RESEARCH ARTICLE

Group A streptococcus cell-associated pathogenic proteins as revealed by growth in hyaluronic acid-enriched media

Meng Zhang[1], Fiona M. McDonald[2], Shane S. Sturrock[2], Simon J. Charnock[1], Ian Humphery-Smith[2] and Gary W. Black[1]***

[1] Biomolecular and Biomedical Research Centre, School of Applied Sciences, Northumbria University, Newcastle upon Tyne, UK
[2] Biosystems Informatics Institute, Bioscience Centre, International Centre for Life, Newcastle upon Tyne, UK

Received: July 31, 2006
Revised: January 16, 2007
Accepted: January 24, 2007

Group A streptococcus (GAS), also know as *Streptococcus pyogenes*, is a human pathogen and can cause several fatal invasive diseases such as necrotising fasciitis, the so-called flesh-eating disease, and toxic shock syndrome. The destruction of connective tissue and the hyaluronic acid (HA) therein, is a key element of GAS pathogenesis. We therefore propagated GAS in HA-enriched growth media in an attempt to create a simple biological system that could reflect some elements of GAS pathogenesis. Our results show that several recognised virulence factors were up-regulated in HA-enriched media, including the M1 protein, a collagen-like surface protein and the glycolytic enzyme glyceraldehyde-3-phosphate dehydrogenase, which has been shown to play important roles in streptococcal pathogenesis. Interestingly, two hypothetical proteins of unknown function were also up-regulated and detailed bioinformatics analysis showed that at least one of these hypothetical proteins is likely to be involved in pathogenesis. It was therefore concluded that this simple biological system provided a valuable tool for the identification of potential GAS virulence factors.

Keywords:
Group A streptococcus / Hyaluronic acid-enriched growth media / Pathogenesis / Proteomic analysis / Signature peptide search

1 Introduction

Group A streptococcus (GAS) is a Gram-positive human pathogen. This bacterium is specific to humans and human disease. It is estimated that between 5–15% of normal individuals harbour the bacteria, usually in the respiratory tract, without signs of disease [1]. However, when the bacteria are able to penetrate vulnerable tissues and when the host defences are compromised, a variety of types of suppurative infection can occur. Commonly, GAS is responsible for pharyngitis, scarlet fever, impetigo or cellulitis. Myositis, necrotizing fasciitis and streptococcal toxic shock syndrome are caused by invasive infections. It has been recently estimated that severe GAS infections affect more than 18 million people *per* year with more than 0.5 million deaths [2].

Correspondence: Dr. Meng Zhang, School of Applied Sciences, Northumbria University, Ellison Building, Newcastle upon Tyne NE1 8ST, UK
E-mail: meng.zhang@northumbria.ac.uk
Fax: +44-19-1227-3519

Abbreviations: ΔC_n, difference in normalised cross correlation score; **GAPDH**, glyceraldehyde-3-phosphate dehydrogenase; **GAS**, group A streptococcus; **HA**, hyaluronic acid; **O/E**, observed/expected; **SpeB**, streptococcal pyrogenic exotoxin B; X_{corr}, cross correlation score

* Present address: Megazyme International Ireland Limited, Bray, County Wicklow, Republic of Ireland.
** Additional corresponding author: Professor Gary W. Black
E-mail: gary.black@northumbria.ac.uk

The degradation of connective tissue by GAS is a key element of invasive infection [3]. Hyaluronic acid (HA) is a major component of the extracellular matrix of connective tissue and is a negatively charged high molecular mass polysaccharide (~M_r of 1 000 000) consisting of alternating residues of *N*-acetylglucosamine and glucuronic acid. It is produced by all mammals and is distributed in various tissues, with the highest concentrations found in soft connective tissues [4]. To invade the connective tissue of the human host, GAS produces hyaluronate lyases that depolymerise HA [5]. The result of the enzymatic depolymerisation of HA is a decrease in the viscosity of the extracellular matrix, and therefore increased permeability of the connective tissues, and potentially an increase in pathogen aggression, and diffusion of its toxins, through the connective tissue [6]. In addition to hyaluronate lyases, many other factors are known to contribute to streptococcal virulence, such as M proteins, lipoteichoic acid, the HA capsule that surrounds the pathogen, pyrogenic exotoxins, streptolysins and C5a peptidase [7]. All these virulence factors play a role in streptococcal infection and disease. M proteins, lipoteichoic acid and C5a peptidase are required for successful adherence and protection of GAS from constitutive host defences. Streptolysins and pyrogenic exotoxins function to destroy tissues as toxins and proteases.

Despite the large number of studies that have analysed the expression of the GAS virulence factors, the molecular basis of many GAS infections remains largely unknown. Classical approaches, such as the generation and characterization of isogenic mutants, are difficult to apply to an in-depth analysis of the vast number of bacterial gene products potentially involved in virulence [8]. Therefore, to examine simultaneous changes in multiple virulence factors, proteomic analysis of GAS is more appropriate. For instance, Lei *et al.* [9] compared the culture supernatants of wild-type GAS strains and isogenic mutant strains and identified previously undescribed extracellular proteins, including several that are immunogenic in the course of host–pathogen interactions. Thongboonkerd *et al.* [10] showed that fluoride caused the decreased expression of proteins used to respond to stress, virulence factors, and proteins implicated in several GAS non-suppurative diseases. Chaussee *et al.* [11] compared proteome maps of a GAS wild-type strain and a *rgg* mutant strain and concluded that Rgg is a global regulatory factor that contributes to growth phase-dependent synthesis of proteins associated with secondary metabolism and oxidative and thermal stress responses. In addition, proteome analysis has been used to confirm findings from transcriptomics experiments and to distinguish between transcriptional and post-transcriptional regulatory events [8].

Our approach was to grow GAS in growth media enriched in HA, in an attempt to create a simple biological system that could reflect some elements of GAS pathogenesis. High-resolution separation of the resulting proteins by 2-DE, in-gel trypsin digestion and static nanospray IT MS were used to characterise proteins expressed by the pathogen. Further functional analysis of these proteins was performed using the signature peptide search tool, TSSview™. We show that several proteins involved in pathogenesis were differentially expressed and two hypothetical proteins of unknown function were up-regulated. Interestingly, detailed bioinformatics analysis showed that at least one hypothetical protein is likely to be involved in GAS pathogenesis.

2 Materials and methods

2.1 Bacteria and growth conditions

The M1 serotype of *Streptococcus pyogenes*, strain SF370 (ATCC 700294 [12]), was employed in this study. The bacteria were kept on Columbia Agar supplemented with 5% v/v horse blood, and were grown either in Brain Heart Infusion broth (Oxoid, UK) supplemented with 0.5% w/v sodium hyaluronate (Fisher Scientific, UK), referred to hereafter as HA-enriched media, or in unsupplemented Brain Heart Infusion broth, referred to hereafter as control media. Incubations were performed at 37°C overnight until the culture reached stationary phase, *i.e.* an optical density at 660 nm between 1.0 and 1.2.

