis profiles of GAS culture supernatant proteins. Supernatant proteins (340 μ g) were subject to isoelectrofocusing in a pH range of 3 to 10 and then resolved by SDS-PAGE with a 15% (A and B) or 12% (C and D) gel. The proteins were stained with GelCode Blue Stain Reagent (Pierce). The numbered and lettered triangles indicate the protein spots whose amino-terminal amino acid sequences were determined by Edman degradation. (A and B) Proteins obtained from culture supernatants of *speB* mutants of MGAS 315 (serotype M3) (A) and MGAS 5005 (serotype M1) (B) grown to exponential phase. (C and D) Proteins obtained from stationary-phase culture supernatants of *speB* mutants of MGAS 315 (C) and MGAS 5005 (D).

TABLE 2. Proteins identified in the culture supernatants of the *speB* mutants of GAS serotype M1 MGAS 5005 and serotype M3 MGAS 315

Spot no. ^a	Amino-terminal amino acid sequence ^b	GenBank accession or ORF no.	Identical or homologous protein	Molecular mass (kDa) ^c	pI ^c
1	2-TEMLKGIAAS	RST01348	Phosphotransferase enzyme I	63.1	4.48
2	1-MkKRVKIVATL	RST00804	Pyruvate kinase	54.5	4.86
3	1-TAQTPIHVYSE	RST01429	Arginine deiminase	46.6	4.93
4	2-TATKQHKKVVIL	RST00350	Lactate dehydrogenase	35.3	5.07
5	1-MKRIAIVLTSGG	RST00803	6-Phosphofructokinase	35.7	5.36
6	2-VKLVFARHGES	RST01563	Phosphoglyceromutase	26.1	5.01
7	45-GGAEEAAKDSF	RST00830	Ribosomal protein L7/L12	12.2	4.40
8	2-AKDIKFSADAR	RST00771	GroEL	57.1	4.65
9	2-AKQYKNLVNGE	RST01349	NADP ⁺ -dependent glyceraldehyde-3-phosphate dehydrogenase	50.3	4.99
10	2-SIITDVYAREV	RST00465	Enolase	47.3	4.64
11	2-AEITAKLVKEL	RST00510	Elongation factor EF-Ts	37.2	4.76
12	2-AKLTVKDVLK	RST00513	Phosphoglycerate kinase	37.2	4.76
13	28-QDVMLETHKA	RST00809	Immunogenic secreted protein homologue	51.5	5.67
14	35-ESNKQNTASTE	RST00389	Streptolysin O	60.4	6.11
15 ^d	42-DARSVNGEFP	U40231	M3 protein	55.9	7.01
16	38-VSGKENKSDV	RST00387	NAD ⁺ -glycohydrolase	46.5	7.92
17	2-TQVFQGRSFL	RST01431	Ornithine carbamoyltransferase	37.9	5.16
18	2-VVKVGINGFGR	RST00038	Plasmin receptor/glyceraldehyde-3-phosphate dehydrogenase	35.9	5.33
19	2-TKQKIVxALG	RST01436	Carbamate kinase	33.2	4.60
20	1-MLVYIAGSSGAM	RST00882	Dehydropantoate 2-reductase	33.8	4.77
21	2-AIVSAEKVFQ	RST00186	Aldolase	31.2	4.81
22	30-DSFSANQEIR	RST01222	Human Mac-1	34.9	6.41
23	38-ETIASSSTRHF	RST01220	Autolysin	21.6	4.23
24	35-DELSTMSEPT	RST00815	Serine protease	178	6.15
25 ^d	[DSG][APV][RP]SVNGEFP	U40231	M3 protein		
26 ^d	23-QQDPDPSQLHR	U40453	Streptococcal pyrogenic exotoxin A	25.8	5.45
27	28-HDNIDEKGVH	RST01215	Unknown function	24.5	7.61
28	2-TFDAIDQLAV	RST01509	Transketolase	71.3	4.81
29	2-AREFLAKTRN	RST00093	Elongation factor G	76.5	4.73
30	31-ADNASVTNKADF	RST00768	Cyclodextrin glucosyltransferase	76.2	5.41
31	42-NGDGNPREVIE	RST01579	M1 protein	49.9	5.51
32	2-TTIDFKAEVDK	RST00833	Dipeptidase	51.3	4.77
33	38-VSGKENKSDV	RST00387	NAD ⁺ -glycohydrolase	46.5	7.92
34	66-LRhenKDLKARL	RST01579	M1 protein	47.3	5.71
35	25-EDLSTKIAK	RST01497	<i>L. lactis</i> secreted 45-kDa protein	39.6	6.65
36	33-ETYTSXXXX	RST00460	Streptococcal inhibitor of complement	31.2	4.17
37	2-SRKPIIAGN	RST00099	Triosephosphate isomerase	26.6	4.44
38	44-QTQVSNDDVVLN	RST01835	Mitogenic factor (DNase)	25.4	7.38
39	43-RQTQVSNDDVVL	RST01835	Mitogenic factor (DNase)	25.5	8.25
40	217-SSEKEQLTIEK	RST01579	M1 protein	29.4	9.04
41	27-IAGYGWLPDRPP	RST00593	Streptokinase	47.1	6.21
I ^d	SIRSLFGLRE	— ^d	— ^d		
II ^d	KKIDRVALETL	— ^d	— ^d		
III	2-STSFENKATNR	RST00259	RopA	47.1	4.29
IV ^d	AIEASKLQAG	— ^d	— ^d		
V	2-ASKDFHIVAET	RST01347	Phosphocarrier protein	9.0	4.63
VI	2-AeYEILYIIRP	RST01606	Ribosome protein S6	11.1	4.89
VII	2-ANKQDLIAKV	RST00136	Histone-like protein	9.6	9.04
VIII ^d	MHQPHQPLPP	— ^d	— ^d		
IX	25-SGPKVPYTQE	RST01706	Class B acid phosphatase	24.5	9.10
X	2-SKIIGIDLGTT	RST00474	DnaK protein	64.9	4.49
XI	2-ANAIITAKE	RST00322	Ribosome recycling factor	20.6	5.66
XII	2-SRIGNKVITM	RST00748	Ribosomal protein L6	19.4	9.57

^a Spots I to XII are not in Fig. 1.

^b The number indicates the position of the first detected amino acid residue in the amino acid sequence inferred on the basis of the ORF. Lowercase letters indicate amino acids that do not match the residues in the amino acid sequences derived from the M1 genome sequence. X indicates amino acid residues that could not be determined.

^c Theoretical values calculated on the basis of the mature protein, presuming an intact carboxy terminus.

^d Protein whose gene was not found in the available M1 genome sequence.

quencing Project, University of Oklahoma [http://genome.ou.edu/strep.html]), and some of the nucleotides were altered to create restriction enzyme sites for cloning. PCR products obtained with MGAS 5005 DNA were digested with the appropriate restriction enzymes and cloned into pET-21b or pET-21d to obtain recombinant plasmids (Table 1). The resulting cloned genes and gene fragments were sequenced to rule out introduction of spurious mutations.

To assess expression of the cloned genes or gene fragments, recombinant *E. coli* BL21 was grown at 37°C for 10 h in 3 ml of Luria-Bertani broth supplemented with 100 mg of ampicillin per liter. The OD₆₀₀ of the cultures was measured. Cells were pelleted by centrifugation and resuspended in 1× SDS-

PAGE loading buffer at a ratio of 100 μl of buffer per 1 OD₆₀₀ unit per ml. The samples were boiled for 3 min before being loaded for protein separation with SDS-PAGE.

