Identification and Validation of Vaccine Candidate Proteins by 2-Dimensional Mass Spectrometry Analysis of Group A Streptococcus

Lori Stansberry, Angela Payne, Julie Skinner, Loren Schultz, Jon Heinrichs, Joseph Joyce, Eberhard Durr

Merck Research Laboratories, West Point, Pennsylvania

Overview

- Surface proteins of bacteria are involved in the adhesion and invasion of host cells
- We characterized the surface proteome of group A Streptococcus using multidimensional protein identification technology.
- The relative abundance of identified proteins was determined using peak height calculations from ion intensity.
- The surface accessibility of identified proteins is determined by a comparative analysis of surface digests and total protein digests, as well as PSORTB, a cell localization prediction algorithm.
- 302 proteins were identified in a surface protein digestion of the M1 Strain. A total of 311 proteins were identified in a surface protein digestion of the M3 strain.
- M1 Protein type 1 (SPy2018) was identified in both strains as the highest abundance protein, a fact known to be true.
- These proteomics results are validated by literature, PSortB, and biological data.
- Proteins identified by this method include the most protective antigens in literature.
- We identified 12 proteins of interest for potential vaccine targets.

Introduction

Surface proteins of bacteria are involved in adhesion and invasion of host cells. Since surface proteins likely interact with the host immune system, they are important for vaccine discovery research.

We employed on-line multidimensional protein identification technology to characterize the surface proteome of group A Streptococcus (GAS). Surface proteins were selectively digested with trypsin after mutanolysin and lysozyme treatments of intact bacteria. The obtained surface proteome was compared to the whole cell proteome obtained by analyzing a total cell lysate. PSortB, a protein prediction algorithm for subcellular localization, was applied to demonstrate enrichment of proteins predicted to be surface localized. We show that identified surface proteins include 2 protective antigens reported in literature and 10 newly discovered potential vaccine candidates.

Methods

GAS total protein digest:

GAS, Strain M1 was lysed using pressure cycling technology (Pressure Biosciences Barocycler NEP2320) with a detergent free buffer included in PB kit. Protein was taken up in 8M Urea and reduced with 10mM tris (2-carboxyethyl)phosphine. The reduction was followed by alkylation with 25 mM iodoacetamide. Reduced and alkylated sample was digested with LysC (1:50 wt/wt; 4h ; 37°C). A second digest was performed with Trypsin (1:100 wt/wt; 14h ; 37°C) after diluting the LysC digest to 1.5M Urea. Peptides were cleaned up with a C18 solid phase extraction cartridge.

GAS surface digest:

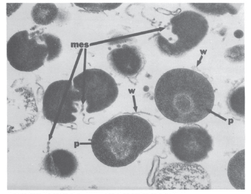
Intact GAS cells were brought up in 20% sucrose in Tris-EDTA pH8.0. They were treated with 300U/ml mutanolysin and 3 mg/ml lysozyme, followed by trypsin digestion (1:50 wt/wt; 0.5hr; 37°C). A reduction step was performed with 100mM tris(2-carboxyethyl)phosphine (20min; 50°C) followed by an alkylation step with 25mM iodoacetamide. Released peptides were recovered in the supernatant after a 3500xg spin for 10 minutes. Peptides were cleaned up with a C18 solid phase extraction cartridge.

MS/MS Measurements:

Data was acquired on a LTQ ion trap instrument (Thermo Fisher) operated in profile mode. Using Dionex Ultimate 3000, 5µg of peptides were separated on-line by strong cation exchanger chromatography followed by a C4 trap column, then C18 reversed phase nano-LC MS/MS. Peptide fractions were displaced from the cation exchanger resin using 10 discrete salt steps. The analysis time for one sample was approximately 20 hours. MS/MS scans were performed on the top 7 pre-cursor ions in reverse order.