2.2 Extraction of proteins

To isolate protein extracts from strain SF370, 80 mL of culture was harvest by centrifugation at $4000 \times g$ at 4°C for 30 min. The culture supernatant was then removed, and the cell pellet resuspended and washed in 5 mL of ice-cold PBS (140 mM NaCl, 10 mM Na_3PO_4, 1.8 mM K_3PO_4 and 2.7 mM KCl, pH 7.4) three times. Then the pellet was air-dried and resuspended in a sample buffer containing 8 M urea, 4% w/v CHAPS and 2% v/v IPG buffer (Amersham Bioscience, UK). The sample was then sonicated, using a Soniprep 150 (Sanyo, UK) fitted with an exponential probe, on ice for 3×10 s. The protein mixture was centrifuged at $14\,000 \times g$ at 4°C for 30 min and the supernatant transferred to fresh microcentrifuge tubes, where the protein extracts were purified using the PlusOne 2-D Clean-Up kit (Amersham Bioscience) according to the manufacturer's directions. Purified protein extract was then quantitated using the PlusOne 2-D Quant kit (Amersham Bioscience) according to the manufacturer's specifications.

2.3 2-DE

Proteins were separated by electrophoresis in the first-dimension with a Multiphor™ II IEF system (Amersham Bioscience) using 18-cm Immobiline dry strips (Amersham Bioscience) with a linear pH range, as described by the manufacturer. Various pH ranges of IPG strips were used and rehydrated with 100 to 120 μg of protein in 350 μL of rehydration solution (8 M urea, 4% w/v CHAPS, 2% v/v IPG buffer, 2.8 mg/mL DTT and 0.002% w/v bromophenol blue)

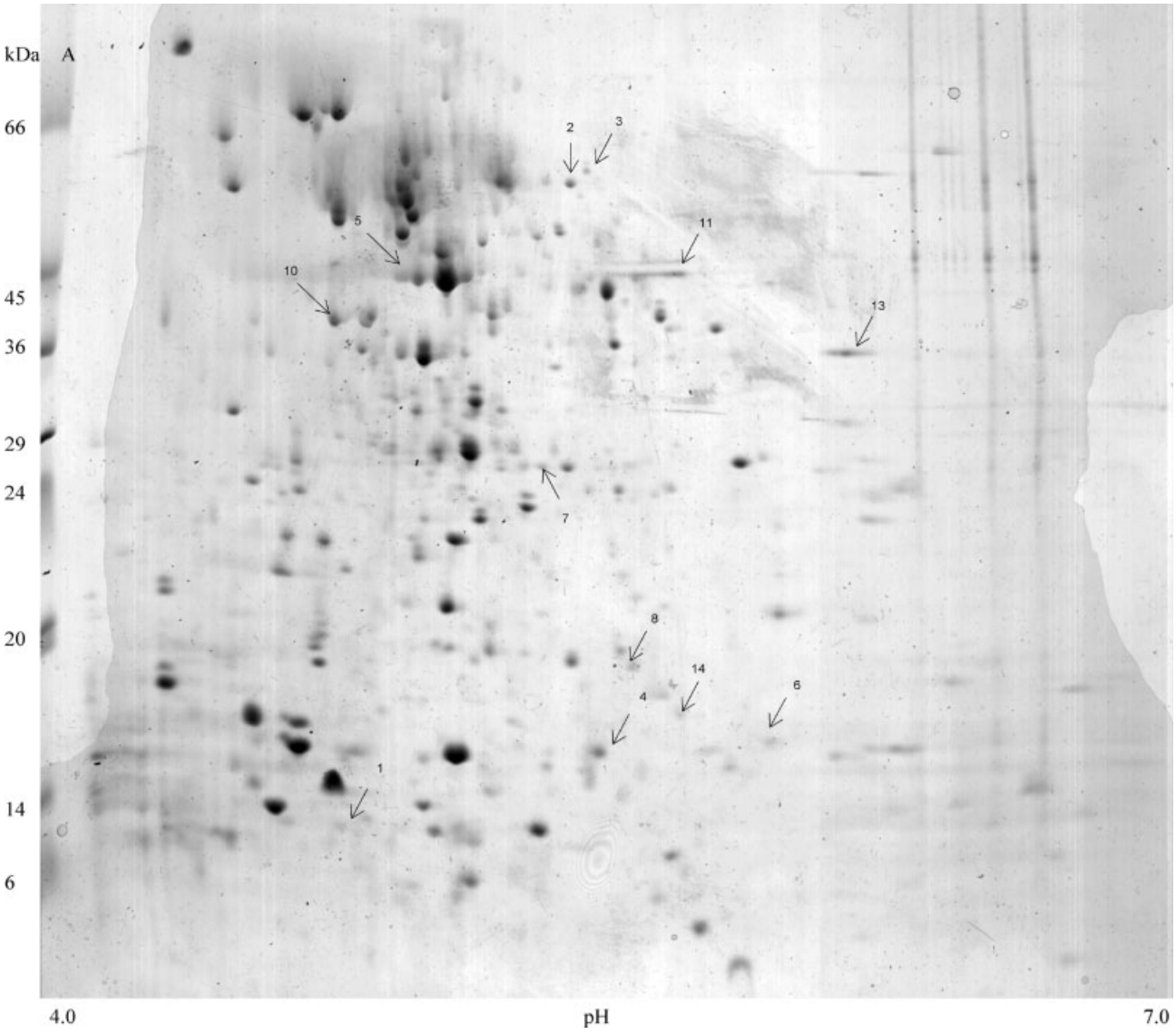
for at least 16 h at room temperature. IEF was done in three steps consisting of 500 V for 1 min, a linear gradient of 500–3500 V for 90 min, and 3500 V for 6.5 h at 20°C. After focusing was completed, the strips were equilibrated with buffer containing 6 M urea, 130 mM DTT, 30% v/v glycerol, 50 mM Tris-Cl pH 8.8, 2% w/v SDS, and 0.002% w/v bromophenol blue for 15 min, and then with buffer containing 6 M urea, 135 mM iodoacetamide, 30% v/v glycerol, 50 mM Tris-Cl pH 8.8, 4% w/v SDS, and 0.002% w/v bromophenol blue for 15 min. The strips were then placed on wet filter paper for 10 min and applied to SDS 14% w/v self-cast polyacrylamide gels ($200 \times 160 \times 1.5$ mm), and electrophoresis was carried out in a Protean II xi cell (BioRad, UK) at 40 mA *per* gel for 4 h at 10°C with 250 mM glycine, 25 mM Tris-Cl and 0.1% w/v SDS as the running buffer. Low-molecular-weight markers (Sigma, UK) were applied next to the acidic end of the IPG strips.

2.4 Visualization and analysis of the protein spots

Gels were fixed in 50% v/v methanol, 12% v/v acetic acid for 1 h followed by staining with colloidal CBB stain (20% v/v methanol, 10% w/v $(NH_4)_2SO_4$, 2% v/v phosphoric acid, and 0.1% w/v CBB G-250). After 24 h, the gels were washed several times with water and scanned using a GS-710 calibrated imaging densitometer (BioRad). Analysis of the gels, including protein spot detection and quantitation, was done with PDQuest™ V6.2 software (BioRad). The background was subtracted and the intensities of the image pixels inside the boundary of each protein spot were quantified. Gels were normalized based on the total density in each gel of the matchset. A reference pattern was constructed and each gel in the matchset was matched to the reference gel. Numerous proteins that were uniformly expressed in all patterns were used as landmarks to facilitate rapid gel matching. Differentially expressed proteins were selected according to the Wilcoxon two-sample test ($p<0.05$), *i.e.* the protein spot must be more intense (or less intense) on at least six gels in each matchset [13].