RESULTS

Two-dimensional electrophoresis profiles of culture supernatant proteins. We chose to focus this analysis on serotype M1 and M3 GAS strains, 2 of greater than 100 distinct M types

MNKNKLLRVAMLLSLLAPTAESMTVL <i>↓</i> AQ	SP13
MKDM SNKKTFFKKYSRVAGLLTAAL IIGNLV TANAE	SP14
MAKNNTNRHYS LRK LKTGTASVAVALTVLGTGLVAGQTYKAD	SP15
MRNKKVTLAHIVAKT SVAIALAGAMGSSLLANSTTYAV	SP16
MRKRCYSTSAAVLAAVTLFVLSVDRGV:AD	SP22
MTKKKGKLVLSLFLVLAACLGAYSAMPQSHKTSNVSAE	SP23
MEKKQRFSLRKYKSGTFSVLIGSVFLVMTTVAAD	SP24
MENNNKVLKMMVFFVLVTFGLTISQEVFAQ	SP26
MKRCNKYLFTSLLAASMLFSSYKSVHAH	SP27
MRELHIKTYKLLTKSAVLLGLISFPLTVSAA	SP30
MAKNNTNRHYS LRK LKTGTASVAVALTVLGAGFANQTEVKAN	SP33
MKKRILSAVLVSGVILGAATTVGAE	SP35
MNIRNKIENSKTLLFTSLVAVALLGATQPVSAE	SP36
MNLLGSRRVFSK KCR LVKFSMVALVSATMAVTTVTLENTALAR	SP39
MKNYLSIGVIALLFALTFGTVKSVQAI	SP41
MKKEFTSILFTVSFCGIIALPVEAS	SPIX

FIG. 2. Putative secretion signal sequences. The first amino acids of the mature proteins (rightmost residue) are the residues after the italicized alanine residue. The arrow indicates the processing cleavage site. The amino acid sequences of the cleaved peptides for 13 of the 16 proteins were derived from the cognate ORFs in an available M1 genome sequence. The amino acid sequences of SP15, SP26, and SP36 were inferred from the nucleotide sequences deposited in GenBank under accession numbers X80186, U40453, and AF232306. The single-letter amino acid abbreviations are used.

that have been described. Because genetic heterogeneity exists among strains expressing these serotypes (16, 17, 33, 34), we selected strains that have been extensively characterized genetically and phenotypically and are known to represent the abundantly occurring M1 and M3 subclones of each serotype infecting humans globally. These strains also are readily genetically manipulated and are virulent for mice. The major drawback of use of the chosen strains is that the genome sequences are not yet available for these organisms. Isogenic derivatives in which the gene encoding the major extracellular protease had been insertionally inactivated were used for three main reasons. First, we thought that use of the *speB* mutants would enhance the probability of successful identification of stationary-phase proteins, since degradation of culture supernatant proteins would be greatly reduced. Second, we thought that use of the *speB* mutants would enhance identification of immunogenic proteins, also due to decreased target protein degradation. Third, decreased proteolysis of culture supernatant proteins would also result in the need to identify fewer protein spots and would be more efficient.

Two-dimensional electrophoresis of the culture supernatant proteins from the *speB* mutants of serotype M1 MGAS 5005 and serotype M3 MGAS 315 grown to exponential and sta-

tionary phases identified about 50 abundant spots in each sample (Fig. 1). (For purposes of description and uniformity, we will refer to these spots as protein spots throughout this paper.) The supernatant protein profiles were similar for the two strains in each growth phase, but some differences were detected. For example, protein spots a, 15, 25, and 27 in Fig. 1C (MGAS 315 *speB*) are absent in Fig. 1D (MGAS 5005 *speB*), and conversely, protein spots 31, 34, 36, and 41 in Fig. 1D are absent in Fig. 1C. In addition, certain protein spots were present uniquely or predominantly in culture supernatant proteins in a growth phase-dependent fashion. For example, protein spots 17, 27, and 34 were preferentially present at stationary phase (Fig. 1). Numerous protein spots had neighboring protein spots with very similar apparent molecular weights but distinct pIs.

Protein spot identification. The amino-terminal 10 or 12 amino acid residues for a total of 66 protein spots were determined by Edman degradation, and the data were used to interrogate available databases, including a GAS M1 genome sequence (<http://genome.ou.edu/strep.html>). Fifty-two of the 53 unique amino-terminal sequences obtained were unambiguous, and one had ambiguous identification of the first three residues (Table 2). Forty-six of the 53 amino-terminal amino

TABLE 3. Identified streptococcal proteins that contain secretion signal sequences

Function and characterization status	Extracellular proteins
Previously characterized	NAD ⁺ -glycohydrolase (SP16), streptolysin O (SP14), M1 (SP31) and M3 (SP15) protein, mitogenic factor (DNase) (SP38), streptococcal pyrogenic exotoxin A (SP26), streptococcal inhibitor of complement (SP36), streptokinase (SP41)
Homologous to protein with known function.....	Class B acid phosphatase (SPIX), serine protease (SP24), cyclodextrin glucosyltransferase (SP30), autolysin (SP23), SP22
Homologous to known protein but no known function.....	Immunogenic secreted protein homologue (SP13), <i>L. lactis</i> putative secreted protein (SP35)
No known homologue and no known function.....	SP27

TABLE 4. Presumed or potential cytosolic proteins identified in the GAS culture supernatants

Protein class	Protein(s)
Glycolysis	6-Phosphofructokinase (SP5), aldolase (SP21), triose-phosphate isomerase (SP37), phosphoglyceromutase (SP6), glyceraldehyde-3-phosphate dehydrogenase (SP18), phosphoglycerate kinase (SP12), enolase (SP10), pyruvate kinase (SP2)
Chaperonin	GroEL (SP8), DnaK (SPX), RopA (SPIII)
Ribosomal	L6 (SPXII), L7/12 (SP7), S6 (SPVI)
DNA binding	Histone-like protein (SPVII)
Protein synthesis	Elongation factor G (SP29), elongation factor EF-Ts (SP11), ribosome recycling factor (SPXI)
Urea cycle	Carbamate kinase (SP19), ornithine carbamoyltransferase (SP17), arginine deiminase (SP3)
Sugar transport	Phosphotransferase enzyme I (SP1), phosphocARRIER protein (SPV)
Other enzymes	Lactate dehydrogenase (SP4), NADP ⁺ -dependent glyceraldehyde-3-phosphate dehydrogenase (SP9), transketolase (SP28), dipeptidase (SP32), dehydropanoate 2-reductase (SP20)

acid sequences were 100% identical to the inferred sequence of an open reading frame (ORF) present in the available M1 genome sequence. Four of the 53 amino-terminal sequences had one or two mismatched amino acid residues and had a cognate ORF in the genome. It is unknown if the amino acid mismatches were due to naturally occurring polymorphisms or nucleotide sequencing errors. The genes encoding 47 protein spots were identified from an available GAS M1 genome se-

quence. The *emm3* and *speA* genes, encoding protein spots 15 and 28, respectively, were obtained by analysis of the GenBank database. The identities of three MGAS 315 protein spots and one MGAS 5005 protein spot could not be discerned from available database information.