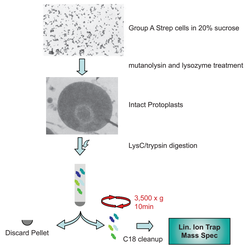
Our goal was to enrich for proteins found on the surface of GAS cells. It was important to stabilize the cells with 20% sucrose to prevent lysing. Using lysozyme and mutanolysin, we separated the cell wall from the intact protoplast to expose the entire surface of GAS cells.

Methods Cont.

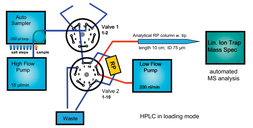


Picture from Calandra et al, 1975[1] depicting protoplasts by CEM
mes: Mesosomes
w: Cell Fragments
p: Intact Protoplasts

Proteolytic digest of surface exposed proteins



Chromatographic separation and MS² analysis:



Data analysis:

Data analysis was performed using the Elucidator Software Package (Rosetta Biosoft). In brief: tandem mass spectra were analyzed with the search algorithm, SEQUEST.

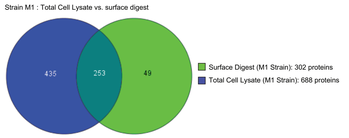
Label free quantitation was performed by calculating the peak height of the extracted ion chromatogram (XIC) for each identified peptide in each salt step. Peptide XIC values were obtained by summation of XIC values of all charge states in each salt step.

Protein XIC values were obtained by summation of underlying peptide XIC values. Average protein XIC values were automatically calculated across all technical replicates. The PSORTb algorithm was applied to predict the subcellular location of identified proteins.

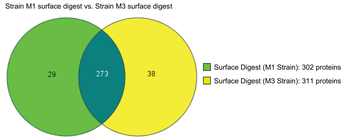
Proteomics results were compared to antigens of interest reported in literature for validation.

Results

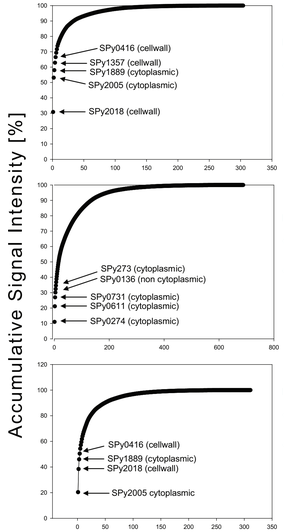
Overall comparison of the proteins identified in total cell lysate or surface GAS digestion



Overall comparison of the proteins identified in Strains M1 and M3 surface GAS analysis



Accumulative GAS protein abundance of the complete or surface GAS analysis



M1 Surface analysis:

- 9 proteins account for 75% of the total signal intensity
- Of these 9 proteins, 3 are designated as cellwall, 5 cytoplasmic, and 1 unknown

M1 Total Cell Lysate analysis:

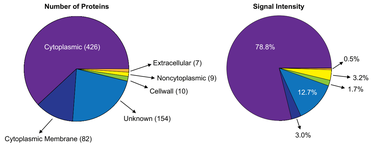
- 64 proteins account for 75% of the total signal intensity
- Of these 64 proteins, 82% were designated as cytoplasmic proteins by PSortB

M3 Surface analysis:

- 18 proteins account for 75% of the total signal intensity
- Of these 18 proteins, 10 are designated to be cytoplasmic proteins, 5 are of unknown localization, and 3 are cellwall proteins

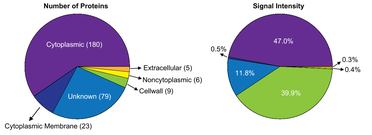
Proteins identified in the GAS total cell lysate

- 688 proteins were identified in the GAS total cell digest
- 426 proteins of the total cell lysate were identified as Cytoplasmic by PSortB
- These 426 cytoplasmic proteins accounted for 78.8% of the MS signal intensity.