2.5 In-gel tryptic digestion and protein identification

The protein spots of interest were excised with a clean scalpel into ~1-mm cubes and each transferred to a 1.2-mL siliconized microcentrifuge tube, washed three times for 30 min in 100 μL of 100 mM NH_4HCO_3 and 60 μL of ACN at room temperature and then incubated in 100 μL ACN for at least 5 min at room temperature. The ACN was removed *via* aspiration and gels were dried in a centrifugal evaporator for 30 min. The dried gel pieces were hydrated with 25 μL of 50 mM NH_4HCO_3 containing 20 μg/mL of sequencing grade trypsin (Promega, UK) on ice for 30 min and then 30 μL of 50 mM NH_4HCO_3 added (enough to cover the gel pieces) and incubated at 37°C overnight. The peptides extracted after overnight incubation were recovered *via* aspiration, and additional peptides were recovered from the gel by extracting with 50% v/v ACN with 5% v/v TFA for 30 min and 83% v/v ACN with 0.1% v/v TFA for a further 30 min. The extracted peptides were recovered *via* aspiration, dried in a centrifugal evaporator, and suspended in 10 μL of 0.1% v/v TFA. Samples were desalted using C18 ZipTips (Millipore, UK). First, the Zip Tips were rehydrated using 10 μL of 50% v/v ACN, and equilibrated with 10 μL of 0.1% v/v TFA. The samples were each loaded onto a ZipTip by pipetting the peptides extract up and down ten times, and then washed with 10 μL of 0.1% v/v TFA three times. The tryptic fragments were then eluted using 50% v/v methanol, 0.1% v/v acetic acid. Mass analysis was performed on an LCQ Advantage IT instrument equipped with a static nanospray source (Thermo Electron, UK). Peptide samples were each loaded into an EconoTip emitter (Presearch, UK). Static nanospray was performed at the following conditions: positive ionisation mode; spray voltage, 1.8 kV; capillary voltage, 28 V; capillary temperature, 180°C. Data was collected in the full scan and data-dependent MS/MS modes; three microscans were performed, with the maximum ion injection time of 200 ms. In the full-scan mode, ions were collected in the m/z range 300–2000 and the MS/MS collision energy was set to 35%. Data analysis obtained from the MS was performed using the TurboSEQUEST™ program in the BioWorks 3.2 software suite (Thermo Electron). The following parameters were used when creating Dta files from raw files: precursor mass tolerance 1.4 m/z; minimum number of different ions 15; maximum number of intermediate scans 25; minimum number of grouped scans 1; minimum total ion count 5×10^5, parent ion mass range 300–2000 Da. A database containing the complete genome sequences of all sequenced streptococcal species (*S. agalactiae* NEM316, *S. pyogenes* MGAS8232, *S. agalactiae* A909, *S. pyogenes* MGAS5005, *S. pyogenes* MGAS6180, *S. thermophilus* CNRZ1066, *S. thermophilus* LMG 18311, *S. pyogenes* MGAS10394, *S. mutans* UA159, *S. agalactiae* 2603V/R, *S. pyogenes* SSI-1, *S. pyogenes* MGAS315, *S. pneumoniae* R6, *S. pneumoniae* TIGR4, *S. pyogenes* M1 GAS) was created by downloading, on 7th April 2006, the appropriate sequences from the National Center for Biotechnology Information (NCBI) (http://www.ncbi.nih.gov/). Proteins were then identified by using the Dta files to search the database according to the following parameters: fragment ion mass tolerance of 0 Da; peptide mass tolerance of 1 Da; average mass; maximum number of internal cleavage sites of 2. We evaluated the results from TurboSEQUEST™ using the following filter parameters: proteins must be from the SF370 strain of the M1 serotype of GAS; and peptides must have a cross correlation score (X_{corr}) >1.5 for singly charged peptides, >2.0 for double-charged peptides, and >2.5 for triple-charged peptides, with at least two peptides successfully matched; and/or must have a difference in normalised cross correlation score (ΔC_n) >0.2 with at least two peptides successfully matched [14]. The best hit (according to the peptide probability value) as provided by TurboSEQUEST™ and that which fulfilled the filter parameters was then used for protein characterisation.



2.6 Further functional analysis of peptides and proteins

TSSview™ v1.0 (Turbinia, UK) was used to identify regions of evolutionary conservation in proteins as determined by high-sensitivity data-mining conducted on 9th May 2006 on ~three million UniProt entries and their corresponding coding and non-coding regions in each of the six reading frames. Peptides that fulfil the TurboSEQUEST™ X_{corr} and ΔC_n values stated in Section 2.5 were analysed by TSSview™ (including those peptides not from the SF370 strain of the M1 serotype of GAS). Observed/expected (O/E) values $\leq 1.0 \times 10^{-4}$ were considered significant with the level of probability varying according to signature peptide length. The full-length sequences of all identified proteins were also analysed by TSSview™. O/E values $\leq 1.0 \times 10^{-4}$ were considered as highly significant.

3 Results and discussion

3.1 Identification of differentially expressed proteins using 2-DE

Six pairs of 2-DE gels were analysed. The proteins present on each of these gels were isolated from independently grown cultures. An average of 210 ± 20 and 179 ± 20 spots of cell-associated proteins from bacteria cultured in HA enriched media and control media, respectively, were detected by PDQuest™ software on each 2-D gel (Fig. 1). About 80% of the protein spots were matched between the two sets of gels and the correlation coefficient between the two patterns was 0.815. The 18 protein spots highlighted in Fig. 1 represent those that were up- or down-regulated according to the Wilcoxon two-sample test [13]. Among the 18 proteins, 12 proteins were up-regulated or were present only on the 2-D gels