Several distinct protein spots had amino-terminal amino acid sequences indicating that they were related proteins. For example, two protein spots (spots 38 and 39 [Fig. 1]) identified as mitogenic factor apparently arose by alternative proteolytic cleavage of the signal peptide. Protein spots 31, 34, and 40 had amino-terminal amino acid sequences corresponding to sequences beginning at amino acid residues 42, 66, and 217, respectively, of M1 protein. Protein spot 31 was the expected mature M1 protein whereas protein spots 34 and 40 apparently arose due to internal proteolysis by an enzyme other than the SpeB cysteine protease. Protein spots 16 and 33 each had the identical amino-terminal amino acid sequence appropriate for NAD⁺-glycohydrolase (3, 14, 24) (referred to as NADase in this report) but had different pIs (Fig. 1). The apparent molecular weight of protein spot 33 was slightly less than that of protein spot 16, and the spot may have arisen as a consequence of proteolytic cleavage toward the carboxy terminus of the protein. In the aggregate, 66 protein spots were analyzed, 53 unique amino-terminal amino acid sequences were obtained, and 44 unique proteins were identified (Table 2).

Proteins with apparent secretion signal sequences. The amino-terminal sequences of 17 proteins did not start with the first or second inferred amino acid residue of the ORF and apparently were products of a proteolytic process. The inferred amino-terminal amino acids of 16 of the 17 proteins have characteristics typical of secretion signal sequences, including charged residues followed by a stretch of hydrophobic residues and cleavage after an alanine residue (Fig. 2). Ribosomal protein L7/L12 (protein spot 7) was excluded from fur-

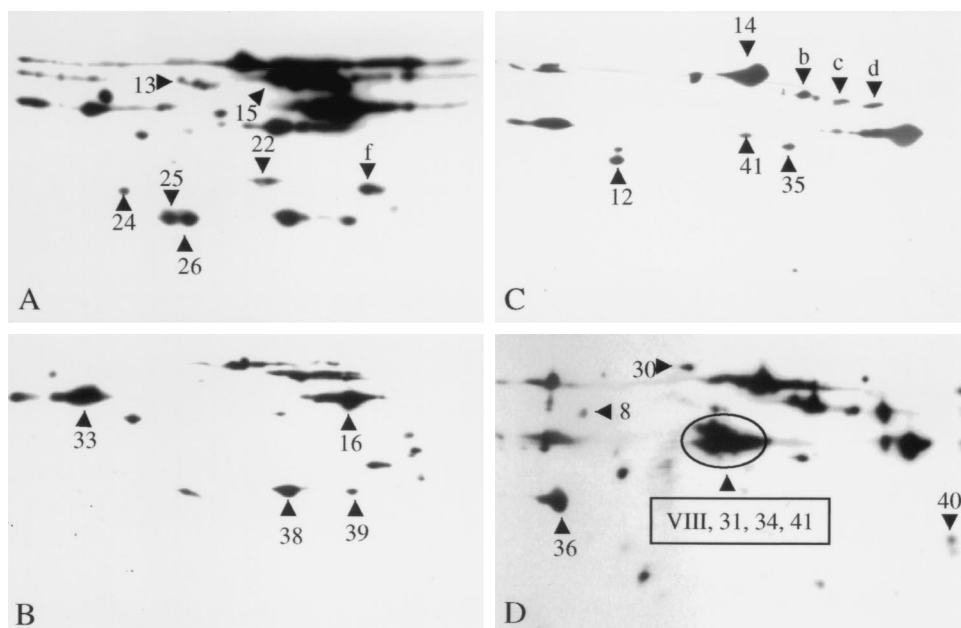


FIG. 3. Two-dimensional Western blot analysis of stationary-phase GAS culture supernatant proteins with convalescent-phase sera obtained from mice with soft-tissue infection. Concentrated culture supernatant proteins (60 μ g) were separated in the first dimension by isoelectric focusing (pH 3 to 10) and in the second dimension by SDS-PAGE with 12% (A and B) or 10% (C and D) gels. Concentrated culture supernatants from *speB* mutants of serotype M3 MGAS 315 (A and B) and serotype M1 MGAS 5005 (C and D) were used in the analysis. The separated proteins were transferred to nitrocellulose membranes, and Western blotting was performed with convalescent-phase sera obtained from mice infected with MGAS 315 (diluted 1:500) (A and C) and MGAS 5005 (diluted 1:1,000) (B and D). The numbered and lettered triangles indicate the proteins reactive in the Western blots that correspond to the protein spots (numbered identically) in Fig. 1.

TABLE 5. Immunogenic proteins identified in the culture supernatants of GAS strains MGAS 315 and MGAS 5005

Spot no.	Identical or homologous protein	Antibody in ^a :				
		Mouse sera		Human sera		
		MGAS 315	MGAS 5005	Pharyngitis	ARF	STSS
8	GroEL	+	+	+	+	+
12	Phosphoglycerate kinase	+	+	-	-	-
13	Immunogenic secreted protein homologue	+	+	+	+	+
14	Streptolysin O	+	+	+	+	+
15	M3 protein	+	-	-	-	-
16, 33	NAD ⁺ -glycohydrolase	+	+	+	+	+
18	Glyceraldehyde-3-phosphate dehydrogenase	-	-	-	+	-
22	Human Mac-1	+	+	+	+	+
24	Serine protease	+	+	+	+	+
26	Streptococcal pyrogenic exotoxin A	+	-	-	-	+
30	Cyclodextrin glucosyltransferase	-	+	-	-	-
31, 34, 40	M1 protein	-	+	-	+	+
35	<i>L. lactis</i> 45-kDa secreted protein	+	+	+	+	+
36	Streptococcal inhibitor of complement	-	+	-	-	-
38, 39	Mitogenic factor (DNase)	+	+	+	+	+
41	Streptokinase	+	+	+	+	+

^a Detection of antigen-antibody reactivity for each protein was based on representative Western blot data, as represented in Fig. 3 to 5. The Western blot results were obtained with sera from mice infected subcutaneously with MGAS 315 or MGAS 5005 and from humans with pharyngitis, acute rheumatic fever (ARF), or streptococcal toxic shock syndrome (STSS). +, antibody reactivity detected; -, no antibody reactivity detected.

ther consideration because it lacks an apparent secretion signal sequence. Eight of these 16 extracellular proteins have been previously identified in GAS and characterized, including NADase, streptolysin O, M1 and M3 proteins, mitogenic factor, streptococcal pyrogenic exotoxin A, streptococcal inhibitor of complement, and streptokinase (8, 9, 23) (Table 3). Eight new proteins were identified, including homologues of class B acid phosphatase (spot IX [not shown in Fig. 1], serine protease (spot 24), and cyclomaltodextrin glucanotransferase (spot 30). Protein spots 13 and 35 are homologues of two proteins with unknown functions, i.e., GAS immunogenic secreted protein (31) and a putative secreted protein made by *Streptococcus mutans* (GenBank accession no. U78607), respectively. Protein spot 22 lacks a known bacterial homologue but has regions of identity with a human protein known as Mac-1 (2, 7, 53). Protein spot 23 is homologous to several peptidoglycan hydrolases (43). Protein spot 27 has no homologue in available protein databases and is of unknown function.

Proteins that lack apparent secretion signal sequences. Twenty-seven of the 43 identified proteins had no apparent secretion signal sequence. Among them, 2 proteins started with the first inferred amino acid residue (methionine), 24 proteins started with the second inferred amino acid residue, and, as noted above, one protein (ribosomal protein L7/L12) started with inferred amino acid 45. Consistent with the lack of a secretion signal sequence, these proteins are presumed to be intracellular proteins, including eight glycolytic pathway metabolic enzymes, three chaperonins, three proteins involved in translation, two sugar transport proteins, three urea cycle proteins, three ribosomal proteins, one DNA binding protein, and five other enzymes (Table 4). The other protein, referred to as antitumor glycoprotein or arginine deiminase (spot 3), may be an extracellular protein that lacks a secretion signal sequence (8, 9, 23).