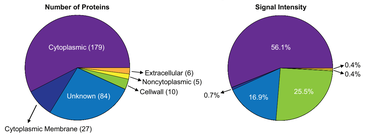
Proteins identified in the GAS M1 surface enrichment

- 302 proteins were identified in the GAS M1 strain surface enrichment
- A significant enrichment (as determined by MS signal intensity) of cellwall proteins was observed when compared to GAS total cell lysate
- The 9 cellwall proteins account for 39.9% of the total MS signal intensity

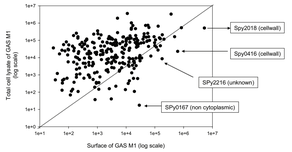


Proteins identified in the GAS M3 surface enrichment

- 311 proteins were identified in the GAS M3 strain surface enrichment
- A significant enrichment (as determined by MS signal intensity) of cellwall proteins was observed when compared to GAS total cell lysate
- The 10 cellwall proteins account for 25.5% of the total MS signal intensity

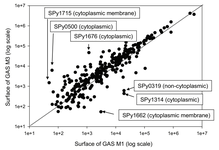


Logarithmic plot of MS signal intensities of total cell lysate and surface digest of GAS strain M1



- Significant deviations below the regression line are expected for proteins thatare found on the surface of GAS cells.

Logarithmic plot of MS signal intensities of surface digests of GAS strains M1 and M3



- Most identified proteins fall along the unity line
- The 6 proteins displayed on the graph showed large differences between MS signal intensities for M1 and M3 surface digestions.
- Significant deviations below the unity line are expected for proteins found on the surface of GAS cells.
- Most of the proteins found below the unity line are proteins not located in the cytoplasm
- The 4 proteins that deviate the furthest below the unity line are of particular interest because they are membrane associated proteins and in high abundance in the surface enriched samples.

TIGR ID	Protein Description (Streptococcus pyogenes SF370 serotype M1)	Spy Identifier	PSortB Cell Localization	% of Total MS signal intensity	Fold change (M1/total cell lysate)	Fold change (M3 surface/ total cell lysate)
NT01SP1648	CsbD-like family	SPy2005	Cytoplasmic	20.4%	Only identified in surface samples	
NT01SP1656	M1 protein precursor	SPy2018	Cellwall	18.0%	9.3x	6.3x
NT01SP0336	cell envelope proteinase (partial)	SPy0416	Cellwall	4.4%	24.5x	36.2x
NT01SP0140	streptolysin O precursor	SPy0167	Non-cytoplasmic	0.17%	1495x	3403x
NT01SP1118	GRAB precursor	SPy1357	Cellwall	2.3%	Only identified in surface samples	
NT01SP1041	CsbD-like family	SPy1261	Unknown	2.3%	Only identified in surface samples	
NT01SP0138	nicotine adenine dinucleotide glycohydrolase precursor	SPy0165	Unknown	0.6%	61.5x	159x
NT01SP1620	conserved hypothetical protein	SPy1971	Unknown	0.71%	Only identified in surface samples	
NT01SP1546	5-nucleotidase, lipoprotein e(P4) family	SPy1882	Unknown	0.02%	Only identified in surface samples	
NT01SP1653	c5a peptidase precursor (partial)	SPy2010	Cellwall	0.37%	Only identified in surface samples	
NT01SP1817	Unknown	SPy2216	Unknown	1.7%	18.9x	36.7x

- Proteins on this list were chosen because they were found only in the surface preparation or because of the large fold increase in the surface enrichment sample when compared to the total cell lysate. Strains M1 and M3 did not show any differences in the top proteins of interest.