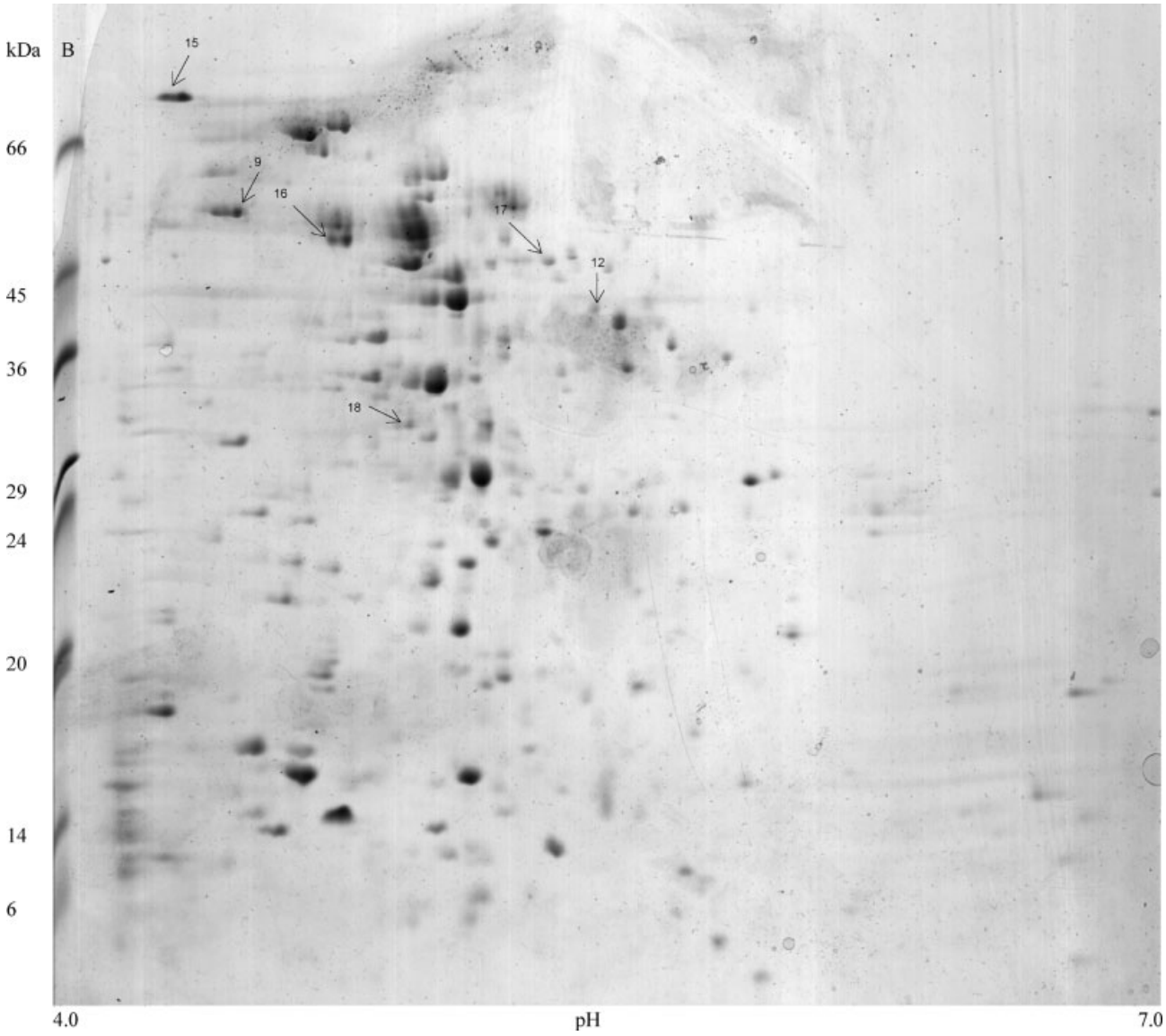


Figure 1. 2-D gel analysis of cell-associated proteins from GAS. (A) Cells cultured in HA-enriched media. The numbers indicate the proteins that were up-regulated in HA-enriched media. (B) Cells cultured in control medium. The numbers indicate the proteins that were down-regulated in HA-enriched media.

prepared from cells grown in HA-enriched media and 6 proteins were down-regulated or were absent only on the 2-D gels prepared from cells grown in HA-enriched media. The spots corresponding to those 18 proteins were excised from gels and subjected to in-gel trypsin digestion. The resulting peptides were analysed by MS.

3.2 Identity of the differentially expressed proteins using MS

An example of a typical full scan of a protein mass spectrum, the 89.46-kDa putative phosphotransferase system phosphohistidine-containing protein, and the MS/MS spectrum of a precursor ion (m/z ratio of 564.98) of this protein can be seen in Fig. 2. Of the 18 proteins analysed, spots 1–11 were identified upon searching the streptococcal genome sequence database with the experimental MS/MS spectra using TurboSEQUEST™ (Table 1). Sequence, charge state and filter parameter values for individual peptides from the identified proteins are shown in Table 2. The remaining 7 proteins could not be identified by TurboSEQUEST™ as they did not satisfy the filter parameters defined in the Section 2.5. Three of these protein spots (spots 14, 16 and 17) gave poor quality MS data and therefore could not be analysed further. However, the remaining four proteins spots (spots 12, 13, 15 and 18) were analysed using detailed bioinformatic screening *via*

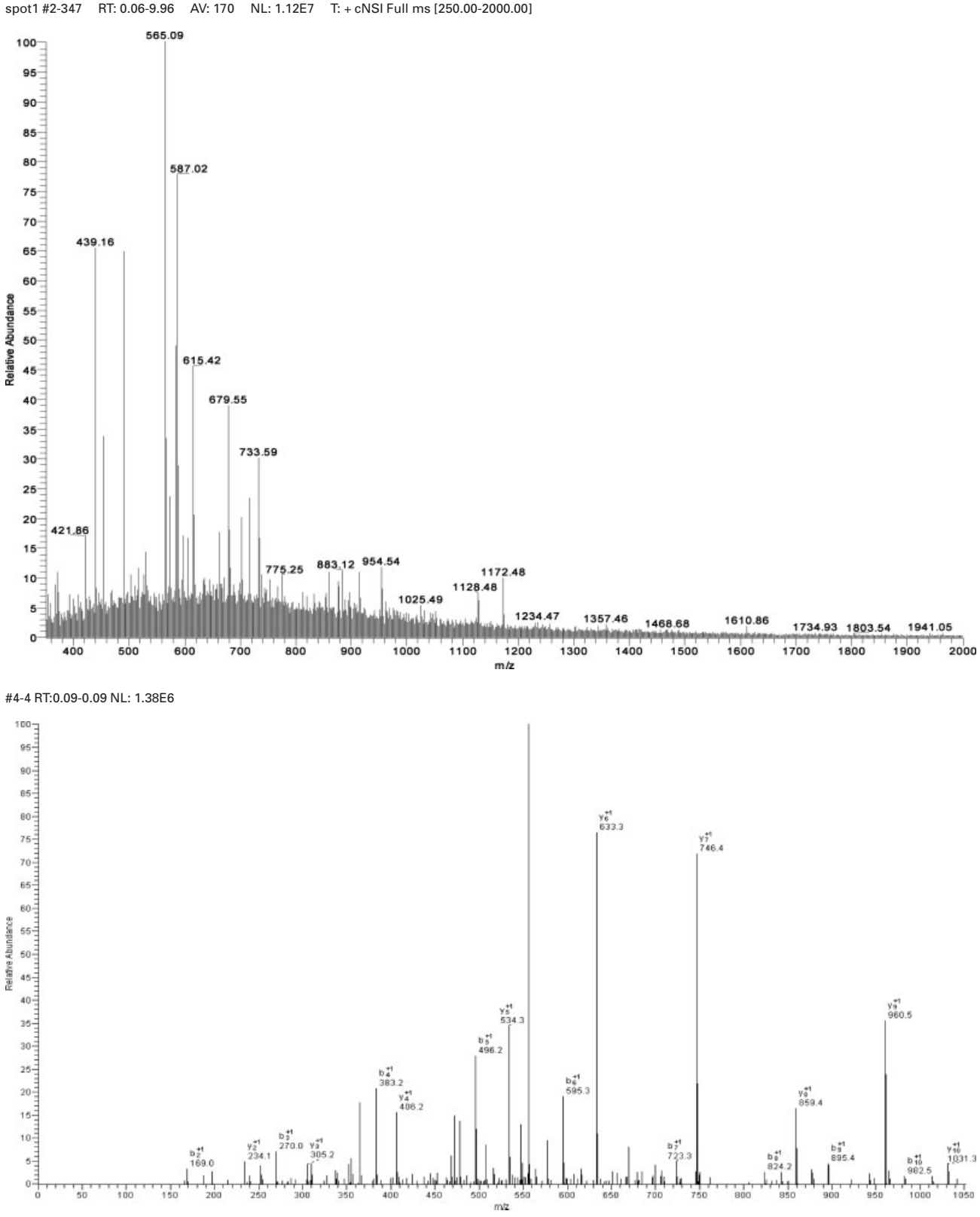


Figure 2. Static nanospray MS. (A) Mass spectrum of a tryptic digestion of protein spot No.1, which was identified as a putative phosphotransferase system phosphohistidine-containing protein using an LCQ Advantage IT MS. (B) MS/MS spectrum of the precursor ion of the protein with an *m/z* ratio of 564.98.