Two-dimensional Western blotting of stationary-phase culture supernatant proteins with sera from infected mice. Culture supernatant proteins obtained from the *speB* mutants of MGAS 315 and MGAS 5005 grown to stationary phase were analyzed by two-dimensional Western blotting with convalescent-phase sera obtained from mice with experimental soft

tissue infections caused by the wild-type parental organisms of each of these two strains. Stationary-phase cultures were used because all abundant proteins present in exponential-phase culture supernatants were present in stationary-phase preparations, but the converse was not true. The analysis was conducted with both the homologous and heterologous protein-serum pairs, and representative Western blots are shown in Fig. 3. Nineteen protein spots (SP8, SP12, SP13, SP14, SP15, SP16, SP22, SP24, SP26, SP30, SP31, SP33, SP34, SP35, SP36, SP38, SP39, SP40, and SP41) that were reactive in the Western blots matched proteins identified in Fig. 1. These 19 protein spots represent 15 distinct proteins (Table 5). SP8 (GroEL) and SP12 (phosphoglycerate kinase) are generally considered to be cytosolic proteins in bacteria. The other 13 proteins have apparent secretion signal sequences (Fig. 2) and are presumed to be proteins that are actively secreted. The identity of protein spot VIII could not be determined because it lacked homology with known proteins inferred from the available M1 genome sequence. Amino-terminal sequencing of protein spots b, c, d, and f repeatedly failed, and hence these protein spots could not be identified.

Although most of the protein spots identified in the analysis were present in both M1 and M3 strains, several proteins were uniquely found in the culture supernatants of one of the organisms. For example, protein spots 31, 34, 40, and VIII were identified only in the supernatant of the serotype M1 MGAS 5005 strain, whereas protein spots 15 and 25 were restricted to the supernatant of the serotype M3 MGAS 315 strain. Several of these protein spots correspond to M1 or M3 protein.

Two-dimensional Western blot analysis of culture supernatant proteins with human sera. Western blot analysis was also performed for stationary-phase culture supernatant proteins of the *speB* mutant of serotype M1 MGAS 5005 with sera from humans with various types of GAS infection. Protein-antibody reactivity was observed with serum obtained from a pharyngitis patient, serum from a patient with acute rheumatic fever, and convalescent-phase sera from a human with streptococcal toxic shock syndrome caused by an M1 GAS strain. Representative immunologic data are shown in Fig. 4.

In the aggregate, the sera used had reactive antibodies

against the following protein spots: SP8, SP13, SP14, SP16, SP33, SP35, SP38, a, c, d, and e. Protein spot h is absent in Fig. 4C. Protein spots 31, 34, 40, and m were present in Fig. 4B and C but absent in Fig. 4A. The occurrence of reactivity with three protein spots representing M1 protein suggests that the patients with toxic shock syndrome and acute rheumatic fever were infected by serotype M1 GAS strains, whereas the pharyngitis patient was infected with a GAS strain other than serotype M1. Antibody reactivity to SP18 was unique to the serum from the acute rheumatic fever patient (Fig. 4B). Convalescent-phase serum from the STSS patient had reactivity with protein spots i and j (Fig. 4C).

Western blot analysis of recombinant GAS proteins. The antigenicities of SP13 (immunogenic secreted protein homologue), SP22 (Mac-1-like protein), SP24 (serine protease homologue), and SP35 (homologue of a *Lactococcus lactis* secreted protein) were verified by Western blot analysis with recombinant proteins expressed in *E. coli* (Fig. 5A). The Western blotting was conducted with pooled mouse sera obtained from animals infected subcutaneously with MGAS 5005 (serotype M1) and human sera taken from patients with pharyngitis, acute rheumatic fever, or streptococcal toxic shock syndrome. The Western blot analysis confirmed that antibodies against these proteins were present in infected hosts (Fig. 5). For example, the mouse and human sera were reactive with the two fragments of SP24 (Fig. 5). Similarly, the three human sera reacted with recombinant SP13, SP22, and SP35 (Fig. 5). Control mouse and human sera lacked reactivity with the four recombinant proteins.

DISCUSSION

In this study we conducted an initial analysis of proteins found in relative abundance in the culture supernatants of two GAS strains grown in vitro at 37°C. The study focused on *speA*-positive strains of serotypes M1 and M3 which are common causes of human invasive infections globally (33–35). We used isogenic *speB* mutants to assist the identification of proteins expressed in the stationary phase, a time when *SpeB* production is greatly upregulated in many strains, resulting in degradation of many GAS proteins. In the aggregate, amino-terminal amino acid sequence data were generated for 66 protein spots and 43 distinct proteins were identified. Eight of the 16 proteins with apparent signal sequences involved in active secretion have not been previously described for GAS (Table 3). These eight proteins include homologues of peptidoglycan hydrolase (SP23) (18), class B phosphatase (SPIX) (42, 47), serine protease (SP24) (50), cyclodextrin glucosyltransferase (SP30) (4), a putative secreted protein made in abundance by *L. lactis* (SP35) (48), an immunogenic secreted protein made by GAS (SP13) (31), a protein with regions of homology with human Mac-1 (SP22) (2, 7, 53), and a protein with no known homologue (SP27).

Previously uncharacterized and apparently secreted GAS proteins. Five of these eight previously uncharacterized GAS proteins that are apparently secreted have homologues with known functions, whereas three proteins do not. SPIX is 38% identical in amino acid sequence with *Morganella morganii* class B acid phosphatase, and the two proteins have similar molecular masses (46). In addition, the GAS protein contains the sequence F-D-I-D-D-T-L-L-F-T-S-Q, which is very similar to a conserved structural motif, F-D-I-D-D-T-V-L-F-S-S-P, present in class B phosphatases. A second structural motif (Y-G-D-[AS]-D-[DNA]-D-[IV]) that is conserved in class B acid phosphatases also is present in SpIX (Y-G-D-S-D-E-D-I) (42, 47). Class B acid phosphatases are secreted enzymes that hydrolyze