SPy2018

- M1 protein Type 1
- Identified by PSortB as a cell wall protein.
- Independently identified as a protein of interest by collaborators using a method of biological screening of serum from convalescent human sera against GAS-specific bacterial surface display libraries.
- 9.3 fold increase from total cell lysate to surface protein enrichment
- Mouse polyclonal antibodies against the recombinant form of SPy2018 were incubated with GAS cells and found to have a 15.1 fold change between cells stained with immune sera vs. pre-immune sera using FACS. (Rodriguez-Ortega et al2)
- SPy2018 Is not a suitable target due to perceived potential cross-reactivity with human heart tissues

SPy2216

- Putative serine protease, DegP
- Identified as a membrane associated protein by pSortB.
- Mouse polyclonal antibodies against the recombinant form of SPy2216 were incubated with GAS cells and found to have a 4.2 fold change between cells stained with immune sera vs. pre-immune sera using FACS. (Rodriguez-Ortega et al2)

SPy1357

- Protein GRAB (protein G-related alpha 2M-binding protein)
- Identified by PSortB as a cell wall protein
- Independently identified as a protein of interest by collaborators using a method of biological screening of serum from convalescent human sera against GAS-specific bacterial surface display libraries.
- Mouse polyclonal antibodies against the recombinant form of SPy1357 were incubated with GAS cells and found to have a 8.9 fold change between cells stained with immune sera vs. pre-immune sera using FACS. (Rodriguez-Ortega et al2)

SPy0167

- Streptolysin O precursor.
- 1460 fold increase from GAS M1 total cell lysate to M1 surface protein enriched.
- Identified by PSortB as a secreted protein
- Mouse polyclonal antibodies against the recombinant form of SPy0167 were incubated with GAS cells and found to have a 2.8 fold change between cells stained with immune sera vs. pre-immune sera using FACS. (Rodriguez-Ortega et al2)

SPy0416

- Putative cell envelope proteinase
- Identified by PSortB as a cellwall protein
- Independently identified as a protein of interest by collaborators using a method of biological screening of serum from convalescent human sera against GAS-specific bacterial surface display libraries.
- Mouse polyclonal antibodies against the recombinant form of SPy0416 were incubated with GAS cells and found to have a 4.8 fold change between cells stained with immune sera vs. pre-immune sera using FACS. (Rodriguez-Ortega et al2)
- SPy0416 was found to be protective in mice (Rodriguez-Ortega et al[2])

- SPy2005 was the only Cytoplasmic protein identified as highly abundant surface enriched
- Four cellwall proteins were enriched in the surface digestion, as expected: SPy2018, SPy0416, SPy1357, and SPy2010
- SPy0167 was the only Non-cytoplasmic protein identified as highly abundant surface enriched
- There were 5 proteins of unknown localization by PSortB that could be of interest. SPy1261, SPy0165, SPy1971, SPy1882, and SPy2216.

Conclusions

- The total cell lysate of GAS M1 and the surface proteome of GAS M1 and M3 have been determined by tandem mass spectrometry.
- A total of 747 proteins have been identified with a predicted false positive rate of 0.5%.
- A total of 311 proteins were identified in the surface digestion of M3, and 302 in the surface digestion of M1.
- Using lysozyme and mutanolysin enzymes, we were able to remove the cellwall and expose the protoplast, enriching for surface proteins of GAS
- 12 Antigens of interest for potential vaccine targets were identified by the described method.
- M1 protein type 1 (SPy2018) is the most abundant protein in the surface digestion of GAS M1 and M3, and accounts for approximately 30% and 18% of the total protein signal intensity, respectively.
- SPy0416 was found to be highly abundant in both M1 and M3 surface digestion samples. This cellwall protein is of particular interest, because it has been identified in literature to be protective against GAS in mice.
- SPy0167, SPy1357, and SPy2216 are also of particular interest because they are highly abundant proteins with corroborating published FACS data and/or biological data.
- Five proteins classified as “unknown” localization by PSortB could be of particular interest for novel vaccine targets.

References

(1) Calandra et a.. 1975. Biochim Biophys Acta. 409: 1-25

(2) Rodriguez-Ortega et al. 2006 Nature Biotechnology. 24(2):191

Copyright ©2011 Merck & Co., Inc., Whitehouse Station, New Jersey, USA, All Rights Reserved