Table 1. Identity of differentially expressed proteins from the M1 serotype of GAS, strain SF370

Spot no.[a]	Protein ID	Accession ID[b]	M_r/p*I*[c]	Matched peptides[d]	Protein coverage (%)[e]	Alteration in HA-enriched media	Level of expression (OD area x 10^6)[f]	
							in HA-enriched media[g]	in control media[g]
1	Putative phosphotransferase system phosphohistidine-containing protein[h]	gi|15675305	8946.6/4.74	4	41.38	Up-regulated	124.0 (215.0)	5.0 (15.0)
2	Putative NADP-dependent glyceraldehyde-3-phosphate dehydrogenase[h]	gi|15675303	50338.4/5.06	2	8.84	Up-regulated	392.1 (347.2)	6.0 (14.9)
3	M1 protein precursor[i]	gi|311758	54203.6/6.53	2	13.22	Up-regulated	127.1 (75.2)	0.0 (17.0)
4	Hypothetical protein SPy1143[i, j]	gi|15675120	21620.7/5.71	2	14.80	Up-regulated	207.8 (226.4)	23.1 (52.5)
5	Histidine protein kinase[i, j]	gi|15674919	30833.2/5.86	3	15.27	Up-regulated	220.5 (261.5)	16.8 (46.9)
6	tRNA modification GTPase[i, j]	gi|15675063	50509.4/4.67	2	6.77	Up-regulated	91.5 (47.7)	0.0 (22.7)
7	Hypothetical protein SPy1262[h]	gi|15675451	19944.4/4.93	7	45.25	Up-regulated	130.0 (42.5)	14.5 (21.0)
8	50 S ribosomal protein L10[h]	gi|15675064	17555.4/5.19	5	30.72	Up-regulated	111.5 (86.3)	35.6 (13.6)
9	Rop A[h]	gi|13622926	47096.3/4.41	5	14.52	Down-regulated	0.0 (47.1)	136.0 (121.2)
10	Collagen-like surface protein[i]	gi|15675773	36260.7/6.91	2	8.62	Up-regulated	409.7 (304.0)	96.1 (10.4)
11	Putative carbamoyl-phosphates synthase small subunit[i]	gi|15674869	39757.9/5.57	2	7.78	Up-regulated	132.4 (40.9)	0.0 (0.0)
12	Transcriptional repressor CodY[i, k]	gi|15675618	28633.7/4.85	1	7.31	Down-regulated	25.2 (48.0)	105.5 (194.9)
13	Putative formate-tetrahydrofolate ligase[i, k]	gi|15675177	59531.1/5.69	2	5.94	Up-regulated	189.1 (59.1)	0.0 (0.0)

a) Spot number refers to the proteins labelled in Fig. 1.
b) Accession ID of each protein is the GenInfo number in the NCBI protein database.
c) M_r and p*I* as calculated by PEPTIDEMASS [54].
d) Number of nonredundant peptides identified for each protein.
e) Percent amino acid coverage of entire protein.
f) Sum of the intensities of the image pixels inside the boundary of the spot $\times 10^6$.
g) The value shown is the median of the dataset with the range in parentheses.
h) Proteins identified according to an X_{corr} of >1.5 for single-charged peptides, >2.0 for double-charged peptides, and >2.5 for triple-charged peptides.
i) Proteins identified according to a ΔC_n >0.2 with at least two peptides successfully matched.
j) Proteins where the theoretical M_r and p*I* differed significantly from their experimental M_r and p*I*.
k) Proteins identified by TSSview™.

TSSView™. Protein spot 12, that did not satisfy the filter-parameters as only one peptide was successfully matched, was added to the list of identified proteins (Table 1) as the successfully matched peptide was shown using TSSview™ to come from a highly-conserved region of the transcription repressor CodY (Table 2). This single peptide is therefore highly significant for assigning function. Protein spot 13, that did not satisfy the filter parameters as it is from *S. mutans* and not *S. pyogenes*, was also added to the list of identified proteins (Table 1) as the two successfully matched peptides were shown using TSSview™ to come from highly-conserved regions of a putative formate-tetrahydrofolate ligase (Table 2). A BlastPAM30 search of these peptides confirmed that highly-identical peptides are present in the same protein from the SF370 strain of the M1 GAS serotype (Table 2). The remaining 2 proteins (spots 15 and 18) that did not satisfy the filter parameters could not be identified using TSSview™ and BlastPAM30 because the peptides from these proteins are located in conserved regions of proteins that are not present in any homologues of the SF370 strain of the M1 serotype of GAS.

3.3 Transport and biomass/energy production

We observed the up-regulation of proteins in HA-enriched media that are involved in (i) protein and DNA synthesis, *i.e.* 50 S ribosomal protein L10, the tRNA modification GTPase, a putative formate-tetrahydrofolate ligase and the carbamoyl-phosphate synthase small subunit, (ii) transport, *i.e.* the phosphohistidine-containing protein, HPr, a component of the phosphotransferase system (PTS) [15] and (iii) glycolysis, *i.e.* glyceraldehyde-3-phosphate dehydrogenase (GAPDH) (Table 1). GAS grown in HA-enriched media reached higher optical densities compared to that grown in control media,

Table 2. Sequence, charge state and filter parameter values for peptides from identified proteins

Spot no.	Protein	Peptides sequence	Charge	X_{corr}	ΔC_n
1	Putative phosphotransferase system phosphohistidine-containing protein	K.DFHIVAETGIHAR.P	3	5.070	0.55
		K.FASDITLDYK.G	2	3.198	0.694
		R.PATLLVQTASK.F	2	3.814	0.602
		K.FASDITLDYKGK.A	3	2.768	0.549
		K.FASDITLDYKGK.A	2	3.714	0.707
		R.PATLLVQTASK.F	1	2.189	0.541
2	Putative NADP-dependent glyceraldehyde-3-phosphate dehydrogenase	K.NIVAGAFGYSGQR.C	2	2.847	0.685
		K.AFGIAEQLEVGTVHLNNK.T	2	2.043	0.348
3	M1 protein precursor	K.ALELAIDQASQDYNR.A	2	1.245	0.338
		K.ARLENAMEVAGRDFK.R	2	1.118	0.366
4	Hypothetical protein SPy1143	R.AVILRQGAITK.E	2	1.535	0.255
		K.TGPLIGPSANLSGKASGR.V	2	1.603	0.428
5	Histidine protein kinase	K.DLVFLLRHFEK.A	2	1.000	0.581
		R.HFEKAVRNPLAHLIK.P	2	1.000	0.664
		R.REPFYFDDMNAIIK.K	2	0.527	0.357
6	tRNA modification GTPase	R.VDLTQAEAVMDIIR.A	2	1.945	0.250
		K.TGTIIDEVMVSVMLAPK.T	3	2.047	0.248
7	Hypothetical protein SPy1262	R.MTDLDVIEVNVK.V	2	4.304	0.600
		K.LVNTESVR.D	2	1.882	0.480
		R.GQLTYDDK.V	2	2.516	0.496
		R.GQLTYDDKVIEK.I	3	1.538	0.204
		K.SIVEEEVKR.M	2	1.292	0.433
		R.STSEFTSHQVENVK.A	2	3.243	0.557
		K.DLTSAIR.G	2	0.648	0.315
8	50S ribosomal protein L10	K.KAEQVELIAEK.M	2	4.582	0.565
		R.GLTVDQDTVLRR.S	2	2.122	0.422
		K.AAASIVIVDSR.G	2	3.835	0.693
		R.SLRESGVEFK.V	2	2.052	0.417
		K.VINDFTK.T	1	1.181	0.455
9	RopA	K.AVEVITSTASVK	2	2.313	0.645
		R.VKTNLVIEAIAK.A	2	2.409	0.574
		K.GKEWTLSAEVVTKPEVK.L	3	4.790	0.570
		R.GVITFTISQDK.I	2	1.750	0.523
		K.TNLVIEAIAK.A	2	2.674	0.361
10	Collagen-like surface protein	R.GETFAQGPVGPQGEK.G	2	1.490	0.222
		R.GEQGIQGKAGEKGER.G	2	1.604	0.229
11	Putative carbamoyl-phosphates synthase small subunit	K.GIPGISGIDTRALTK.I	2	1.439	0.252
		K.MTFGHRGFNHAVR.E	1	1.330	0.275
12	Transcriptional repressor CodY	K.DTYPGGLTTIAPIYGGGMR.L[a)]	2	1.826	0.361
13	Putative formate-tetrahydrofolate ligase	K.LVLVTAINPTPAGEGK.S[a, b)]	3	1.685	0.541
		K.STITIGLADALNKIGK.K[a, c)]	2	1.484	0.392