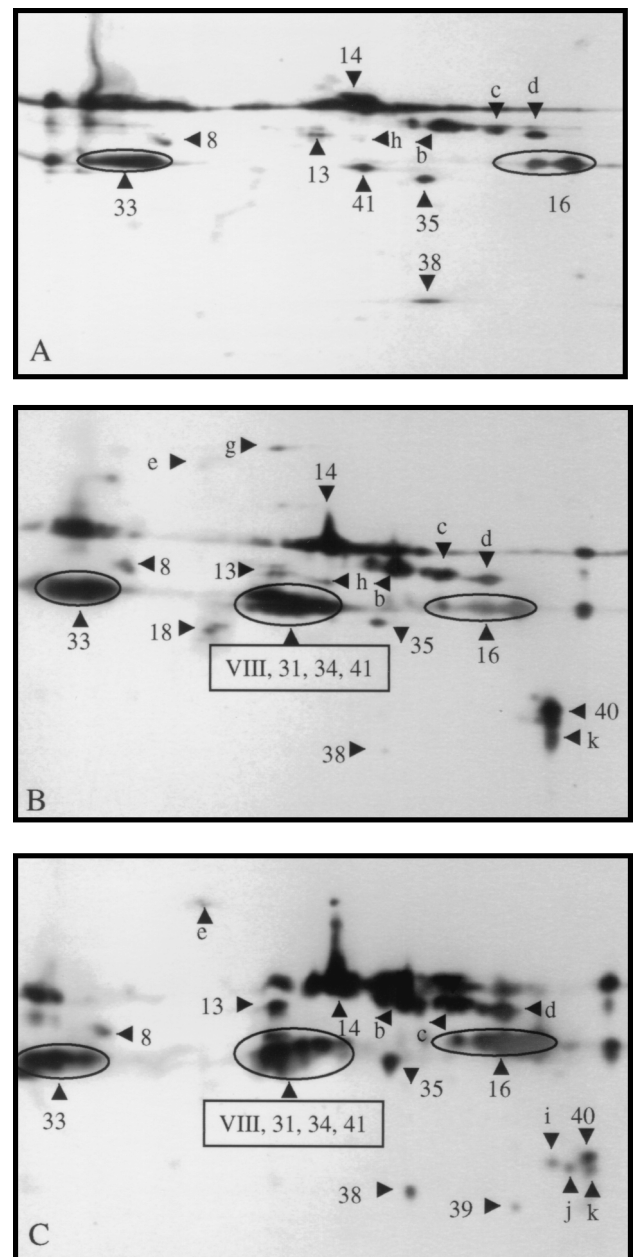


FIG. 4. Two-dimensional Western blot analysis of stationary-phase GAS culture supernatant proteins with human and mouse sera. Concentrated culture supernatant proteins (60 μ g) obtained from the *speB* mutant of MGAS 5005 (serotype M1) were separated in the first dimension by isoelectric focusing (pH 3 to 10) and in the second dimension by SDS-PAGE with 10% gels. (A) Convalescent-phase serum (diluted 1:20,000) obtained from a human with GAS pharyngitis; (B) serum obtained from a patient with acute rheumatic fever; (C) convalescent-phase serum (diluted 1:20,000) obtained from a human with streptococcal toxic shock syndrome caused by an M1 GAS strain. Control sera obtained from these patients were essentially unreactive when used at the same dilutions. The numbered and lettered triangles indicate the proteins reactive in the Western blots that correspond to the protein spots (numbered identically) in Fig. 1.

organic phosphate monoesters and also transfer the orthophosphoryl group from organic phosphoric acid esters to nucleosides and other compounds with free hydroxyl groups. The exact physiologic role of these enzymes has not been clarified, and it is not known if they participate in host-pathogen interactions.

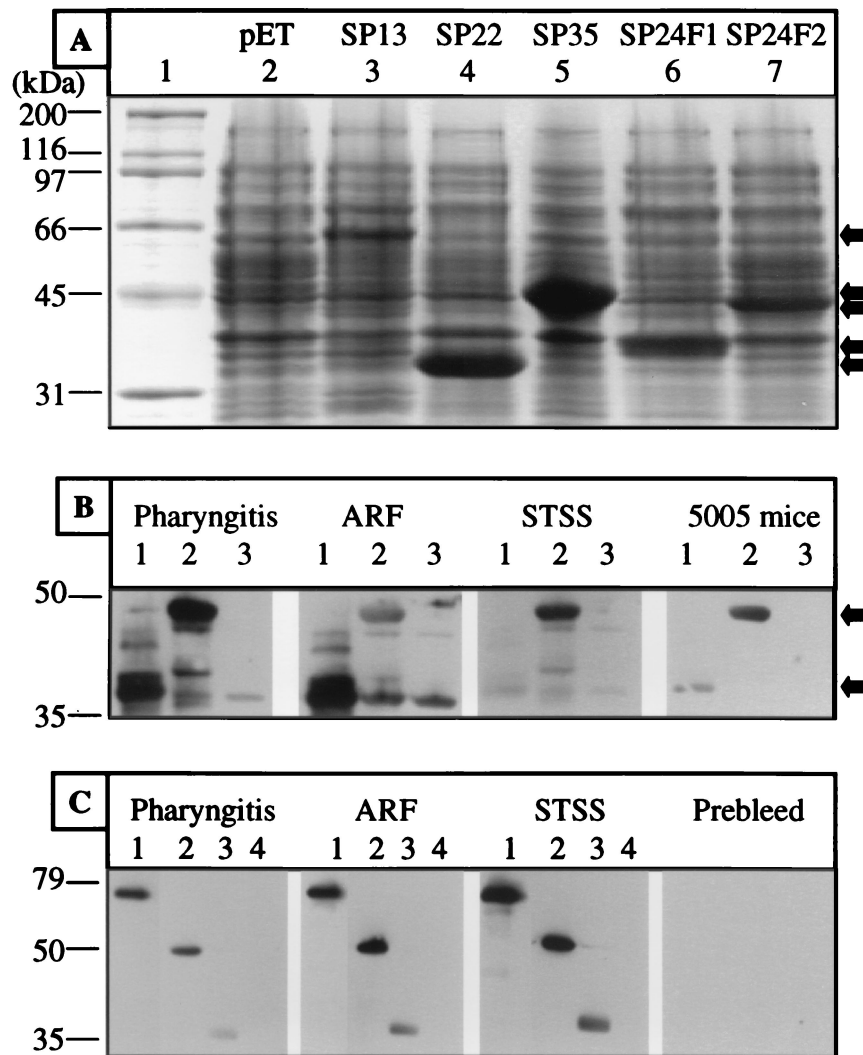


FIG. 5. Western blot analysis of recombinant GAS proteins. SP13, SP22, SP35, and two fragments of SP24 were used in the analysis. (A) SDS-PAGE of lysates of *E. coli* BL21 expressing the recombinant proteins. Lane 1, molecular mass markers; lane 2, empty vector; lanes 3 to 7, *E. coli* lysates containing recombinant proteins as indicated. (B) Western blots showing reactivity of sera with the two recombinant fragments of SP24. Sera (diluted 1:2,000) were obtained from humans with pharyngitis, acute rheumatic fever (ARF), and streptococcal toxic shock syndrome (STSS) and from mice infected subcutaneously with MGAS 5005. Lanes 1, SP24 fragment 1 (amino acid residues 57 to 282); lanes 2, SP24 fragment 2 (amino acid residues 1029 to 1366); lanes 3, vector control. (C) Western blots showing reactivity of sera with recombinant SP22 and SP35. Sera (diluted 1:10,000) were obtained from humans with pharyngitis, acute rheumatic fever, and streptococcal toxic shock syndrome and from mice prior to infection with MGAS 5005. Lanes 1, SP13; lanes 2, SP22; lanes 3, SP35; lanes 4, vector control.

SP23 has 46% amino acid sequence identity with *Streptococcus thermophilus* peptidoglycan hydrolase (18). Both of these streptococcal proteins lack the domain located in the carboxy terminus of the *L. lactis* major autolysin that is involved in cell wall binding.

SP30 has 30 to 50% amino acid identity and a very similar inferred molecular mass to bacterial cyclodextrin glycosyltransferases (4). These enzymes are secreted proteins and catalyze the degradation of starch into cyclodextrins that are further metabolized by the bacterium.

The inferred amino acid sequence of the protein corresponding to SP24 is homologous to those of several serine proteases. The best sequence matches were identified with serine proteases that cleave and inactivate human complement protein C5 and are made by GAS and the related pathogen group B *Streptococcus* (8, 15, 27). The amino acid sequences of SP24 and C5a peptidase inferred from the available M1 genome sequence are 31.4% identical in a region containing 936

amino acid residues. SP24 is also homologous to a *L. lactis* serine protease located on the cell surface. The *L. lactis* SK11 protease (50) and SP24 precursors have 1,962 and 1,647 amino acid residues, respectively, and the mature forms have 1,775 and 1,613 amino acid residues, respectively.

Importantly, SP24 has regions of homology with several serine proteases in areas containing the catalytic triad amino acid residues (aspartic acid, histidine, and serine) (Fig. 6). SP24 also has an inferred LPXTG motif located at the carboxy terminus that is characteristic of gram-positive cell surface-anchored proteins (12). Interestingly, SP24 was detected as a protein spot with an apparent molecular mass of 31 kDa, indicating that the full-length protein undergoes posttranslational cleavage. The SP24 protein fragment begins after the apparent secretion signal peptide and is composed of 280 amino acid residues that contain the aspartic acid and histidine residues of the catalytic triad, but not the serine residue. These results suggest that, unlike the *L. lactis* SK11 protease (50), the

A			↓	
SP	136	WDK G Y K G Q G K V V A V I D T G I D P A H Q S M R I S D V	166	
LP	202	S N Y K Y K G E G T V V S V I D S G I D P T H K D M R L S D D	232	
SC	122	Q A Q G F K G A N V K V A V L D T G I Q A S H P D L N V V G G	152	
PK	129	Y Y D E S A G Q G S C V Y V I D T G I E A S H P E F E G R A Q	159	
B			↓	
SP	264	D I D W T Q T D D D T K Y E S H G M H V T G I V A G N S K E A	294	
LP	266	A D N N D T I T D D K V D E Q H G M H V A G I I G A N G T G D	346	
SC	153	A S F V A G E A Y N T D G N G H G T H V A G T V A A L D N T T	183	
PK	159	Q M V K T Y Y S S R D G N G H G T H C A G T V G S R T Y G V	189	
C			↓	
SP	605	Y N D N H Y G S Q T G T S M A S P Q I A G A S L L V K Q Y L E	635	
LP	608	Q N N G Y T N M S G T S M A S P F I A G S Q A L L K Q A L N	638	
SC	313	Y P T S T Y A T L N G T S M A S P H V A G A A L I L S K H P	343	
PK	317	W I G G S T R S I S G T S M A T P H V A G L A A Y L M T L G K	347	

FIG. 6. Homologous regions of several bacterial serine proteases and an inferred streptococcal serine protease. Shown are regions containing the catalytic triad amino acid residues (arrows) aspartic acid (A), serine (B), and histidine (C). The shaded amino acid residues match the consensus sequence. The numbers at the left and right sides of each sequence indicate amino acid positions. SP, serotype M1 MGAS 5005 serine protease homologue; LP, *L. lactis* SK11 serine protease (GenBank accession number J04962); SC, subtilisin Carlsberg (GenBank accession number X03341); PK, proteinase K (GenBank accession number X14689).

streptococcal serine protease homologue does not have a conventional propeptide region. It is unknown whether the 31-kDa SP24 was generated by nonspecific proteolysis or by a targeted and orchestrated GAS cleavage process that has biological significance.

SP13 is homologous to a GAS immunogenic secreted protein of unknown function termed Isp and described by McIver et al. (31). The amino acid sequences of SP13 and Isp are 49% identical. The SP13 Isp homologue also has 43.7% amino acid sequence identity in a 113-amino-acid region with the TraG putative transfer protein encoded by a gene in staphylococcal conjugative multiresistance plasmid pSK41.

SP35 has 32.8% identity in a region of 340 amino acid residues with Usp 45, a major secreted 45-kDa protein of *L. lactis* with unknown function (48). Schubert et al. (43) recently reported that a homologous extracellular protein made by *L. monocytogenes* has peptidoglycan lytic activity.

SP22 does not have a known bacterial homologue, but it does have two features that might provide clues regarding its interaction with the host. The SP22 protein has 23% amino acid identity in a 203-residue region with the alpha subunit of human leukocyte adhesion receptor Mac-1 (1, 7, 53). Mac-1 is a complement receptor and participates in inflammation. Monoclonal antibodies directed against the Mac-1 alpha subunit inhibit neutrophil aggregation and adhesion to endothelial cells. A second feature that may provide insight regarding function is that SP22 has an arginine-glycine-aspartic acid tripeptide (RGD) which can mediate the interaction of RGD-containing proteins with human integrins (22, 45). On the basis of these structural clues, we speculate that SP22 interferes with the normal host function of Mac-1. Studies to address the role of the SP22 protein in several aspects of host-pathogen interactions are under way.

SP27 has no known prokaryotic or eukaryotic homologue and, hence, is of unknown function.

Proteins generally considered to be located in the cytosol. More than one-half of the culture supernatant proteins that we identified do not have apparent secretion signal sequences and are traditionally considered to be cytosolic proteins involved in intracellular processes. However, several of these proteins have been previously reported to be located extracellularly in streptococci, including histone-like protein (44), glyceraldehyde-3-phosphate dehydrogenase (37, 38, 51, 52), α -enolase

(39), arginine deiminase (antitumor glycoprotein) (9, 23), and phosphocarrier protein (11). Glycolytic enzymes also have been detected in association with the cell surfaces of several other microbial pathogens (13, 32). We note that our analysis identified a total of eight glycolytic enzymes in the culture supernatants of the GAS strains studied. It is not clear if the proteins we identified that are generally considered to be cytosolic proteins are actively secreted by an unknown mechanism or passively released from cells undergoing autolysis. However, several of the proteins were present in abundance in the supernatants of mid-log-phase cultures, suggesting that the extracellular location of these proteins may not be solely due to passive release. Regardless of the mechanism responsible for their presence in the culture supernatant of GAS, their occurrence may influence the course of host-pathogen interactions. In this regard, we note that it has been reported that α -enolase binds to plasminogen in vitro (39), and binding of glyceraldehyde-3-phosphate dehydrogenase to human pharyngeal cells induces phosphorylation of host cell proteins (40). Hence, this enzyme appears to modulate host cell signaling events.

Immunogenicity of the culture supernatant proteins in infected hosts. A general finding was that proteins with apparent secretion sequences were more commonly immunogenic in the few infected hosts studied than proteins lacking these sequences. Unlike the proteins without the apparent secretion signal sequences, 12 of the 16 proteins with secretion signal sequences had host antibody reactivity, and 8 of them were reactive with all five sera tested (Table 5). As expected, serotype-specific M3 protein, M1 protein, and streptococcal inhibitor of complement were not immunoreactive with some of the sera tested.

Although many antigenic proteins were identified by the two-dimensional Western blot analysis, in this study we chose to clone and overexpress only four of them, mainly to confirm their identity and immunogenicity. These four proteins were selected because they had apparent secretion signal sequences and were previously uncharacterized GAS molecules. In addition, inasmuch as two GAS extracellular proteases are known virulence factors (27, 29, 30), it is reasonable to speculate that the serine protease homologue also contributes to host-pathogen interactions. However, this issue clearly requires additional analysis. Two factors that contributed to the selection of SP22 were its regions of identity with the human Mac-1-like

protein and the presence of an RGD motif. RGD motifs present in proteins made by microbial pathogens mediate binding to human integrins, thereby triggering downstream cell-signaling effects. Several microbial integrin binding proteins with the RGD motif are proven virulence factors, including the extracellular cysteine protease (SpeB) made by serotype M1 organisms (22, 45). We will report elsewhere on extensive characterization of the SP22 molecule, including cell biologic and immunologic aspects (B. Lei et al., unpublished data).

ACKNOWLEDGMENTS

We thank I. Abdi for technical assistance; J. C. Smoot and M. S. Chaussee for assistance and helpful discussions; M. Garfield for protein sequencing; H. Hill, L. G. Veasy, and D. Low for providing sera; and N. P. Hoe for critical reading of the manuscript.