a) TSSview™ O/E ratio of ∞.
b) BlastPAM30 E value of 1.0×10^{-5}.
c) BlastPAM30 E value of 1.4×10^{-2}.

but there was not a statistically significant difference in the growth rates (data not shown). It was therefore not surprising that we observed the up-regulation of proteins involved in transport and biomass/energy production. These observations are in line with those of Voyich *et al.* [16] who also observed the up-regulation of many of these proteins following analysis of a phagocytic interaction of GAS with human polymorphonuclear leukocytes.

3.4 Virulence factors

3.4.1 LPXTG motif-containing proteins

3.4.1.1 General remarks

Many of the proteins that were shown to be up-regulated in HA-enriched media were LPXTG motif-containing cell-wall

anchored proteins. Such proteins have been shown to play important roles in pathogenesis, such as resistance to phagocytosis, adherence to plasma and extracellular matrix proteins, and degradation of host proteins [17]. As LPXTG motif-containing proteins are covalently attached to the cell wall, we were surprised to see these proteins in the cell-associated protein fraction, as we did not expect them to be released from the cell wall during protein preparation. However, the sonication step used during protein preparation may have released these proteins from the cell wall, as seen elsewhere by Dryla *et al.* [18]. Alternatively, it is possible that the precursor forms of these proteins were derived from the cytoplasm and not the cell wall.

3.4.1.2 M1 protein

We observed an increase in the intensity of a protein spot in HA-enriched media that was identified as the major GAS virulence factor, *i.e.* the M1 protein [7], an LPXTG motif-containing protein. M1 proteins play several roles in GAS pathogenesis. They protect the bacteria from phagocytic clearance by interfering with complement-mediated uptake [19], they can attach to keratinocytes and cause skin infection [20], and, during the throat infection, they promote bacteria-bacteria interaction following attachment to tonsillar epithelial cells [21]. Functional analysis by TSSview™ of the sequence of the M1 protein showed that this protein has several conserved regions (Fig. 3). As shown in Fig. 3, several conserved areas of M-like proteins were discerned. M-like proteins share a similar structure with M proteins and are encoded by *emmL* genes. Both M protein and M-like proteins have been shown to have the ability to bind to the Fc region of immunoglobulin [22–24], and this ability can be further subdivided into those proteins that bind immunoglobulin A or proteins that bind to immunoglobulin G [25–28]. The ability of M1 proteins to bind immunoglobulin plays an important role in anti-phagocytosis and interestingly the conserved regions that facilitate immunoglobulin binding were shown by TSSview™ to be present in the GAS M1 protein sequence (Fig. 3). In addition, the M1 protein sequence was shown to have several conserved regions for plasminogen-binding function (Fig. 3). This function has been shown to facilitate the accumulation of plasmin on the GAS surface, and can therefore help break host barriers and promote migration of GAS, which is important for the invasive GAS infections [29]. The M1 protein also possessed conserved regions for fibrinogen-binding activity (Fig. 3). This activity can precipitate fibrinogen that prevents the killing of GAS in host blood [30]. As all of these M1 protein activities have been experimentally determined, it confirms that TSSview™ is a powerful tool in aiding the functional characterisation of proteins.

3.4.1.3 Collagen-like surface protein

We also observed the up-regulation in HA enriched media of another LPXTG motif-containing cell-wall anchored proteins, *i.e.* a collagen-like surface protein. Lukomski and

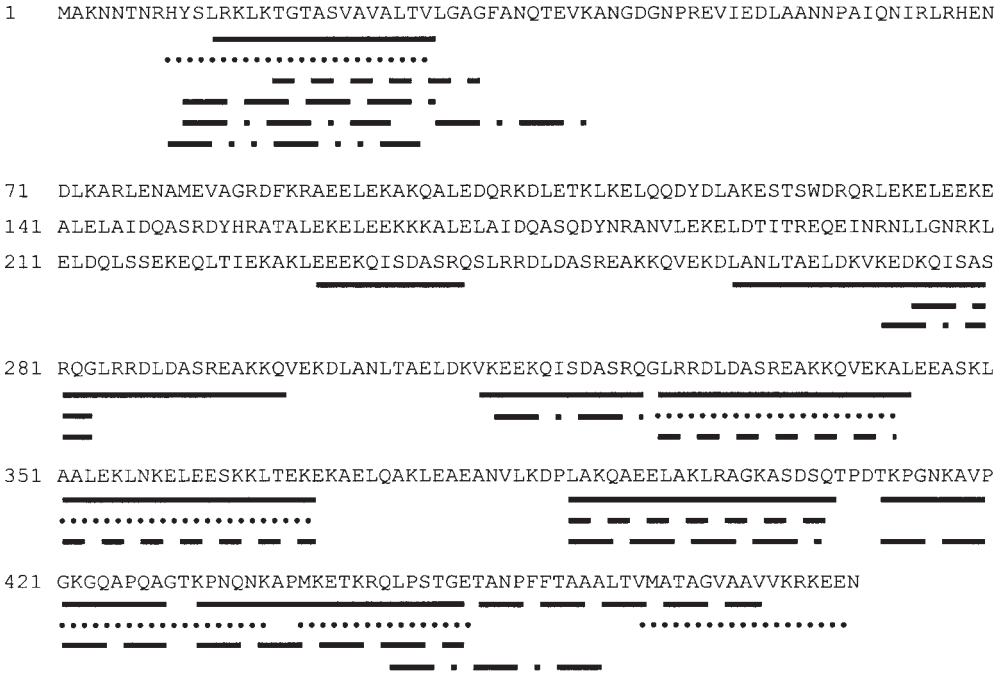


Figure 3. Identity of regions of unusual conservation in the M1 protein from the SF370 strain of the M1 serotype of GAS, as determined by TSSview™. Plasminogen-binding activity, ————; M-like proteins, — — —; Fibrinogen-binding activity, ·····; Immunoglobulin G binding activity, —— ——; Immunoglobulin A-binding activity, —··—; Fc-gamma receptor, —·—·.

Nakashima [31] demonstrated that a collagen-like surface protein mutant possessed a significantly reduced ability to adhere to human epithelial cells grown in culture and was significantly less pathogenic in a mouse model of soft tissue infection, which suggested that collagen-like surface protein is a GAS virulence factor that participates in the adherence of GAS to host cells and soft tissue pathology.