REFERENCES

- Akesson, P., A. G. Sjöholm, and L. Björck. 1996. Protein SIC, a novel extracellular protein of *Streptococcus pyogenes* interfering with complement function. *J. Biol. Chem.* **271**:1081–1088.
- Arnaout, M. A. 1990. Structure and function of the leukocyte adhesion molecules CD11/CD18. *Blood* **75**:1037–1050.
- Bernhard, G. C., and G. H. Stollerman. 1959. Serum inhibition of streptococcal diphosphopyridine nucleotidase in uncomplicated streptococcal pharyngitis and in rheumatic fever. *J. Clin. Invest.* **38**:1942–1949.
- Binder, F., O. Huber, and A. Böck. 1986. Cyclodextrin-glycosyltransferase from *Klebsiella pneumoniae* M5a1: cloning, nucleotide sequence and expression. *Gene* **47**:269–277.
- Carroll, J. A., C. F. Garon, and T. G. Schwan. 1999. Effects of environmental pH on membrane proteins in *Borrelia burgdorferi*. *Infect. Immun.* **67**:3181–3187.
- Cash, P. 2000. Proteomics in medical microbiology. *Electrophoresis* **21**:1187–1201.
- Corbi, A. L., T. K. Kishimoto, L. J. Miller, and T. A. Springer. 1988. The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CD11b) a subunit. *J. Biol. Chem.* **263**:12403–12411.
- Cunningham, M. W. 2000. Pathogenesis of group A streptococcal infections. *Clin. Microbiol. Rev.* **13**:470–511.
- Degnan, B. A., M. C. Fontaine, A. H. Doebereiner, J. J. Lee, P. Mastroeni, G. Dougan, J. A. Goodacre, and M. A. Kehoe. 2000. Characterization of an isogenic mutant of *Streptococcus pyogenes* Manfredo lacking the ability to make streptococcal acid glycoprotein. *Infect. Immun.* **68**:2441–2448.
- Diederich, L., A. Roth, and W. Messer. 1994. A versatile plasmid vector system for the regulated expression of genes in *Escherichia coli*. *BioTechniques* **16**:916–923.
- Dubreuil, J. D., M. Jacques, D. Brochu, M. Frenette, and C. Vadeboncoeur. 1996. Surface location of HPr, a phosphocarrier of the phosphoenolpyruvate:sugar phosphotransferase system in *Streptococcus suis*. *Microbiology* **142**:837–843.
- Fischetti, V. A. 2000. Surface proteins on gram-positive bacteria, p. 11–24. *In* V. A. Fischetti, R. P. Novick, J. J. Ferretti, D. A. Portnoy and J. I. Rood (ed.), *Gram-positive pathogens*. ASM Press, Washington, D.C.
- Gozalbo, D., Gil-Navarro, I. Azorin, J. Renau-Piqueras, J. P. Martinez, and M. L. Gil. 1998. The cell wall-associated glyceraldehyde-3-phosphate dehydrogenase of *Candida albicans* is also a fibronectin and laminin binding protein. *Infect. Immun.* **66**:2052–2059.
- Grushoff, P. S., S. Shany, and A. W. Bernheimer. 1975. Purification and properties of streptococcal nicotinamide adenine dinucleotide glycohydrolase. *J. Bacteriol.* **122**:599–605.
- Hill, H. R., J. F. Bohnsack, E. Z. Morris, N. H. Augustine, C. J. Parker, P. P. Cleary, and J. T. Wu. 1988. Group B streptococci inhibit the chemotactic activity of the fifth component of complement. *J. Immunol.* **141**:3551–3556.
- Hoe, N., K. Nakashima, D. Grigsby, X. Pan, S. J. Dou, S. Naidich, M. Garcia, E. Kahn, D. Bergmire-Sweat, and J. M. Musser. 1999. Rapid molecular genetic subtyping of serotype M1 Group A *Streptococcus* strains. *Emerg. Infect. Dis.* **5**:254–263.
- Hoe, N. P., K. Nakashima, S. Lukomski, D. Grigsby, M. Liu, P. Kordari, S. J. Dou, X. Pan, J. Vuopio-Varkila, S. Salmenlinna, A. McGeer, D. E. Low, B. Schwartz, A. Schuchat, S. Naidich, D. De Lorenzo, Y.-X. Fu, and J. M. Musser. 1999. Rapid selection of structural variants of group A *Streptococcus* complement-inhibiting protein in serotype M1 epidemic waves. *Nat. Med.* **5**:924–929.
- Husson-Kao, C., J. Mengaud, L. Benbadis, and M.-P. Chapot-Chartier. 2000. Murl1, a *Streptococcus thermophilus* peptidoglycan hydrolase devoid of a specific cell wall binding domain. *FEMS Microbiol. Lett.* **187**:69–76.
- Jungblut, P. R., D. Bumann, G. Haas, U. Zimny-Arndt, P. Holland, S. Lamer, F. Siejak, A. Aebischer, and T. F. Meyer. 2000. Comparative proteome analysis of *Helicobacter pylori*. *Mol. Microbiol.* **36**:710–725.
- Jungblut, P. R., G. Grabher, and G. Stoffer. 1999. Comprehensive detection of immunorelevant *Borrelia garinii* antigens by two-dimensional electrophoresis. *Electrophoresis* **20**:3611–3622.
- Jungblut, P. R., U. E. Schaible, H. J. Mollenkopf, U. Zimny-Arndt, B. Raupach, J. Mattow, P. Halada, S. Lamer, K. Hagens, and S. H. E. Kaufmann. 1999. Comparative proteome analysis of *Mycobacterium tuberculosis* and *Mycobacterium bovis* BCG strains: towards functional genomics of microbial pathogens. *Mol. Microbiol.* **33**:1103–1117.
- Kagawa, T. F., J. C. Cooney, H. M. Baker, S. McSweeney, M. Liu, S. Gubba, J. M. Musser, and E. N. Baker. 2000. Crystal structure of the zymogen form of the streptococcal virulence factor SpeB: an integrin-binding cysteine protease. *Proc. Natl. Acad. Sci. USA* **97**:2235–2240.
- Kanaoka, M., C. Kawanaka, T. Negoro, Y. Fukita, K. Taya, and H. Agui. 1987. Cloning and expression of the antitumor glycoprotein gene of *Streptococcus pyogenes* Su in *Escherichia coli*. *Agric. Biol. Chem.* **51**:2641–2648.
- Kellner, A., E. B. Freeman, and A. S. Carlson. 1958. Neutralizing antibodies to streptococcal diphosphopyridine nucleotidase in the serum of experimental animals and human beings. *J. Exp. Med.* **108**:299–309.
- Langen, H., B. Takacs, S. Evers, P. Berndt, H.-W. Lahm, B. Wipf, C. Gray, and M. Fountoulakis. 2000. Two-dimensional map of the proteome of *Haemophilus influenzae*. *Electrophoresis* **21**:411–429.
- LeBlanc, D. J., L. N. Lee, and J. M. Inamine. 1991. Cloning and nucleotide base sequence analysis of a spectinomycin adenylyltransferase Aad(9) determinant from *Enterococcus faecalis*. *Antimicrob. Agents Chemother.* **35**:1804–1810.
- Li, Y., L. McLandsborough, A. Kondagunta, and P. P. Cleary. 1996. C5a peptidase alters clearance and trafficking of group A streptococci by infected mice. *Infect. Immun.* **64**:503–510.
- Lukomski, S., N. P. Hoe, I. Abdi, J. Rurangirwa, P. Kordari, M. Liu, D. Shu-Jun, G. G. Adams, and J. M. Musser. 2000. Nonpolar inactivation of the hypervariable streptococcal inhibitor of complement gene (*sic*) in serotype M1 *Streptococcus pyogenes* significantly decreases mouse mucosal colonization. *Infect. Immun.* **68**:535–542.
- Lukomski, S., C. A. Montgomery, J. Rurangirwa, R. S. Geske, J. P. Barrish, G. J. Adams, and J. M. Musser. 1999. Extracellular cysteine protease produced by *Streptococcus pyogenes* participates in the pathogenesis of invasive skin infection and dissemination in mice. *Infect. Immun.* **67**:1779–1788.
- Lukomski, S., S. Sreevatsan, C. Amberg, W. Reichardt, M. Woischnik, A. Podbielski, and J. M. Musser. 1997. Inactivation of *Streptococcus pyogenes* extracellular cysteine protease significantly decreases mouse lethality of serotype M3 and M49 strains. *J. Clin. Invest.* **99**:2574–2580.
- McIver, K. S., S. Subbarao, E. M. Kellner, A. S. Heath, and J. R. Scott. 1996. Identification of *isp*, a locus encoding an immunogenic secreted protein conserved among group A streptococci. *Infect. Immun.* **64**:2548–2555.
- Modun, B., and P. Williams. 1999. The staphylococcal transferrin-binding protein is a cell wall glyceraldehyde-3-phosphate dehydrogenase. *Infect. Immun.* **67**:1086–1092.
- Musser, J. M., A. R. Hauser, M. H. Kim, P. M. Schlievert, K. Nelson, and R. K. Selander. 1991. *Streptococcus pyogenes* causing toxic shock-like syndrome and other invasive diseases: clonal diversity and pyrogenic exotoxin expression. *Proc. Natl. Acad. Sci. USA* **88**:2668–2672.
- Musser, J. M., V. Kapur, J. Szeto, X. Pan, D. Swanson, and D. Martin. 1995. Genetic diversity and relationships among *Streptococcus pyogenes* strains expressing serotype M1 protein: recent intercontinental spread of a subclone causing episodes of invasive disease. *Infect. Immun.* **63**:994–1003.
- Musser, J. M., and R. M. Krause. 1998. The revival of group A streptococcal diseases with a commentary on staphylococcal toxic shock syndrome, p. 185–218. *In* R. M. Krause and A. Fauci (ed.), *Emerging infections*. Academic Press, San Diego, Calif.
- Musser, J. M., K. Stockbauer, V. Kapur, and G. W. Rudgers. 1996. Substitution of cysteine 192 in a highly conserved *Streptococcus pyogenes* extracellular cysteine protease (interleukin β convertase) alters proteolytic activity and ablates zymogen processing. *Infect. Immun.* **64**:1913–1917.
- Pancholi, V., and V. A. Fischetti. 1992. A major surface protein on group A streptococci is a glyceraldehyde-3-phosphate dehydrogenase with multiple binding activity. *J. Exp. Med.* **176**:415–426.
- Pancholi, V., and V. A. Fischetti. 1993. Glyceraldehyde-3-phosphate dehydrogenase on the surface of group A streptococci is also an ADP-ribosylating enzyme. *Proc. Natl. Acad. Sci. USA* **90**:8154–8158.
- Pancholi, V., and V. A. Fischetti. 1998. α -Enolase, a novel strong plasmin(ogen) binding protein on the surface of pathogenic streptococci. *J. Biol. Chem.* **273**:14503–14515.
- Pancholi, V., and V. A. Fischetti. 1997. Regulation of the phosphorylation of human pharyngeal cell proteins by group A streptococcal surface dehydrogenase: signal transduction between streptococci and pharyngeal cells. *J. Exp. Med.* **186**:1633–1643.
- Rosenkrands, I., K. Weldingh, S. Jacobsen, C. Veggerby Hansen, W. Florio, I. Gianetri, and P. Andersen. 2000. Mapping and identification of *Mycobacterium tuberculosis* proteins by two-dimensional gel electrophoresis, microsequencing and immunodetection. *Electrophoresis* **21**:935–948.
- Rossolini, G. K., S. Schippa, M. L. Riccio, F. Berlutti, L. E. Macaskie, and M. C. Thaller. 1998. Bacterial nonspecific acid phosphohydrolases: physiol-

- ogy, evolution and use as tools in microbial biotechnology. *Cell. Mol. Life Sci.* **54**:833–850.
43. Schubert, K., A. M. Bichlmaier, E. Mager, K. Wolff, G. Ruhland, and F. Fiedler. 2000. P45, an extracellular 45 kDa protein of *Listeria monocytogenes* with similarity to protein p60 and exhibiting peptidoglycan lytic activity. *Arch. Microbiol.* **173**:21–28.
 44. Stinson, M. W., R. McLaughlin, S. H. Choi, Z. E. Juarez, and J. Barnard. 1998. Streptococcal histone-like protein: primary structure of *hlpA* and protein binding to lipoteichoic acid and epithelial cells. *Infect. Immun.* **66**:259–265.
 45. Stockbauer, K. E., L. Magoun, M. Liu, E. H. Burns, Jr., S. Gubba, S. Renish, X. Pan, S. C. Bodary, E. Baker, J. Coburn, J. M. Leong, and J. M. Musser. 1999. A natural variant of the cysteine protease virulence factor of group A *Streptococcus* with an arginine-glycine-aspartic acid (RGD) motif preferentially binds human integrins $\alpha_5\beta_3$ and $\alpha_{IIb}\beta_3$. *Proc. Natl. Acad. Sci. USA* **96**:242–247.
 46. Thaller, M. C., G. Lombardi, F. Berlutti, S. Schippa, and G. M. Rossolini. 1995. Cloning and characterization of the NapA acid phosphatase/phosphotransferase of *Morganella morganii*: identification of a new family of bacterial acid-phosphatase-encoding genes. *Microbiology* **141**:147–154.
 47. Thaller, M. C., S. Schippa, and G. M. Rossolini. 1998. Conserved sequence motifs among bacterial, eukaryotic, and archaeal phosphatases that define a new phosphohydrolase superfamily. *Prot. Sci.* **7**:1647–1652.
 48. van Asseldonk, M., G. Rutten, M. Oteman, R. J. Siezen, W. M. de Vos, and G. Simons. 1990. Cloning of *usp45*, a gene encoding a secreted protein from *Lactococcus lactis* subsp. *lactis* MG1363. *Gene* **95**:155–160.
 49. Veasy, L. G., L. Y. Tani, and H. R. Hill. 1994. Persistence of acute rheumatic fever in the intermountain area of the United States. *J. Pediatr.* **124**:9–16.
 50. Vos, P., G. Simons, R. J. Siezen, and W. M. de Vos. 1989. Primary structure and organization of the gene for a procaryotic, cell envelope-located serine proteinase. *J. Biol. Chem.* **264**:13579–13585.
 51. Winram, S. B., and R. Lottenberg. 1996. The plasmin-binding protein Plr of group A streptococci is identified as glyceraldehyde-3-phosphate dehydrogenase. *Microbiology* **142**:2311–2320.
 52. Winram, S. B., and R. Lottenberg. 1998. Site-directed mutagenesis of streptococcal plasmin receptor (Plr) identifies the C-terminal Lys334 as essential for plasmin binding, but mutation of the *plr* gene does not reduce plasmin binding to group A streptococci. *Microbiology* **144**:2025–2035.
 53. Wright, S. D., J. S. Weitz, A. J. Huang, S. M. Levin, S. C. Silverstein, and J. D. Leike. 1988. Complement receptor type three (CD11b/CD18) of human polymorphonuclear leukocytes recognizes fibrinogen. *Proc. Natl. Acad. Sci. USA* **85**:7734–7738.
 54. Zevering, Y., L. Jacob, and T. F. Meyer. 1999. Naturally acquired human immune responses against *Helicobacter pylori* and implications for vaccine development. *Gut* **45**:465–474.

Editor: E. I. Tuomanen