3.4.2 Other virulence factors

Interestingly, the glycolytic enzyme, GAPDH [32, 33] and the transport protein, HPr [34], that we have shown to be up-regulated in HA-enriched media (see Section 3.3), have been identified at the cell surface of pathogenic streptococci, with GAPDH shown to be implicated in pathogenesis [35–37]. It has been demonstrated that GAPDH was able to bind to fibronectin, lysozyme and cytoskeletal proteins and this multiple binding capacity of the GAPDH has be postulated to play a role in the colonization, internalization and the subsequent proliferation of GAS [35]. Additionally, GAPDH produced by GAS has been found to have adenosine diphosphate-ribosylating activity [35]. Since adenosine diphosphate-ribosylation is an important component of intracellular signalling events, GAPDH has been shown to have a role in mediating cell-to-cell communication between GAS and pharyngeal cells [36]. Recently, it has also been found that GAPDH helped GAS to bind to the complement component, C5a [38]. Interestingly, two other proteins, that were up-regulated in HA-enriched media, have been shown to be involved in signal transduction, the transport protein, HPr [37], and histidine protein kinase, which works as a transmembrane sensor detecting extracellular environmental changes to regulate the expression of several GAS virulence factors to enhance survival [39, 40]. These enzymes may therefore be important in facilitating the sensing of the external environment by GAS during infection. However, it should be noted that the sum of surface and cytosolic forms of the enzymes were analysed in this study.

3.5 Down-regulation of two regulators of virulence factors

A regulator of streptococcal virulence factors, RopA, was down-regulated in HA-enriched media (Table 1). RopA is a regulator of streptococcal pyrogenic exotoxin B (SpeB) and assists SpeB in translocation *via* the secretory pathway and functions as a molecular chaperone to achieve an active conformation of SpeB [41]. The down-regulation of RopA suggests that the level of expression of SpeB may have decreased in our system, yet we have not identified SpeB in our studies, as it is secreted and we have only analysed cell-associated proteins. However, the function of SpeB is uncertain, as it can hydrolyze M proteins and C5a peptidase [42], which are important in anti-phagocytosis. In fact, several studies have shown that increased expression of SpeB resulted in decreased virulence of GAS [43, 44]. Both Kansal *et al.* [45] and Raeder *et al.* [42] reported the inverse relationship between SpeB production and expression of M protein, and Kansal *et al.* [45] hypothesized that the reduced expression or activity of SpeB may be advantageous to the bacteria in cases of streptococcal toxic shock syndrome.

We also observed the down regulation of a transcriptional repressor, CodY, as was also seen by Graham *et al.* [46] when wild-type GAS was exposed to human blood. Interestingly, Malke *et al.* [47] have shown that CodY is a growth phase-dependent positive transcriptional regulator of Mga, a positive multiple regulator, and CovR, which is a member of the two-component global regulatory system CovRS. The positive regulation of Mga and CovR by CodY positively affect the expression of six virulence factors (hyaluronate synthase A, immunoglobulin G degrading endopeptidase, pyrogenic exotoxin H, C5a peptidase, a cell surface proteinase, and a collagen-like surface protein). However, the same report showed that the regulation of Mga and CovR by CodY has been shown to occur at the mid-exponential phase of cell growth [47], while our cells were harvested at early stationary phase. Interestingly, at early stationary phase CodY has also been shown to act as a guanosine 5-triphosphate-binding protein [48] that senses the intracellular guanosine 5-triphosphate concentration as an indicator of nutritional limitations. It is therefore possible that the observed down-regulation of CodY is due to a reduction in the levels of guanosine 5-triphosphate in our system at early stationary phase.

It should be noted that in addition to the two streptococcal regulators we observed, GAS has a series of global regulators of virulence factors [8]. The previously mentioned CovR, for example is involved in the regulation of the synthesis of the streptococcal HA capsule [49]. Although we did not detected altered expression of CovR in HA-enriched media, further post-genomic analyses will add to our understanding of the contribution of carbohydrate metabolism to GAS pathogenesis.

3.6 Hypothetical proteins

TurboSEQUEST™ identified two hypothetical proteins, the functions of which have not been annotated in the genome database. By analysis of these proteins using TSSview™, we found that hypothetical protein SPy1143 has conserved regions that are also present in SUA5 protein (Table 3). SUA5 is required for the normal growth of yeast cells [50], and similar sequence have been found in the M28 GAS serotype [51]. However, the function of SUA5 is not yet known, therefore it is possible that this protein, due to its up-regulation in our system, is involved in GAS pathogenesis. Another hypothetical protein, SPy1262 has been shown by TSSview™ to include some of the conserved regions of the general stress protein, Gls24 (Table 3). This protein, in *Enterococcus faecalis*, has been found to be implicated in virulence as well as stress response [52, 53] and may therefore have a similar role in GAS. The observed up-regulation of this protein is consistent

Table 3. Functional analysis of two hypothetical proteins using TSSview™

Hypothetical protein	Functionality	Peptide sequence	Conserved regions
Hypothetical protein SPy1143	SUA5	VLPTETVYGLFAKALDEKAV	15–34
	Necessary for normal growth	AFLPGPLTIILKANDQVP	76–93
		TGPLIGPSANLSG	120–132
Hypothetical protein SPy1262	Gls24	DLTSAIRGQLTYDDKVIEK	13–30
	General stress protein	GVNVEVGKKQVAVDL	66–80

with the result of the transcriptome study performed by Graham *et al.* [46] who found that SPy1262 was up-regulated in the late stage of GAS exposure to human blood.

4 Concluding remarks

We have used 2-DE in conjunction with static nanospray MS to identify proteins from a GAS M1 serotype that have altered expression as a result of propagation in HA-enriched media. Some of these proteins are major virulence factors of GAS, such as the M1 protein and a collagen-like surface protein, some are regulators of virulence factors, such as RopA, and some are proteins that are not commonly associated with virulence, *e.g.* the glycolytic enzyme GAPDH, but have been shown to play important roles in streptococcal pathogenesis. Interestingly, we also identified a hypothetical protein, which may be involved in GAS pathogenesis. Unlike many models of streptococcal pathogenicity, we have shown that GAS grown in HA-enriched media is readily amenable to proteomic analysis due to the relatively large amounts of biomass propagated. The fact that we have observed the up-regulation of several recognised virulence factors indicates that this simple biological system reflects several elements of GAS pathogenesis in humans and is therefore a valuable tool in the identification of potential virulence factors.

We would like to thank Iain Sutcliffe (Northumbria University) for his helpful comments on the manuscript and John Perry (Freeman Hospital, Newcastle upon Tyne) for his help with storage of S. pyogenes SF370. Meng Zhang is a recipient of the Universities UK Overseas Research Students Awards Scheme scholarship.

5 References

[1] Mims, C., Dockrell, H., Goering, R., Roitt, I. *et al.*, *Medical Microbiology*, Mosby, USA 1998.

[2] Carapetis, J. R., Steer, A. C., Mulholland, E. K., Weber, M., *Lancet Infect. Dis.* 2005, *5*, 685–694.

[3] Hynes, W. L., Ferretti, J. J., *Infect. Immun.* 1989, *57*, 533–539.

[4] Laurent, T. C., Fraser, J. R., *FASEB J.* 1992, *6*, 2397–2404.

[5] Smith, N. L., Taylor, E. J., Lindsay, A. M., Charnock, S. J. *et al.*, *Proc. Natl. Acad. Sci. USA* 2005, *102*, 17652–17657.

[6] Hynes, W. L., Walton, S. L., *Microbiol. Lett.* 2000, *183*, 201–207.

[7] Cunningham, M. W., *Clin. Microbiol. Rev.* 2000, *13*, 470–511.

[8] Graham, M. R., Smoot, L. M., Lei, B., Musser, J. M., *Curr. Opin. Microbiol.* 2001, *4*, 65–70.

[9] Lei, B., Mackie, S., Lukomski, S., Musser, J. M.*Infect. Immun.* 2000, *68*, 6807–6818.

[10] Thongboonkerd, V., Luengpailin, J., Cao, J., Pierce, W. M. *et al.*, *J. Biol. Chem.* 2002, *277*, 16599–16605.

[11] Chaussee, M. A., Callegari, E. A., Chaussee, M. S., *J. Bacteriol.* 2004, *186*, 7091–7099.

[12] Ferretti, J. J., McShan, W. M., Ajdic, D., Savic, D. J. *et al.*, *Proc. Natl. Acad. Sci. USA*2001, *98*, 4658–4663.

[13] Houtman, R., Krijgsveld, J., Kool, M., Romijn, E. P. *et al.*, *Proteomics* 2003, *3*, 2008–2018.

[14] Eng, J. K., McCormack, A. L., Yates, J. R.*J. Am. Soc. Mass Spectrom.* 1994, *5*, 976–989.

[15] LiCalsi, C., Crocenzi, T. S., Freire, E., Roseman, S., *J. Biol. Chem.* 1991, *266*, 19519–19527.

[16] Voyich, J. M., Sturdevant, D. E., Braughton, K. R., Kobayashi, S. D. *et al.*, *Proc. Natl. Acad. Sci. USA* 2003, *100*, 1996–2001.

[17] Hynes, W., *Front. Biosci.* 2004, *9*, 3399–3433.

[18] Dryla, A., Prustomersky, S., Gelbmann, D., Hanner, M. *et al.*, *Clin. Diagn. Lab. Immunol.* 2005, *12*, 387–398.

[19] Husmann, L. K., Yung, D. L., Hollingshead, S. K., Scott, J. R., *Infect. Immun.* 1997, *65*, 1422–1430.

[20] Okada, N., Pentland, A. P., Falk, P., Caparon, M. G., *J. Clin. Invest.* 1994, *94*, 965–977.

[21] Caparon, M. G., Stephens, D. S., Olsen, A., Scott, J. R., *Infect. Immun.* 1991, *59*, 1811–1817.

[22] Guyre, P. M., Campbell, A. S., Kniffin, W. D., Fanger, M. W.*J. Clin. Invest.* 1990, *86*, 1892–1896.

[23] La Penta, D., Zhang, X. P., Cleary, P. P., *Mol. Microbiol.* 1994, *12*, 873–879.

[24] Sjobring, U., Ringdahl, U., Ruggeri, Z. M., *Blood* 2002, *100*, 4470–4477.

[25] Frithz, E., Heden, L. O., Lindahl, G., *Mol. Microbiol.* 1989, *3*, 1111–1119.

[26] Gomi, H., Hozumi, T., Hattori, S., Tagawa, C. *et al.*, *J. Immunol.* 1990, *144*, 4046–4052.

[27] Burova, L. A., Nagornev, V. A., Pigarevsky, P. V., Gladilina, M. M. *et al.*, *Apmis* 2005, *113*, 21–30.

[28] Podbielski, A., Schnitzler, N., Beyhs, P., Boyle, M. D., *Mol. Microbiol.* 1996, *19*, 429–441.

[29] Berge, A., Sjobring, U.*J. Biol. Chem.* 1993, *268*, 25417–25424.

[30] Johansson, H. M., Morgelin, M., Frick, I. M. *Microbiology* 2004, *150*, 4211–4221.

[31] Lukomski, S., Nakashima, K., Abdi, I., Cipriano, V. J. *et al.*, *Infect. Immun.* 2000, *68*, 6542–6553.

[32] Pancholi, V., Fischetti, V. A., *J. Exp. Med.* 1992, *176*, 415–426.

[33] Wilkins, J. C., Homer, K. A., Beighton, D., *Appl. Environ. Microbiol.* 2002, *68*, 2382–2390.

[34] Dixon, S., Haswell, M., Harrington, D., Sutcliffe, I. C., *Syst. Appl. Microbiol.* 2001, *24*, 486–489.

[35] Pancholi, V., Fischetti, V. A., *Proc. Natl. Acad. Sci. USA* 1993, *90*, 8154–8158.

[36] Pancholi, V., Fischetti, V. A., *Adv. Exp. Med. Biol.* 1997, *418*, 499–504.

[37] Wilkins, J. C., Beighton, D., Homer, K. A., *Appl. Environ. Microbiol.* 2003, *69*, 5290–5296.

[38] Terao, Y., Yamaguchi, M., Hamada, S., Kawabata, S., *J. Biol. Chem.* 2006, *281*, 14215–14223.

[39] Senadheera, M. D., Guggenheim, B., Spatafora, G. A., Huang, Y. C. *et al.*, *J. Bacterial.* 2005, *187*, 4064–4076.

[40] Musser, J. M., DeLeo, F. R., *Am. J. Pathol.* 2005, *167*, 1461–1472.

[41] Collin, M., Olsen, A., *Infect. Immun.* 2003, *71*, 2983–2992.

[42] Raeder, R., Woischnik, M., Podbielski, A., Boyle, M. D., *Res. Microbiol.* 1998, *149*, 539–548.

[43] Raeder, R., Harokopakis, E., Hollingshead, S., Boyle, M. D., *Infect. Immun.* 2000, *68*, 744–751.

[44] Ashbaugh, C. D., Warren, H. B., Carey, V. J., Wessels, M. R., *J. Clin. Invest.* 1998, *102*, 550–560.

[45] Kansal, R. G., McGeer, A., Low, D. E., Norrby-Teglund, A., Kotb, M., *Infect. Immun.* 2000, *68*, 6362–6369.

[46] Graham, M. R., Virtaneva, K., Porcella, S. F., Barry, W. T. *et al.*, *Am. J. Pathol.* 2005, *166*, 455–465.

[47] Malke, H., Steiner, K., McShan, W. M., Ferretti, J. J., *Int. J. Med. Microbiol.* 2006, *296*, 259–275.

[48] Ratnayake-Lecamwasam, M., Serror, P., Wong, K. W., Sonenshein, A. L., *Genes Dev* 2001, *15*, 1093–1103.

[49] Levin, J. C., Wessels, M. R., *Mol. Microbiol.* 1998, *30*, 209–219.

[50] Na, J. G., Pinto, I., Hampsey, M., *Genetics* 1992, *131*, 791–801.

[51] Green, N. M., Zhang, S., Porcella, S. F., Nagiec, M. J. *et al.*, *J. Infect. Dis.* 2005, *192*, 760–770.

[52] Teng, F., Nannini, E. C., Murray, B. E., *J. Infect. Dis.* 2005, *191*, 472–480.

[53] Nannini, E. C., Teng, F., Singh, K. V., Murray, B. E., *Infect. Immun.* 2005, *73*, 7772–7774.

[54] Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S. *et al.*, in: Walker, J. M. (Ed.), *The Proteomics Protocols Handbook*, Humana Press, Totowa, NJ 2005, pp. 571–607.

